

## BAB II

### TINJAUAN PUSTAKA DAN KERANGKA TEORI

#### A. Tinjauan Pustaka

Berkaitan dengan penelitian ini, maka penulis akan memaparkan tinjauan pustaka terhadap beberapa penelitian yang membahas mengenai analisis butir item soal dalam evaluasi kegiatan pembelajaran, antara lain sebagai berikut:

Penelitian pertama yang serupa disusun oleh Tri Yuli Astuti, Mahasiswa Jurusan Pendidikan Agama Islam Fakultas Agama Islam Universitas Muhammadiyah Yogyakarta, dengan judul Validitas dan Reliabilitas Mata Pelajaran Pendidikan Agama Islam Kelas V Di SD Negeri Bintaran Bantul. Hasil penelitian ini menunjukkan bahwa butir-butir soal tes mata pelajaran Pendidikan Agama Islam Kelas V semester satu tahun pelajaran 2013/2014 di SD N Bintaran Bantul dapat dinyatakan sebagai berikut, ditinjau dari validitas itemnya pada tes objektif 23% valid dan 77% tidak valid. Ditinjau dari reliabilitasnya tes bentuk objektif dengan menggunakan rumus *C. Hyot* besar koefisien reliabilitas tes yaitu sebesar 0,6047, jadi dapat dinyatakan tes tersebut tidak reliabel. Sedangkan tes subjektif dengan menggunakan rumus *Alpha* besar koefisien reliabilitas tes yaitu sebesar 0,55125, jadi dapat dinyatakan tes tersebut tidak reliabel. Ditinjau dari validitas logis pada tes bentuk objektif dari 30 soal terdapat 7 soal atau sebesar 23% yang tidak memiliki validitas logis, pada tes bentuk

subyektif semua soal belum memiliki validitas logis ditinjau dari segi konstruksi.

Penelitian kedua disusun oleh Minarni, Mahasiswa Jurusan Pendidikan Agama Islam Fakultas Agama Islam Universitas Muhammadiyah Yogyakarta, dengan judul Analisis Butir Soal Ulangan Akhir Semester Gasal Mata Pelajaran Aqidah Akhlak di MI Muhammadiyah Al-Islam Kaliangkrik Magelang. Hasil penelitian ini menunjukkan bahwa validitas empirik pada soal pilihan ganda kelas IV, V, dan VI mempunyai tingkat kesukaran yang rendah, daya beda yang tinggi dan pengecoh yang efektif atau berfungsi dibawah 50 % serta reliabilitas dibawah atau lebih kecil dari 0,70 sehingga dapat dikatakan bahwa soal pilihan ganda kelas IV, V, dan VI belum memiliki reliabilitas. Kemudian validitas logis pada soal pilihan ganda, soal isian singkat maupun soal uraian kelas IV, V dan VI memiliki validitas logis diatas 50% ditinjau dari segi materi, konstruksi, dan bahasa yang baik.

Penelitian ketiga disusun oleh Dwi Fajar Sulistyanyingsih, Mahasiswa Jurusan Pendidikan Agama Islam Fakultas Agama Islam Universitas Muhammadiyah Yogyakarta, dengan judul Kualitas Tes Ujian Akhir Semester (UAS) Gasal Kelas VIII Mata Pelajaran Pendidikan Agama Islam (PAI) Tahun Ajaran 2013/2014 di SMP Se-Kecamatan Pandak (Analisis Butir Soal). Hasil penelitian ini menunjukkan bahwa tingkat kesukaran butir soal 86,7 % termasuk dalam kategori sedang atau sebanyak 39 butir soal dan 13,3 % termasuk dalam kategori mudah atau sebanyak 6 butir soal. 26,7 % butir soal memiliki daya beda sedang, 62,2 % butir soal memiliki daya beda

rendah dan 11,1 % butir soal tidak mampu membedakan kemampuan siswa. 77,8 % butir soal yang distraktornya menjalankan fungsinya dengan baik, 8,9 % butir soal yang distraktornya perlu untuk direvisi dan 13,3 % butir soal distraktor tidak berfungsi dengan baik sehingga harus diganti. 80 % butir soal yang diterima, 8,9 % butir soal yang direvisi, dan 11,1 % butir soal yang dibuang/diganti. Besarnya reliabilitas tes sebesar 0,724 termasuk dalam kategori rendah. Dilihat dari validitas konstruksinya yaitu kesesuaian dengan kisi-kisi sebanyak 45 butir soal sesuai dengan kompetensi dasar dan indikatornya. Sedangkan penulisan dan pemakaian kalimat yang efektif terdapat beberapa kesalahan yaitu butir soal nomor 1,11,12,19,21,32.

Penelitian terakhir disusun oleh Ahmad Fikri Aji Pamilu, Mahasiswa Jurusan Pendidikan Agama Islam, Fakultas Tarbiyah dan Keguruan, Universitas Islam Negeri Sunan Kalijaga Yogyakarta, dengan judul Analisis Butir Soal Pada Ulangan Akhir Semester Gasal Mata Pelajaran Aqidah Akhlak Kelas X MAN Yogyakarta II Tahun Pelajaran 2013/2014. Hasil penelitiannya menunjukkan bahwa validitas soal Ulangan Akhir Semester Gasal termasuk pada kategori baik, karena 22 butir soal dinyatakan valid (57,7%). Nilai reliabilitas untuk soal pilihan ganda diketahui koefisien reliabilitasnya ( $r_{11}$ ) sebesar 0,469 sedangkan untuk soal uraian diketahui koefisien reliabilitas ( $r_{11}$ ) sebesar 0,112 sehingga untuk soal UAS tersebut dinyatakan belum reliabel. Tingkat kesukaran soal termasuk pada kategori soal yang mudah, karena 23 butir soal termasuk kategori mudah (51%). Daya pembeda butir soal yang termasuk buruk sebanyak 21 butir soal, 10 butir soal

termasuk sedang, 10 butir soal termasuk baik, dan 4 butir soal termasuk dalam kategori negatif. Fungsi pengecoh belum dikatakan baik karena 40,66% distraktor yang mampu menjalankan fungsinya dengan baik.

Berdasarkan penelitian-penelitian diatas, dapat disimpulkan persamaan penelitian terletak dalam melakukan pengukuran validitas yang meliputi: validitas logis dan validitas butir instrumen serta reliabilitas terhadap sebuah instrumen tes. Sedangkan perbedaan penelitian terletak pada tes mata pelajaran yang akan dianalisis, tempat penelitian, serta jenjang pendidikan.

## **B. Kerangka Teori**

### **1. Validitas**

#### **a. Pengertian Validitas**

Validitas dapat diartikan sebagai ketepatan penafsiran yang dihasilkan dari skor tes atau instrumen evaluasi (Pramono, 2014: 224). Instrumen evaluasi dikatakan valid apabila instrumen tersebut dapat dengan tepat mengukur apa yang hendak diukur. Dengan kata lain validitas berkaitan dengan “ketepatan” dengan alat ukur (Widoyoko, 2015: 128).

Dalam bahasa Indonesia istilah “valid” sangat sukar dicari kata penggantinya sehingga diganti menjadi istilah “sahih” sedangkan “validitas” diganti menjadi istilah “kesahihan”, namun ada yang

menterjemahkan berbeda yaitu “tepat” dan “ketepatan” atau “cermat” dan “kecermatan” (Kamus Besar Bahasa Indonesia).

## **b. Macam-Macam Validitas**

Validitas secara garis besar dapat dibedakan menjadi dua macam validitas, yaitu validitas internal (*internal validity*) dan validitas eksternal (*external validity*) (Widoyoko, 2015: 129).

### **1) Validitas Internal**

Validitas Internal (*internal validity*) disebut juga dengan validitas logis (*logical validity*). Istilah “validitas logis” mengandung kata “logis” berasal dari kata “logika” yang berarti penalaran atau rasional. Dengan makna demikian maka validitas logis untuk sebuah instrumen evaluasi menunjuk pada kondisi sebuah instrumen yang memenuhi persyaratan valid yang berdasarkan hasil penalaran dan rasional (Arikunto, 2012: 80). Validitas logis untuk sebuah instrumen evaluasi mengarah kepada kondisi bagi sebuah instrumen yang memenuhi persyaratan valid berdasarkan hasil penalaran. Kondisi valid tersebut dipandang terpenuhi karena instrumen yang bersangkutan sudah dirancang dengan baik, mengikuti teori dan ketentuan yang ada. (Pramono, 2015: 225)

Terdapat dua macam validitas logis yang dapat dicapai oleh sebuah instrumen, yaitu: validitas isi (*content validity*) dan validitas konstruk (*construct validity*).

a) Validitas Isi (*content validity*)

Menurut Elis dan Rusdiana (2015: 169) validitas isi adalah kesesuaian butir-butir soal dalam tes dengan deskripsi bahan yang diajarkan. Validitas isi dari suatu tes hasil belajar adalah validitas yang diperoleh setelah dilakukan analisis, penelusuran atau pengujian terhadap isi yang terkandung dalam tes hasil belajar tersebut. Sudijono (2015: 164) menyatakan validitas isi adalah validitas yang dilihat dari segi isi tes itu sendiri sebagai alat pengukur hasil belajar telah dapat mewakili secara *representatif* terhadap keseluruhan materi atau bahan pelajaran yang seharusnya diujikan.

Untuk menguji validitas isi dari suatu instrumen tes dapat dilakukan dengan membandingkan antara isi instrumen dengan materi pelajaran yang telah diajarkan. Penyusunan instrumen tes harus disusun berdasarkan materi pelajaran yang telah dipelajari oleh siswa. Dengan demikian instrumen tes tersebut dapat dinyatakan telah memiliki validitas isi.

b) Validitas Konstruksi (*construct validity*)

Secara *etimologis*, kata “konstruksi” mengandung arti susunan, kerangka atau rekaan. Adapun secara *terminologis*, yaitu suatu tes hasil belajar dapat dinyatakan sebagai tes yang telah memiliki validitas konstruksi, apabila tes hasil belajar tersebut ditinjau dari segi susunan, kerangka, atau rekaannya

telah dapat mencerminkan suatu konstruksi dalam teori psikologis (Sudijono, 2015: 166). Teori psikologis yang dimaksud adalah teori yang digunakan dalam mengukur aspek-aspek berfikir siswa. Benjamin S. Bloom telah merincinya dalam tiga aspek kejiwaan yaitu aspek *kognitif*, aspek *afektif*, dan aspek *psikomotorik*.

Untuk menguji validitas konstruk dari suatu instrumen tes yaitu dengan melakukan analisis secara rasional serta menunjukkan bahwa aspek-aspek berfikir yang diungkap melalui butir-butir soal sudah tepat mencerminkan aspek-aspek berfikir yang menjadi tujuan instruksional (Arikunto, 2012: 83). Dengan demikian, maka tes tersebut dapat dikatakan telah memiliki validitas konstruk.

## 2) Validitas Eksternal

Validitas Eksternal (*external validity*) disebut juga dengan validitas empirik (*empirical validity*). Yang dimaksud dengan validitas empirik adalah validitas yang diperoleh atas dasar pengamatan lapangan. Suatu tes hasil belajar dapat dikatakan mempunyai validitas empirik apabila hasil analisis yang dilakukan dalam pengamatan lapangan terbukti bahwa tes hasil belajar itu dengan secara tepat telah dapat mengukur hasil belajar yang harus diukur lewat tes hasil belajar tersebut (Sudijono, 2015: 168).

Menurut Arifin (2014: 249) menyatakan bahwa validitas ini biasanya menggunakan validitas statistik, yaitu analisis korelasi. Hal ini disebabkan validitas empirik mencari hubungan antara skor tes dengan suatu kriteria tertentu yang menjadi tolok ukur di luar tes yang bersangkutan. Namun, kriteria itu harus relevan dengan apa yang hendak diukur.

Untuk dapat menentukan apakah tes hasil belajar sudah memiliki validitas empirik ataukah belum, dapat dilakukan penelusuran dari tiga segi, yaitu dari segi daya ketepatan meramalnya, daya ketepatan bandingannya, serta daya ketepatan yang dimiliki butir instrumen.

#### a) Validitas Ramalan

Meramal artinya memprediksikan mengenai suatu hal yang akan terjadi pada masa yang akan datang, yang saat ini belum terjadi. Arikunto dalam Elis dan Rusdiana (2015: 171) mengatakan bahwa: “sebuah tes dikatakan memiliki validitas ramalan apabila memiliki kemampuan untuk meramalkan sesuatu yang akan terjadi pada masa yang akan datang.”

Validitas ramalan diperoleh apabila pengambilan skor kriteria tidak bersamaan dengan pengambilan skor tes. Setelah subjek dikenakan tes yang akan diuji validitas prediksinya, lalu diberikan tenggang waktu tertentu sebelum skor kriteria diambil dari subjek yang sama. Prosedur validitas prediksi memerlukan



waktu yang lama dan biaya yang besar karena prosedur ini pada dasarnya bukan pekerjaan yang dianggap *selesai* setelah sekali melakukan analisis, melainkan berlangsung terus-menerus.

Jadi, validitas ramalan ditandai dengan adanya kesejajaran, kesesuaian, atau kesamaan arah antara nilai-nilai hasil tes yang diperoleh masa kini dengan nilai-nilai hasil belajar mereka kelak. Dalam mencari validitas ramalan sebuah tes yaitu dengan menggunakan teknik analisis korelasional *Product Moment* dari Karl Pearson (Elis dan Rusdiana, 2015: 171)

#### b) Validitas Bandingan

Sebuah instrumen dikatakan memiliki validitas bandingan apabila tes tersebut dalam kurun waktu yang sama dengan tepat telah mampu menunjukkan adanya hubungan yang searah, antara tes pertama dengan tes berikutnya (Sudijono, 2015: 177).

Validitas bandingan juga sering dikenal dengan istilah validitas ada sekarang. Arikunto (2012: 87) menjelaskan bahwa dikatakan sebagai validitas ada sekarang, sebab setiap kali menyebut istilah pengalaman, maka istilah itu selalu mengenai hal yang telah lampau sehingga data pengalaman tersebut sekarang sudah ada.

Dalam rangka menguji validitas bandingan, data yang mencerminkan pengalaman yang diperoleh pada masa lalu itu,

kita bandingkan dengan data hasil tes yang diperoleh sekarang ini. Jika hasil tes yang ada sekarang ini mampu mempunyai hubungan searah dengan hasil tes berdasarkan pengalaman yang lalu, maka tes yang memiliki karakteristik seperti itu dapat dikatakan telah memiliki validitas bandingan.

#### c) Validitas Butir Instrumen

Validitas butir instrumen yaitu daya ketepatan yang dihasilkan dari butir-butir soal tes, untuk menguji suatu butir instrumen dapat dilakukan dengan meliputi: uji validitas item, uji derajat kesukaran item, uji daya pembeda item, uji fungsi distraktor, serta uji reliabilitas tes.

##### (1) Uji Validitas Item

Sebutir item dapat dikatakan telah memiliki validitas yang tinggi atau dapat dinyatakan valid, jika skor-skor pada butir item yang bersangkutan memiliki kesesuaian kesejajaran arah dengan skor totalnya atau dalam bahasa statistik terdapat korelasi positif yang signifikan antara skor item dengan skor totalnya (Sudijono, 2015: 184).

Rumus yang digunakan untuk mengetahui validitas item tes bentuk objektif yaitu menggunakan teknik korelasi *point biserial* yaitu sebagai berikut :

$$r_{pbi} = \frac{M_p - M_t}{SD_t} \times \sqrt{\frac{p}{q}}$$

Keterangan :

$r_{pbi}$  = koefisien korelasi biserial atau disebut juga sebagai koefisien validitas item.

$M_p$  = skor rata-rata yang dimiliki peserta tes yang telah dijawab dengan benar.

$M_t$  = skor rata-rata dari skor total.

$SD_t$  = deviasi standar dari skor total.

$p$  = proporsi peserta tes yang jawabannya benar pada soal yang sedang diuji validitasnya.

$q$  = proporsi peserta tes yang jawabannya salah pada soal yang sedang diuji validitasnya. (Sudijono, 2015: 185).

Pemberian *interpretasi* validitas item tes bentuk

objektif ditentukan dengan patokan sebagai berikut :

Apabila  $r_{pbi} \geq r_{tabel}$  berarti item yang diuji adalah valid.

Apabila  $r_{pbi} < r_{tabel}$  berarti item yang diuji adalah invalid.

Sedangkan rumus untuk mengetahui validitas item tes bentuk subjektif menggunakan teknik korelasi *product moment* yaitu sebagai berikut :

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}}$$

Keterangan :

$r_{xy}$  = koefisien korelasi antara variabel X dan variabel Y.

$\sum XY$  = jumlah hasil kali skor total X.

$\sum X$  = jumlah skor total X

$\sum Y$  = jumlah skor total Y

$\sum X^2$  = jumlah kuadrat skor total X

$\sum Y^2$  = jumlah kuadrat skor total Y

$N$  = jumlah peserta tes (Sudijono, 2011: 206).

Pemberian *interpretasi* validitas item tes bentuk objektif ditentukan dengan patokan sebagai berikut, apabila :

Pemberian *interpretasi* validitas item tes bentuk subjektif ditentukan dengan patokan sebagai berikut :

Apabila  $r_{pbi} \geq r_{tabel}$  berarti item yang diuji adalah valid.

Apabila  $r_{pbi} < r_{tabel}$  berarti item yang diuji adalah invalid.

## (2) Uji Derajat Kesukaran Item

Arifin (2014: 266) menyatakan bahwa perhitungan derajat kesukaran item merupakan pengukuran seberapa besar tingkat kesukaran suatu soal. Jika suatu soal memiliki tingkat kesukaran seimbang (proposional) maka butir item soal dapat dikatakan baik. Suatu tes hendaknya tidak terlalu terlalu sukar dan tidak terlalu mudah. Butir item soal yang terlalu mudah tidak merangsang siswa untuk mempertinggi usaha memecahkannya. Sebaliknya butir item soal yang terlalu sukar akan menyebabkan siswa menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi di luar jangkauannya (Arikunto, 2012: 222).

Bilangan yang menunjukkan sukar dan mudahnya sesuatu tes disebut indeks kesukaran (*difficulty index*). Besar indeks kesukaran antara 0,0 sampai 1,0. Indeks kesukaran ini menunjukkan taraf kesukaran tes. Tes dengan indeks kesukaran 0,0 menunjukkan bahwa soal itu terlalu



berkemampuan tinggi dengan siswa yang berkemampuan rendah sehingga dapat diketahui bahwa siswa yang memiliki kemampuan tinggi untuk menjawab butir soal lebih banyak menjawab dengan benar, sementara siswa yang memiliki kemampuan rendah untuk menjawab butir soal sebagian besar tidak dapat menjawab item dengan benar.

Angka yang menunjukkan besarnya daya pembeda disebut indeks diskriminasi, disingkat D. Seperti halnya indeks kesukaran, indeks diskriminasi ini berkisar antara 0,00 sampai 1,00. Hanya bedanya, indeks kesukaran tidak mengenal mengenal tana negatif (-), tetapi pada indeks diskriminasi ada tanda negatif (-). Terdapat tida titik pada daya pembeda, yaitu :

$$-1,00 \longleftrightarrow 0,00 \longleftrightarrow 1,00$$

daya pembeda (-) (daya pembeda rendah) (daya pembeda tinggi)

Rumus yang digunakan untuk mencari D (indeks diskriminasi) yaitu sebagai berikut:

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B} = P_A - P_B$$

Keterangan :

$J$  = jumlah peserta tes.

$J_A$  = banyaknya peserta tes kelompok atas.

$J_B$  = banyaknya peserta tes kelompok bawah.

$B_A$  = banyaknya peserta tes kelompok atas yang menjawab soal dengan benar.

$B_B$  = banyaknya peserta tes kelompok bawah yang menjawab soal dengan benar.

$P_A$  = proporsi peserta kelompok atas yang menjawab dengan benar.

$P_B$  = proporsi peserta kelompok bawah yang menjawab dengan benar. (Arikunto, 2012: 228).

Pemberian *interpretasi* terhadap daya pembeda item, diklasifikasikan sebagai berikut:

a)  $D : 0,71 - 1,00 =$  baik sekali (*excellent*)

b)  $D : 0,41 - 0,70 =$  baik (*good*)

c)  $D : 0,21 - 0,40 =$  cukup (*satisfactory*)

d)  $D : 0,00 - 0,20 =$  buruk (*poor*)

e)  $D : \text{negatif } (-) =$  buruk sekali, jadi semua butir soal yang mempunyai nilai  $D$  negatif sebaiknya di buang saja.

(Arikunto, 2013: 232).

#### (4) Uji Fungsi Distraktor

Uji fungsi distraktor dilakukan hanya pada butir item soal berbentuk pilihan ganda. Setiap butir item soal pilihan ganda yang dikeluarkan dalam sebuah tes hasil belajar telah dilengkapi dengan beberapa kemungkinan jawaban atau yang sering dikenal dengan istilah *option* atau alternatif. *Option* atau alternatif jawaban itu jumlahnya berkisar antara tiga sampai lima buah, dan dari kemungkinan jawaban yang terpasang pada setiap butir item itu, salah satu diantaranya adalah merupakan jawaban betul sedangkan sisanya adalah merupakan jawaban yang salah.

Jawaban-jawaban yang salah itu biasa dikenal dengan istilah distraktor (Sudijono, 2015: 409).

Arikunto (2012: 234) mengungkapkan bahwa distraktor dinyatakan telah dapat menjalankan fungsinya dengan baik apabila distraktor tersebut paling sedikit sudah dipilih oleh 5% peserta tes. Sebagai tindak lanjut atas hasil analisis terhadap fungsi distraktor tersebut maka distraktor yang sudah dapat menjalankan fungsinya dengan baik dapat diterima dan dipakai lagi pada tes-tes yang akan datang, sedangkan distraktor yang belum dapat berfungsi dengan baik sebaiknya diperbaiki atau diganti dengan distraktor yang lain.

## **2. Reliabilitas**

Kata reliabilitas dalam bahasa Indonesia diambil dari kata *reliability* dalam bahasa Inggris, yang berasal dari kata *reliable* artinya dapat dipercaya. Widoyoko (2014: 144) menyatakan bahwa instrumen tes dikatakan dapat dipercaya jika memberikan yang tetap atau ajek (konsisten) apabila diteskan berkali-kali. Sebuah instrumen tes yang valid pada umumnya sudah reliabel, namun instrumen yang reliabel belum tentu valid.

### **a. Macam-Macam Reliabilitas**

Sukardi (2008) sebagaimana dikutip oleh Pramono (2014: 236) menyatakan bahwa terdapat beberapa tipe reliabilitas, diantaranya



adalah reliabilitas dengan tes-retes, reliabilitas dengan ekuivalen, dan reliabilitas internal.

- 1) Reliabilitas dengan tes-retes, reliabilitas ini merupakan derajat yang mengarah kepada konsistensi hasil sebuah tes dari waktu ke waktu. Tes-retes menunjukkan variasi skor yang diperoleh dari penyelenggaraan satu tes evaluasi yang dilaksanakan dua kali atau lebih. Dapat diketahui bahwa skor yang didapatkan relatif sama dari setiap tes yang dilakukan.
- 2) Reliabilitas dengan ekuivalen, yaitu penyusunan dua instrumen yang hampir sama, kemudian diuji cobakan pada sekelompok responden yang sama dengan mengerjakan dua instrumen tersebut. Instrumen ekuivalen merupakan dua instrumen yang memiliki karakter yang sama, seperti: mengukur variabel yang sama, jumlah item yang sama, serta struktur dan tingkat kesulitan yang sama, tetapi butir-butir pertanyaan/pernyataan berbeda.
- 3) Reliabilitas internal, merupakan reliabilitas yang mengukur terhadap konsistensi internal. Maksud dari konsistensi internal adalah salah satu tipe reliabilitas yang didasarkan pada keajegan dalam setiap item tes evaluasi. Reliabilitas internal diperoleh dengan cara pengumpulan data dari satu kali pengumpulan data atau hanya satu kali dalam pelaksanaannya. Terdapat dua metode dalam analisis reliabilitas internal, yaitu instrumen skor diskrit untuk item tes

bentuk objektif dan instrumen skor non diskrit untuk item tes bentuk subjektif.

#### **b. Faktor-Faktor yang Mempengaruhi Reliabilitas**

Gronlund (1985) dalam Arifin (2014: 258) mengemukakan bahwa terdapat empat faktor yang dapat mempengaruhi reliabilitas, yaitu: panjang tes, sebaran skor, tingkat kesukaran, dan objektivitas.

- 1) Panjang tes (*length of test*), panjang tes berarti banyaknya soal tes. Semakin banyak soal dalam tes, maka akan semakin tinggi tingkat reliabilitas suatu tes dikarenakan terdapat banyak sampel yang diukur dan proporsi jawaban yang benar semakin banyak.
- 2) Sebaran skor (*spread of scores*), besarnya sebaran skor akan membuat tingkat reliabilitas menjadi lebih tinggi, karena koefisien reliabilitas yang lebih besar diperoleh ketika siswa tetap pada posisi yang relatif sama dalam satu kelompok pengujian ke pengujian berikutnya.
- 3) Tingkat kesukaran (*difficulty indeks*), tingkat tes yang sukar dan tes yang mudah untuk siswa cenderung akan menghasilkan tingkat reliabilitas yang rendah.
- 4) Objektivitas (*objectivity*), menunjukkan skor tes kemampuan yang sama antara siswa yang satu dengan siswa yang lainnya.

#### **c. Perhitungan Reliabilitas Tes**

Tujuan utama untuk menguji reliabilitas sebuah tes adalah untuk mengetahui tingkat ketepatan dan keajegan skor tes. Indeks

reliabilitas antara 0-1. Semakin tinggi koefisien reliabilitas suatu tes mendekati angka 1, semakin tinggi pula ketepatan atau keajegannya. Untuk mengetahui reliabilitas tes bentuk objektif dapat dilakukan dengan menggunakan lima jenis formula, antara lain: 1) *Formula Spearman Brown*, 2) *Formula Flanagan*, 3) *Formula Rulon*, 4) *Formula Kuder-Richardson*, dan 5) *Formula C. Hyoth* (Sudijono, 2015: 214). Sedangkan untuk mengetahui reliabilitas tes bentuk subjektif dapat dilakukan dengan menggunakan rumus *Alpha*. Adapun rumus *Alpha* yang dimaksud adalah :

$$r_{11} = \left( \frac{n}{n-1} \right) \left( 1 - \frac{\sum s_i^2}{s_t^2} \right)$$

Keterangan :

$r_{11}$  = koefisien reliabilitas tes  
 $n$  = banyaknya butir item yang dikeluarkan dalam tes  
 $1$  = bilangan konstan  
 $\sum s_i^2$  = jumlah varian skor dari tiap-tiap butir item  
 $s_t^2$  = varian total (Sudijono, 2015: 208)

### 3. Tes

#### a. Pengertian Tes

Tes merupakan salah satu alat untuk melakukan pengukuran, yaitu alat untuk mengumpulkan informasi karakteristik suatu objek (Pramono, 2013: 45). Menurut Arifin (2014: 118) tes merupakan suatu teknik atau cara yang digunakan dalam rangka melaksanakan kegiatan pengukuran, yang didalamnya terdapat berbagai pertanyaan, pernyataan, atau serangkaian tugas yang harus dikerjakan atau dijawab oleh siswa.

## **b. Bentuk-Bentuk Tes**

Secara umum, tes dapat dibedakan menjadi dua, tes uraian atau subjektif dan tes objektif. Dalam sejarahnya, tes uraian lebih dulu berkembang daripada tes objektif.

### 1) Tes Objektif

Tes objektif merupakan salah satu jenis tes hasil belajar yang terdiri dari butir-butir soal yang dapat dijawab oleh siswa dengan memilih salah satu diantara kemungkinan jawaban yang telah disediakan dalam soal tes (Sudijono, 2015: 106). Tes objektif terdiri atas beberapa bentuk, yaitu benar-salah, pilihan ganda, menjodohkan, dan melengkapi atau jawaban singkat. Sebagaimana dikemukakan Witherington (1952) dalam Arifin (2014: 135) bahwa, *“There are many varieties of there new test, but four kinds are in most common use, true-false, multiple choice, completion, and matching.”*

#### a) Benar salah (*true false*)

Widoyoko (2013: 51) menjelaskan tes tipe benar salah adalah tes yang butir soalnya terdiri dari pernyataan yang disertai dengan alternatif jawaban yaitu jawaban atau pernyataan yang benar atau salah. Fungsi dari bentuk tes ini adalah untuk mengukur kemampuan siswa dalam membedakan antara fakta dan pendapat (Arifin, 2014; 136).

Tugas siswa hanya diminta untuk menandai masing-masing jawaban atau pernyataan itu dengan melingkari ataupun memberi tanda silang pada huruf “B” jika jawaban atau pernyataan dianggap benar atau “S” jika jawaban atau pernyataan dianggap salah. Penyusunan bentuk tes ini tidak hanya menggunakan kalimat pernyataan, tetapi dapat berbentuk gambar, tabel, dan diagram.

b) Menjodohkan (*matching*)

Tes objektif bentuk *matching* sering dikenal dengan istilah menjodohkan, tes mencari pasangan, tes menyesuaikan, dan tes mencocokkan (Sudijono, 2015: 111). Bentuk soal tes menjodohkan berfungsi untuk mengukur kemampuan siswa dalam mengidentifikasi informasi berdasarkan hubungan yang sederhana dan kemampuan menghubungkan antara dua hal (Arifin, 2014: 144).

Butir soal tes bentuk menjodohkan ditulis dalam dua kolom. Kolom pertama disebelah kiri adalah pertanyaan/pernyataan yang dapat disebut juga sebagai premis. Kolom kedua disebelah kanan adalah kelompok jawaban. Tugas siswa adalah mencari atau menjodohkan jawaban-jawaban, sehingga sesuai atau cocok dengan pertanyaan/pernyataan (Widoyoko, 2013: 55).

c) Pilihan ganda (*multiple choice*)

Arikunto (2013: 183) mengatakan bahwa tes bentuk pilihan ganda terdiri atas suatu keterangan atau pemberitahuan dan bagian kemungkinan jawaban atau alternatif (*options*). Kemungkinan jawaban terdiri dari satu jawaban yang benar yaitu kunci jawaban dan jawaban lainnya adalah pengecoh atau *distraktor*.

Mengenai jumlah alternatif jawaban biasa berkisar dua sampai lima *options*. Tipe tes ini adalah yang paling populer dan banyak digunakan kelompok tes objektif karena banyak sekali materi yang dapat dicakup (Widoyoko, 2013: 59). Bentuk tes pilihan ganda ini banyak telah dimodifikasi mulai dari pilihan ganda bentuk biasa hingga bentuk pilihan ganda analisis hubungan antar-hal, analisis kasus, kompleks, asosiasi, serta bentuk diagram, grafik, tabel dan sebagainya.

d) Melengkapi (*completion*)

Tes bentuk melengkapi sering dikenal dengan istilah tes melengkapi atau menyempurnakan. *Completion test* terdiri atas kalimat-kalimat yang ada bagian-bagiannya yang dihilangkan (Arikunto, 2012: 190). Tugas siswa adalah mengisi atau menjawab bagian yang hilang dengan singkat baik jawaban berupa kata, frase, nama tempat, nama tokoh, lambang, dll sehingga dapat menjadi sebuah kalimat yang utuh dan padu.

## 2) Tes Subjektif

Tes subjektif, pada umumnya berbentuk uraian (*essai*). Tes bentuk uraian adalah pertanyaan yang menuntut siswa dalam menjawab tes dalam bentuk menguraikan, menjelaskan, mendiskusikan, membandingkan, memberi alasan, dsb (Elis dan Rusdiana, 2014: 121). Ciri khas tes uraian adalah jawaban terhadap soal tersebut tidak disediakan oleh penyusun soal, tetapi harus disusun oleh siswa. Butir soal tipe uraian hanya terdiri pertanyaan atau tugas dan jawaban sepenuhnya harus dipikirkan oleh siswa. Ciri-ciri pertanyaan didahului dengan kata-kata seperti: uraikan, jelaskan, bandingkan, mengapa, bagaimana, simpulkan, dan sebagainya (Arikunto 2012: 177).

Berdasarkan tingkat kebebasan siswa untuk menjawab soal tes uraian, secara umum tes uraian dapat dibagi menjadi dua bentuk, yaitu: tes uraian bebas atau uraian terbuka (*extended response*) dan tes uraian terbatas (*restricted response*).

### a) Tes Uraian Bebas (*extended esponse test*)

Tes uraian bebas merupakan bentuk tes uraian yang memberi kebebasan kepada siswa seluas-luasnya dalam merumuskan, mengorganisasikan, dan menyajikan jawabannya (Sudijono, 2015: 100). Jawaban siswa bersifat terbuka, fleksibel, dan tidak terstruktur. Bentuk soal uraian bebas ini berfungsi untuk mengukur kemampuan siswa secara lebih kompleks, seperti

kemampuan menghasilkan, menyusun dan menyatakan ide-ide, memadukan beberapa hasil belajar dari berbagai bidang studi, serta menilai arti atau makna suatu ide (Arifin (2009) dalam Pramono, 2014: 65) .

Kelemahan dari tes ini adalah guru sulit menilai hasil jawaban siswa dikarenakan jawaban sangat bervariasi dan guru sulit menentukan kriteria penilaian sehingga sangat subjektif karena bergantung pada guru yaitu sebagai penilai.

b) Tes Uraian Terbatas (*restricted response test*)

Tes uraian terbatas merupakan bentuk tes uraian yang memberi batasan-batasan atau rambu-rambu tertentu kepada siswa dalam menjawab soal tes. Batasan atau rambu tersebut mencakup format, isi dan ruang lingkup jawaban. (Widoyoko, 2013: 80). Sedangkan Pramono (2014: 56) menjelaskan bahwa tes uraian bebas adalah bentuk tes yang butir-butir soalnya mempunyai himpunan jawaban yang lebih pasti, sehingga guru dapat melakukan proses penskoran secara lebih objektif.

**c. Fungsi Tes**

Arikunto (2012: 166) mengidentifikasi fungsi-fungsi tes yang ditinjau dari tiga hal yaitu fungsi tes untuk kelas, fungsi tes untuk bimbingan, dan fungsi tes untuk administrasi. Ketiga fungsi tes tersebut akan peneliti paparkan dalam bentuk tabel dibawah ini:



**Tabel 1**  
**Fungsi Tes**

Fungsi untuk kelas	Fungsi untuk bimbingan	Fungsi untuk administrasi
Mengadakan diagnosis terhadap kesulitan belajar siswa.	Menentukan arah pembicaraan dengan orang tua tentang anak-anak mereka.	Memberi petunjuk dalam pengelompokkan siswa.
Mengevaluasi celah antara bakat dengan pencapaian.	Membantu siswa dalam menentukan pilihan.	Penempatan siswa baru.
Menaikkan tingkat prestasi.	Membantu siswa dalam mencapai tujuan pendidikan.	Membantu siswa memilih kelompok
Mengelompokkan siswa dalam kelas pada waktu metode kelompok.	Memberi kesempatan pada guru dan orang tua dalam memahami kesulitan anak.	Menilai kurikulum.
Merencanakan kegiatan proses belajar mengajar untuk siswa secara perorangan.		Memperluas hubungan masyarakat ( <i>public relation</i> )
Menentukan siswa yang memerlukan bimbingan khusus		
Menentukan tingkat pencapaian untuk setiap siswa.		

#### **d. Komponen-Komponen Tes**

Arikunto (2012: 173) menyatakan komponen atau kelengkapan sebuah tes terdiri atas :

- 1) Buku tes, yakni lembaran atau buku yang memuat butir-butir soal yang harus dikerjakan oleh siswa

- 2) Lembar jawaban tes, yaitu lembaran yang disediakan penilaian bagi siswa untuk mengerjakan tes.
- 3) Kunci jawaban tes berisi jawaban-jawaban yang dikehendaki. Kunci jawaban ini dapat berupa huruf-huruf yang dikehendaki atau kata/kalimat. Ide daripada adanya kunci jawaban ini agar:
  - a) Pemeriksaan tes dapat dilakukan oleh orang lain
  - b) Pemeriksaannya benar
  - c) Dilakukan dengan mudah
  - d) Mengurangi masuknya unsur subjektif.
- 4) Pedoman penilaian, pedoman penilaian atau pedoman skoring berisi keterangan perincian tentang skor atau angka yang diberikan kepada siswa bagi soal-soal yang telah dikerjakan

#### **e. Pengembangan Tes**

Terdapat sembilan langkah yang perlu ditempuh dalam mengembangkan tes hasil belajar. Djemari dalam Widoyoko (2015: 88) menyebutkan kesembilan langkah tersebut diantaranya adalah :

- 1) Menyusun spesifikasi tes, langkah awal ini yaitu berisi uraian yang menunjukkan keseluruhan karakteristik yang harus dimiliki suatu tes. Widoyoko, (2015: 88) menjelaskan dalam penyusunan spesifikasi tes mencakup kegiatan tes sebagai berikut:
  - a) Menentukan tujuan tes, ditinjau dari segi tujuannya terdapat empat macam tes yang banyak digunakan di lembaga pendidikan, yaitu tes penempatan, tes diagnostik, tes formatif, dan tes sumatif.

- (1) Tes penempatan bertujuan ketika suatu tes dilaksanakan pada awal pelajaran. Hasil tes ini berguna untuk mengetahui tingkat kemampuan yang telah dimiliki siswa.
- (2) Tes diagnostik berfungsi untuk mengetahui kesulitan belajar yang dihadapi siswa, termasuk kesalahan pemahaman konsep. Tes ini dilakukan apabila diperoleh informasi bahwa sebagian besar siswa gagal dalam mengikuti proses pembelajaran. Hasil tes ini memberikan informasi tentang konsep-konsep yang belum dipahami dan sudah dipahami.
- (3) Tes formatif bertujuan untuk memperoleh masukan tentang tingkat keberhasilan pelaksanaan proses pembelajaran. Masukan ini berguna untuk memperbaiki strategi mengajar. Tes ini dilakukan secara periodik sepanjang semester. Materi tes dipilih berdasarkan tujuan pembelajaran tiap pokok bahasan atau sub-pokok bahasan. Dalam pengalaman di sekolah, tes formatif disamakan dengan ulangan harian atau ujian tengah semester.
- (4) Tes sumatif diberikan di akhir suatu pelajaran, atau akhir semester. Hasilnya untuk menentukan keberhasilan belajar siswa untuk pelajaran tertentu. Tingkat keberhasilan ini dinyatakan dengan skor atau nilai, pemberian sertifikat, dan sejenisnya. Dalam pengalaman di sekolah, tes sumatif

disamakan dengan ulangan umum yang biasa dilaksanakan pada akhir semester.

- b) Menyusun kisi-kisi tes, kisi-kisi atau biasa disebut juga sebagai tabel spesifikasi tes merupakan tabel *matriks* yang berisi spesifikasi soal-soal yang akan dibuat. Kisi-kisi ini merupakan acuan bagi penulis soal, sehingga siapapun yang menulis soal akan menghasilkan soal yang isi dan tingkat kesulitannya yang relatif sama. *Matriks* kisi-kisi soal terdiri dari dua jalur, yaitu kolom dan baris. Kolom menyatakan kompetensi dasar dan indikator pokok dan sub pokok bahasan, serta uraian materi. Sedangkan baris menyatakan tujuan yang akan diukur dalam tes.
  - c) Memilih bentuk tes, pemilihan bentuk tes yang tepat ditentukan oleh tujuan tes, jumlah peserta tes, waktu yang tersedia untuk memeriksa lembar jawaban tes, cakupan materi, dan karakteristik mata pelajaran yang diujikan.
  - d) Menentukan panjang tes, penentuan panjang tes didasarkan pada cakupan materi ujian dan kelelahan siswa. Pada umumnya tes tertulis menggunakan waktu 90 sampai 150 menit, untuk tes praktik bisa lebih dari itu. Penentuan panjang tes berdasarkan pengalaman saat melakukan tes.
- 2) Menulis soal tes, penulisan soal tes merupakan langkah menjabarkan indikator menjadi pertanyaan-pertanyaan yang karakteristiknya sesuai dengan perincian pada kisi-kisi yang telah dibuat. Langkah ini

perlu dilakukan secara hati-hati agar keseluruhan tes dapat berkualitas baik. Kualitas tes secara keseluruhan sangat dipengaruhi oleh tingkat kebaikan dari masing-masing butir soal. Pertanyaan perlu dikembangkan dan dibuat dengan jelas dan simpel. Dengan demikian, setiap pertanyaan sedemikian rupa jelas pertanyaannya dan jelas pula jawaban yang diharapkan.

- 3) Menelaah soal tes, setelah soal dibuat, perlu dilakukan telaah atas soal tersebut. Hal ini perlu dilakukan untuk memperbaiki soal jika ternyata dalam pembuatannya masih ditemukan kekurangan dan kesalahan. Telaah soal ini sebaiknya dilakukan oleh orang lain yang terdiri dari para ahli, bukan si pembuat soal.
- 4) Melakukan uji coba tes, uji coba digunakan untuk sebagai saran memperoleh data empirik tentang tingkat kebaikan soal yang telah disusun. Melalui uji coba dapat diperoleh data tentang reliabilitas, validitas, tingkat kesukaran, pola jawaban, efektivitas pengecoh daya beda, dan lain-lain. Jika memang soal yang disusun belum memenuhi kualitas yang diharapkan, berdasarkan hasil uji coba tersebut kemudian dilakukan pembenahan atau perbaikan.
- 5) Menganalisis butir soal tes, setelah melakukan uji coba tes perlu kiranya untuk dilakukan analisis butir soal yang telah disusun. Melalui analisis butir ini dapat diketahui antara lain: tingkat kesulitan butir soal, daya pembeda, dan efektifitas pengecoh.

- 6) Memperbaiki tes, langkah ini berupa melakukan perbaikan-perbaikan tentang bagian soal yang masih belum sesuai dengan yang diharapkan. Butir soal yang sudah baik tidak perlu direvisi, dan beberapa butir soal mungkin perlu direvisi, dan beberapa butir soal yang lain harus dibuang atau diganti karena tidak memenuhi standar kualitas yang diharapkan.
- 7) Merakit tes, langkah ini berupa merakit butir-butir soal tersebut menjadi satu kesatuan tes yang terpadu. Dalam merakit soal, hal-hal yang dapat mempengaruhi validitas soal seperti nomor urut soal, pengelompokan bentuk soal, *lay out* dan sebagainya harus diperhatikan. Hal-hal ini sangat penting karena walaupun butir-butir soal yang disusun telah baik, tetapi jika penyusunannya sembarangan dapat menyebabkan soal tersebut menjadi tidak baik.
- 8) Melaksanakan tes, pelaksanaan tes dilakukan sesuai dengan waktu yang telah ditentukan. Dalam pelaksanaan tes ini memerlukan pemantauan atau pengawas agar tes tersebut benar-benar dikerjakan oleh peserta tes dengan jujur dan sesuai dengan ketentuan yang telah digariskan. Namun, peserta tes yang mengerjakan tes tidak boleh sampai terganggu oleh kehadiran pengawas. Hal ini akan berakibat tidak akuratnya hasil tes yang diperoleh. Oleh karena itu, pelaksanaan tes perlu dilakukan secara hati-hati agar tujuan tes tersebut benar-benar dapat tercapai.

9) Menafsirkan hasil tes, hasil tes menghasilkan data kuantitatif yang berupa skor/nilai. Skor ini kemudian ditafsirkan sehingga menjadi nilai, yaitu rendah, sedang, atau tinggi. Tinggi rendahnya nilai ini selalu dikaitkan dengan acuan penilaian. Nilai merupakan alat yang sangat berguna untuk memotivasi siswa belajar dan guru yang mengajar agar lebih baik. Nilai juga merupakan informasi mengenai keberhasilan guru dalam melaksanakan proses pembelajaran. Oleh karena itu, pencapaian hasil belajar atau prestasi belajar siswa merupakan indikator keberhasilan belajar dan keberhasilan guru mengajar.

#### **f. Ciri-Ciri Tes yang Baik**

Linn dan Gronlund (1995) menyatakan bahwa tes yang baik harus memenuhi tiga karakteristik, yaitu: validitas, reliabilitas, dan usabilitas. Validitas artinya ketepatan interpretasi hasil prosedur pengukuran, reliabilitas artinya konsistensi hasil pengukuran, dan usabilitas artinya praktis prosedurnya (panduan analisis butir soal *pdf* dalam <http://gurupembaharu.com>). Arifin (2014: 70) menyatakan bahwa sebuah tes yang dapat dikatakan baik sebagai alat pengukur, harus memenuhi persyaratan tes, yaitu memiliki:

##### 1) Validitas

Alat ukur dikatakan valid apabila alat ukur itu dapat dengan tepat mengukur apa yang hendak diukur. Dengan kata lain validitas berkaitan dengan “ketepatan” dengan alat ukur. Tes sebagai salah

satu alat ukur hasil belajar dapat dikatakan valid apabila tes itu dengan tepat mengukur hasil belajar yang hendak diukur. Dengan tes yang valid akan menghasilkan data hasil yang valid pula (Widoyoko, 2013: 98).

Banyak faktor yang dapat mempengaruhi hasil instrumen tes evaluasi menjadi tidak valid. Beberapa faktor tersebut, secara garis besar dapat dibedakan menurut sumbernya yaitu: faktor internal tes, faktor eksternal tes (Prmono, 2014: 226). Faktor-faktor tersebut akan dijelaskan pada tabel berikut ini:

**Tabel 2**  
**Faktor Internal dan Eksternal Tes**

<b>Faktor Internal Tes</b>	<b>Faktor Eksternal Tes</b>
Arahan tes yang disusun dengan makna yang tidak jelas, sehingga dapat mengurangi validitas tes	Waktu pengerjaan tidak cukup, sehingga siswa memberikan jawaban dalam situasi tergesa-gesa
Kata-kata yang digunakan dalam struktur instrumen evaluasi tidak terlalu sulit	Adanya kecurangan dalam tes
Item tes dikonstruksi dengan jelas	Pemberian petunjuk dari pengawas yang tidak dapat dilakukan pada semua siswa
Tingkat kesulitan item tes tidak tepat dengan materi pembelajaran yang diterima siswa	Teknik pemberian skor yang tidak konsisten
Waktu yang dialokasikan tidak tepat	Siswa tidak adapat mengikuti arahan yang diberikan dalam tes
Jumlah item terlalu sedikit sehingga tidak mewakili sampel	Adanya joki (orang lain bukan siswa) yang masuk dalam menjawab item tes yang diberikan
Jawaban masing-masing item evaluasi dapat diprediksi oleh siswa	



## 2) Reliabilitas

Kata reliabilitas dalam bahasa Indonesia diambil dari kata *reliability* dalam bahasa Inggris, berasal dari kata *reliable* yang artinya dapat dipercaya. Sebuah tes dikatakan dapat dipercaya (*reliable*) jika memberikan hasil yang tetap dan ajeg (*consistent*) apabila diteskan berkali-kali (Arikunto, 2012: 74). Serupa dengan pendapat Arifin (2014: 69) yang mengatakan bahwa sebuah instrumen tes dapat dikatakan reliabel atau handal apabila ia mempunyai hasil yang taat asas (*consistent*). Jadi, jika siswa diberikan tes yang sama dengan waktu yang berlainan, maka setiap siswa akan tetap berada dalam urutan (*ranking*) yang sama atau ajeg dalam kelompoknya.

Terdapat hubungan yang erat antara validitas dan reliabilitas sebuah tes. Sebuah tes yang valid sudah pasti *reliable*, namun tidak demikian sebaliknya (Pramono: 2014 : 235) .

## 3) Objektivitas

Objektivitas berarti tidak adanya unsur pribadi yang mempengaruhinya. Sebuah tes dikatakan memiliki objektivitas apabila dalam melaksanakan tes tidak ada faktor subjektif yang mempengaruhinya, terutama dalam sistem penskoringsnya (Arikunto, 2012: 75) . Oleh karena itu, untuk menghindari atau mengurangi masuknya faktor-faktor subjektif dalam penilaian, maka penilaian harus dilaksanakan secara *kontinu* (terus-menerus) sehingga akan

diperoleh gambaran yang lebih jelas tentang keadaan siswa dan secara *komprensif* (menyeluruh) mengenai keseluruhan materi dan mencakup berbagai aspek berfikir.

#### 4) Praktibilitas

Sebuah tes dikatakan memiliki praktibilitas yang tinggi apabila tes tersebut bersifat praktis, mudah pengadministrasinya. Tes yang praktis tidak identik dengan tes yang mudah, melainkan tes itu mampu disajikan dengan cara praktis sehingga sesuatu yang sulit dapat dipahami, dicerna, dan disimpulkan dengan mudah (Pramono, 2014: 243). Tes yang praktis adalah tes :

- a) Mudah dilaksanakan, artinya tidak menuntut peralatan yang banyak dan memberi kebebasan kepada siswa untuk mengerjakan terlebih dahulu bagian yang dianggap mudah oleh siswa.
- b) Mudah pemeriksaannya, artinya bahwa tes itu dilengkapi dengan kunci jawaban maupun pedoman skoringnya.
- c) Dilengkapi dengan petunjuk-petunjuk sehingga dapat diberikan oleh orang lain (Widoyoko, 2015: 101).

#### 5) Ekonomis

Arti dari ekonomis disini adalah bahwa pelaksanaan tes tersebut tidak membutuhkan biaya yang mahal, tenaga yang banyak dan waktu yang lama (Arikunto, 2012: 77).

#### 6) Relevan

Instrumen tes yang digunakan harus sesuai dengan standar kompetensi, kompetensi dasar, dan indikator yang telah ditetapkan. Kesesuaian tes terhadap aspek-aspek hasil belajar, seperti aspek *kognitif*, *afektif*, dan *psikomotorik* harus relevan.

#### 7) Representatif

Materi instrumen tes harus benar-benar mewakili seluruh materi yang disampaikan pada proses pembelajaran. Hal ini dapat dilakukan dengan penggunaan silabus sebagai acuan pemilihan materi tes.

### **4. Bahasa Arab**

#### **a. Pengertian Bahasa Arab**

Salim (2015: 170) menyatakan bahwa bahasa Arab adalah bahasa yang lengkap dan sempurna bila dibandingkan dengan bahasa-bahasa yang lain. Kesempurnaan dan kelengkapannya itulah yang menjadi keistimewaannya. Salah satu keistimewaan bahasa Arab adalah kemampuannya dalam mengurai sesuatu yang sukar menjadi mudah, yang belum mengerti menjadi dapat dimengerti, dan yang belum baik (indah) menjadi baik (indah), dsb.

Hermawan (2013: 57) menjelaskan bahwa bahasa Arab merupakan bahasa yang diposisikan sebagai bahasa asing karena bahasa Arab digunakan oleh masyarakat luas. Dalam instansi pendidikan, kedudukan bahasa Arab juga sebagai bahasa asing termasuk dalam

kurikulum. Hal ini ditunjukkan dengan indikator keasingannya di sekolah-sekolah bahwa bahasa arab tidak digunakan sebagai bahasa pengantar melainkan sebagai materi pelajaran. Namun pada sekolah-sekolah yang berbasis Islam sebagian besar bahasa arab diposisikan tidak lagi menjadi materi pelajaran tetapi sebagai alat komunikasi sehari-hari bahkan digunakan sebagai pengantar pelajaran.

#### **b. Tujuan Pembelajaran Bahasa Arab**

Hermawan (2013:57) menjelaskan tujuan pembelajaran bahasa Arab yang telah tercantum dalam peraturan Menteri Agama RI nomor 2 tahun 2008 tentang Standar Kompetensi dan Standar Isi Pendidikan Agama Islam dan Bahasa Arab, yaitu sebagai berikut:

- 1) Mengembangkan kemampuan berkomunikasi dalam bahasa Arab, baik lisan maupun tulis, yang mencakup empat kecakapan berbahasa, yakni mendengarkan (*istima*), berbicara (*kalam*), membaca (*qira'ah*), dan menulis (*kitabah*). Depag RI sebagaimana dikutip Rosyidi dan Ni'mah (2012: 142) menyatakan bahwa kemampuan berbahasa Arab secara konvensional dianggap meliputi empat jenis kemampuan yang proses pengajaran bahasanya sudah diatur urutannya, yaitu: a) kemampuan menyimak, untuk memahami bahasa yang digunakan secara lisan, b) kemampuan berbicara, untuk mengungkapkan diri secara lisan, c) kemampuan membaca, untuk memahami bahasa yang diungkapkan secara tertulis, d) kemampuan menulis, untuk mengungkapkan diri secara tertulis

- 2) Menumbuhkan kesadaran tentang pentingnya bahasa Arab sebagai salah satu bahasa asing untuk menjadi alat utama belajar, khususnya dalam mengkaji sumber-sumber ajaran Islam.
- 3) Mengembangkan pemahaman tentang saling keterkaitannya antara bahasa dan budaya serta memperluas cakrawala budaya. Dengan demikian siswa diharapkan memiliki wawasan lintas budaya dan melibatkan diri dalam keragaman budaya.

### **c. Efektifitas Pembelajaran Bahasa Arab**

Fachrurozzi dan Mahyudin (2016: 205) menyatakan bahwa pembelajaran efektif adalah suatu pembelajaran yang ditandai dengan terpenuhinya tujuan pembelajaran dan tercapainya kompetensi pembelajaran setelah proses pembelajaran. Pembelajaran efektif menyiratkan bahwa pembelajaran harus dilakukan sedemikian rupa untuk mencapai semua hasil belajar yang telah dirumuskan dalam bidang pendidikan.

Suatu kegiatan dapat dikatakan efektif tentu saja jika kegiatan itu ada efeknya, dapat membawa hasil dan pengaruh bagi siswa. Efek, hasil, dan pengaruh dalam bidang pembelajaran bahasa Arab dapat dilihat dari dua segi, yaitu segi proses dan segi hasil.

Ditinjau dari segi proses, pembelajaran efektif terlihat ketika; 1) seluruh siswa atau setidaknya sebagian besar (75%) siswa terlihat secara aktif, baik fisik, mental, maupun sosial dalam proses pembelajaran, 2) guru dapat menunjukkan profesionalitasnya dalam

pembelajaran bahasa asing antara lain: memiliki penguasaan materi yang luas dan mendalam, mempunyai visi, misi, orientasi, pendekatan, dan metode secara memadai, memiliki komitmen yang tinggi, selalu berusaha memotivasi dan memajukan peserta didik, mempunyai integritas moral dan keteladanan yang baik, 3) proses pembelajaram bahasa asing berlangsung secara manusiawi (humanis), dinamis, menyenangkan, produktif, dan tidak monoton dan membosankan, sehingga siswa merasa termotivasi , ingin terus belajar, dan terpacu untuk berprestasi.

Ditinjau dari segi hasil, pembelajaran efektif terlihat ketika; 1) sebagian besar (75%) siswa dapat mencapai kompetensi yang telah ditentukan, 2) hasil pembelajarn bahasa asing itu dirasakan bermanfaat bagi pengembangan kepribadian siswa, dapat membantu penguasaan ilmu sesuai dengan kebutuhan siswa.

#### **d. Pengembangan Tes Pembelajaran Bahasa Arab**

Selama ini tes bahasa Arab sering tidak menggambarkan tes komunikatif seperti yang dituntut oleh idealisme pengajaran bahasa Arab dengan menggunakan pendekatan komunikatif (Hermawan, 2013: 284). Muradi (2014: 32) menjelaskan pengertian pendekatan komunikatif adalah pendekatan pembelajaran bahasa yang lebih menekankan pembelajaran pada penguasaan kecakapan bahasa daripada struktur bahasa. Salah satu cara dalam penggunaan pendekatan komunikatif yaitu perlu adanya kompetensi komunikatif. Kompetensi

komunikatif yang dimaksud ialah suatu kompetensi untuk melihat kemampuan siswa tidak hanya kemampuan membentuk kalimat yang benar tetapi juga menggunakannya secara tepat.

Sering kita jumpai bahwa tes bahasa Arab lebih mengedepankan pada aspek *kognitif* yaitu penguasaan tata bahasa dan unsur-unsur bahasa seperti kosakata dan sebagainya. Paradigma ini perlu segera diubah, karena pada hakekatnya belajar bahasa terutama pada jenjang sekolah menengah bukan semata-mata belajar untuk menguasai ilmu bahasa itu, akan tetapi lebih dimaksudkan untuk membantu siswa mampu menggunakan bahasa tersebut sebagai alat komunikasi, baik secara lisan maupun tulisan. Hal ini berarti dalam mengembangkan tes kemampuan berbahasa tidak lagi harus menekankan pada aspek *kognitif*, tetapi juga pada aspek *afektif* dan *psikomotorik*.

Menurut Struckmann dalam Hermawan (2013: 287), evaluasi dalam pengajaran bahasa asing harus berfokus pada fungsi-fungsi berikut ini :

- 1) Fungsi diagnostik, tujuannya adalah menambah kemampuan performansi setiap individu siswa.
- 2) Fungsi terapi, sebagai dasar dukungan remedial terhadap siswa secara individu.
- 3) Fungsi evaluatif, untuk mendapatkan informasi tentang keberhasilan pengajaran.

4) Fungsi informatif, memberikan bekal informasi tentang apa yang telah dicapai siswa.

Dapat disimpulkan tes, penilaian, dan evaluasi dalam pembelajaran bahasa Arab seharusnya dapat menggambarkan kemampuan siswa menggunakan bahasa dalam komunikasi nyata dan dalam kehidupan sehari-hari. Dan tidak hanya berputar pada pembelajaran tata bahasa, ucapan, kosakata secara terpisah, tetapi tes harus dipandang dari segi keterpaduan penggunaannya dalam kehidupan sehari-hari.

#### **e. Materi Tes Mata Pelajaran Bahasa Asing (Arab)**

Syai'ri (2013: 58) menjelaskan bahwa terdapat tiga komponen utama bahan ajar pembelajaran bahasa Arab, yaitu :

1) Komponen utama mencakup informasi atau topik utama yang ingin disampaikan kepada siswa atau harus dikuasai oleh siswa, 2) komponen pelengkap mencakup informasi atau topik tambahan yang terintegrasi dengan bahan ajar utama, atau topik pengayaan wawasan siswa, seperti materi pengayaan, materi tambahan, atau materi pendukung non cetak lainnya, 3) komponen evaluasi hasil belajar mencakup tes dan non tes yang dapat digunakan untuk tes formatif dan sumatif siswa selama proses pembelajaran.

Materi pokok tes mata pelajaran Bahasa Asing kelas X pada UAS ganjil tahun pelajaran 2016/2017 yaitu mengenai ucapan salam dan perkenalan (التحية والتعارف), bilangan (العدد), kosakata (المفردة), kata ganti (اسم الضمير), kata petunjuk (اسم الإشارة), kalimat (الجملة).



التحية والتعارف	
ط1 : السلام عليكم	ط2 : وعليكم السلام
ط1 : اسمي خالد، ما اسمك ؟	ط2 : اسمي احمد، من اين انت ؟
ط1 : انا من باكستان	ط2 : هل انت باكستاني ؟
ط1 : نعم، انا باكستاني، وما جنسيتك ؟	ط2 : انا تركي. انا من تركيا
ط1 : اهلاً و سهلاً	

العدد	
واحد = 1	ستة = 6
اثنان = 2	سبعة = 7
ثلاثة = 3	ثمانية = 8
اربعة = 4	تسعة = 9
خمسة = 5	عشرة = 10

اسم الضمير	
هو = Dia laki-laki	انتما = (2) Kamu laki-laki
هما = Dia laki-laki (2)	انتم = Kalian laki-laki
هم = Mereka laki-laki	انت = Kamu perempuan
هي = Dia perempuan	انتما = (2) Kamu perempuan
هما = Dia perempuan (2)	انتن = Kalian perempuan
هن = Mereka perempuan	انا = Saya
انت = Kamu laki-laki	نحن = Kami

اسم الإشارة	
مؤنث	مذكر
Ini (Perempuan) = هذه	Ini (Laki-laki) = هذا

المفردة	
مؤنث	مذكر
Guru Perempuan = مدرسة	Guru Laki-laki = مدرس
Murid Perempuan = طالبة	Murid Laki-laki = طالب
Dokter Perempuan = طبيبة	Dokter Laki-laki = طبيب
Insinyur Perempuan = مهندسة	Insinyur Laki-laki = مهندس
Saudara Perempuan = اخت	Saudara Laki-laki = اخ
Teman Perempuan = صديقة	Teman Laki-laki = صديق

**Sumber:** Buku mata pelajaran Bahasa Asing (Arab) kelas X  
(الأولنا لوحدت: التحيّة والتعارف)

Mata pelajaran Bahasa Asing (Arab) merupakan mata pelajaran yang tidak termasuk dalam kurikulum ISMUBA, mata pelajaran ini merupakan mata pelajaran tambahan dan sebagai pengantar dari mata pelajaran Bahasa Arab ISMUBA. Bahasa Asing (Arab) ini memang lebih mudah dipahami dan dimengerti oleh siswa dikarenakan materi yang memuat hanya mengenai kosakata dan percakapan yang digunakan dalam kehidupan sehari-hari, berbeda dengan mata pelajaran Bahasa Arab ISMUBA yang telah mempelajari tata bahasa atau dikenal dengan ilmu *nahwu sharaf*.

### C. Kerangka Berfikir

