



SEMINAR NASIONAL KIMIA DAN PENDIDIKAN KIMIA VI
"Pemantapan Riset Kimia dan Asesmen Dalam Pembelajaran
Berbasis Pendekatan Saintifik"
Program Studi Pendidikan Kimia Jurusan PMIPA FKIP UNS
Surakarta, 21 Juni 2014



MAKALAH UTAMA

MAKALAH UTAMA

ISBN : 979363174-0

ASESMEN DALAM PEMBELAJARAN BERBASIS SAINTIFIK

Oleh: Prof. Djemari Mardapi, Ph.D*)

Abstrak

Asesmen atau penilaian merupakan kegiatan untuk menentukan pencapaian belajar peserta didik. Pencapaian belajar yang diharapkan pada Kurikulum 2013 dapat dilihat pada standar kompetensi lulusan. Melalui asesmen akan diperoleh informasi tentang pencapaian belajar peserta didik yang selanjutnya digunakan oleh peserta didik dan pendidik untuk memperbaiki proses pembelajaran. Metode asesmen digunakan berkaitan dengan metode pembelajaran yang digunakan hasil, yang digunakan. Pendekatan pembelajaran pada Kurikulum 2013 seperti pada kurikulum tingkat satuan pendidikan menekankan pada peserta didik. Namun penambahannya adalah menggunakan pendekatan saintifik Pendekatan saintifik pada dasarnya dimulai dengan bertanya, mengamati, menganalisis, dan membuat kesimpulan. Untuk itu asesmen yang digunakan sesuai dengan pendekatan pembelajaran. Asesmen yang bersifat saintifik harus mendorong peserta didik belajar, berpikir lebih luas, dan untuk itu instrumen asesmen yang digunakan harus bersifat divergen. Asesmen otentik merupakan salah bentuk penilaian yang relevan dengan pendekatan saintifik karena menggunakan masalah nyata di lapangan. Persyaratan yang diperlukan agar hasil penilaian akurat adalah instrumen yang digunakan harus memiliki bukti validitas dan reliabilitas. Instrumen penilaian tidak selalu bentuk tes tertulis bisa berupa pedoman observasi, namun tetap harus memiliki bukti validitas dan reliabilitas. Berbagai bentuk cara penilaian yang sesuai dengan pendekatan pembelajaran yang digunakan dibahas pada makalah ini.

Kata kunci: *Asesmen, pendekatan saintifik, validitas, dan reliabilitas*

A. Prinsip Pembelajaran

Lampiran Permen 65 tahun 2013 tentang standar proses menjelaskan sebagai berikut. Karakteristik pembelajaran pada setiap satuan pendidikan terkait erat pada Standar Kompetensi Lulusan dan Standar Isi. Standar Kompetensi Lulusan memberikan kerangka konseptual tentang sasaran pembelajaran yang harus dicapai. Standar Isi memberikan kerangka konseptual tentang kegiatan belajar dan pembelajaran yang diturunkan dari tingkat

kompetensi dan ruang lingkup materi. Sesuai dengan Standar Kompetensi Lulusan, sasaran pembelajaran mencakup pengembangan ranah sikap, pengetahuan, dan keterampilan yang dielaborasi untuk setiap satuan pendidikan.

Ketiga ranah kompetensi tersebut memiliki lintasan perolehan (proses psikologis) yang berbeda. Sikap diperoleh melalui aktivitas "menerima, menjalankan, menghargai, menghayati, dan mengamalkan". Pengetahuan diperoleh melalui aktivitas "mengingat, memahami, menerapkan,

menganalisis, mengevaluasi, mencipta. Keterampilan diperoleh melalui aktivitas “ mengamati, menanya, mencoba, menalar, menyaji, dan mencipta”. Karakteristik kompetensi beserta perbedaan lintasan perolehan turut serta mempengaruhi karakteristik standar proses. Untuk memperkuat pendekatan ilmiah (scientific), tematik terpadu (tematik antarmata pelajaran), dan tematik (dalam suatu mata pelajaran) perlu diterapkan pembelajaran berbasis penyingkapan/penelitian (discovery/inquiry learning). Untuk mendorong kemampuan peserta didik untuk menghasilkan karya kontekstual, baik individual maupun kelompok maka sangat disarankan menggunakan pendekatan pembelajaran yang menghasilkan karya berbasis pemecahan masalah (project based learning).

Penjelasan diatas dapat dimaknai bahwa pembelajaran yang dilakukan di satuan pendidikan menggunakan pendekatan ilmiah. Tujuan semua usaha yang bersifat saintifik adalah untuk menjelaskan, mempreiksi, dan mengendalikan phenomena. Pendekatan dapat dilakukan dengan menggunakan penalaran induktif dan penalaran dedukif. Penalaran induktif pada dasarnya mulai dari melakukan pengamatan, menganalisis, dan menyimpulkan. Penalaran dedukif dilakukan dengan menentukan masalah dan kemudian menentukan hipotesis, mengumpulkan data, menganalisis data, dan membuat kesimpulan, yang kemudian digeneralisasi pada populasi yang diteliti. Jadi pada pendekatan ini peserta didik dituntut aktif bertanya, mengamati, menganalisis, dan menyimpulkan.

A. Asesmen

Asesmen atau penilaian merupakan rangkaian kegiatan dalam melaksanakan pembelajaran di pendidikan dasar dan menengah, dan pendidikan tinggi. Kegiatan pendidik atau dosen dalam melaksanakan pembelajaran melalui tiga tahapan, yaitu persiapan, pelaksanaan, dan penilaian. Untuk melaksanakan kegiatan tersebut dengan baik, pendidik harus memiliki kompetensi merancang proses pembelajaran, melaksanakan proses pembelajaran, dan menilai proses dan hasil pembelajaran.

Ada empat istilah yang sering digunakan berkaitan dengan kegiatan penilaian yaitu pengukuran, pengujian, penilaian, dan evaluasi. Pertama adalah pengukuran, yaitu penetapan angka terhadap suatu objek atau gejala dengan cara yang sistematis (Allen & Yen, 1979). Akurasi penetapan angka ini ditentukan oleh kualitas instrumen dan cara menggunakan instrumen ini, yang kemungkinan memiliki kesalahan pengukuran (Johnson & Johnson: 2, 2003). Kesalahan pengukuran ini bisa disebabkan oleh alat ukur, objek yang diukur, subjek yang mengukur, dan lingkungan pengukuran. Kesalahan ini ada yang bersifat acak dan ada yang sistematis. Kesalahan yang bersifat acak ini dapat diestimasi dengan menggunakan beberapa asumsi, sedang kesalahan yang sistematis sulit diestimasi besarnya, hanya arahnya yang kemungkinan dapat diduga.

Kedua adalah pengujian, yaitu kegiatan untuk mengetahui pencapaian belajar atau kompetensi yang dicapai peserta didik. Pengujian dilakukan dengan menggunakan tes yang terdiri atas sejumlah pertanyaan yang memiliki jawaban benar atau salah. Banyak bentuk tes yang dapat digunakan, untuk pengujian masing-masing memiliki kelebihan dan keterbatasan. Ketiga adalah penilaian, yaitu kegiatan mengumpulkan informasi tentang kualitas atau kuantitas perubahan pada peserta didik. Penilaian juga diartikan sebagai kegiatan menafsirkan data hasil pengukuran, misalnya tinggi, rendah, baik, buruk, indah, jelek, dan sebagainya. Penilaian berfokus pada individu sedang evaluasi berfokus pada kelompok atau kelas. Pembelajaran pada dasarnya adalah kegiatan melakukan perubahan pada peserta didik yang hasilnya harus diketahui. Untuk mengetahui besar dan kualitas perubahan dilakukan penilaian. Jadi penilaian merupakan hal yang penting dalam melaksanakan pembelajaran.

Keempat adalah evaluasi, yaitu merupakan salah satu rangkaian kegiatan dalam meningkatkan kualitas, kinerja, atau produktivitas suatu lembaga dalam melaksanakan programnya. Kegiatan evaluasi sering menggunakan *judgment* terhadap hasil suatu penilaian. Bila hasilnya bagus

terus dan bila tidak baik berhenti, atau walau hasilnya kurang baik program terus dilaksanakan dengan melakukan perbaikan-perbaikan, karena termasuk program prioritas utama.

Melalui evaluasi akan diperoleh informasi tentang apa yang telah dicapai dan mana yang belum, dan selanjutnya informasi ini digunakan untuk perbaikan suatu program. Menurut Tyler (1950), evaluasi adalah proses penentuan sejauh mana tujuan pendidikan telah tercapai. Evaluasi menurut Griffin & Nix (1991) adalah *judgment* terhadap nilai hasil pengukuran atau implikasi dari hasil pengukuran. Tyler menekankan pada pencapaian tujuan suatu program, sedang Griffin & Nix lebih menekankan pada penggunaan hasil penilaian. Kegiatan penilaian dan evaluasi memerlukan data hasil pengukuran

Alat ukur yang sering digunakan untuk mengumpulkan data bisa berupa tes dan nontes. Tes yang baik harus mewakili domain yang diukur dan mengukur tingkat berfikir yang tepat. Domain yang diukur ini adalah bahan ajar yang dapat dilihat pada silabus mata pelajaran. Informasi yang akurat berasal dari sejumlah pengukuran, yaitu yang dilaksanakan sepanjang semester. Bentuknya bisa berupa tugas, ulangan mingguan, dan ulangan pertengahan dan ulangan akhir semester. Dengan demikian pencapaian belajar peserta didik pada suatu bidang studi dapat diketahui.

Alat ukur yang digunakan harus diusahakan agar memberikan data yang sah (valid) dan andal (reliable) Sahih berkaitan dengan sampel bahan ajar yang diujikan, karena waktu untuk ujian sangat terbatas, khususnya ujian pertengahan semester dan akhir semester. Andal berkaitan dengan kesalahan pengukuran, yang sering dinyatakan dengan koefisien keandalan.

Penilaian hasil pembelajaran harus mencakup ranah kognitif, afektif, dan psikomotorik. Sesuai dengan karakteristiknya teknik penilaian untuk ketiga ranah tersebut tidak sama. Untuk ranah kognitif bisa dilakukan dengan tes tertulis, ranah psikomotor dengan tes perbuatan, dan ranah afektif melalui pengamatan, inventori, atau kuesioner. Hasil penilaian ketiga ranah tersebut

akan memberikan informasi tentang kompetensi peserta didik. Informasi ini diperlukan untuk merancang program perbaikan atau remedi

B. Hasil Pembelajaran

Hasil Pembelajaran yang diharapkan dari peserta didik disebut standar kompetensi lulusan. Standar kompetensi ini mencakup tiga aspek, yaitu aspek kognitif, aspek psikomotor, dan aspek afektif. Aspek atau ranah kognitif berkaitan dengan kemampuan berpikir, dan menurut Anderson dan Krathwohl (2001) ada enam kategori dimensi proses kognitif, yaitu pengingatan, pemahaman, aplikasi, analisis, evaluasi, dan kreasi. Pengingatan adalah menghafal pengetahuan yang relevan dalam memori jangka panjang. Pada tingkat pemahaman mahasiswa membangun makna dari pesan pembelajaran melalui lisan, tertulis, dan/atau komunikasi grafik. Aplikasi adalah menerapkan pengetahuan atau menggunakan suatu prosedur dalam suatu situasi baru. Analisis adalah menguraikan materi ke dalam beberapa bagian dan menentukan bagaimana masing-masing bagian berhubungan satu dengan lainnya, dan terhadap keseluruhan struktur. Evaluasi adalah kegiatan membuat *judgment* berdasarkan suatu kriteria atau standar. Terakhir, kreasi adalah kegiatan menyatukan semua elemen untuk membentuk suatu yang koheren dan menjadi suatu atau menyusun elemen-elemen menjadi bentuk pola yang baru atau struktur yang baru.

Kemampuan berpikir peserta didik menurut Marzano dan Kendall (1996) berdasarkan tingkatannya adalah memahami dan menerapkan konsep yang ada dalam suatu mata pelajaran. Kategori yang rendah adalah memahami sedang yang tinggi adalah menerapkan dalam berbagai situasi. Untuk mencapai kemampuan yang lebih tinggi, peserta didik harus melalui tingkat berpikir di bawahnya terlebih dahulu.

Dalam melakukan penilaian tingkat berpikir peserta didik dikaitkan dengan karakteristik mata pelajaran. Pengetahuan dapat dikategorikan

menjadi dua, yaitu pengetahuan deklaratif dan pengetahuan prosedural (Marzano & Kendall, 1996). Pengetahuan deklaratif dinyatakan sebagai informasi dan biasanya melibatkan komponen bagian-bagian. Sebagai contoh, pengetahuan tentang konsep demokrasi mencakup keputusan bahwa tiap orang hanya memiliki satu suara dalam suatu pemilihan. Pengetahuan deklaratif terbatas pada paham akan suatu konsep.

Pengetahuan prosedural menuntut peserta didik mampu menerapkan konsep-konsep yang ada dalam suatu mata pelajaran. Pengetahuan tingkat ini menuntut tingkat berpikir peserta didik yang lebih tinggi dibanding pengetahuan deklaratif. Oleh karena itu dalam mengembangkan sistem pengujian, pengetahuan deklaratif dan pengetahuan prosedural harus tampak.

Menurut Kratwohl (1964) proses internalisasi ranah afektif melalui lima tingkatan, yaitu penerimaan, pemberian respon, penilaian, pengorganisasian, dan karakterisasi. Penilaian pencapaian kompetensi pada ranah ini ditempuh melalui pengamatan sehari-hari. Perubahan pada ranah ini tidak bisa cepat, tetapi melalui proses panjang. Untuk itu pendidik harus menyusun rencana pembelajaran pada aspek afektif ini, kemudian dilakukan pengamatan yang hasilnya digunakan untuk perbaikan.

Berkenaan dengan ranah afektif, kompetensi yang ingin dicapai meliputi tingkatan pemberian respon (*responding*), apresiasi (*appreciating*), penilaian (*valuing*), dan internalisasi (*internalization*). Pengalaman belajar yang relevan dengan berbagai jenis tingkatan afektif tersebut antara lain: berlatih memberikan respon atau reaksi terhadap nilai-nilai yang dihadapkan kepadanya, berlatih menikmati atau menerima nilai, norma, serta objek yang mempunyai nilai etika dan estetika; berlatih menilai ditinjau dari segi baik buruk, adil tidak adil, indah tidak indah terhadap objek studi; berlatih menerapkan/mempraktekkan nilai, norma, etika dan estetika dalam perilaku kehidupan sehari-hari. Secara kongkrit, pengalaman belajar yang perlu dilakukan agar peserta didik

mencapai berbagai tingkatan kompetensi afektif tersebut antara lain dengan mengamati dan menirukan perilaku contoh/model/panutan, mendatangi objek studi yang dapat memupuk pertumbuhan nilai, berbuat atau berpartisipasi aktif sesuai dengan tuntunan nilai yang dipelajari.

Ranah psikomotor berkaitan dengan gerak, dan menurut Simpson proses pembelajarannya melalui 6 tingkatan, yaitu persepsi atau kesadaran, kesiapan penyesuaian, respons terbimbing, mekanis, respons yang kompleks, penyesuaian dan pengembangan. Pembelajaran ranah atau aspek psikomotor pada umumnya dimulai dengan demonstrasi yang dilakukan pendidik. Peserta didik pada kegiatan demonstrasi adalah memperhatikan sambil menirukan gerak. Ada peserta didik yang mampu belajar sampai tingkat tinggi, tetapi ada yang tidak, tergantung pada bakat peserta didik.

Berkenaan dengan ranah psikomotorik, kompetensi yang dicapai meliputi tingkatan gerakan awal, semi rutin, gerakan rutin. Untuk mencapai kompetensi tersebut, pengalaman belajar yang perlu dilakukan antara lain: pada tingkatan penguasaan gerakan awal, peserta didik perlu berlatih menggerakkan sebagian anggota badan, pada tingkatan gerakan semi rutin, peserta didik berlatih, mencoba, atau menirukan gerakan yang melibatkan seluruh anggota badan. Pada tingkatan gerakan rutin peserta didik melakukan gerakan secara menyeluruh dengan sempurna sampai pada tingkatan otomatis. Pengalaman belajar yang umum dilakukan untuk mencapai ketiga tingkatan tersebut adalah berlatih dengan frekuensi tinggi dan intensif (*drill*), menirukan, mensimulasikan, mendemonstrasikan gerakan yang ingin dikuasai. Misalnya, peserta didik sekolah penerbang mensimulasikan cara menerbangkan pesawat dengan menggunakan simulator pesawat.

C. Kualitas Instrumen

Kualitas instrumen alat ukur dapat dilihat pada bukti kesahihan (*validity*), dan besarnya koefisien

keandalan (reliability). Ada tiga kesahihan alat ukur, yaitu kesahihan konstruk, kesahihan isi, dan kesahihan terkait kriteria. Bukti kesahihan konstruk alat ukur dilihat dari teori yang digunakan dan dimensi yang diukur. Pada umumnya dimensi yang diukur adalah satu. Bukti kesahihan isi dilihat dari kesesuaian antara materi yang diujikan dan yang diajarkan, yaitu yang terdapat pada kurikulum Kesahihan terkait kriteria merupakan hubungan antara prediktor dan kriteria. Prediktor pada suatu sistem pengujian adalah tes masuk, sedang kriterianya adalah keberhasilan belajar atau keberhasilan kerja.

Keandalan alat ukur dapat dikategorikan menjadi tiga, yaitu konsistensi internal, stabilitas dan antar penilai. Pada konsistensi internal, alat ukur digunakan sekali kemudian dihitung besarnya koefisien keandalannya. Ada banyak formula yang dapat digunakan untuk mencari besarnya koefisien keandalan. Perbedaan formula disebabkan asumsi yang digunakan berbeda. Oleh karena itu sebelum menghitung koefisien keandalan alat ukur, diteliti dulu asumsi yang digunakan.

Alat ukur yang komparabel adalah yang hasil pengukurannya dapat dibandingkan antar tempat dan waktu. Komparabel bisa dicapai apabila digunakan kisi-kisi sama, tujuan sama, dan skala yang sama. Apabila diperoleh skala yang sama, maka hasil pengukuran tiap peserta didik dapat dibandingkan. Prinsip ini diterapkan dengan menggunakan bank soal. Butir-butir yang ada dalam bank soal telah dikalibrasi, yaitu telah memiliki parameter butir. Parameter butir berupa tingkat kesulitan dan daya beda butir.

Alat ukur yang digunakan harus *feasible*, yaitu dapat diterapkan di satuan pendidikan. Penerapan alat ukur dilihat dari relatif efisiensi suatu alat ukur. Relatif efisiensi dilihat dari informasi yang diperoleh. Apabila jumlah butir dua alat ukur yang mengukur hal yang sama, maka yang dipilih adalah yang lebih sedikit jumlah butirnya tetapi informasinya sama dengan jumlah butir yang banyak. Oleh karena itu dalam mengembangkan sistem penilaian perlu

ditelaah kemungkinan penggunaan instrument di satuan pendidikan.

D. Acuan Norma dan Kriteria

Ada dua acuan yang digunakan untuk menafsirkan hasil tes, yaitu acuan norma dan kriteria. Kedua acuan ini menggunakan asumsi yang berbeda. Penafsiran hasil tes kedua acuan ini berbeda sehingga menghasilkan informasi yang berbeda maknanya. Pemilihan acuan yang tepat ditentukan oleh karakteristik bidang studi dan kurikulum yang digunakan.

Tes acuan norma berasumsi bahwa kemampuan orang berbeda dan dapat digambarkan menurut distribusi normal. Perbedaan ini harus ditunjukkan oleh hasil pengukuran, misalnya setelah mengikuti kuliah selama satu semester peserta didik dites. Hasil tes seseorang dibandingkan dengan kelompoknya, sehingga dapat diketahui posisi seseorang. Acuan ini digunakan terutama pada tes untuk seleksi, karena sesuai dengan tujuannya tes seleksi adalah untuk mengetahui perbedaan kemampuan seseorang.

Acuan kriteria berasumsi bahwa hampir semua orang bisa belajar apa saja namun waktunya yang berbeda. Konsekuensi acuan ini adalah adanya program remedi dan program pengayaan, bagi yang belum mencapai kriteria harus belajar lagi melalui program remedi, sedang yang telah mencapai kriteria mengikuti program pengayaan. Penafsiran skor hasil tes selalu dibandingkan dengan kriteria yang telah ditetapkan lebih dahulu. Hasil tes dinilai lulus atau tidak. Lulus berarti bisa melakukan, tidak lulus berarti belum bisa melakukan. Acuan ini digunakan pada kurikulum berbasis kompetensi, karena jelas apa yang akan diukur.

Asesmen hasil pembelajaran peserta didik dilakukan oleh pendidik untuk memantau proses dan hasil pembelajaran peserta didik, yang selanjutnya digunakan untuk menyusun program perbaikan atau remedi. Asesmen kategori ini termasuk pada asesmen formatif. Menurut Cowie dan Bell (1999), asesmen formatif adalah proses yang digunakan oleh pendidik

dan peserta didik untuk mengetahui dan merespon pencapaian belajar peserta didik dengan tujuan untuk menguatkan selama aktivitas belajar berlangsung. Jadi wewenang dosen adalah melakukan asesmen selama pembelajaran berlangsung dengan tujuan untuk memperbaiki proses pembelajaran agar tujuan pembelajaran atau standar kompetensi tercapai.

Menurut Clarke (2005), asesmen formatif terdiri atas empat elemen dasar, yaitu: (1) tujuan pembelajaran, (2) pertanyaan yang efektif, (3) evaluasi peer-diri, (4) umpan balik yang efektif. Keempat elemen ini terjadi selama pembelajaran berlangsung, agar tujuan atau kompetensi yang ingin dicapai dapat dicapai peserta didik. Dosen atau pendidik dalam melaksanakan proses pembelajaran harus memperhatikan bagaimana peserta didik belajar. Hal ini berarti bahwa pendidik harus selalu mencari strategi mengajar yang memudahkan peserta didik belajar. Proses ini berlangsung selama kegiatan belajar mengajar, yaitu yang merupakan kegiatan asesmen formatif.

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, kompetensi dasar. Jadi tes ini sebenarnya bukan untuk menentukan keberhasilan belajar semata, tetapi untuk mengetahui keberhasilan proses pembelajaran.

Tes sumatif diberikan di akhir suatu pelajaran, atau akhir jenjang pendidikan. Menurut Griffin dan Nix (1991: 269), asesmen sumatif adalah asesmen penentuan final yang umumnya tidak berkaitan dengan penyempurnaan pembelajaran. Sebagai contoh adalah *grade* yang diberikan atas unjuk kerja peserta didik oleh penguji eksternal untuk pengakuan. Asesmen sumatif menggunakan tes sumatif yang hasilnya digunakan untuk menentukan keberhasilan belajar peserta didik untuk mata pelajaran tertentu atau semua mata pelajaran. Tingkat keberhasilan ini

dinyatakan dengan skor atau nilai, pemberian sertifikat, dan sejenisnya. Tingkat kesukaran soal pada tes sumatif bervariasi, sedang materinya harus mewakili bahan yang diajarkan. Hasil tes bisa ditafsirkan sebagai keberhasilan belajar, keberhasilan mengajar, serta keduanya.

Tingkat berpikir yang terlibat dalam sistem ujian mencakup pengetahuan deklaratif dan pengetahuan prosedural. Deklaratif berisi tentang konsep, prinsip, dan fakta-fakta, sedang prosedural mencakup proses, strategi, aplikasi, dan ketrampilan.

E. Bentuk Tes

Bentuk tes yang digunakan di program studi dapat dikategorikan menjadi dua, yaitu tes objektif dan tes nonobjektif. Objektif di sini dilihat dari sistem penskorannya, siapa saja yang memeriksa lembar jawaban tes akan menghasilkan skor yang sama. Tes yang nonobjektif adalah yang sistem penskorannya dipengaruhi oleh subjektivitas pemberi skor. Dengan kata lain dapat dikatakan bahwa tes yang objektif adalah yang sistem penskorannya objektif, sedang tes yang nonobjektif sistem penskorannya dipengaruhi oleh subjektivitas pemberi skor.

Ada beberapa bentuk soal ujian yang dapat digunakan, seperti berikut ini.

1. **Pertanyaan lisan di kelas:** pertanyaan lisan di kelas pada umumnya ditujukan pada kelompok, namun bisa individu dan dilakukan pada saat pembelajaran di kelas berlangsung atau bisa juga di awal pelajaran untuk materi pelajaran yang lalu, atau di akhir pelajaran untuk materi pelajaran hari itu. Peserta didik yang dipilih untuk ditanya harus mewakili karakteristik kemampuan individu di kelas. Bila ada konsep yang belum diketahui sebagian besar peserta didik, maka pendidik harus menjelaskan kembali dan bisa disertai dengan pemberian tugas. Pertanyaan lisan dilakukan dengan menggunakan teknik bertanya yang baik agar semua peserta didik berpikir. Pertanyaan yang diajukan harus mendorong peserta didik

berpikir dan untuk harus bersifat divergen, artinya jawabannya bisa lebih dari satu.

2. **Pilihan ganda:** bentuk ini bisa mencakup banyak materi pelajaran, penskorannya objektif, dan bisa dikoreksi dengan komputer. Namun membuat butir soal pilihan ganda yang berkualitas cukup sulit, dan kelemahan lain adalah peluang kerja sama antar peserta tes sangat besar. Oleh karena itu, bentuk ini dipakai untuk ujian yang melibatkan banyak peserta didik dan waktu untuk koreksi sedikit. Penggunaan bentuk ini menuntut agar pengawas ujian harus teliti dalam melakukan pengawasan saat ujian berlangsung. Tingkat berpikir yang diukur bisa tinggi tergantung pada kemampuan pembuat soal. Tes pilihan ganda bisa bersifat gradasi, artinya tiap jawaban memiliki makna tertentu dan diberi skor sesuai dengan tingkat kebenarannya.
3. **Uraian objektif:** bentuk ini cocok untuk mata pelajaran yang batasnya jelas seperti matematika dan IPA. Agar hasil penskorannya objektif diperlukan pedoman penskoran. Penskoran dilakukan secara analitik, setiap langkah pengerjaan diberi skor. Misalnya menuliskan rumusnya diberi skor, menghitung hasilnya diberi skor, dan menafsirkan atau menyimpulkan hasilnya, juga diberi skor. Penskoran bersifat herarki, sesuai dengan langkah pengerjaan soal. Bobot skor untuk tiap butir soal ditentukan oleh tingkat kesulitan butir soal, yang sulit bobotnya lebih besar dibandingkan dengan yang mudah. Objektif di sini berarti hasil penilaian terhadap suatu lembar jawaban akan sama walau diperiksa oleh orang yang berbeda asal memiliki latar belakang pendidikan sesuai dengan mata ujian. Tingkat berpikir yang diukur bisa sampai pada tingkat yang tinggi. Tes ini bisa mengembangkan kemampuan berpikir peserta didik, apabila cara yang dilakukan oleh peserta didik berbeda dengan cara pada umumnya, namun hasilnya sama.

4. **Uraian nonobjektif / uraian bebas:**

bentuk ini cocok untuk mengukur kemampuan berpikir divergen. Walau hasil penskoran cenderung subjektif, namun bila disediakan pedoman penskoran, sehingga hasilnya diharapkan dapat lebih objektif. Tingkat berpikir yang diukur bisa tinggi, namun cakupan materi yang diujikan terbatas. Bentuk soal ini sangat sesuai untuk pengembangan kreativitas peserta didik, asalkan pertanyaan yang dilakukan tidak memiliki jawaban yang benar hanya satu.

5. **Unjuk kerja:**

bentuk ini cocok untuk mengukur kemampuan seseorang dalam melakukan tugas tertentu, seperti praktek di laboratorium. Peserta tes diminta untuk mendemonstrasikan kemampuan dalam bidang tertentu. Kegaitan saintifik peserta didik dapat dinilai dengan bentuk ini, asalkan masalah yang diberikan agak bebas sesuai dengan minat dari peserta didik asal masih dalam rumpun keilmuan yang dipelajari.

6. **Portfolio:**

bentuk ini cocok untuk mengetahui perkembangan unjuk kerja peserta didik, dengan menilai kumpulan karya-karya, atau tugas-tugas yang dikerjakan peserta didik. Portfolio berarti kumpulan karya atau tugas-tugas yang dikerjakan peserta didik. Karya-karya ini dipilih kemudian dinilai, sehingga dapat dilihat perkembangan kemampuan peserta didik. Cara ini bisa dilakukan dengan baik bila jumlah peserta didik yang dinilai tidak banyak. Penilaian dengan cara ini bisa dilakukan untuk memenilai usaha, dan karya yang telah dihasilkan peserta didik.

F. **Kesahihan dan Keandalan Tes**

Suatu tes yang baik harus memiliki bukti kesahihan dan keandalan, hasilnya dapat dibandingkan, dan ekonomis. Kesahihan tes dapat dikategorikan menjadi tiga, yaitu kesahihan isi, konstruk, dan kriteria. Kesahihan isi dilihat dari bahan yang diujikan, kesahihan konstruk dilihat dari definisi atau teori yang diukur, dan kesahihan kriteria dilihat dari daya prediksinya.

Kesahihan isi dilihat dari kisi-kisi tes, yaitu matrik yang menunjukkan bahan tes serta tingkat berpikir yang terlibat dalam mengerjakan tes. Kesahihan ini ditelaah sebelum tes digunakan. Kesahihan konstruk diperoleh dari hasil analisis faktor, yaitu jumlah faktor yang diukur suatu tes. Bukti kesahihan konstruk diperoleh dari hasil penggunaan tes, yaitu data empirik. Kesahihan prediktif juga memerlukan data empirik untuk dapat menghitung besarnya daya prediksi.

Keandalan suatu tes memberikan informasi tentang besarnya kesalahan pengukuran. Keandalan suatu tes dapat dikategorikan menjadi tiga, yaitu konsisten internal, stabilitas, dan antar penilai. Besarnya indeks konsistensi internal diperoleh dari data hasil uji coba atau data hasil tes. Untuk mencari indeks ini cukup dilakukan satu kali tes. Indeks stabilitas merupakan tingkat kestabilan hasil pengukuran yang dilakukan paling tidak dua kali untuk orang yang sama dalam waktu yang berbeda, dengan asumsi tidak ada efek tes. Keandalan antar penilai diperoleh dari besarnya korelasi hasil pensekoran dari dua orang terhadap lembar jawaban tes yang sama. Besarnya indeks keandalan ini adalah 0 sampai 1,0 dan yang dapat diterima minimum 0,70.

Indek keandalan digunakan untuk menghitung besarnya kesalahan pengukuran. Kesalahan pengukuran ini ada dua, yaitu acak dan sistematis. Acak berarti kesalahan karena kondisi yang diukur dan yang mengukur bervariasi dan pemilihan bahan yang diujikan, sedang yang sistematis karena alat ukurnya atau cara pensekoran yang cenderung murah atau mahal untuk semua peserta didik.

Besarnya kesalahan pengukuran dapat dihitung dengan formula berikut ini:

$$S_e = S_x \sqrt{1 - \rho_{xx}}$$

S_e adalah besarnya kesalahan pengukuran, S_x simpangan baku sekor, dan ρ_{xx} adalah indeks keandalan tes. Formula di atas menunjukkan bahwa apabila indeks keandalan

tes besar maka kesalahan pengukuran kecil, dan sebaliknya bila indeks keandalan tes kecil maka kesalahan pengukuran besar.

DAFTAR BACAAN

- [1] Allen, Mary, J., & Yen, Wendy, J (1979). *Introduction to measurement theory*. California: Brooks/Cole Publishing Company.
- [2] Andersen, L. W., & Krathwohl. (2001). *A Taxonomy for learning, teaching and assessment*. Boston: Longman.
- [3] Astin, W. Alexander. (1993). *Assessment for excellence*. Phonix: The Oryx Press.
- [4] Berk, R. A. (1986, Ed). *Performance assessment*. Baltimore: The JohnHopkinsUniversity Press.
- [5] Djemari Mardapi. (2012). *Pengukuran, penilaian, dan evaluasi pendidikan*. Yogyakarta: Medica Publishing.
- [6] Griffin, Patrix., & Nix, Peter. (1991). *Educational assessment dan reporting*. Sydney: Harcout Brace Javanovich, Publisher.
- [7] Johnson, D. W., & Johnson, R. T. (2002). *Meaningful assessment*. Boston:: Allen & Bacon.
- [8] Popham, W. J. (1996). *Classroom assessment*. Boston: Allyn and Bacon.

n