

EVALUASI PEMBELAJARAN BAHASA ARAB

(Kajian Tentang Instrumen Tes)

Roviin

IAIN Salatiga, Indonesia

roviinrovi@gmail.com

Abstrak

Evaluasi merupakan salah satu unsur yang penting dalam proses pembelajaran. Dengan evaluasi, bisa diketahui apakah tujuan pembelajaran bahasa Arab sudah tercapai atau belum. Oleh karena itu, pengetahuan dan pemahaman tentang kriteria dan prinsip-prinsip penyusunan instrumen evaluasi bahasa Arab harus dimiliki oleh setiap guru bahasa Arab. Dalam kaitannya dengan tes bahasa Arab, agar tes tersebut memiliki kualitas yang baik, maka harus terpenuhi tiga kriteria, yaitu validitas, reliabilitas, dan kepraktisan. Ruang lingkup tes bahasa Arab dapat dikelompokkan menjadi dua, yaitu: (1) tes unsur bahasa yang mencakup: tes *mufradat* dan *qawa'id*; dan (2) tes keterampilan berbahasa mencakup: tes keterampilan menyimak, keterampilan berbicara, keterampilan membaca, dan keterampilan menulis. Tes bahasa Arab dapat dikelompokkan menjadi beraneka ragam jenis tes berdasarkan berbagai kriteria pengelompokannya. Dalam penyusunan tes bahasa Arab, perlu diperhatikan rambu-rambu serta panduan langkah-langkah penyusunan tes agar soal tes yang disusun terjamin kualitasnya.

Kata kunci: evaluasi pembelajaran, bahasa Arab, instrument tes

ملخص

التقويم من العناصر الهامة في عملية التعليم. فبهذا التقويم، يعرف المعلم هل كان هدف تعليم اللغة العربية قد تحقق أم لم يتحقق. لذلك، فإن فهم معايير ومبادئ إعداد أدوات التقويم العربية يجب أن تكون مملوكة لكل معلم اللغة العربية. وفيما يتعلق باختبارات اللغة العربية، لكي يكون الاختبار ذا نوعية جيدة، يجب أن يستوفي ثلاثة معايير، وهي الصلاحية والموثوقية والتطبيقية. ويمكن تقسيم اختبارات اللغة العربية إلى اختبارين هما:

(1) اختبارات العناصر اللغوية التي تشمل اختبار المفردات والقوائد؛ و(2) اختبارات المهارات اللغوية وتشمل مهارة الاستماع، ومهارة التحدث، ومهارة القراءة، ومهارة الكتابة. ويمكن تجميع اختبارات اللغة العربية في مجموعة واسعة من أنواع الاختبارات استنادا إلى معايير تجميعها المختلفة. ومن الضروري في إعداد اختبارات اللغة العربية أن يلاحظ المعلم علامات وتوجيهات من إعداد خطوات الاختبار بحيث يتم ضمان أسئلة الاختبار التي يتم ترتيبها الجودة.

الكلمات الرئيسية : التقويم، اللغة العربية، أدوات الاختبار

A. PENDAHULUAN

Kemahiran dalam menyusun alat evaluasi pembelajaran merupakan salah satu kompetensi yang harus dimiliki oleh guru bahasa Arab, sebab evaluasi merupakan salah satu unsur yang penting dalam proses pembelajaran. Dengan evaluasi, bisa diketahui apakah tujuan pembelajaran bahasa Arab yang telah ditetapkan sudah tercapai atau belum. Selanjutnya, hasil evaluasi itu akan menjadi masukan berupa umpan balik bagi perbaikan atau pengembangan proses pembelajaran berikutnya.

Dalam melakukan evaluasi pembelajaran bahasa Arab diperlukan alat atau instrumen. Alat evaluasi pembelajaran ada dua macam, yaitu tes dan non-tes. Alat evaluasi tes biasanya terdiri dari sejumlah so'al secara lisan dan/atau tertulis, dan peserta tes diminta untuk menjawab so'al tersebut secara lisan dan/atau tertulis pula. Sedangkan alat evaluasi bentuk non-tes terdiri dari skala sikap, questioner, wawancara, dan pengamatan.

Dalam kaitannya dengan tes bahasa Arab, agar tes tersebut memiliki kualitas yang baik, maka harus terpenuhi tiga kriteria, yaitu validitas, reliabilitas, dan kepraktisan. Dengan demikian, ketika guru ingin menyusun alat evaluasi bahasa Arab berupa tes, ia harus memperhatikan tiga kriteria tersebut. Selain itu, ia juga harus mempertimbangkan prinsip-prinsip penyusunan tes bahasa Arab. Tanpa memperhatikan kriteria dan prinsip-prinsip penyusunan tes tersebut, dapat dipastikan hasil yang diperoleh dari tes bahasa Arab memiliki tingkat validitas dan reliabilitas yang rendah. Oleh karena itu, pengetahuan dan pemahaman

tentang kriteria dan prinsip-prinsip penyusunan tes bahasa Arab harus dimiliki oleh setiap guru bahasa Arab.

Berdasarkan latar belakang tersebut, dalam artikel ini, pembahasan difokuskan pada penyusunan alat evaluasi bahasa Arab berupa tes yang mencakup pengertian tes, ruang lingkup tes bahasa Arab, jenis-jenis tes bahasa Arab, karakteristik tes yang baik, dan langkah-langkah penyusunan tes bahasa Arab.

B. PENGERTIAN EVALUASI, PENGUKURAN, DAN TES

Penggunaan istilah evaluasi, pengukuran, dan tes sering disamakan. Oleh karena itu, perlu dijelaskan terlebih dahulu pengertian dan hubungan antara evaluasi, pengukuran, dan tes.

1. Evaluasi

Gronlund dan Linn, dalam M. Ainin dkk. (2006), mengemukakan pengertian evaluasi sebagai berikut: *“Evaluation is a systematic process of collecting, analyzing, and interpreting information to determine the extent to which pupils are achieving instructional objectives.”* (Ainin, 2006: 2)

Pengertian tersebut menjelaskan bahwa evaluasi merupakan suatu proses yang sistematis dalam mengumpulkan, menganalisis, dan menafsirkan data untuk menentukan apakah seorang peserta didik dipandang telah mencapai target pengetahuan atau keterampilan yang dirumuskan dalam tujuan pembelajaran. Dengan kata lain, evaluasi adalah suatu proses penentuan keputusan tentang kualitas suatu objek atau aktivitas dengan melibatkan pertimbangan nilai berdasarkan data dan informasi yang dikumpulkan, dianalisis, dan ditafsirkan secara sistematis.

2. Pengukuran

Istilah pengukuran (*measurement*) didefinisikan oleh Gronlund dan Linn, dalam M. Ainin dkk. (2006) sebagai berikut: *“Measurement is the process of obtaining a numerical description of the degree to which an individual processes a particular characteristic.”* (Ainin, 2006: 5)

Definisi tersebut menjelaskan bahwa pengukuran (*measurement*) adalah suatu proses untuk memperoleh deskripsi dalam bentuk angka-angka mengenai tingkat dari sifat atau kemampuan yang dimiliki seseorang. Jadi, pengukuran merupakan bagian dari evaluasi berupa prosedur pengumpulan data dan informasi numerik yang diperlukan sebagai salah satu pertimbangan pengambilan keputusan dalam evaluasi.

3. Tes

Tes merupakan salah satu jenis alat untuk memperoleh data numerik atau alat melakukan pengukuran yang hasilnya dimanfaatkan sebagai salah satu bahan pertimbangan dalam evaluasi (Ainin, 2006: 7). Tes tersebut terdiri dari sejumlah pertanyaan atau pernyataan yang menuntut jawaban atau tanggapan yang memiliki kemungkinan benar atau salah, dengan tujuan mengukur tingkat kemampuan seseorang atau mengidentifikasi aspek tertentu dari orang yang dikenai tes (Djemari, 2008:67). Dengan demikian, evaluasi merupakan konsep yang lebih umum, yang di dalamnya tercakup konsep pengukuran dan tes.

Uraian berikut dapat lebih memperjelas hubungan antara evaluasi, pengukuran, dan tes. Seorang guru bahasa Arab bermaksud melakukan evaluasi terhadap kemampuan membaca bahasa Arab para peserta didiknya untuk menentukan apakah tujuan pembelajaran membaca bahasa Arab telah tercapai atau belum.

Dalam hal ini, guru menghadapi dua hal penting. *Pertama*, ia harus memiliki data atau informasi untuk mengetahui secara pasti tingkat kemampuan membaca para peserta didiknya. Guru perlu memikirkan apakah data yang dimaksud berupa angka-angka ataukah berupa pernyataan-pernyataan. Jika data berupa angka dipandang lebih akurat dalam menggambarkan tingkat kemampuan membaca, maka guru tersebut perlu melakukan pengukuran. Dalam hal ini, guru berpikir bahwa alat yang dipandang tepat untuk mengukur kemampuan membaca tersebut berupa tes. Oleh karena itu, guru perlu menyiapkan tes yang akan digunakan untuk mengukur kemampuan membaca.

Kedua, guru perlu memiliki kriteria tentang tingkat kemampuan yang dianggapnya telah memenuhi syarat atau belum. Penggunaan kriteria dan pertimbangan-pertimbangan tertentu untuk menafsirkan data angka yang telah diperoleh merupakan kegiatan evaluasi. Hal ini tidak terdapat dalam pengukuran. Pengukuran terbatas pada penggalian data angka, sedangkan evaluasi melibatkan pertimbangan guru dalam menentukan apakah data angka yang telah diperoleh itu dapat dianggap memadai. Sebagai contoh, berdasarkan pertimbangan ketuntasan belajar, para peserta didik yang memperoleh angka 70 ke atas dipandang telah memiliki kemampuan yang cukup untuk melanjutkan program pembelajaran. Sebaliknya, mereka yang memperoleh angka $50 \leq 70$ dianggap belum memadai dan perlu tambahan-tambahan latihan membaca. Adapun yang mendapat angka < 50 dipandang memerlukan pengulangan latihan secara lebih intensif.

C. RUANG LINGKUP TES BAHASA ARAB

Ruang lingkup tes bahasa dapat dikelompokkan menjadi dua, yaitu: (1) tes unsur bahasa yang mencakup: tes *mufradat* dan *qawa'id*; dan (2) tes keterampilan berbahasa mencakup: tes keterampilan menyimak (مهارة الاستماع), keterampilan membaca (مهارة القراءة), keterampilan berbicara (مهارة الكلام), dan keterampilan menulis (مهارة الكتابة) (Ainin, 2006: 129-144).

1. Tes Unsur Bahasa

a. Tes Kosa kata (مفردات)

Tes kosa kata (مفردات) dapat dikelompokkan menjadi tes kosa kata aktif (*active vocabulary*/المفردات النشيطة) dan tes kosa kata pasif (*passive vocabulary*/المفردات الخاملة). *Active vocabulary* (المفردات النشيطة) merupakan kosa kata yang digunakan oleh peserta tes dalam berbicara dan menulis. Sedangkan *passive vocabulary* (المفردات الخاملة) yaitu kosa kata yang dikenal dan dipahami oleh peserta tes ketika mendengar pembicaraan seseorang atau ketika membaca teks (Heaton, 2000:48).

Dengan kata lain, tes *passive vocabulary* (المفردات النشيطة) merupakan tes pemahaman kosa kata dan tes *active vocabulary* (المفردات الخاملة) merupakan tes penggunaan kosa kata. Tes pemahaman lebih ditekankan pada pengukuran kemampuan peserta tes dalam memahami arti kosa kata, sedangkan tes penggunaan lebih dititikberatkan pada kemampuan menggunakan kosa kata dalam kalimat. Khusus untuk tes pemahaman kosa kata, indikator kompetensi yang diukur dapat berupa arti kosa kata, atau padanan kata (مرادف الكلمات), lawan kata (ضدّ الكلمات), kelompok kata, melengkapi kata yang tepat/sesuai, pengertian kata, dan pemilihan kata yang sesuai dengan gambar.

Contoh tes padanan kata:

يستلم حسن رسالة من صديقه في الرياض.

أ – طردا ب – خطابا ج – نقودا د – كتابا

Contoh tes memilih kata yang tepat:

أريد أن أكون ... فأواصل دراستي إلى كلية الهندسة.

أ – طبيبا ب – مدرّسا ج – مهندسا د – فنّانا

b. Tes *Qawa'id* (القواعد النحوية والصرفية)

Sebagaimana tes kosa kata, tes *qawa'id* juga dapat dikelompokkan menjadi tes pemahaman dan tes penggunaan kaidah nahwu-sharaf dalam konteks kalimat atau paragraph. Contoh tes pemahaman sharaf:

نصحي الطبيب بأن ... الرياضة كل يوم.

أ – يمارس ب – تمارس ج – نمارس د – أمارس

Contoh tes pemahaman nahwu:

تباع فواكه ... في السوق، منها برتقال وتفاح وموز.

أ – المتنوّعة ب – المتنوّع ج – متنوّعة د – متنوّع

2. Tes Keterampilan Berbahasa

a. Tes Keterampilan Menyimak (مهارة الاستماع)

Tes keterampilan menyimak dimaksudkan untuk mengukur kemampuan peserta didik menangkap, memahami, dan menanggapi informasi yang terkandung dalam wacana lisan. Sarana yang dipergunakan biasanya berupa media rekaman suara (Burhan, 2010:355).

Pemilihan wacana sebagai bahan tes menyimak haruslah mempertimbangkan beberapa faktor, antara lain: tingkat kesulitan wacana, isi dan cakupan wacana, dan jenis wacana (Burhan, 2010:355). Tingkat kesulitan wacana terkait erat dengan kompleksitas/kerumitan kosa kata dan struktur kalimat yang dipergunakan. Jika kosa kata yang dipergunakan abstrak dan bermakna ganda, jarang digunakan, struktur kalimat yang rumit, maka tingkat kesulitan wacana itu termasuk tinggi.

Selain itu, Jika isi dan cakupan wacana tidak sesuai dengan minat, pengalaman, dan kemampuan peserta didik, hal itu akan menambah tingkat kesulitan wacana. Wacana yang baik untuk tes keterampilan menyimak adalah wacana yang tingkat kesulitannya sedang atau sesuai dengan kemampuan peserta didik.

Jenis tes yang digunakan untuk tes keterampilan menyimak bisa berbentuk tes obyektif pilihan ganda. Dilihat dari cara kerja peserta tes dan koreksi jawaban, jenis tes ini lebih praktis, cara penilaian atau pemberian skornya pun lebih obyektif. Ditambah lagi, jenis tes ini dapat mencakup macam-macam wacana dan banyak soal, walaupun pembuatan soalnya lebih sulit dan lebih lama.

Jenis wacana yang ditekankan dapat berupa pertanyaan atau pernyataan singkat, dialog, dan wacana narasi. Sedangkan pilihan jawabannya disediakan dalam lembar jawaban tersendiri.

Contoh soal keterampilan menyimak dengan memahami pernyataan singkat:

صوت الشريط : وضعت فاطمة رسالة في صندوق البريد.

أ – فتحت فاطمة رسالة

ب – أرسلت فاطمة رسالة

ج – فقدت فاطمة رسالة

د – كتبت فاطمة رسالة

Contoh soal keterampilan menyimak dengan memahami dialog:

صوت رجل 1: هل استمتعت بهذه السهرة؟

صوت رجل 2: نعم، ولكن الفصل الأول كان مملاً.

صوت رجل 3: أين أمضى المتحدثان السهرة؟

أ – في النادي

ج – في الملعب

ب – في المعمل

د – في المسرح

b. Tes Keterampilan Membaca (مهارة القراءة)

Tes keterampilan membaca dimaksudkan untuk mengukur kemampuan peserta didik memahami isi informasi yang terdapat dalam wacana tertulis. Pemilihan wacana sebagai bahan tes membaca hendaknya mempertimbangkan berbagai faktor, antara lain: tingkat kesulitan wacana, isi dan cakupan wacana, panjang-pendek wacana, dan jenis wacana. (Burhan, 2010:371).

Sebagaimana wacana dalam tes menyimak, tingkat kesulitan wacana dalam tes membaca ini juga terkait erat dengan tingkat kerumitan kosa kata dan struktur kalimat yang dipergunakan, serta isi dan cakupan wacana. Wacana yang baik untuk tes keterampilan membaca adalah wacana yang tingkat kesulitannya sedang atau sesuai dengan kemampuan peserta didik.

Di samping itu, wacana yang ditekankan hendaknya tidak terlalu panjang. Sebaiknya menggunakan wacana pendek berkisar satu atau dua paragraph, atau kira-kira 50 sampai 100 kata. Jenis wacana yang dipergunakan sebagai bahan tes keterampilan membaca dapat berupa wacana jenis prosa non-fiksi, dialog, tabel, diagram, iklan, dan lain-lain.

Soal yang ditanyakan dalam tes keterampilan membaca ini umumnya mencakup: mengungkapkan kembali fakta, menemukan tema, gagasan pokok, gagasan pendukung, makna tersurat dan tersirat, bahkan juga makna istilah dan ungkapan. Jadi, tes kosa kata dapat pula disisipkan di sini. Soal tes membaca dapat juga hanya terdiri dari satu atau dua kalimat atau pernyataan, kemudian disediakan pilihan jawaban yang sesuai dengan pernyataan dalam soal.

Jenis tes yang digunakan untuk tes keterampilan menyimak bisa berbentuk tes obyektif pilihan ganda. Dilihat dari cara kerja peserta tes dan koreksi jawaban, jenis tes ini lebih praktis, cara penilaian atau pemberian skornya pun lebih obyektif. Ditambah lagi, jenis tes ini dapat mencakup macam-macam wacana dan banyak soal, walaupun pembuatan soalnya lebih sulit dan lebih lama.

c. Tes Keterampilan Berbicara (مهارة الكلام)

Di antara tes kemampuan berbicara dapat berupa tes berbicara berdasarkan gambar, wawancara, dialog, menceritakan peristiwa atau pengalaman, dan pidato.

Berdasarkan tingkat kemampuan peserta didik, Abu Hamid (2006:42-45) membagi tes keterampilan berbicara menjadi tiga macam:

1) Tes berbicara terikat (اختبار الكلام المقيد)

Jenis tes ini sesuai untuk pembelajar tingkat pemula (المستوى المبتدئ). Contoh tes jenis ini adalah tes berbicara berdasarkan gambar.

2) Tes berbicara terbimbing (اختبار الكلام الموجه)

Jenis tes ini sesuai untuk pembelajar tingkat intermidate (المستوى المتوسط). Jenis tes ini memberikan kebebasan lebih dari jenis tes pertama, namun belum bebas sepenuhnya, masih menggunakan bimbingan atau arahan. Misalnya tes berbicara dalam situasi tertentu:

ماذا تقول إذا ذكر لك صديقك أنه سيزورك هذا المساء ولكنك مشغول؟

3) Tes berbicara bebas (اختبار الكلام الحر)

Jenis tes ini sesuai untuk pembelajar tingkat lanjut/advance. Peserta tes diberi kebebasan penuh untuk mengungkapkan pikiran, gagasan, opini, dan perasaannya tentang sebuah tema atau persoalan tertentu. Misal:

هل توافق على جعل الدراسة بالنظام الفصلي أو بنظام الساعات؟ ولماذا؟

4) Tes Keterampilan Menulis (مهارة الكتابة)

Dalam mengukur kemampuan menulis, ada beberapa aspek yang tercakup di dalamnya, antara lain:

- a) *Al-Kitabah al-Syakliyah*, yaitu teknis penulisan huruf-huruf hijaiyah dan tanda baca sesuai dengan kaidah penulisan (القواعد الإملائية).
- b) *Al-Mufradat*, yaitu pemilihan dan penggunaan kosa kata dalam konteks kalimat yang sesuai.
- c) *Al-Shiyagh al-Sharfiyah wa al-Tarakib al-Nahwiyah*, yaitu penggunaan bentuk kata dan struktur kalimat yang tepat.
- d) *Al-Mustawa al-Lughawiy*, yaitu pemilihan dan penggunaan gaya bahasa formal atau informal sesuai dengan konteks dan situasi komunikasi.
- e) *Al-Mafahim al-Tsaqafiyah wa al-Hadhariyah*, yaitu penggunaan bahasa sesuai dengan konteks budaya penutur bahasa target atau bahasa yang dipelajari. (Abu, 2006:62-63)

Berdasarkan tingkat kemampuan peserta didik, tes keterampilan menulis dikelompokkan menjadi empat jenis tes tertulis, yaitu:

- a) Tes imla' (الاختبار الإملائي), yaitu tes menulis atau dikte untuk mengukur kemampuan menulis huruf-huruf dan menggunakan tanda baca dengan benar.
- b) Tes menulis terikat (اختبار الكتابة المقيدة), jenis tes ini sesuai untuk pembelajar tingkat pemula. Contoh tes ini antara lain: menuliskan pertanyaan yang sesuai dengan pernyataan, merangkai urutan kata-kata agar menjadi kalimat sempurna, merangkai urutan kalimat agar menjadi cerita sederhana, membuat kalimat sederhana berdasarkan gambar, dan lain-lain.
- c) Tes menulis terbimbing (اختبار الكتابة الموجهة), jenis tes ini sesuai untuk pembelajar tingkat intermidate. Contoh tes ini antara lain: merubah kalimat berita menjadi kalimat pertanyaan, kalimat positif menjadi kalimat negatif, kalimat dengan *fi'il madhiy* menjadi *mudhari'* atau sebaliknya, dan lain-lain.

- d) Tes menulis bebas (اختبار الكتابة الحرة), Jenis tes ini sesuai untuk pembelajar tingkat lanjut/advance. Peserta tes diberi kebebasan penuh untuk mengungkapkan pikiran, gagasan, opini, dan perasaannya tentang sebuah tema atau persoalan tertentu (Abu,2006:66-77).

D. JENIS TES BAHASA ARAB

Tes bahasa dapat dikelompokkan menjadi beraneka ragam jenis tes berdasarkan berbagai kriteria pengelompokannya. Misalnya, berdasarkan kriteria tujuan pelaksanaan tes, waktu pelaksanaan, pendekatan tes, cara menjawab soal, cara penilaian atau pemberian skor, dan lain-lain (Soenardi,1996:16-36). Pembagian jenis tes dalam tulisan ini terbatas pada kriteria cara menjawab soal dan cara penilaian atau pemberian skor.

1. Berdasarkan Kriteria Menjawab Soal

Dilihat dari cara mengerjakan soal atau cara menjawabnya, tes bahasa Arab dapat dibedakan menjadi tes tertulis dan tes lisan (Ainin, 2006:115-116). Tes tertulis merupakan tes yang cara menjawab pertanyaan atau mengerjakan soal dilakukan secara tertulis. Sedangkan cara memberikan pertanyaannya bisa dalam bentuk tertulis dan bisa juga dalam bentuk lisan. Tes tertulis ini dapat diberlakukan baik untuk tes pemahaman unsur bahasa Arab (seperti bunyi, kosa kata dan struktur kalimat), maupun untuk tes keterampilan berbahasa (seperti menyimak, membaca, dan menulis).

Adapun tes lisan adalah tes yang cara menjawab pertanyaan atau mengerjakan soalnya dilakukan secara lisan. Tes lisan ini sangat tepat untuk mengukur kemampuan berbicara, baik dari aspek ketepatan qawaid dan pilihan mufradatnya, kefasihan dalam melafalkan kata, ketepatan dalam memberikan informasi atau merespon informasi, serta intonasi. Selain itu, tes lisan dapat pula digunakan untuk mengukur kemampuan membaca, mencakup pemahaman isi teks, kelancaran membaca, kefasihan dalam melafalkan kata, intonasi dan kelancaran membaca. Bisa juga untuk mengukur kemampuan menyimak, terkait dengan kemampuan merespon secara lisan tentang pertanyaan atau wacana lisan.

2. Berdasarkan Kriteria Penilaian

Selanjutnya, dilihat dari kriteria cara penilaian atau pemberian skor, tes bahasa dapat dibedakan menjadi tes subyektif (الاختبار الذاتي) dan tes obyektif (الاختبار الموضوعي). Tes subyektif (الاختبار الذاتي) adalah tes yang penilaian atau pemberian skor terhadap jawaban dipengaruhi oleh kesan dan pendapat pribadi penilai (Ainin, 2006:128). Penilai yang

berbeda kemungkinan akan menghasilkan skor yang berbeda pula sesuai dengan cara pandang penilai itu sendiri. Bentuk tes yang mengimplikasikan cara penilaian subyektif ini adalah tes essay (الاختبار المقالي). Hal ini dikarenakan jawaban dalam tes essay berupa uraian bebas mengenai gagasan, ide, dan pikiran peserta didik (peserta tes) yang dikemukakan sebagai respon terhadap pertanyaan/soal.

Kelebihan tes essay ini antara lain: (a) penyusunan soalnya lebih mudah dibandingkan tes pilihan ganda; (b) dapat mengukur kemampuan berpikir peserta didik secara kritis, analitis, dan komprehensif; (c) dapat menghindarkan peserta didik dari kemungkinan memberikan jawaban secara spekulatif (untung-untungan/tebak-tebakan); dan (d) dapat mengukur hasil pembelajaran yang kompleks.

Sementara itu, kelemahan tes essay ini antara lain: (a) reliabilitas penskorannya kurang terjamin karena subyektivitas pemberi skor sulit dihindari; (b) waktu yang dibutuhkan untuk mengoreksi jawaban relatif lama, apalagi jika jumlah peserta didik (peserta tes) besar; dan (c) bahan/materi yang diteskan kurang merepresentasikan bahan/materi yang diajarkan secara keseluruhan (Ainin, 2006:119).

Untuk mengurangi kadar subyektivitas dalam pemberian skor atau untuk meningkatkan tingkat keterandalan pemberian skor yang bersifat subyektif, dapat dilakukan dengan sebagai berikut:

- a. perlu disusun rambu-rambu jawaban, sehingga penilai lebih konsisten dalam pemberian skor pada setiap butir soal;
- b. perlu ada pembobotan skor untuk masing masing soal, mengingat masing-masing butir soal untuk tes essay dimungkinkan memiliki tingkat kesulitan dan kompleksitas yang berbeda;
- c. perlu diperhatikan representasi materi yang dipilih sebagai bahan tes. Materi yang dipilih tersebut bukan didasarkan pada pertimbangan materi yang paling disenangi oleh pembuat soal, bukan pula atas pertimbangan materi yang paling mudah atau sulit, melainkan materi yang merepresentasikan bahan ajar secara menyeluruh.
- d. Sebaiknya penilaian dilakukan oleh lebih dari satu orang, selanjutnya skor dari masing-masing penilai digabungkan.
- e. Penilaian dilakukan lebih dari sekali dan skor dari penilaian pertama digabungkan dengan skor penilaian berikutnya (Ainin, 2006:128).

Adapun tes obyektif (الاختبار الموضوعي) adalah tes yang penilaian terhadap jawaban tidak terpengaruh oleh kesan dan pendapat pribadi penilai (Ainin, 2006:128). Dengan demikian, siapapun penilainya, kapanpun dan di manapun penilaian itu dilakukan, skor yang dihasilkan oleh tes obyektif adalah sama, dengan catatan adanya kepastian kebenaran kunci jawaban yang ada. Contoh tes obyektif adalah, tes menjodohkan, tes benar – salah (اختبار الصواب والخطأ) dan tes pilihan ganda (اختبار الاختيار من متعدد).

Kelebihan tes obyektif antara lain: (a) dapat mencakup bahan/materi tes yang representatif dan komprehensif (mencakup berbagai kemampuan berbahasa dan unsur bahasa); (b) penskorannya lebih mudah dan hasil penskorannya obyektif; dan (c) cocok untuk tes bahasa dengan jumlah peserta tes yang besar.

Sedangkan kelemahannya antara lain: (a) ada peluang bagi peserta tes untuk memberikan jawaban secara tebakan/untung-untungan; (b) penyusunan soalnya lebih lama dan lebih sulit, membutuhkan ketelatenan, ketelitian, dan profesionalitas yang memadai karena kompleksnya perihal yang harus diperhatikan, seperti: ketepatan dan keproporsionalan dalam penyusunan pilihan jawaban dan pengecohnya.

Untuk mengatasi kelemahan tersebut, maka dalam penyusunan soal pilihan ganda perlu diperhatikan rambu-rambu sebagai berikut (Sumarna, 2005: 194-195)(Hidayat, 2015):

a. Materi Soal

- 1) Soal harus sesuai dengan indikator pencapaian;
- 2) Pengecoh (pilihan jawaban) harus homogen, logis, dan berfungsi;
- 3) Setiap soal harus mempunyai satu jawaban yang benar/paling tepat;
- 4) Hindari menanyakan persoalan di luar kebahasaan.

b. Konstruksi Soal

- 1) Pokok soal dalam stem (butir soal) harus jelas dan tegas, tidak menimbulkan pengertian ganda, dan hanya mengandung satu persoalan untuk setiap butir;
- 2) Stem dan option (pilihan jawaban) harus merupakan pernyataan yang diperlukan saja;
- 3) Hindari pada stem penggunaan kata atau ungkapan yang persis sama dengan yang terdapat pada option;
- 4) Hindari penggunaan option (pilihan jawaban) seperti: جميع الإجابات السابقة صحيحة

- 5) Option yang berupa angka perlu disusun berdasarkan urutan, mulai dari yang besar ke yang kecil atau sebaliknya;
- 6) Hindari penggunaan kata atau ungkapan yang tidak pasti, seperti:

أكثرها – كثيرا – غالبا – أحيانا – عادة – تقريبا – الأحسن – أظنّ.

Kecuali jika bertujuan untuk menguji penggunaan kata-kata tersebut.

- 7) Option disusun dengan panjang kalimat atau ungkapan yang relatif sama;
- 8) Setiap butir soal berdiri sendiri dan tidak bergantung pada jawaban soal sebelumnya;
- 9) Gambar dan media lain yang menyertai soal harus jelas dan berfungsi;
- 10) Hindari ungkapan yang sama dalam option;
- 11) Letak jawaban yang benar disusun secara acak. Hindari penyusunan jawaban yang berpola, seperti: *aa, bb, cc, dd*, atau *ab, ac, ad*, dan seterusnya.

c. Bahasa Soal

- 1) Untuk mengukur pemahaman mufradat dan susunan kalimat, maka pada stem dan option digunakan bahasa yang benar. Sedangkan untuk mengukur *qawa'id*, maka tiga option memang sengaja “dibuat salah”;
- 2) Gunakan kalimat lengkap (*jumlah mufidah*). Jadi stem sendiri merupakan jumlah mufidah atau stem dan optionnya membentuk jumlah mufidah;
- 3) Dalam “tes lisan”/rekaman untuk menguji kemampuan menyimak dengan memahami, stem yang diperdengarkan harus merupakan pertanyaan yang lengkap;
- 4) Dalam tes untuk mengetahui hasil belajar (*achievement test*), digunakan bahasa yang sesuai dengan tingkat penguasaan peserta didik itu sendiri.

E. KARAKTERISTIK TES BAHASA ARAB YANG BAIK

Sebagaimana telah dijelaskan di atas, tes adalah salah satu jenis alat untuk memperoleh data numerik atau alat melakukan pengukuran yang hasilnya dimanfaatkan sebagai salah satu bahan pertimbangan dalam evaluasi. Namun, sebelum digunakan untuk mengukur suatu objek, perlu dipastikan terlebih dahulu kualitas tes tersebut. Suatu tes dipandang memiliki kualitas yang baik jika mempunyai tingkat validitas, reliabilitas, dan kepraktisan yang tinggi.

1. Validitas

Menurut Djiwandono (1996), validitas dapat diartikan adanya kesesuaian antara tes dengan apa yang ingin diukur dengan menggunakan tes itu. Istilah validitas sebenarnya bukan ciri yang terkait dengan tesnya sebagai alat, melainkan lebih pada kesesuaian hasilnya, yang dapat diinterpretasikan sesuai dengan tujuan penyelenggaraan tesnya. Jadi

tidak tepat jika diajukan pertanyaan: apakah tes itu valid, melainkan apakah hasilnya dapat diinterpretasikan sesuai dengan tujuan diselenggarakan tes itu. Selain itu, validitas bukanlah ukuran yang bersifat dikhotomis: valid – tidak valid, melainkan ditunjukkan dalam bentuk rentangan atau tingkatan: sangat tinggi, tinggi, sedang, rendah, sangat rendah (Sumarna, 2005: 22).

Hal senada diungkapkan oleh Mardapi (2008) bahwa pengertian validitas suatu tes mengacu tingkat kebenaran penafsiran skor tes. Penafsiran ini berdasarkan pada tujuan penggunaan tes. Dalam proses validasi, sebenarnya kita tidak bertujuan melakukan validasi tes tetapi melakukan validasi terhadap interpretasi data yang diperoleh melalui prosedur tertentu (Djemari, 2008:17)

Meskipun demikian, dalam penggunaan sehari-hari, secara umum dan demi kemudahan, validitas lebih sering dianggap sebagai ciri yang terkait dengan tes daripada interpretasi hasil tes. Ungkapan tes yang valid lebih umum dan lebih sering terdengar daripada hasil tes yang valid.

Jadi, hasil suatu tes dikatakan valid apabila hasil tes tersebut benar-benar menggambarkan kemampuan yang diukur/diteskan. Misalnya, jika seorang guru bermaksud mengukur kemampuan membaca peserta didik, maka ia menyusun tes yang terdiri atas teks bacaan dan pertanyaan-pertanyaan tentang kemampuan membaca. Melalui pertanyaan-pertanyaan tersebut, ia berusaha menanyakan isi teks bacaan, seperti (a) ide pokok; (b) ide penunjang; dan (c) fakta. Dengan tes tersebut, akan dapat diperoleh hasil tes dengan tingkat validitas yang relatif tinggi.

2. Reliabilitas

Istilah reliabilitas bermakna keajegan (konsisten) hasil/skor tes yang diperoleh orang yang sama ketika diuji ulang dengan tes yang sama pada situasi yang berbeda atau dari satu pengukuran ke pengukuran lainnya (Sumarna, 2005:89)(Soenardi, 1996:98). Jadi, istilah reliabilitas juga merupakan ciri yang terkait dengan hasil tes, bukan tes sebagai alat. Di samping itu, reliabilitas juga tidak bersifat dikhotomis: reliable – tidak reliable, melainkan ditunjukkan dalam rentangan atau tingkatan: sangat tinggi, tinggi, sedang, rendah, dan sangat rendah.

Tingkat reliabilitas paling tinggi secara statistik dituliskan dengan angka 1,00 yang menandakan adanya keajegan mutlak tanpa perbedaan atau penyimpangan sedikitpun. Reliabilitas mutlak ini umumnya hanya bersifat teoritis karena pada kenyataannya hampir tidak ada hasil tes yang mutlak ajeg, tanpa perbedaan, lebih-lebih tes dalam bidang yang memiliki banyak aspek, seperti pembelajaran bahasa Arab. Di dalam praktek dan kenyataannya, tingkat reliabilitas yang ditemukan selalu lebih rendah dari reliabilitas mutlak, dengan koefisien korelasi di bawah 1,00 seperti 0,95; 0,57; 0,31; dan sebagainya.

Ada beberapa faktor yang mempengaruhi reliabilitas, namun secara garis besar, dapat dikelompokkan menjadi dua, yaitu: (1) faktor instrumen, mencakup jumlah butir soal, homogenitas isi butir soal, dan tingkat kesulitan soal; dan (2) faktor subjek/individu, mencakup heterogenitas kemampuan individu, kemampuan memahami cara mengerjakan soal, dan motivasi, kesehatan, dan kelelahan individu (Ainin, 2006: 35-50).

3. Kepraktisan

Selain validitas dan reliabilitas, tes yang baik juga harus memiliki kriteria kepraktisan. Suharsimi Arikunto (2005) menyebutkan empat kriteria tes yang praktis sebagai berikut:

- a. Mudah dilaksanakan, misalnya tidak memerlukan peralatan yang banyak dan memberi kebebasan kepada peserta tes untuk mengerjakan terlebih dahulu bagian yang dianggap mudah oleh peserta tes.
- b. Mudah pemeriksaannya, artinya tes tersebut dilengkapi dengan kunci jawaban atau pedoman penskorannya dan lembar jawaban.
- c. Dilengkapi dengan petunjuk cara mengerjakan soal tes dengan jelas.
- d. Ekonomis, pelaksanaan tes tersebut tidak membutuhkan biaya yang mahal, tenaga yang banyak, dan waktu yang lama (Suharsimi, 2005:62).

F. LANGKAH PENYUSUNAN TES BAHASA ARAB

Dalam menyusun tes bahasa Arab, ada beberapa langkah atau tahapan yang perlu ditempuh, agar tes yang disusun memiliki kualitas yang baik dan dapat dipertanggungjawabkan. Langkah-langkah tersebut sebagai berikut:

1. Menelaah silabus/kurikulum

Langkah pertama yang harus ditempuh sebelum menyusun soal tes adalah menelaah silabus/kurikulum, yang mencakup kompetensi dasar, indikator, dan pokok/sub pokok bahasan, sebagai acuan dalam penulisan soal tes.

2. Membuat kisi-kisi soal tes

- a. Menuliskan kompetensi dasar dan indikator capaian hasil belajar yang terdapat dalam silabus/kurikulum.
- b. Menuliskan daftar pokok/sub pokok bahasan yang akan diujikan.
- c. Menentukan jumlah butir soal setiap pokok/sub pokok bahasan.

Jumlah soal hendaknya representatif untuk setiap pokok/sub pokok bahasan yang diujikan dengan pertimbangan pentingnya pokok/sub pokok bahasan tersebut. Selain itu, dalam menentukan jumlah soal juga perlu mempertimbangkan waktu yang disediakan untuk pelaksanaan tes.

d. Menentukan bentuk soal tes

Dalam menentukan bentuk tes, perlu mempertimbangkan karakteristik materi yang hendak diukur. Jika tes itu untuk mengukur pemahaman *muradat*, *qawa'id*, *istima'*, dan *qira'ah*, maka bisa digunakan bentuk tes pilihan ganda. Namun jika tes itu untuk mengukur kemampuan menulis dan berbicara, maka lebih tepat jika menggunakan bentuk tes unjuk kerja seperti mengarang, dialog, wawancara, menceritakan (gambar, peristiwa, atau pengalaman), pidato, dan lain-lain.

3. Menuliskan butir soal dan pedoman penilaian/pemberian skor

Penulisan butir soal harus mengacu dan sesuai dengan kisi-kisi soal yang telah dibuat sebelumnya. Di samping itu, penulisan soal juga harus memperhatikan rambu-rambu penulisan soal agar tidak banyak revisi.

4. Menelaah butir soal

Telaah soal ini perlu dilakukan untuk memperbaiki kekeliruan atau kesalahan yang mungkin masih ditemukan dalam penyusunan soal, sehingga dapat dilakukan revisi yang diperlukan. Telaah soal ini akan lebih baik jika dilakukan oleh orang yang ahli di bidangnya atau teman sejawat agar lebih cermat dan obyektif. Jika dimungkinkan, telaah soal sebaiknya dilakukan oleh lebih dari satu orang agar dapat saling melengkapi dan lebih meyakinkan.

5. Melaksanakan uji coba soal tes

Uji coba soal tes diperlukan agar soal tes tersebut benar-benar baik dan dapat dipertanggungjawabkan, misalnya akan dipergunakan untuk tes standar atau untuk penelitian. Namun, untuk keperluan pengujian di kelas sendiri, misalnya untuk ujian tengah semester dan akhir semester, uji coba tersebut tampaknya tidak dilakukan oleh guru.

6. Menganalisis butir soal dan jawaban

Berdasarkan data dari uji coba soal tes tersebut, selanjutnya dilakukan analisis butir soal dan jawaban dengan menggunakan rumus uji validitas dan reliabilitas hasil tes, serta uji tingkat kesukaran.

7. Memperbaiki dan merakit butir soal

Berdasarkan analisis soal tersebut, jika memang soal yang telah disusun belum memenuhi kualitas yang diharapkan, maka perlu diperbaiki atau direvisi seperlunya. Selanjutnya, butir-butir soal itu dirakit agar menjadi sebuah tes yang siap digunakan. Bentuk soal yang sejenis disusun dalam satu kelompok, dan butir soal diurutkan berdasarkan tingkat kesulitannya. Butir soal yang tingkat kesulitannya rendah diletakkan di nomor-nomor awal, dan yang tingkat kesulitannya tinggi ditempatkan di nomor-nomor akhir.

8. Melaksanakan tes

Agar tes tersebut dapat memberikan hasil yang benar, dikerjakan dengan jujur, dan sesuai dengan ketentuan yang telah digariskan, pelaksanaannya harus dilakukan sebaik mungkin dengan pengawasan yang cermat, tetapi tidak mengganggu konsentrasi peserta tes.

9. Menafsirkan hasil tes

Hasil tes berupa data kuantitatif (skor) perlu diinterpretasikan sehingga menjadi nilai yang merupakan informasi mengenai ketercapaian hasil pembelajaran.

G. SIMPULAN

Evaluasi pembelajaran bahasa Arab tidak dapat dilakukan tanpa adanya proses pengukuran, dan untuk melakukan proses pengukuran diperlukan instrumen atau alat evaluasi, salah satunya adalah tes. Tes bahasa Arab yang berkualitas baik adalah tes yang memenuhi kriteria validitas, reliabilitas, dan kepraktisan.

Ruang lingkup tes bahasa dikelompokkan menjadi dua, yaitu: (1) tes unsur bahasa yang mencakup: tes *mufradat* dan *qawa'id*; dan (2) tes keterampilan berbahasa mencakup: tes keterampilan menyimak (مهارة الاستماع), keterampilan membaca (مهارة القراءة), keterampilan berbicara (مهارة الكلام), dan keterampilan menulis (مهارة الكتابة).

Tes bahasa dapat dikelompokkan menjadi beraneka ragam jenis tes berdasarkan berbagai kriteria pengelompokannya. Misalnya, berdasarkan kriteria tujuan pelaksanaan tes, waktu pelaksanaan, pendekatan tes, cara menjawab soal, cara penilaian atau pemberian skor, dan lainnya.

Dalam penyusunan tes bahasa Arab, perlu diperhatikan rambu-rambu serta panduan langkah-langkah penyusunan tes agar soal tes yang disusun terjamin kualitasnya dan dapat dipertanggungjawabkan.

DAFTAR PUSTAKA

- Ainin, M., dkk., 2006, *Evaluasi dalam Pembelajaran Bahasa Arab*, Malang: Misykat.
- Al-Khuliy, Muhammad 'Ali, 2000, *al-Ikhtibarat al-Lughawiyah*, al-Urdun: Dar al-Falah.
- Arikunto, Suharsimi, 2005, *Dasar-Dasar Evaluasi Pendidikan*, Jakarta: Bumi Aksara.
- Djiwandono, M. Soenardi, 1996, *Tes Bahasa dalam Pengajaran*, Bandung: ITB.
- Heaton, J.B., 1988, *Writing English Language Tests*, New York: Longman Inc.
- Hidayat, D. 2011, *Menyusun Materi Tes Pilihan Ganda*, Program Pasca Sarjana Pendidikan Bahasa Arab UIN Maulana Malik Ibrahim Malang.
- Mardapi, Djemari, 2008, *Teknik Penyusunan Instrumen Tes dan Nontes*, Jogjakarta: Mitra Cendikia.
- Nurgiantoro, Burhan, 2010, *Penilaian Pembelajaran Bahasa Berbasis Kompetensi*, Yogyakarta: BPFE-Yogyakarta.
- Surapranata, Sumarna, 2005, *Analisis, Validitas, Reliabilitas, dan Intrepetasi Hasil Tes: Implementasi Kurikulum 2004*, Bandung: Remaja Rosdakarya.
- , 2005, *Panduan Penulisan Tes Tertulis*, Bandung: Remaja Rosdakarya.