

ASSESSMENT PEMBELAJARAN IPA

**Tarzan Purnomo
Dyah Ayu Fitriana
Achmad Rante Suparman
Suwarno
Cahya Amalia Chusna
Evi Ristiana
Yuli Tamar Filindity**



GETPRESS INDONESIA

ASSESSMENT PEMBELAJARAN IPA

Penulis : Tarzan Purnomo
Dyah Ayu Fitriana
Achmad Rante Suparman
Suwarno
Cahya Amalia Chusna
Evi Ristiana
Yuli Tamar Filindity

ISBN :

Editor : Ari Yanto, M.Pd.
Penyunting: Yuliatr Novita, M.Hum.
Desain Sampul dan Tata Letak : Tri Putri Wahyuni, S.Pd.

Penerbit: GETPRESS INDONESIA
Anggota IKAPI No. 033/SBA/2022

Redaksi:
Jl. Palarik RT 01 RW 06, Kelurahan Air Pacah
Kecamatan Koto Tengah, Padang, Sumatera Barat
website: www.getpress.co.id
email: adm.getpress@gmail.com

Cetakan pertama, Februari 2024

Hak cipta dilindungi undang-undang
Dilarang memperbanyak karya tulis ini dalam bentuk
dan dengan cara apapun tanpa izin tertulis dari penerbit.

KATA PENGANTAR

Dengan mengucapkan puji syukur kehadirat Allah SWT, atas limpahan rahmat dan hidayahNya, maka Penulisan Buku dengan Judul *Assessment Pembelajaran IPA* dapat diselesaikan dengan baik atas kerjasama tim penulis. Buku ini berisi mengenai makna pengukuran, evaluasi dan assesment, penyusunan kisi-kisi tes, uji coba analisis validitas dan reliabilitas soal, revisi butir soal dan menafsirkan hasil tes, pengembangan tes sikap dan rubrik penilaiannya, penilaian berbasis kelas, penilaian kinerja.

Buku ini masih banyak kekurangan dalam penyusunannya. Oleh karena itu, kami sangat mengharapkan kritik dan saran demi perbaikan dan kesempurnaan buku ini selanjutnya. Kami mengucapkan terima kasih kepada berbagai pihak yang telah membantu dalam proses penyelesaian Buku ini. Semoga Buku ini dapat menjadi sumber referensi dan literatur bagi Perguruan Tinggi yang mudah dipahami.

Padang, Februari 2024
Penulis

DAFTAR ISI

KATA PENGANTAR.....	i
DAFTAR ISI	ii
DAFTAR GAMBAR.....	iv
DAFTAR TABEL	v
BAB 1 MAKNA PENGUKURAN, EVALUASI DAN ASSESMENT	1
1.1 Pendahuluan.....	1
1.2 Pengukuran (<i>Measurement</i>).....	4
1.3 Evaluasi.....	11
1.4 Penilaian (<i>Assesment</i>)	19
DAFTAR PUSTAKA	28
BAB 2 PENYUSUNAN KISI-KISI TES	29
2.1 Pendahuluan.....	29
2.2 Karakteristik Penilaian Pendidikan	33
2.3 Pengertian Kisi-kisi.....	34
2.4 Kesimpulan	39
DAFTAR PUSTAKA	40
Bab 3 UJI COBA ANALISIS VALIDITAS DAN RELIABILITAS	
SOAL.....	41
3.1 Definisi.....	41
3.2 Validitas Soal Secara Klasik.....	43
3.3 Reliabilitas Soal Secara Klasik.....	50
3.4 Validitas dan Reliabilitas Soal Secara Modern	54
DAFTAR PUSTAKA	57
BAB 4 REVISI BUTIR SOAL DAN MENAFSIRKAN HASIL TES	59
4.1 Pendahuluan.....	59
4.2 Validitas.....	62
4.3. Reliabilitas.....	64
4.3 Tingkat Kesukaran.....	66
4.4 Daya pembeda	68
4.5 Analisis Distraktor/pengecoh	71
4.6 Contoh Analisis Soal.....	72
4.7 Penutup	74
DAFTAR PUSTAKA	76
BAB 5 PENGEMBANGAN TES SIKAP DAN RUBRIK	
PENILAIANNYA.....	79
5.1 Pengertian Tes Sikap.....	79
5.2 Lingkup Penilaian Sikap	80
5.3 Perumusan Indikator Penilaian Sikap	81

5.4 Teknik, Instrumen dan Rubrik Penilaian Sikap	87
DAFTAR PUSTAKA.....	100
BAB 6 PENILAIAN BERBASIS KELAS	101
6.1 Pendahuluan.....	101
6.2. Tujuan Penilaian Berbasis Kelas	104
6.3. Prinsip-prinsip PBK.....	107
6.3 Teknik Penilaian Berbasis Kelas	108
6.4 Bentuk Penilaian Berbasis Kelas.....	113
DAFTAR PUSTAKA.....	115
BAB 7 PENILAIAN KINERJA	117
7.1 Penilaian Kinerja	117
7.2 Bentuk Penilaian Kinerja.....	118
7.3 Pengembangan Penilaian Kinerja.....	123
DAFTAR PUSTAKA.....	134
BIODATA PENULIS	

DAFTAR GAMBAR

Gambar 1. 1. Fungsi evaluasi pendidikan	15
Gambar 3. 1. Analisis korelasi <i>product moment</i> dengan SPSS.....	49
Gambar 3. 2. Hasil korelasi <i>product moment</i> dengan R Program	50
Gambar 3. 3. Hasil formula Spearman Brown dengan R Program	54
Gambar 5. 1. Diagram Penilaian Teman Sejawat.....	90

DAFTAR TABEL

Tabel 3. 1. Hasil uji coba tes pilihan ganda	45
Tabel 3. 2. Perhitungan butir nomor 5	46
Tabel 3. 3. Interpretasi nilai koefisien korelasi.....	47
Tabel 3. 4. Hasil pembagian butir ganjil dan butir genap	51
Tabel 3. 5. Interpretasi koefisien reliabilitas	52
Tabel 3. 6. Hasil penelitian penggunaan fit dan tidak fit.....	56
Tabel 5. 1. Tingkatan Sikap.....	80
Tabel 5. 2. Contoh Rumusan Indikator Penilaian Sikap.....	82
Tabel 7. 1. Tabel Penilaian praktek.....	119
Tabel 7. 2. Penilaian produk.....	120
Tabel 7. 3. Contoh format rubrik.....	124

BAB 1

MAKNA PENGUKURAN, EVALUASI DAN ASSESMENT

Oleh Tarzan Purnomo

1.1 Pendahuluan

Pembelajaran adalah upaya sadar yang dilakukan secara terencana untuk merubah pengetahuan, keterampilan, dan atau sikap peserta didik. Untuk itu dalam menyelenggarakan pendidikan harus mempunyai kurikulum sebagai pedoman dan pengarah tujuan, sumberdaya manusia sebagai pelaksana program, dan sarana prasarana sebagai sarana pendukung. Pembelajaran sebagai implementasi dari program pendidikan merupakan jantungnya pendidikan. Untuk mengetahui sejauh mana keberhasilan dari lembaga pendidikan, salah satunya dapat dilihat dari keberhasilan proses pembelajaran melalui evaluasi.

Pengukuran, evaluasi, dan assesment merupakan bagian integral dari kurikulum yang berfungsi untuk mengetahui apakah keterlaksanaan program pendidikan sesuai kurikulum, efektivitas dan efisiensi pelaksanaan program, kinerja sumberdaya manusia, dan tujuannya sudah tercapai. Untuk itu diperlukan alat ukur, instrumen evaluasi, dan teknik assesment untuk menetapkan nya.

Komponen dasar pembelajaran terdiri dari perencanaan pembelajaran, kegiatan belajar mengajar (KBM) sebagai implementasi dari rencana pembelajaran, dan evaluasi untuk mengetahui efektivitas proses pembelajaran dan mengukur kompetensi peserta didik sebagai hasil belajar. Secara normatif agar mampu menjalankan tugasnya secara profesional, seorang

guru harus memiliki 3 kompetensi dasar yaitu mampu membuat perencanaan pengajaran, menerapkan model dan strategi belajar mengajar, dan mengevaluasi proses dan hasil pembelajarannya. Oleh karena itu kompetensi profesional seorang guru meliputi kemampuan dalam merencanakan pembelajaran meliputi membuat RPP atau Modul Pembelajaran, membuat Lembar Kerja Peserta Didik (LKPD), membuat atau memilih media pembelajaran, dan menyusun soal evaluasi mutlak dipersyaratkan. Dalam melaksanakan kegiatan belajar mengajar agar efektif, sudah seharusnya seorang guru dituntut menguasai berbagai metode, strategi, pendekatan, dan model pembelajaran. Selanjutnya untuk mengukur efektivitas kegiatan pengajaran yang diselenggarakannya, seorang guru harus mampu melakukan evaluasi proses dan hasil pembelajarannya melalui kegiatan pengukuran, evaluasi, dan *assesment*.

Pengukuran (*measurement*) merupakan proses pendeskripsian secara kuantitatif (numerik) terhadap pencapaian kompetensi peserta didik. Dari pengukuran diperoleh data kuantitatif untuk menentukan kuantitas kompetensi peserta didik. Dengan demikian terbatas pada hasil belajar atau kompetensi yang bersifat kuantitatif.

Evaluasi (*evaluation*) adalah penentuan kualitas (kualifikasi) proses pembelajaran berdasarkan kriteria tertentu, yang dilakukan secara sistematis dan berkelanjutan untuk pengambilan keputusan (*value judgement*). Aspek yang dievaluasi meliputi ketercapaian program, keberhasilan peserta didik, berharga atau tidak, dan efisiensinya. Oleh karena itu, yang dievaluasi meliputi kurikulum, kebijakan, metode pembelajaran, dan kinerja pendidik. Hasil akhir dari evaluasi adalah rekomendasi untuk mengambil keputusan.

Penilaian (*assessment*) adalah cara untuk mengetahui sejauh mana capaian atau kompetensi (*kognitif, psikomotorik, afektif*) peserta didik. Penilaian digunakan untuk menilai kompetensi sebagai wujud hasil belajar. Hasil penilaian dapat bersifat kuantitatif (numerik) atau kualitatif (deskriptif) yang

menggambarkan kompetensi peserta didik. Penilaian (*assesment*) ini dilakukan oleh badan, lembaga, organisasi atau institusi penyelenggara pendidikan. Dengan demikian penilaian (*assessment*) adalah prosedur evaluasi kemajuan dan kompetensi peserta didik menggunakan metode dan aturan tertentu (melalui observasi atau tes tertulis).

Assesment dan evaluasi harus dilakukan untuk mengukur efektivitas pembelajaran. Guru harus menguasai keterampilan melakukan *assesment* dan evaluasi untuk mengukur ketercapaian pembelajarannya. Agar evaluasi terlaksana secara tepat maka harus disertai dengan proses *assesment* untuk mengungkap kendala yang dihadapi peserta didik dalam mencapai kompetensi yang ditetapkan. *Assesment* berperan penting untuk mengevaluasi efektivitas proses pembelajaran sebagai balikan (*follow up*) yang digunakan untuk perbaikannya. Dalam melakukan *assesment* harus mencakup semua kompetensi, dilakukan *assesment perkembangan* meliputi aspek perkembangan kognitif, sosial, emosi, dan motorik peserta didik.

Urutan tahapan melakukan *assesment* adalah pengumpulan data menggunakan *alat ukur* yang memiliki kriteria sesuai dengan variabel kompetensi yang diukur, menggunakan prosesur yang sesuai. Hasilnya adalah sekumpulan data mentah yang perlu diinterpretasikan. Data besaran fisis dapat secara langsung diukur, namun data hasil pengukuran skala-skala psikometri sikap (*atitute*) dan perilaku (*psikomotorik*) harus diinterpretasikan dengan skala interpretasi sehingga menghasilkan *justifikasi lulus atau gagal*. Proses interpretasi inilah yang disebut sebagai *assesment*.

Dalam *assesment* pembelajaran IPA, Guru tidak hanya sekedar menghasilkan kesimpulan (*justifikasi*) lulus atau gagal kepada peserta didik. Karena Pembelajaran merupakan proses perubahan yang berkesinambungan, maka Guru harus memberi rekomendasi dan masukan konstruktif kepada peserta didik berdasarkan hasil kajian yang dituangkan dalam dokumen hasil *assesment*. Proses ini

mencakup 1) menentukan alat ukur, 2) melaksanakan *measurement*, 3) melakukan *assessment*, dan 4) mengevaluasi sehingga dapat menghasilkan makna yang holistik.

1.2 Pengukuran (*Measurement*)

1.2.1 Pengertian Pengukuran

Pengukuran (*measurement*) adalah proses pendeskripsian terhadap kompetensi peserta didik secara numerik (berupa nilai angka). Hasil pengukuran bersifat kuantitatif (Arikunto, 2005); Gronlund (1971, dalam Hasan, 1988). Untuk memperoleh hasil yang representatif dan mencerminkan kompetensi peserta didik yang sesungguhnya maka digunakan alat ukur tertentu, berdasarkan formula tertentu, dan hasilnya diinterpretasikan berdasarkan kriteria tertentu sebagai pembanding.

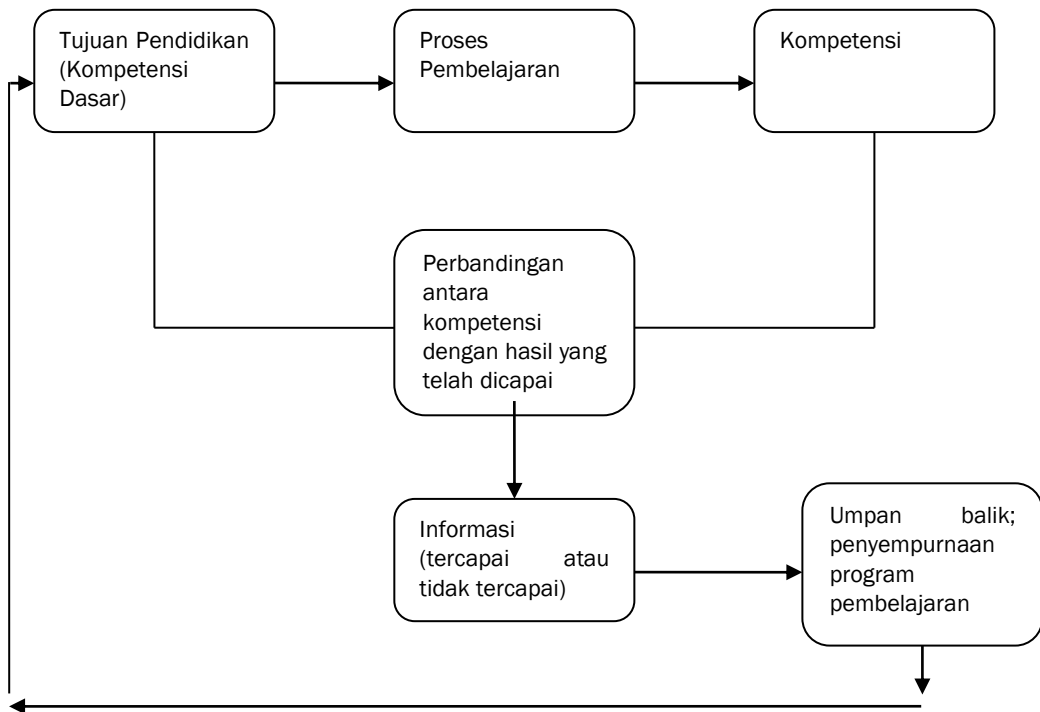
Dalam pembelajaran IPA, kegiatan pengukuran dilakukan untuk mengetahui tingkat pemahaman, penguasaan, dan keterampilan peserta didik sebagai manifestasi kompetensinya terhadap konsep-konsep, konten, dan keterampilan IPA. Oleh karena itu pengukuran merupakan indikator untuk mengukur keberhasilan suatu program pembelajaran. Kompetensi sebagai hasil belajar dapat bersifat kuantitatif maupun kualitatif. Karena itu hasil belajar dapat diukur secara kuantitatif (berupa angka) maupun kualitatif (berupa deskripsi) dan dapat dilakukan menggunakan tes atau nontes (kinerja). Hasilnya angka (*score*) mentah, kemudian berdasar kriteria (Hamzah & Satria, 2012) diinterpretasikan sehingga dapat diambil keputusan (naik atau tidak naik, atau lulus tidak lulus, atau remidi).

Secara substansif ada tiga macam pengukuran, yaitu: (1) *tidak menguji sesuatu*, misalnya seorang nelayan mengenai panjang jala, lebar jala, size lubang jala dan lain-lain. (2) *menguji sesuatu*, misalnya menguji intensitas cahaya terhadap laju fotosintesis, dan (3) *menguji untuk menilai sesuatu*, misalnya uji kompetensi untuk

nilai rapor. Dalam pendidikan yang biasa diterapkan adalah yang terakhir (Sudiyono, 1996).

Pengukuran adalah menentukan karakteristik individu berkaitan dengan kompetensinya, untuk itu dalam melakukan pengukuran harus cermat dan memiliki kesalahan seminimal mungkin. Dalam IPA kesalahan pengukuran lebih banyak disebabkan oleh ketidakakuratan alat ukurnya. Sementara itu dalam ilmu sosial kesalahan tersebut dapat karena alat ukur, teknik, dan objek yang diukur (Arikunto, 2005).

Agar hasilnya presisi, validitas dan reliabilitas alat ukur yang digunakan harus terjamin, baik dari aspek substansi yang diukur, konstruksi soal, dan bahasa yang digunakan. Untuk itu perlu disusun kisi-kisi soal sebagai alat ukur, yang memuat substansi yang diukur (merupakan konsep esensial, sebagai indikator hasil belajar, bernilai aplikatif, sinambung, dan sebagai prasyarat untuk mempelajari materi selanjutnya), soal (konstruksi atau bentuk soal, jumlah, *degree*), alokasi waktu, dan *scoring*. Sebagai praktek baik pengukuran, maka kisi-kisi soal harus dikomunikasikan kepada peserta didik.



Gambar 1.1 Kedudukan Pengukuran, Evaluasi, dan Assesment

Untuk meminimalkan kesalahan, alat ukur yang digunakan harus handal. Berdasarkan sifatnya kesalahan pengukuran dapat *bersifat acak* yang disebabkan adanya variasi situasi saat ujian, kondisi fisik (tidak fit, sakit, sehat), mental (emosi) yang diukur dan yang mengukur. Atau *bersifat sistematis* dapat dari alat ukur, objek, dan pengukur. Soal tes sebagai alat ukur jika sangat mudah atau sangat sulit, akan menghasilkan *underestimate* atau *overestimate* dari kompetensi sesungguhnya. Peserta didik sebagai *testee* memiliki rasa kecemasan yang variatif. Jika saat dites sangat cemas, hasilnya cenderung *underestimate* dari kompetensi sesungguhnya. Kesalahan sistematis Guru terutama pada saat *skoring*, ada yang "pemurah" dan yang "mahal". Kesalahan sistematis terjadi jika hal ini berlaku terhadap semua peserta didik. Namun, jika hanya terhadap peserta didik tertentu, maka akan terjadi bias.

Mengingat sifatnya sangat kompleks, untuk itu pengukuran harus mengacu pada teori sebagai pijakan. Teori pengukuran ada 2, yaitu *teori tes klasik* dan *teori tes modern*. 2 asumsi yang mendasari *Teori tes klasik*, yaitu (1) hasil pengukuran yang didapat peserta didik dapat dibagi menjadi skor sesungguhnya dan skor kesalahan. (2) antara skor yang sesungguhnya dengan skor kesalahan tidak berhubungan. Dari kedua asumsi tersebut, kemudian dibuat berbagai rumus indeks validitas dan indeks reliabilitasnya. Kelemahannya, adalah sifatnya tidak ajeg, karena butir soal tergantung pada karakteristik kelompok peserta didik. Akibat variasi soal antar kelompok sehingga sulit membandingkan kemampuan antar kelompok dan antar individu. Asumsi utama (asumsi independen) *Teori tes modern* adalah probabilitas peserta didik dalam menjawab benar tidak tergantung dalam menjawab soal lainnya. Sehingga estimasi kompetensi peserta didik tidak dipengaruhi oleh alat ukur yang digunakan. Di bidang IPA dan teknologi alat ukur sifatnya ajeg. Inilah sebabnya mengapa teori tes modern banyak diadopsi.

1.2.2 Tujuan dan Fungsi Pengukuran

Dalam rangka melakukan pengukuran melibatkan *testing* yaitu alat ukur/instrumen dan prosedur tes yang digunakan dalam melakukan pengukuran, *testee* yaitu peserta didik yang dites, dan *tester* yaitu guru yang melakukan tes. Berdasarkan karakteristik jawabannya pengukuran ada 2, yaitu tes dan non-tes. Tes adalah pengukuran yang responnya berkriteria *benar* dan *salah*. Non tes responnya berupa pendapat terhadap objek yang diukur.

1. Tes

Tes adalah soal yang harus dijawab atau tugas yang harus dikerjakan peserta didik untuk mengukur kompetensinya. Aspek yang diukur dapat berupa pemahaman, penguasaan materi ajar, keterampilan, atau sikap terhadap sistem nilai tertentu. Tes merupakan salah satu *assesment* (Hamzah & Satria, 2012). Tes

merupakan alat ukur mengenai perubahan perilaku peserta didik sebagai manifestasi hasil belajar, dapat berupa aspek psikologis tertentu (perilaku) dari peserta didik (Anastari, 1982; Arikunto (2005). Dengan demikian hasilnya dideskripsikan dalam sistem kategori skala numerik.

a. Jenis tes

Sebagai alat ukur, tes terdiri beberapa jenis yaitu:

1) Jenis Tes berdasarkan tujuannya

a) *Tes Seleksi*

Tes seleksi dipakai menyeleksi calon peserta didik (penerimaan peserta didik baru). Materi tes berupa butir-butir soal yang mempunyai tingkat kesulitan tinggi, sebagai prasyarat mengikuti program pendidikan, sehingga hanya calon-calon yang memenuhi kualifikasi yang mampu menjawab.

b) *Tes Awal (Pre test)*

Bertujuan mengukur kompetensi awal peserta didik terkait materi ajar yang akan dipelajari, dan dilaksanakan sebelum kegiatan pembelajaran.

c) *Tes Akhir (post test)*

Bertujuan mengukur capaian hasil belajar dari suatu tahapan pembelajaran. Soalnya relatif sama dengan soal *pre test*.

d) *Tes Diagnostik*

Bertujuan mengetahui kendala dalam mempelajari materi ajar, sehingga dapat diupayakan solusinya. Selain itu untuk mengetahui penguasaan pengetahuan prasyarat mempelajari materi selanjutnya. Bentuk tes lisan, tertulis, atau ujiketrampilan.

e) *Tes Formatif*

Bertujuan untuk mengetahui kompetensi peserta didik setelah pembelajaran. Dilaksanakan setelah selesai kegiatan pembelajaran. Materi tes terdiri dari soal kategori mudah, sedang, maupun sulit.

f) *Tes Sumatif*

Bertujuan untuk menentukan status atau kedudukan nilai sebagai bentuk kompetensi peserta didik, untuk menentukan:

- ranking peserta didik dalam kelompoknya,
- bisa tidaknya peserta didik mengikuti program selanjutnya,
- capaian peserta didik untuk naik atau tidak naik, lulus atau tidak lulus.

2) Jenis tes berdasarkan individu yang dites

Terdiri dari (1) *tes individual*; yang di tes hanya seorang peserta didik, dan (2) *tes kelompok*, yang dites kelompok peserta didik.

3) Jenis tes berdasarkan jawabannya

Terdiri dari; (1) *Tes verbal*; jawaban berupa lisan atau tertulis, dan (2) *Tes ketrampilan*; jawaban berupa tindakan, tingkah laku, atau ketrampilan gerakan otot (*skill*).

b. Bentuk tes

Berdasarkan bentuknya, yaitu tes subyektif (*esai*) dan tes objektif.

1) *Tes Subjektif (esai)*

Menghendaki jawaban berupa uraian menggunakan kalimat sendiri. Peserta didik bebas berfikir untuk menyusun jawabannya berdasarkan kemampuan analisis, sintesis, dan evaluasinya.

2) *Tes Objektif*

Opsi jawaban sudah tersedia, peserta didik tinggal memilih. Berdasarkan bentuknya ada 5, yaitu a) benar salah (*true-false test*); b). menjodohkan (*matching test*); c). melengkapi (*completion test*); d). isian singkat (*fill-in test*); dan e). pilihan ganda (*multiple choice test*). Bentuk tes benar salah, isian singkat, dan menjodohkan hanya mengukur kemampuan berpikir rendah (mengingat atau pengetahuan). Tes objektif pilihan ganda memiliki kelebihan yaitu dapat digunakan untuk menilai kemampuan mengingat dan memahami dengan cakupan materi yang luas. Namun kelemahannya (1) tidak dapat mengukur atau mengungkapkan proses berpikir tingkat tinggi, (2) memungkinkan peserta didik berspekulasi dalam menjawab.

1.2.3 Teknik Pengukuran

1. *Testing*

Menurut Anastari (1982) tes merupakan pengukuran yang obyektif dan standard karena menggunakan alat yang baku dan prosedur yang sistematis dalam rangka mendeskripsikan kompetensi peserta didik dengan skala numerik atau system kategori. Sistematis, karena skornya jelas dan rinci, serta berlaku untuk semua peserta didik.

Untuk mengukur kompetensi peserta didik dapat menggunakan tes buatan guru atau tes standar. Saat ini yang dominan dan banyak digunakan adalah tes buatan guru dengan pertimbangan fungsinya dalam menentukan prestasi peserta

didik dan keberhasilan PBM. Oleh karena itu guru harus mampu menyusun tes dengan baik.

1.3 Evaluasi

1.3.1 Pengertian Evaluasi (*Evaluation*)

Evaluasi pendidikan merupakan proses penggunaan informasi yang diperoleh melalui pengukuran (*measurement*) dan penilaian (*assesment*) untuk pengambilan keputusan. Evaluasi dapat dilakukan, melalui pengukuran atau tanpa pengukuran. Evaluasi adalah proses inklusif dari pengukuran untuk mengukur ketercapaian tujuan atau efektivitas pelaksanaannya sehingga bisa diputuskan nilainya (*value judgement*). Dengan demikian esensi dari evaluasi adalah dalam rangka untuk memutuskan. Oleh karena itu yang perlu dievaluasi mencakup kurikulum, kebijakan, sumber belajar, metode pembelajaran, dan kinerja guru.

Evaluasi dilakukan untuk kepentingan pengambilan keputusan, sehingga cakupannya lebih luas meliputi pengukuran, testing, dan penilaian (Arikunto, 2005). Dari evaluasi akan diperoleh gambaran dan informasi yang berguna untuk menilai efektivitas intervensi pendidikan sehingga dapat ditetapkan keberhasilan atau kualitas suatu program. Untuk itu dalam melakukan evaluasi harus mencakup 3 kegiatan, yaitu penentuan kualitas kinerja harus menggunakan standar baku, menggunakan instrumen yang relevan, dan menerapkan prosedur yang sesuai.

Standar atau kriteria yang diacu dalam evaluasi akan digunakan untuk menerima menolak mutu kinerja. Pemberian makna terhadap hasil pengukuran dengan kriteria tertentu yang ditetapkan sebagai patokan (Hamzah & Satria, 2012) inilah sebagai ukuran capaian prestasi pembelajaran setiap peserta didik (Aries, 2011).

Menurut *Educational System Evaluation Model* (Arifin, 2012), berdasarkan *Context, Input, Process, dan Product* (CIPP) evaluasi dibagi menjadi 3 jenis (Mohaimin, 2013) yaitu:

1) *Evaluasi Input*

Input evaluation, structuring decision. Bertujuan untuk memutuskan terkait penggunaan sumberdaya. Evaluasi meliputi analisis personal, tujuan program, identifikasi dan kapabilitas sistem, desain prosedur implementasi, pembiayaan dan penjadwalan. Gunanya untuk memilih strategi penggunaan sumberdaya. Tujuannya agar penggunaan sumberdaya optimal dan program terlaksanan dengan efektif dan efisien.

2) *Evaluasi Proses*

Evaluasi ini bertujuan untuk melihat kesesuaian pelaksanaan program dengan strategi yang dirancang. Tujuan evaluasi proses yaitu mengidentifikasi proses pelaksanaan, ketidaksesuaian prosedur atau implementasinya, dan menyediakan informasi untuk memperbaiki program, sehingga: a) diketahui kelemahan dan kekuatannya, b) dapat diambil keputusan yang akan ditetapkan, dan c) diimplementasikan.

Evaluasi proses dapat dilakukan melalui pengukuran pengetahuan dan ketrampilan (*pre-test* dan *pos-test*), perilaku peserta didik (observasi), evaluasi diri (*self-report*), *performance* (portofolio), partisipatory (*self-study*), kendala yang ditemui.

3) *Evaluasi Product (Outcome)*

Untuk mengukur capaian program. Tujuan lainnya menilai dan menginterpretasikan kelayakan dan keberhargaan luaran (*outcome*) terhadap tujuan, konteks, input, dan proses berdasarkan definisi operasional, kriteria, nilai dan ketrampilan (*performing*) peserta didik baik secara kuantitatif maupun kualitatif. Produk atau hasilnya (berupa skor tes,

data), dibandingkan dengan tujuan yang ditetapkan dalam rancangan lalu dianalisis secara kualitatif dikaitkan dengan tujuan-tujuan yang lebih detail.

1.3.2 Tujuan, Fungsi, dan Sasaran Evaluasi

1. Tujuan Evaluasi

Evaluasi memiliki tujuan umum yaitu: a). *memperoleh data dan informasi* kompetensi peserta didik, dan b). *mengetahui tingkat efektivitas* proses pembelajaran. Sedangkan secara khusus, evaluasi bertujuan untuk; a). *memotivasi peserta didik*, dan b). *menemukan faktor-faktor penyebab* kegagalan peserta didik, dan mencari solusinya. Dengan demikian evaluasi dalam pembelajaran bertujuan mendeskripsikan kompetensi peserta didik, efektivitas pembelajaran, menentukan tindak lanjut, dan sebagai pertanggung jawaban (*accountability*).

2. Fungsi Evaluasi

Evaluasi berfungsi untuk (a) mengetahui kemajuan peserta didik, (b) menyusun rencana, dan, (c) penyempurnaan rencana. Sedangkan secara khusus, evaluasi berfungsi (a) psikologis; bagi peserta didik untuk mengetahui kompetensinya didalam kelasnya (termasuk kategori mana). Memberikan kepastian sejauh manakah efektivitas pembelajaran yang dilakukan guru dan apa yang harus diperbaiki. Fungsi (b) pedagogis-didaktik; memotivasi untuk memperbaiki (bila nilainya jelek), meningkatkan (bila kurang puas dengan nilainya), dan mempertahankan kompetensinya (bila nilainya bagus). Sedangkan bagi guru, secara didaktik evaluasi berfungsi a).*diagnostik*: sebagai landasan menilai kompetensi peserta didik, b). *penempatan*: sebagai informasi posisi setiap peserta didik di dalam kelasnya, c). *selektif*: sebagai acuan penetapan status peserta didik, d). *bimbingan*: sebagai solusi, dan e). *instruksional*: petunjuk pencapaian program pengajaran. (c) Fungsi administratif evaluasi pendidikan a). *sebagai laporan*

perkembangan peserta didik, b). *data* untuk pengambilan keputusan (naik atau tidak naik, atau lulus tidak lulus), dan c). *memberikan gambaran* materi apa saja yang kompetensi masih kurang, atau sudah baik. Secara komprehensif fungsi evaluasi pendidikan, dapat dilihat dalam Gambar 1.2.

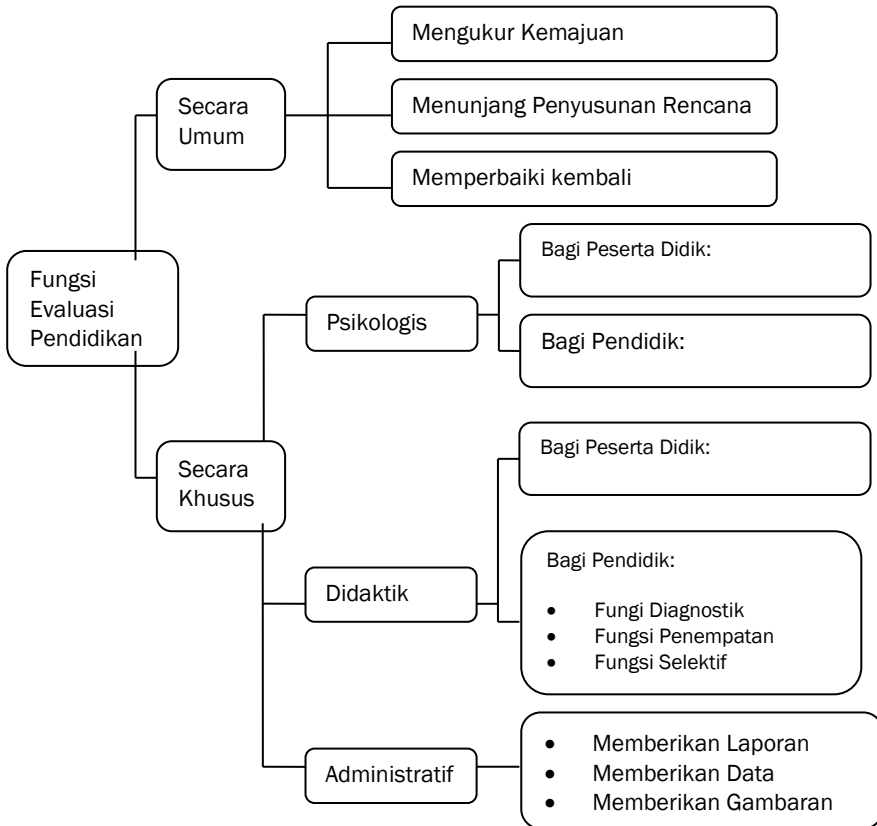
3. Aspek atau Sasaran Evaluasi

Evaluasi mencakup tiga aspek yaitu aspek (ranah/domain) pengetahuan, sikap, dan keterampilan.

Ranah Kognitif

Aspek kognitif mencakup semua aktivitas mental (otak), terdiri dari 6 jenjang kognitif, yaitu:

- 1) *Pengetahuan* adalah kemampuan mengingat kembali (*recall*) nama, istilah, ide, gejala, rumus-rumus, dan lain-lain. Contoh: menyebutkan nama ilmiah padi.
- 2) *Pemahaman* adalah kemampuan memahami dari berbagai perspektif. Contoh: menjelaskan proses terjadinya fotosintesis secara benar.
- 3) *Penerapan* atau aplikasi adalah kemampuan menerapkan metode atau rumus secara kongkrit. Contoh: menerapkan metode pengukuran kadar oksigen terlarut dalam perairan.
- 4) *Analisis* adalah kemampuan memerinci komponen atau unsur dan hubungan antar penyusun suatu sistem. Contoh: mampu mencari penyebab tidak berfungsinya alarm akibat listrik padam.



Gambar 1. 1. Fungsi evaluasi pendidikan

- 5) Sintesis adalah kemampuan merangkai antar bagian secara logis, sehingga terbentuk pola baru. Contoh: kemampuan menemukan solusi terhadap permasalahan rendahnya hasil panen.
- 6) Penilaian atau evaluasi merupakan kemampuan menilai, menimbang dan memilih sesuai dengan kriteria. Contoh: memilih jenis pupuk untuk meningkatkan hasil panen tanaman tertentu.

Ranah afektif

Ranah ini berkaitan dengan sikap dan nilai yang mengalami perubahan setelah proses belajar. Terdiri 5 jenjang, yaitu:

- 1) *Receiving* atau *attending* adalah kemampuan menerima stimulus eksternal (masalah, situasi, peraturan), secara sadar. Contoh: taat pada aturan.
- 2) *Responding* atau menanggapi adalah kemampuan menanggapi dan memberikan reaksi. Contoh: tumbuh hasratnya untuk mempelajari lebih detail.
- 3) *Valuing* artinya menghargai, menilai konsep atau fenomena, baik-buruk, dan diinternalisasikan sebagai sistem nilai. Contoh: mampu membedakan hal-hal positif dan negatif.
- 4) *Organization* artinya mampu menghubungkan antar nilai sehingga terbentuk nilai baru dan memantapkan nilai yang telah dimilikinya. Contoh: mendukung penegakkan hukum.
- 5) *Characterization by a value or value complex* yakni kemampuan memadukan sistem nilai dan membentuk pola kepribadian dan tingkah laku yang konsisten dan mempengaruhi emosinya sehingga membentuk karakteristik perilakunya. Contoh: menerapkan ilmu pengetahuan dan teknologi yang dipelajari dalam kehidupan sehari-hari untuk meningkatkan kualitas hidup yang sehat.

Ranah Psikomotorik

Ranah psikomotorik merupakan keterampilan (*skill*) yang dikuasai setelah proses pembelajaran. Kompetensi psikomotorik peserta didik dapat berupa:

- a. terampil menggunakan alat ukur (Thermometer, pH meter).
- b. mampu melakukan rangkaian percobaan IPA sesuai prosedur.
- c. menerapkan pola hidup bersih dan sehat.

Syarat-syarat evaluasi yang baik adalah valid, handal, objektif, seimbang, membedakan, normatif, fair, dan praktis (Sukardi, 2008). Evaluasi pembelajaran mempunyai 6 tujuan (Sukardi, 2008) yaitu:

1. menilai ketercapaian (*attainment*) tujuan pembelajaran,
2. mengukur semua hasil belajar, baik aspek kognitif, psikomotorik, maupun afektif,
3. sebagai sarana (*means*) untuk mengukur pengetahuan, pemahaman, dan pengalaman peserta didik,
4. mampu memotivasi belajar peserta didik,
5. sebagai bahan untuk bimbingan,
6. dasar pertimbangan perubahan kurikulum.

1.3.3 Prinsip-prinsip Evaluasi

Agar hasil evaluasi sesuai yang diharapkan maka harus memperhatikan:

1. *Keterpaduan*. Sebagai bagian integral, selain capaian pembelajaran, materi, dan metode pembelajaran (segitiga Tyler), untuk itu evaluasi harus disusun saat menyusun rencana pembelajaran (Modul Pembelajaran/RPS) agar selaras dengan Capaian Pembelajaran dan materi ajar.
2. *Keterlibatan siswa*. Peserta didik harus dilibatkan dalam evaluasi.
3. *Koherensi*. Evaluasi harus mengukur ranah kompetensi yang sudah diajarkan.
4. *Pedagogis*. Hasil evaluasi harus dapat dipakai sebagai alat untuk memotivasi bagi yang gagal dan ganjaran (*reward*) atau penghargaan bagi yang berhasil.
5. *Akuntabilitas*. Hasil evaluasi dilaporkan kepada wali siswa, komite sekolah, lembaga pendidikan, dan pemerintah (bagi sekolah negeri).

1.3.4 Teknik Evaluasi

1. Teknik non-tes

a. Skala bertingkat (*rating scale*)

Nilai yang diperoleh berbentuk angka dan disajikan dalam bentuk skala.

b. Kuesioner

Kuesioner (*questionnaire*) berupa daftar pertanyaan untuk mengetahui pendapat responden.

1). Berdasarkan repondennya:

- a. *Kuesioner langsung*; jika diisi langsung oleh responden.
- b. *Kuesioner tidak langsung*; diisi bukan oleh responden (boleh keluarganya).

2). Berdasarkan cara menjawab:

- a. *Kuesioner tertutup*: telah tersedia pilihan jawaban, responden hanya tinggal memilih.
- b. *Kuesioner terbuka*; belum tersedia jawaban, responden bebas memberikan jawaban.
- c. *Daftar cocok (check list)*

Berupa pertanyaan singkat dan responden tinggal membubuhkan tanda cocok pada tempat yang tersedia.

d. Wawancara (*interview*)

Bisa melalui 2 cara:

- *Interviu bebas*; responden bebas menyampaikan pendapatnya.
- *Interviu terpimpin*; pertanyaan sudah disusun, responden tinggal memilih jawaban yang tersedia.

e. Pengamatan (*observation*)

Dapat melalui 3 cara:

- *Observasi partisipan*; yaitu observer mengikuti aktivitas yang sedang diamati.
- *Observasi sistematis*; yaitu yang diamati sudah dipilih sesuai tujuannya.
- *Observasi eksperimental*; observer mengendalikan dan mengatur sesuai tujuan evaluasi.

f. *Riwayat hidup*

Berdasar riwayat hidupnya dapat disimpulkan kompetensi dan sikapnya.

2. Teknik tes

Tes merupakan prosedur sistematis dan objektif untuk mengetahui kompetensi peserta didik secara tepat dan cepat. Berdasar kegunaannya, tes terdiri dari:

- a. *Tes diagnostic*; digunakan untuk mengukur kemampuan awal peserta didik, sebagai pertimbangan agar guru dapat membantunya.
- b. *Tes formatif*; untuk mengukur kompetensi peserta didik setelah mengikuti suatu pembelajaran, dan tes diberikan diakhir pembelajaran (ulangan harian).
- c. *Tes sumatif*; untuk mengukur kompetensi peserta didik diakhir program (ulangan umum, catur wulan atau akhir semester).

1.4 Penilaian (*Assesment*)

1.4.1 Pengertian Penilaian (*Assesment*)

Secara historis kegiatan penilaian (*assesment*) pertama kali dilakukan tahun 206 sebelum masehi di China untuk menyeleksi pegawai Kerajaan dinasti Han. Berdasarkan peraturan pemerintah No.19/2005, penilaian (*assesment*) adalah proses pengumpulan data dan informasi tentang capaian hasil belajar peserta didik yang dilakukan oleh pendidik, satuan pendidikan, dan pemerintah kemudian diinterpretasikan untuk pengambilan keputusan (*judgment*) (Sudjana, 1989).

Assessment dilakukan dengan beragam cara dan alat penilaian (Sudrajat, 2008) dan menuntut peserta didik untuk mendemonstrasikan apa yang sudah diketahui, pahami, dan mampu lakukan. Penilaian pembelajaran meliputi semua aspek

pengajaran dan pembelajaran, dan dilakukan oleh guru sebagai bagian dari pengajaran.

Assesment adalah upaya meningkatkan mutu pembelajaran, karena penilaian dan pembelajaran merupakan satu kesatuan. Untuk itu penilaian (*assessment*) dilakukan oleh guru dengan menggunakan berbagai metode dan alat penilaian, yaitu melalui ujian, produk, observasi, portofolio, unjuk kerja, dan interview. Proses *assesment* mencakup pengumpulan bukti-bukti kompetensi peserta didik, baik melalui tes, pengamatan, atau laporan diri (*self report*), oleh karena itu dapat dilakukan melalui 4 cara, yaitu:

- a. *Penelusuran*: apakah proses pembelajaran sudah sesuai rencana.
- b. *Pengecekan*: terhadap kompetensi peserta didik.
- c. *Pencarian*: mencari dan menemukan berbagai kendala proses pembelajaran guna mencari solusi untuk mengatasinya.
- d. *Penyimpulan*: menyimpulkan tingkat kompetensi peserta didik sebagai laporan kepada sekolah, orang tua, maupun pemerintah.

1.4.2 Tujuan dan Fungsi Penilaian (*Assesment*)

Menurut Darma (2008) tujuan penilaian adalah:

1. mendeskripsikan kompetensi peserta didik dalam kelompoknya,
2. mengukur keberhasilan proses pembelajaran,
3. *follow up* untuk perbaikannya.
4. sebagai pertanggungjawaban (*accountability*) pendidik kepada orangtua, masyarakat, dan pemerintah.

Bagi peserta didik penilaian memiliki 6 tujuan (Sudrajat, 2008), yaitu untuk:

1. *grading*; menentukan atau membedakan kompetensi antar peserta didik.
2. *alat seleksi*; untuk menggolongkan peserta didik dalam kategori tertentu.
3. *Tingkat penguasaan kompetensi* peserta didik .
4. *bimbingan*; memutuskan langkah selanjutnya.
5. *alat diagnosis*; kesulitan belajar yang dialami peserta didik atau kompetensi yang bisa dikembangkan.
6. *alat prediksi*; untuk memilih jenis pendidikan berikutnya atau profesi yang sesuai. Contoh: tes bakat skolastik atau tes potensi akademik.

Sedangkan fungsi penilaian adalah (Arikunto, 2005):

1. *Fungsi selektif*; hasil penilaian digunakan guru untuk menyeleksi siswanya (diterima, naik ke kelas atau tingkat, mendapat beasiswa, atau lulus).
2. *Fungsi diagnostic*; hasil penilaian digunakan guru mengadakan diagnosis tentang kesulitan belajar peserta didik sehingga dapat diatasi.
3. *Fungsi penempatan*; menentukan peserta didik masuk di kelompok mana.
4. *Fungsi pengukur keberhasilan*; mengukur keberhasilan program.

Menurut Darma (2008), penilaian berfungsi untuk:

1. mengukur ketercapaian pembelajaran,
2. *follow up* perbaikan proses pembelajaran,
3. laporan *progress* peserta didik.

1.4.3 Ruang Lingkup Penilaian (*Assesment*)

1. Aspek Perkembangan Kognitif

Merupakan aspek yang berperan dalam proses pembentukan konsep dan perkembangan pengetahuan peserta didik meliputi:

a. *Perkembangan Bahasa*

Berdasar kemampuan berbahasa peserta didik digolongkan menjadi dua, yaitu kemampuan berbahasa pasif (*reseptif*) yaitu mampu memahami pikiran, perasaan, dan kehendak orang lain, dan kemampuan berbahasa aktif (*ekspresif*) yaitu mampu mengekspresikan pikiran, perasaan dan kehendaknya kepada orang lain. Dalam hal ini assesment untuk mengukur kemampuan berbahasa aktif maupun pasif.

b. *Perkembangan Persepsi*

Assesment perkembangan persepsi merencanakan program pembelajaran, meliputi: (a) *persepsi visual*; warna, hubungan spasial, pembeda visual, pembeda bentuk dan latar, *visual closure*, dan pengenalan objek (*object recognition*), (b) *persepsi auditif*; kesadaran fonologis, pembeda auditif, ingatan auditif, urutan auditif, dan perpaduan auditif, (c) *persepsi kinestetik*; gerak, dan (d) *persepsi taktil*; perabaan.

c. *Perkembangan Motorik*

Assesment perkembangan motorik untuk mengetahui perkembangan motorik mencakup kemampuan melakukan: (a) gerakan kasar (*gross motor*); (b) gerakan halus (*fine motor*); (c) keseimbangan (*balance*); dan (d) koordinasi (*coordination*).

2. Aspek Perkembangan Sosial

Merupakan kemampuan peserta didik berinteraksi dalam situasi tertentu, seperti beradaptasi, menilai situasi, mengikuti aturan, dan menghargai hak milik.

3. Aspek Perkembangan Emosi

Merupakan kemampuan berekspresi (gembira, marah, sedih, takut, dan berani).

4. Aspek Perkembangan Motorik

Kemampuan melakukan gerakan kasar, gerakan halus dan keseimbangan.

1.4.4 Prinsip-prinsip Penilaian (*Assesment*)

Dalam melakukan *assesment* ada tujuh prinsip yang harus diikuti, yaitu:

a. *Memperbaiki belajar peserta didik*

Assesment harus memberikan informasi tentang capaian tujuan belajar peserta didik dan hambatannya, sehingga dapat digunakan untuk memperbaiki cara belajar peserta didik.

b. *Mendukung belajar peserta didik*

Berdasar laporan hasil *assesment* digunakan untuk perbaikan dan *akuntabilitas* sekolah dalam mendukung keberhasilan belajar peserta didik.

c. *Objektif bagi semua peserta didik*

Jenis *assesment* yang sama diberlakukan untuk semua peserta didik.

d. *Kolaborasi profesional*

Dilakukan secara profesional oleh *stakeholders* (Lembaga pendidikan dan Pemerintah).

e. *Partisipasi Komite Sekolah dalam Pengembangan Assesment*

Assesment harus melibatkan orang-orang yang peduli dengan pendidikan (orangtua, komite, dan pakar pendidikan).

f. *Keteraturan dan Kejelasan Komunikasi Mengenai Assesment*

Tujuan, metode, dan hasil *assesment*, baik mengenai sumberdaya, prestasi sekolah, dan kompetensi peserta didik, harus dilakukan secara teratur dan hasilnya dikomunikasikan.

g. *Peninjauan Kembali dan Perbaikan Assesment*

Assesment harus dikaji dan diperbaiki secara berkesinambungan karena kondisi selalu berubah dan pengetahuan terus berkembang, termasuk analisis biaya manfaat (*cost-benefit analysis*) dengan melibatkan pihak-pihak yang berkepentingan (*Stakeholders*).

1.4.5 Teknik Penilaian (*Assesment*)

Penilaian dapat menggunakan teknik tes (tertulis atau lisan), observasi, wawancara, atau tugas rumah. Metode dan alat *assesment* meliputi: lembar observasi, *assesment* mandiri, tugas praktek harian, contoh hasil pekerjaan, tes tertulis, proyek, laporan tertulis, review kinerja, dan *assesment* portofolio. Dalam menilai kompetensi peserta didik guru mengumpulkan informasi menggunakan teknik-teknik penilaian. Standar untuk tugas-tugas harus jelas, kompetensi yang harus didemonstrasikan harus terstandar, performansi peserta didik yang akan diamati harus dikomunikasikan. *Assesment* harus terbuka dan adil bagi semua peserta didik. Komponen yang harus diperhatikan dalam melakukan *assesment* peserta didik yaitu:

1. pengetahuan dan proses yang dituntut dalam tugas harus jelas,
2. aspek-aspek yang di amati tertuang di lembar observasi.
3. performansi peserta didik terdeskripsi secara detail.
4. harus ada performansi *role model* yang dijadikan acuan.

Assesment pembelajaran pada dasarnya ada 4 jenis, yaitu:

a. *Assesment formatif dan sumatif*

Assesment formatif dan sumatif disebut dengan *assesment belajar*. *Assesment* formatif dilaksanakan diakhir proses pembelajaran, sebagai *feedback*. Salah satu bentuknya

adalah *assesment* diagnostic. *Assesment* sumatif dilaksanakan di akhir program, untuk kenaikan kelas peserta didik.

b. *Assesment objektif dan subjektif*

Assesment objektif berupa pertanyaan dengan satu jawaban yang benar (benar dan salah, *multiple choice*, menjodohkan, dan jawaban ganda). *Assesment* subjektif soal dengan lebih dari satu jawaban benar, membutuhkan uraian jawaban yang luas.

c. *Assesment acuan patokan dan acuan normatif*

Assesment acuan patokan, menggunakan tes acuan patokan, untuk mengukur kemampuan peserta didik berdasarkan kriteria tertentu atau *standar absolut*. *Assesment* acuan normatif, atau penentuan rangking berdasarkan kurva normal. *Assesment* ini sangat efektif untuk membandingkan kemampuan antar peserta didik, dan biasanya digunakan untuk ujian masuk sekolah.

d. *Assesment formal dan informal*

Assesment formal berbentuk tes tertulis, hasilnya skor atau rangking peserta didik. Sedangkan *assesment* informal lebih terbuka, misalnya melalui observasi, inventori, partisipasi, evaluasi diri dan teman sebaya, serta diskusi.

Berdasarkan sifatnya dapat berupa:

1. *Assesment Autentik (Assesment Kinerja)*

Penilaian autentik (*authentic assessment*) adalah penilaian yang melibatkan peserta didik dalam menilai kompetensinya. Bentuk *assesment* autentik meliputi: Portofolio, Performance, Proyek, Respon tertulis, Wawancara lisan, Tugas *problem solving* kelompok, Merancang, Membuat presentasi, Penelitian, Menulis/Essei, Merevisi, Mendiskusikan masalah, Analisis lisan, Ceklist, Simulasi, Demonstrasi, Presentasi, Evaluasi, dll. (Corebima, dalam Suraprana dan Hatta, 2004).

Assesment berbasis kinerja (*performance based assesment*) menuntut peserta didik menjawab pertanyaan, membuat produk, mendemonstrasikan ketrampilan, atau menampilkan pengetahuan tingkat tinggi dalam memecahkan masalah realistik atau *autentik*.

2. *Assesment Portofolio*

Penilaian portofolio dapat sebagai alat formatif maupun sumatif karena dapat digunakan untuk memantau kemajuan peserta didik dari hari ke hari dan merefleksi pembelajarannya. Beberapa karakteristik esensial portofolio (Barton dan Collins dalam Suraprana dan Hatta , 2004), yaitu:

- a) *Multi sumber*; dapat untuk menilai berbagai macam evidence.
- b) *Authentic*; sebagai bukti (*evidence*) *autentik*, (*context and evidence are directly linked*).
- c) *Dinamis*; portofolio mencakup perkembangan dan perubahan peserta didik.
- d) *Eksplisit*; kompetensi dasar dan indikatornya harus dinyatakan secara jelas.
- e) *Integrasi*; selalu berkaitan dengan realita nyata.
- f) *Kepemilikan*; harus berkorelasi antara *evidence* kompetensi peserta didik dengan standar kompetensi, kompetensi dasar, atau indikator pencapaian belajar.
- g) *Beragam tujuan*; mencakup beberapa indikator pencapaian hasil belajar

Assesment portofolio berbentuk *file* atau *folder* yang berisi koleksi karya peserta didik dan merupakan bentuk evaluasi kinerja paling populer. *Assesment* portofolio memiliki berbagai keuntungan karena dapat:

- a) mendemonstrasikan keterampilan dan kompetensi peserta didik.
- b) diintegrasikan dengan kegiatan pembelajaran.

- c) mengukur kemampuan berpikir tingkat tinggi dan kompleks, pemahaman mendalam dan aplikatif.
- d) mendorong peserta didik menjadi pembelajar reflektif dan mampu mengukur perkembangan diri sendiri.
- e) meningkatkan kompetensi peserta didik.

DAFTAR PUSTAKA

- Akhmad Sudrajat. 2008. *Penilaian Hasil Belajar*. <http://akhmadsudrajat.wordpress.com/2008/05/01/penilaian-hasil-belajar/>. Diakses 25 Februari 2009
- Aries, Erna Febru. 2011. *Asesmen dan Evaluasi*. Malang: Aditya Media Publishing.
- Arifin, Zainal. 2012. *Evaluasi pembelajaran*. Jakarta: Direktorat Pendidikan Tinggi Islam.
- Arikunto, Suharsimi. 2005. *Dasar-dasar Evaluasi Pendidikan*. Jakarta : Erlangga.
- Mohaimin, Mohammad. 2013. *Model evaluasi CIPP (context, input, process, product)*. Tersedia pada http://www.kompasiana.com/muhaiminmoh/model-evaluasi-cipp-context-input-process-product_552ab300f17e611530d62496. (Diakses pada tanggal 2 April 2016).
- Sukardi. 2008. *Evaluasi pendidikan*. Jakarta: Bumi Aksara
- Supranata dan Hatta. 2004. *Penilaian Portofolio*. Bandung: Remaja Rosdakarya
- Uno, Hamzah, B., dan Koni, Satria. 2012. *Assessment Pembelajaran*. Jakarta: PT Bumi Aksara.

BAB 2

PENYUSUNAN KISI-KISI TES

Oleh Dyah Ayu Fitriana

2.1 Pendahuluan

Kegiatan pengukuran dan evaluasi merupakan suatu hal yang tidak bisa dihindari siapapun yang pernah mengenyam pendidikan formal maupun informal. Dua istilah dalam menilai suatu hasil belajar peserta didik tersebut saling berkaitan satu sama lain tetapi memiliki makna yang berbeda. Kedua istilah yang dimaksud yaitu pengukuran dan evaluasi.

Secara teori, untuk menentukan hasil belajar secara kuantitatif berdasarkan standar/ alat ukur yang telah ditentukan merupakan suatu kegiatan pengukuran. Hasil luaran dari suatu pengukuran yang berbentuk kuantitatif inilah seringkali dalam bentuk skor. Alat ukur digunakan untuk melakukan suatu pengukuran.

Pengukuran sering kali dianggap setara dengan pengujian karena sebagian besar kegiatan pengukuran dilakukan dengan menggunakan alat uji. Kajian kali ini merupakan upaya untuk mengetahui aspek kualitatif berdasarkan kriteria tertentu dari hasil suatu pengukuran yang telah ditentukan alat ukurnya. Namun, tidak semua penilaian perlu diukur terlebih dahulu (pengujian partisipan).

KBK dan KTSP juga sangat mendorong pelaksanaan penilaian kelas, yaitu penilaian dengan menggunakan berbagai teknik, tidak hanya sebatas menggunakan tes seperti sebelumnya. PP No 19 Tahun 2005 Pasal 1 ayat 17 dengan jelas menyebutkan bahwa

untuk mengukur hasil belajar peserta didik dalam mengumpulkan proses informasi selama pembelajaran merupakan kegiatan penilaian.

Untuk memperoleh data/informasi penilaian yang autentik (yang dapat menggambarkan keterampilan nyata), pendidik hendaknya menggunakan teknik penilaian yang berbeda-beda secara saling melengkapi tergantung pada harga keterampilan yang dinilai. Menggabungkan metode penilaian yang berbeda akan memberikan informasi yang lebih akurat tentang kemajuan siswa dan hasil belajar dibandingkan menggunakan tes sebagai satu-satunya metode penilaian.

Penilaian kelas adalah suatu proses di mana pendidik mengumpulkan informasi tentang pertumbuhan siswa dan pencapaian hasil belajar dengan menggunakan berbagai teknik yang dapat mendemonstrasikan, mendemonstrasikan, atau mendemonstrasikan dengan cara yang tepat bahwa keterampilan tersebut benar-benar diperoleh/diimplementasikan, disebut penilaian kelas atau Penilaian Berbasis Kelas (PBK).

Dalam kehidupan sehari-hari, ujian dan penilaian seringkali menentukan nasib siswa. Namun yang menentukan nasib peserta didik dan peserta tes bukanlah hasil tes atau pengukuran itu sendiri melainkan interpretasi (makna) dari hasil pengukuran dan alat ukur tersebut. Hal ini karena orang yang berbeda memberikan interpretasi dan makna yang berbeda terhadap informasi yang sama. Penafsiran yang berbeda ini menempatkan siswa pada risiko dan harus dihindari.

Hal ini juga sama berbahayanya jika informasi yang diberikan tidak akurat akibat kesalahan penggunaan pengujian atau pengukuran yang dilakukan. Hal ini menunjukkan pentingnya pengujian dan pengukuran dalam dunia pendidikan. Jadi, sebagai pendidik, kita tidak bisa mengabaikan bagaimana tes dibuat dan bagaimana tes tersebut digunakan dan diinterpretasikan.

Saat ini pengambilan keputusan terhadap siswa berdasarkan hasil pengukuran dan penilaian pada dasarnya mudah namun tidak kalah sulitnya. Ini relatif mudah jika Anda tidak memikirkan bagaimana keputusan Anda tidak hanya berdampak pada diri Anda sendiri tetapi juga orang lain. Di sisi lain, segala sesuatunya menjadi sulit ketika kita menyadari tanggung jawab atas konsekuensi keputusan kita. Berikut tabel perbedaan pengukuran, penilaian, dan evaluasi.

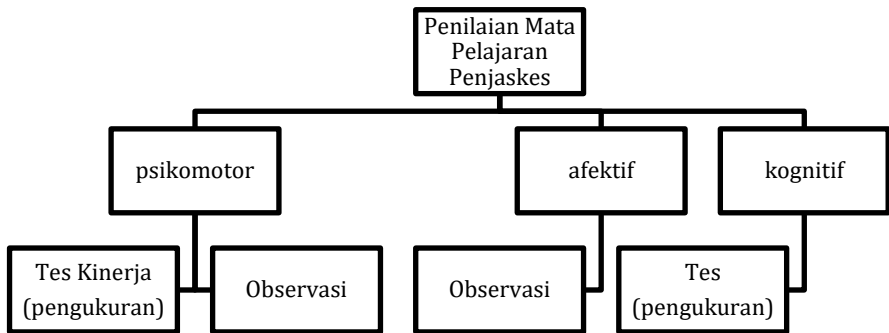
No	Kegiatan	Tujuan	Proses	Hasil
	Pengukuran	Prosedur untuk memberikan angka pada atribut atau karakteristik tertentu dari subyek.	Menggunakan alat ukur/ instrumen untuk memperoleh data: (Leonora dan Eleanor) 1. Mengidentifikasi dan mendefinisikan kualitas atau atribut yang akan diukur 2. Menentukan sekumpulan kegiatan yang dengan instrumen dapat dibuat nyata dan dapat dipahami 3. Menetapkan definisi untuk menerjemahkan pengamatan ke dalam pernyataan kuantitatif tentang gelar atau jumlah	Angka (<i>score</i>) (kuantitatif)

No	Kegiatan	Tujuan	Proses	Hasil
	Penilaian	Mengumpulkan informasi yang digunakan untuk membuat keputusan tentang siswa, kurikulum, program, dan sekolah, serta kebijakan pendidikan	Membandingkan data dengan kriteria tertentu	Nilai (<i>decision</i>) (kualitatif)
	Evaluasi	Proses membuat keputusan tentang kelayakan produk atau performa siswa	<i>Delineating, obtaining, providing</i> Terintegrasi dengan pengalaman evaluator	Keputusan (<i>judgement</i>) dan atau Rekomendasi (kualitatif)

Sebagai contoh,

Dilakukan evaluasi model pembelajaran kooperatif pada mata pelajaran kesehatan, jasmani, dan olahraga di SMA kelas VIII. Untuk melakukan evaluasi maka dilakukan penilaian yang meliputi penilaian psikomotor, afektif, dan kognitif.

1. **Penilaian psikomotor** dilakukan dengan tes kinerja (melakukan pengukuran) dan pengamatan.
2. **Penilaian afektif** dilakukan melalui pengamatan menggunakan lembar observasi.
3. **Penilaian kognitif** dilakukan melalui tes (melakukan pengukuran)



Selanjutnya hasil pengukuran dan pengumpulan data tersebut diolah kemudian dibandingkan dengan kriteria tertentu. Berikut ini cara mengolah skor hasil pengukuran dan kriteria yang digunakan untuk menghasilkan nilai.

2.2 Karakteristik Penilaian Pendidikan

Untuk lebih memahami pentingnya mengukur dan mengevaluasi hasil pembelajaran, maka perlu terlebih dahulu diperhatikan ciri-ciri dan ciri-ciri evaluasi dalam pendidikan:

1. Evaluasi (pengukuran) dalam dunia pendidikan tidak dapat dilakukan secara langsung melainkan hanya berdasarkan gejala yang terlihat. Misalnya mengukur ataupun menilai hasil belajar siswa berdasarkan kemampuannya dalam menjawab soal tes hasil belajar yang disiapkan oleh guru.
2. Evaluasi biasanya diawali dengan pengukuran kegiatan, dan kuantitatif merupakan hasil luarannya, yaitu dengan menggunakan simbol-simbol numerik. Hasil pengukuran diartikan sebagai ukuran kualitatif berupa simbol angka.

3. Dalam melakukan pengukuran kegiatan untuk tujuan evaluasi, kesalahan pengukuran selalu terjadi karena sumber-sumber sebagai berikut: (1) Alat pengukur (tidak valid atau tidak dapat diandalkan), (2) Pemeriksa (faktor subjektif, kecenderungan skor disukai atau tinggi, perasaan pribadi seorang pendidik terhadap peserta ujian, pengaruh hasil sebelumnya, suasana hati pemeriksa); (3) Kondisi mental dan fisik peserta ujian, (4) Kesalahan yang disebabkan oleh suasana ujian (suasana bising, pengawasan kurang, dan lain-lain).

2.3 Pengertian Kisi-kisi

Kisi-kisi memiliki bentuk berupa matriks yang berisi informasi yang digunakan sebagai pedoman untuk menulis soal atau mengelompokkan soal ke dalam suatu tes. Menata kisi-kisi merupakan langkah penting yang harus dilakukan sebelum menulis soal. Kisi-kisi tersebut disusun sesuai dengan tujuan penggunaan tes. Dengan cara ini, berbagai jenis jaring dapat diperoleh. Kisi-kisi suatu tes dimaksudkan untuk menyesuaikan atau menyiapkan soal-soal untuk mendiagnosis kesulitan belajar siswa berbeda dengan kisi-kisi tes yang dimaksudkan untuk menyiapkan soal-soal tentang kinerja akademik.

Kisi-kisi yang digunakan untuk menyusun tes penempatan juga berbeda dengan kisi-kisi yang digunakan untuk menyusun ujian kompetitif. Kisi-kisi yang digunakan untuk mempersiapkan ujian umum juga berbeda dengan kisi-kisi yang digunakan untuk mempersiapkan ujian akhir nasional. Perlu dicatat bahwa tidak ada satu kisi-kisi yang dapat digunakan untuk semua tujuan pengujian.

Contoh format pembuatan kisi-kisi untuk penulisan soal:

Kisi-kisi Penulisan Soal

Nama Sekolah :
Mata Pelajaran :
Kelas dan Semester :
Kurikulum :

SK	KD	Materi	Pengalaman Belajar	Indikator	Penilaian	
					Kognitif	No Soal

2.2.1 Fungsi kisi-kisi

Kisi-kisi tes sebagai panduan untuk menulis pertanyaan dan menyusun tes. Dengan menggunakan panduan ini, penulis soal dapat membuat pertanyaan yang selaras dengan tujuan tes, dan penulis tes dapat dengan mudah merakit alat tes. Oleh karena itu, jika tersedia rubrik yang baik, penulis soal yang berbeda akan mampu menghasilkan rangkaian pertanyaan yang relatif sama, baik dari segi kedalaman maupun cakupan materi yang dibutuhkan. Berikut perbandingan fungsi pengujiannya:

- a. Fungsi untuk kelas
 - 1) Melakukan diagnosis terhadap kesulitan belajar siswa
 - 2) Menilai kesenjangan antara bakat dengan kesuksesan.
 - 3) Meningkatkan tingkat keberhasilan.
 - 4) Mengelompokkan siswa dalam kelas dengan menggunakan metode kelompok.
 - 5) Merencanakan kegiatan proses belajar mengajar kepada setiap individu

- 6) Menentukan peserta didik (siswa) mana yang memerlukan pengajaran khusus dikarenakan tertinggal atau besar kekurangannya.
 - 7) Tentukan besarnya keberhasilan dari masing-masing individu
- b. Fungsi untuk bimbingan.
- 1) Mengawasi arah pembicaraan tentang peserta didik dengan wali murid.
 - 2) Membantu dan mengarahkan peserta didik dalam mengambil suatu keputusan.
 - 3) Membantu dan mengarahkan peserta didik dalam mencapai tujuan dari pendidikan dan keprofesionalan
 - 4) Memberikan waktu kepada konselor, guru, dan wali murid untuk memahami dan menemukan kesulitan yang dialami maupun dijumpai pada peserta.
- c. Fungsi untuk administrasi.
- 1) Mengarah siswa dalam kelompok
 - 2) Mengatur siswa baru
 - 3) Mendukung siswa dalam memilih kelompok
 - 4) Evaluasi kurikulum.
 - 5) Mengembangkan hubungan masyarakat
 - 6) Memberikan informasi kepada organisasi selain sekolah

2.2.2 Syarat Kisi-kisi

Memperhatikan perbedaan variasi dari kisi-kisi yang disajikan, maka dapat disimpulkan bahwa kisi-kisi tersebut harus memenuhi beberapa syarat, yaitu:

1. Penyajian soal-soal yang telah disesuaikan dengan kurikulum yang akan diujikan
2. Komponen yang mudah dipahami, terperinci, dan jelas

3. Format soal dan indikator telah sesuai dengan buti masing-masing pertanyaan yang dibuat

2.2.3 Elemen Kisi-kisi

Elemen-elemen yang diperlukan dalam grid ditentukan oleh tujuan tes yang sedang disiapkan. Komponen-komponen tersebut dapat dikelompokkan menjadi dua kelompok, yaitu kelompok identitas dan kelompok matriks. Kelompok identitas dicantumkan pada bagian atas matriks, sedangkan kelompok matriks dicantumkan pada kolom-kolom sesuai dengan tujuan pengujian. Faktor-faktor yang umum digunakan dalam pengembangan kisi-kisi tes prestasi akademik adalah sebagai berikut:

- a. Nama Sekolah
- b. Mata Pelajaran
- c. Kelas dan Semester
- d. Kurikulum
- e. Standar Kompetensi
- f. Kompetensi Dasar
- g. Materi Pokok
- h. Pengalaman Belajar
- i. Indikator
- j. Tingkat Kognitif
- k. Nomer Butir Soal

Idealnya, selama proses pembelajaran yang diujikan dalam suatu kelas telah mencakup seluruh komponen keterampilan dasar dan indikator kurikulum yang berlaku pada masing-masing satuan ajar. Namun diantara komponen-komponen tersebut di atas, khususnya ujian umum, ujian peningkatan, ujian akhir nasional, komponen kompetensi dasar dan indeks merupakan komponen yang wajib dimiliki, diseleksi secara mendalam. Sebenarnya, ini soal memilih siapa yang akan diuji. Pilihan ini diambil karena dalam satu tes tidak mungkin menguji seluruh keterampilan dasar dan indikator yang ada dalam kurikulum pada waktu singkat. Pemilihan keterampilan dasar dilakukan dengan memperhatikan:

- a. Urgensi, artinya secara mutlak harus dikuasai oleh siswa dalam menguasai indikator maupun kompetensi dasar
- b. Kontinuitas, artinya indikator atau kompetensi utama pada tingkat yang sama atau antar tingkat dalam mengembangkan lebih lanjut dari satu atau lebih kompetensi utama
- c. Relevansi, artinya kompetensi inti atau indikator yang dipilih hendaknya yang diperlukan untuk mempelajari atau memahami bidang studi lain.
- d. Penerapan, keterampilan atau indikator utama harus menjadi keterampilan inti dan indikator yang mempunyai aplikabilitas tinggi pada kehidupan sehari-hari dari peserta didik..

Untuk memilih kompetensi inti dan indikatornya, selain memperhatikan kriteria seleksi di atas juga perlu memperhatikan bahwa penguasaan materi dan indikator kompetensi inti yang dipilih harus diukur dengan menggunakan format pertanyaan yang telah ditentukan. Misalnya, jika seseorang memutuskan untuk mengikuti tes pilihan ganda, penguasaan keterampilan dasar dan metrik dapat diukur dengan menggunakan pertanyaan pilihan ganda. Sebaliknya jika diputuskan untuk mengikuti tes deskriptif, maka tingkat penguasaan keterampilan dasar atau indikator yang dipilih juga harus dapat diukur dengan tes deskriptif.

Semua komponen kisi-kisi yang disebutkan sebelumnya merupakan komponen penting untuk tata letak kisi-kisi. Namun, tidak ada instruksi atau kewajiban untuk menggunakan semua bahan tersebut. Penggunaan komponen-komponen tersebut disesuaikan seperlunya tergantung pada jenis dan tujuan pengujian yang disiapkan. Setelah menentukan elemen mana yang akan dimasukkan ke dalam kisi-kisi, langkah selanjutnya adalah memasukkan semua elemen tersebut ke dalam suatu format atau matriks.

2.4 Kesimpulan

Kisi-kisi adalah suatu bentuk matriks yang memuat informasi yang digunakan sebagai pedoman penulisan soal atau kelompok soal pada suatu tes. Sebelum menulis soal, terdapat langkah penting yang harus terlebih dahulu dilakukan yaitu membuat tata letak kisi-kisi.

Kegiatan menulis adalah kegiatan membuat huruf (angka) dengan menggunakan alat tulis pada alat tulis atau media untuk menyampaikan gagasan, pemikiran, dan komunikasi secara tertulis kepada pihak lain. Penulis merupakan seseorang yang melakukan pekerjaan menulis. Proses pengumpulan data oleh individu atau organisasi bisnis dilakukan melalui metode yang tepat dan sistematis.

Kotak tes berfungsi sebagai panduan untuk menulis soal dan mempersiapkan tes. Penggunaan panduan ini memungkinkan penulis soal membuat pertanyaan yang memenuhi tujuan dari suatu tes, dan penulis dari suatu tes dapat dengan cepat dan tepat membuat alat tes. Rubrik yang baik dapat memungkinkan penulis pertanyaan yang berbeda untuk membuat pertanyaan yang relatif seragam baik isi maupun telah disesuaikan dengan cakupan materi yang dibutuhkan.

Faktor-faktor yang digunakan dalam penyusunan garis besar tes adalah nama sekolah, mata pelajaran, kelas/semester, kurikulum, standar kompetensi, keterampilan dasar, mata pelajaran khusus, pengalaman pengalaman belajar, indeks, penilaian tingkat kognitif, jumlah soal..

DAFTAR PUSTAKA

- Akhadiah Sabarti. 1988. Evaluasi Dalam Pengajaran Bahasa. Jakarta: Departemen Pendidikan dan Kebudayaan.
- Arikunto Suharsimi. 2005. Dasar-dasar Evaluasi Pendidikan. Jakarta: bumi aksara.
- Basroi & Siskandar. 2012. Evaluasi Belajar Berbasis Kinerja. Bandung: Karyaputra Darwati.
- Penulisan.pdf. 2023. Tersedia di Wikipedia.org (Diakses tanggal 21 November 2023)
- Penyusunan.pdf. 2023. Tersedia di Felib.unikom.ac.id (Diakses tanggal 21 November 2023)
- Sudijono Anas. 2012. Pengantar Evaluasi Pendidikan. Jakarta: Kharisma Putra Utama Offset.
- Syrapranata Sumarna. 2005. Panduan Penulisan Tes Tertulis Implementasi Kurikulum 2004. Bandung: Remaja Rosdakarya

Bab 3

UJI COBA ANALISIS VALIDITAS DAN RELIABILITAS SOAL

Oleh Achmad Rante Suparman

3.1 Definisi

Reliabilitas dan validitas tes merupakan dua sifat utama yang melekat pada pengembangan suatu tes. Reliabilitas dan validitas merupakan penentu kualitas dari suatu tes. Validitas dan reliabilitas merupakan dua fitur terpenting dari sebuah tes karena berkaitan dengan proses evaluasi kesesuaian tes untuk tujuan pengguna. Bab ini memberikan penjelasan tentang uji coba berupa validitas dan reliabilitas. Penjelasan ini akan membantu dalam memahami informasi reliabilitas dan validitas yang dilaporkan dalam manual tes dan tinjauan serta menggunakan informasi tersebut untuk mengevaluasi kesesuaian tes yang akan digunakan.

Validitas mengacu pada seberapa baik suatu tes yang dibuat dapat mengukur apa yang menjadi tujuan pengukuran, sedangkan reliabilitas merupakan sejauh mana tes yang dibuat menghasilkan hasil yang stabil dan konsisten (Cohen, Manion and Morrison, 2018). Menurut Sumintono and Widhiarso (2015), seberapa tepat pengukuran suatu instrumen yang dibuat dapat mengukur atribut yang seharusnya diukur sesuai dengan tujuan yang telah ditetapkan disebut dengan validitas. Pendapat ini sejalan dengan Azwar (2013) yang menyatakan bahwa validitas merupakan pertimbangan yang harus diutamakan dalam proses mengevaluasi kualitas suatu tes sebagai instrumen ukur. Kebermaknaan, kelayakan dan kebermanfaatan intervensi tertentu merupakan

acuan dalam konsep validasi yang dapat didasarkan pada skor hasil tes.

Validitas memberi arti pada nilai tes serta memberikan gambaran sejauh mana penarikan kesimpulan atau prediksi spesifik tentang seseorang berdasarkan nilai tes yang diperoleh. Reliabilitas dapat didefinisikan sebagai proporsi varians dalam skor tes yang diamati yang diperhitungkan oleh konstruk yang diukur oleh instrumen (Finch and French, 2019). Konsistensi mempunyai dampak atau pengaruh dalam penentuan reliabilitas karena konsistensi berhubungan erat dengan tingkat kesalahan hasil suatu tes yang berupa skor. Suatu koefisien yang memperlihatkan keajekan atau konsistensi dari hasil pengukuran suatu tes disebut sebagai reliabilitas (Mardapi, 2018b). Reliabilitas mengacu pada akurasi atau presisi dari prosedur pengukuran. Indeks reliabilitas memberikan indikasi sejauh mana skor yang dihasilkan oleh prosedur pengukuran tertentu konsisten dan dapat direproduksi (Thorndike and Chris, 2014).

Memahami perbedaan antara reliabilitas dan validitas merupakan suatu hal penting harus dipahami oleh pengembang soal. Validitas dapat memberi informasi tentang bagusny suatu tes sedangkan reliabilitas akan memberi informasi tentang tingkat kepercayaan skor pada tes tersebut. Seorang peneliti tidak dapat menarik kesimpulan yang valid dari suatu nilai tes kecuali meyakini bahwa tes tersebut dapat diandalkan. Meskipun suatu tes dapat diandalkan, tes tersebut mungkin tidak valid. Hal ini mengharuskan agar berhati-hati agar tes yang dibuat dapat diandalkan dan valid untuk situasi yang diharapkan.

Tujuan dikembangkannya suatu tes sangat erat kaitannya dengan validitas, karena hanya akan valid sesuai dengan tujuan utamanya. Misalnya, tes yang digunakan untuk mengetahui tentang kemahiran seseorang dalam menyetarakan suatu reaksi kimia, mungkin tidak valid untuk memprediksi kemampuan seseorang dalam penentuan bentuk molekul. Hal ini mengarah pada prinsip

penilaian berikutnya. Demikian pula, validitas tes ditentukan dengan mengacu pada kelompok tertentu. Kelompok-kelompok ini disebut kelompok referensi. Tes ini mungkin tidak valid untuk kelompok yang berbeda.

Reliabilitas pada soal menghasilkan informasi yang dapat diandalkan, dapat diulang, dan konsisten terhadap hasilnya. Reliabilitas dapat digunakan untuk menafsirkan nilai tes secara bermakna dan membuat keputusan yang berguna terkait tujuan yang diharapkan. Sejauh mana nilai tes tidak terpengaruh oleh kesalahan pengukuran (*measurement error*) merupakan indikasi reliabilitas yang baik. Jika tidak terdapat kesalahan pengukuran yang acak, individu tersebut akan mendapatkan skor tes yang sama, yaitu skor "benar" setiap saat. Beberapa kesalahan pengukuran dapat disebabkan oleh: (1) kinerja tes; (2) faktor perbedaan lingkungan pengujian; (3) jenis tes memiliki lebih dari satu versi atau bentuk; (4) banyaknya penilai.

Validitas dan reliabilitas dalam pengembangan tes dapat dibedakan menjadi 2 yaitu validitas dan reliabilitas secara klasik dan validitas dan reliabilitas secara modern. Validitas dan reliabilitas secara klasik adalah perhitungan menentukan valid atau reliabel menggunakan teori klasik, sedangkan pada validitas dan reliabilitas secara modern adalah penentuan valid dan reliabel menggunakan teori tes modern.

3.2 Validitas Soal Secara Klasik

Validitas secara umum dibagi menjadi dua, yaitu validitas logis dan validitas empiris (Arikunto, 2018). Sesuai dengan namanya, validitas logis menggunakan logika untuk menetapkan bukti validitas. Meskipun beberapa argumen kuat dapat dibuat dengan menggunakan logika untuk menunjukkan validitas suatu instrumen, validitas logis jauh lebih lemah dibandingkan validitas empiris, meskipun validitas logis hampir selalu menjadi

pendekatan pertama yang digunakan untuk mengatasi kekhawatiran mengenai validitas instrumen apa pun. Validitas empiris berdasarkan kepada bukti-bukti, dimana data dikumpulkan untuk menunjukkan bahwa skor yang diperoleh dari suatu instrumen berperilaku dapat diprediksi. Semakin banyak bukti yang diberikan untuk menunjukkan bahwa skor tersebut benar-benar mewakili apa yang seharusnya diwakilinya, semakin banyak pula bukti validitas skor yang diperoleh dari instrumen tersebut (Schaughency *et al.*, 2012).

Validitas dibagi menjadi 4 kategori besar, yaitu *content validity*, *construct validity*, *concurrent validity* dan *predictive validity* (Istiyono, 2018). Pada bab ini akan dibahas salah satu uji coba dengan menggunakan rumus korelasi *product moment*. Rumus *product moment* sebagai berikut (Arikunto, 2018):

$$r_{xy} = \frac{N\Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{\{N\Sigma X^2 - (\Sigma X)^2\} \{N\Sigma Y^2 - (\Sigma Y)^2\}}} \dots\dots\dots 1)$$

Penggunaan rumus *product moment* ini dapat digunakan pada saat akan menentukan validitas dari suatu butir tes dalam suatu soal. Misalnya seorang peneliti akan menentukan validitas dari 24 nomor tes soal pilihan ganda yang telah dibuatnya. Setelah 24 butir tes telah dibuat, maka selanjutnya 24 butir soal dapat diberikan kepada beberapa peserta tes untuk dikerjakan. Hasil dari pengerjaan tes tersebut, selanjutnya dibuat dalam suatu tabel seperti contoh pada Tabel 6.1 tentang nilai hasil uji coba.

Penentuan valid tidaknya butir tes pada soal yang dikembangkan selanjutnya dihitung dengan meenggunakan rumus 1, dimana X adalah skor butir tes yang akan ditentukan validitasnya dan Y adalah skor total seluruh perolehan setiap siswa. Dalam perhitungan dengan menggunakan rumus *product moment* ini, harap diperhatikan dengan baik antara ΣY^2 dan $(\Sigma Y)^2$ serta ΣX^2 dan $(\Sigma X)^2$ karena sering terjadi kekelirian dalam perhitungan secara manual disebabkan kesalahan dalam mengartikannya.

Sesuai Tabel 3.1 akan dihitung validitas butir nomor 5, apakah butir nomor 5 valid atau tidak.

Hal pertama yang harus dilakukan adalah dengan menganggap butir nomor 5 sebagai X dan total perolehan peserta tes sebagai Y. Contoh penggunaan X dan Y sesuai dengan Tabel 3.2.

Tabel 3. 1. Hasil uji coba tes pilihan ganda

Siswa	Nomor Soal																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Angga	1	1	1	1	0	1	1	1	1	1	0	1	0	1	1	1	0	1	1	1	1	1	0	0
Abdi	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	0
Asrii	1	1	1	1	0	1	1	0	1	1	1	0	0	0	1	0	0	1	1	1	1	1	1	0
Badi	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	0	0	1	0	1	1	1	1	0
Basri	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
Biyah	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1
Cerika	1	1	1	1	0	0	1	0	1	1	1	1	0	1	1	0	1	0	1	1	1	1	1	0
Cendi	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1
Capu	1	1	1	1	1	0	1	0	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1
Delima	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0
Delta	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	0	1	0	0	1	0	1	1
Donda	1	1	1	1	0	0	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	0
Engelina	1	1	1	1	0	0	1	1	0	1	0	1	1	0	1	0	1	1	1	1	1	1	0	1
Enre	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1
Flor	1	1	1	1	1	1	0	1	1	1	1	1	0	0	1	0	0	1	0	1	1	1	1	0
Fredrika	1	1	1	1	1	0	1	1	0	0	1	0	1	0	0	0	0	1	0	0	1	1	0	0
Gustafo	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
Gindal	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	0
Halima	1	1	1	1	0	1	1	1	1	1	1	0	1	0	1	1	0	1	0	0	1	1	1	0
Husni	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1
Proton	1	1	1	1	1	0	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
Prenda	1	1	1	1	1	0	0	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	0
Qris	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1
Qiqy	1	1	1	1	0	0	1	0	1	1	0	1	0	1	1	0	0	1	0	1	1	0	1	0
Raul	1	1	1	1	1	0	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	0
Risma	1	1	1	1	1	1	0	1	1	1	0	1	0	1	1	1	0	1	1	1	1	1	1	1
Tata	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	0	1	0	1	1	1	1	1	0
Titi	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1
Umar	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
Udin	1	1	1	1	1	0	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1

Tabel 3. 2. Perhitungan butir nomor 5

Siswa					X	Nomor Soal																								Y
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	Total					
Angga	1	1	1	1	0	1	1	1	1	1	0	1	0	1	1	1	0	1	1	1	1	1	0	0	0	18				
Abdi	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	0	0	20				
Asrii	1	1	1	1	0	1	1	0	1	1	1	0	0	0	1	0	0	1	1	1	1	1	1	1	0	16				
Badi	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	0	0	1	0	1	1	1	1	1	0	17				
Basri	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	21				
Biyah	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	22				
Cerika	1	1	1	1	1	0	0	1	0	1	1	1	1	0	1	1	0	1	0	1	1	1	1	1	1	0	17			
Cendi	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	21				
Capu	1	1	1	1	1	0	1	0	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	20				
Delima	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	21				
Delta	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	0	1	0	0	1	0	1	0	1	18				
Donda	1	1	1	1	1	0	0	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	0	17				
Engelina	1	1	1	1	0	0	1	1	0	1	0	1	1	0	1	0	1	1	1	1	1	1	0	1	0	16				
Enre	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	21				
Flor	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0	1	0	0	1	0	1	1	1	1	0	17				
Fredrika	1	1	1	1	1	1	0	1	1	0	0	1	0	1	0	0	0	0	1	0	0	1	1	0	0	12				
Gustafo	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	21				
Gindal	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	0	20				
Halima	1	1	1	1	0	1	1	1	1	1	1	0	1	0	1	1	0	1	0	0	1	1	1	1	0	17				
Husni	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	21				
Proton	1	1	1	1	1	0	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	20				
Prenda	1	1	1	1	1	0	0	0	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	0	19				
Qris	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	21				
Qiqy	1	1	1	1	1	0	0	1	0	1	1	0	1	0	1	1	0	0	1	0	1	1	0	1	0	14				
Raul	1	1	1	1	1	1	0	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	0	18				
Risma	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	0	1	1	1	1	1	1	1	1	20				
Tata	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	0	1	0	1	1	1	1	1	0	18				
Titi	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	22				
Umar	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	22				
Udin	1	1	1	1	1	1	0	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	19				

Pada ujicoba dengan data pada Tabel 3.2 dimasukkan pada rumus 1, maka akan diperoleh hasil:

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

$$r_{xy} = \frac{(30)(430) - (22)(566)}{\sqrt{\{(30)(22) - (22)^2\} \{(30)(10854) - (566)^2\}}}$$

$$r_{xy} = 0,465$$

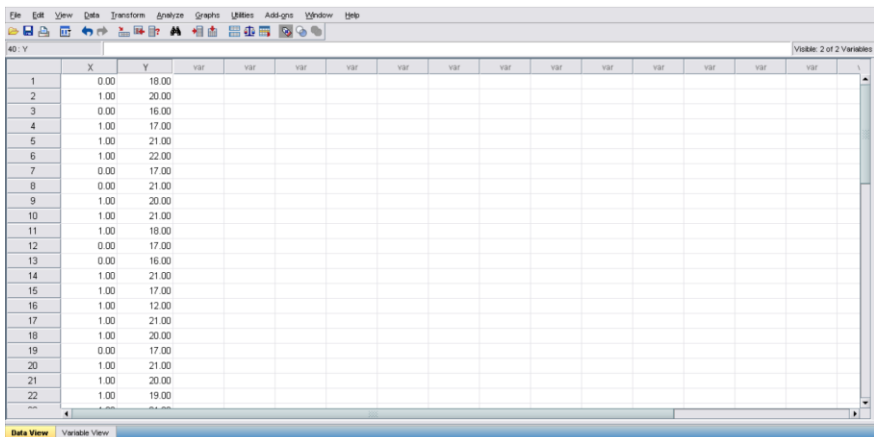
Nilai yang diperoleh selanjutnya disesuaikan dengan kriteria yang dinyatakan oleh Arikunto (2018) sesuai Tabel 3.3.

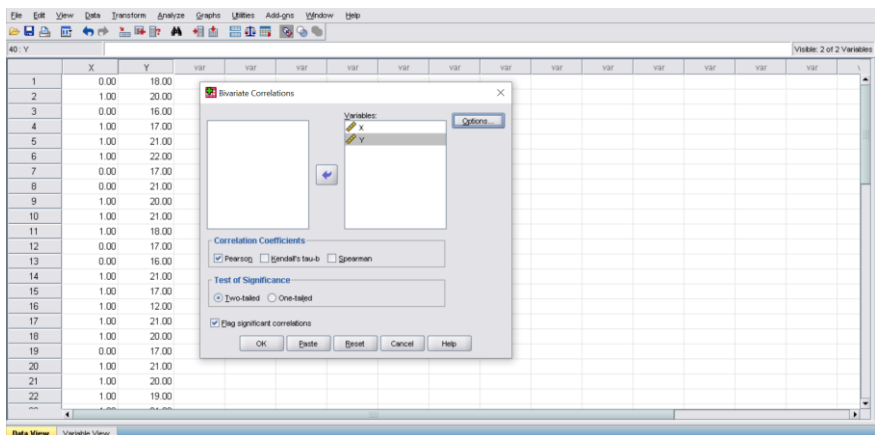
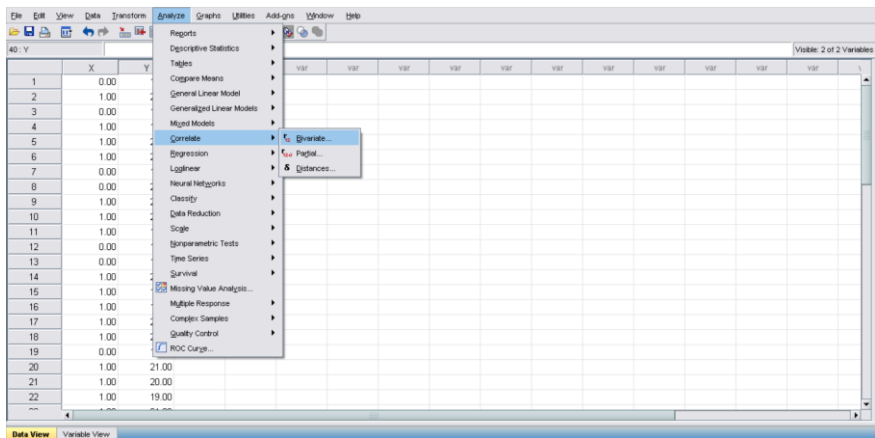
Tabel 3. 3. Interpretasi nilai koefisien korelasi

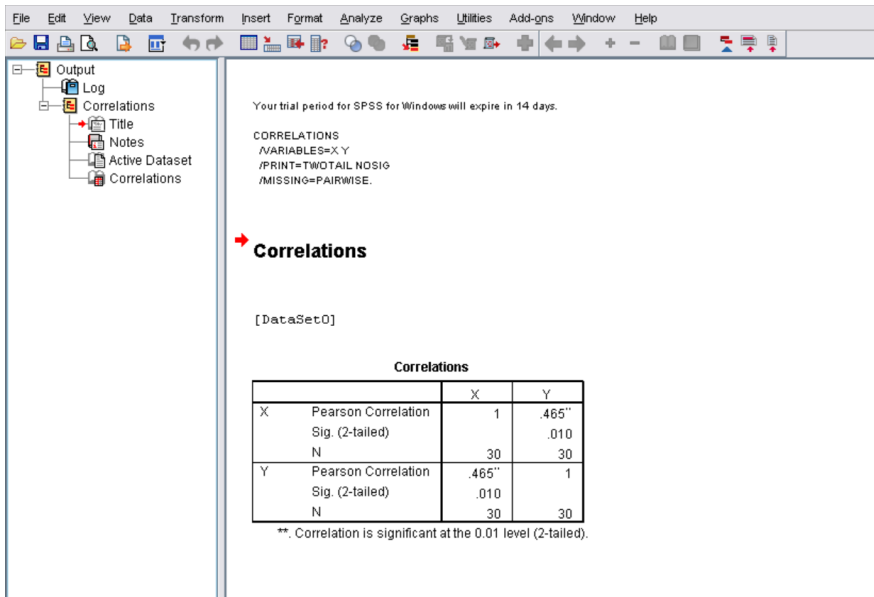
Nilai koefisien korelasi	Keterangan
0,00 – 0,20	Sangat rendah
>0,20 – 0,40	Rendah
>0,40 – 0,60	Cukup
>0,60 – 0,80	Tinggi
>0,80 – 1,00	Sangat tinggi

Berdasarkan nilai korelasi pada Tabel 3.3, maka nilai yang diperoleh pada perhitungan dengan nilai $r_{xy} = 0,465$ masuk pada kriteria cukup, sehingga dapat disimpulkan kriteria validnya cukup. Dengan menggunakan cara yang sama untuk butir lainnya, maka dapat diperoleh nilai r_{xy} sehingga dapat ditentukan kriterianya.

Penggunaan rumus 1 apabila digunakan dalam jumlah butir yang banyak tentu akan membutuhkan waktu yang lumayan lama. Cara lain yang dapat digunakan adalah dengan menggunakan bantuan SPSS. Adapun langkah-langkah yang dapat digunakan dalam SPSS sesuai Gambar 3.1.







Gambar 3. 1. Analisis korelasi *product moment* dengan SPSS

Pada Gambar 3.1 diperoleh bahwa nilai korelasi yang diperoleh adalah sebesar 0,465, nilai ini sama dengan nilai yang diperoleh dengan perhitungan manual. Selain menggunakan bantuan SPSS, nilai korelasi ini juga dapat menggunakan bantuan R Program. Pada R program menggunakan sintaks sebagai berikut:

```
data_x <- c(0, 1, 0, 1, 1, 1, 0, 0,
            1, 1, 1, 0, 0, 1, 1, 1,
            1, 1, 0, 1, 1, 1, 1, 0,
            1, 1, 1, 1, 1, 1)
data_y <- c(18, 20, 16, 17, 21, 22, 17, 21,
            20, 21, 18, 17, 16, 21, 17, 12,
            21, 20, 17, 21, 20, 19, 21, 14,
            18, 20, 18, 22, 22, 19)

# Menghitung korelasi Pearson
correlation_coefficient <- cor(data_x, data_y)

print(paste("Koefisien Korelasi Pearson:", correlation_coefficient))
```

The screenshot displays the RStudio interface with the following components:

- Menu Bar:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Toolbar:** Includes icons for saving, running, and other standard RStudio functions.
- Source Editor:** Contains R code for data entry and correlation calculation.


```

2021
2022
2023
2024
2025
2026
2027 data_x <- c(0, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1, 1,
2028           1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1)
2029 data_y <- c(18, 20, 16, 17, 21, 22, 17, 21, 20, 18, 17, 16, 21, 17, 12, 21,
2030           20, 17, 21, 20, 19, 21, 14, 18, 20, 18, 22, 22, 19)
2031
2032 # Menghitung korelasi Pearson
2033 correlation_coefficient <- cor(data_x, data_y)
2034
2035 print(paste("koefisien Korelasi Pearson:", correlation_coefficient))
2036
      
```
- Environment Pane:** Shows the 'Values' of the objects created.

Object	Value
correlation_coe_	0.465440038458647
data_x	num [1:30] 0 1 0 1 1 1 0 0 1 1 ...
data_y	num [1:30] 18 20 16 17 21 22 17 21 20 18 ...
- Console:** Displays the execution output.


```

R 4.3.1 ~>
> data_x <- c(0, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1, 1,
1, 1, 0, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1,
+ 1, 1, 1, 1)
> data_y <- c(18, 20, 16, 17, 21, 22, 17, 21, 20, 18, 17, 16, 21, 17, 12, 21,
+ 20, 17, 21, 20, 19, 21, 14, 18, 20, 18, 22, 22, 19)
> # Menghitung korelasi Pearson
> correlation_coefficient <- cor(data_x, data_y)
> print(paste("koefisien Korelasi Pearson:", correlation_coefficient))
[1] "koefisien Korelasi Pearson: 0.465440038458647"
      
```

Pada Gambar 3.2 terlihat bahwa hasil yang diperoleh adalah 0,465440038458647 atau setara dengan 0,465. Nilai yang diperoleh dengan perhitungan manual, SPSS dan R program semua memberikan nilai yang sama yaitu 0,465 yang menandakan bahwa validatas butir 5 cukup untuk digunakan.

Hubungan antara validitas dan reliabilitas masih menjadi sorotan utama para pakar pengukuran. Singkatnya, akan sulit mempunyai validitas tanpa reliabilitas, sebab validitas berarti skor dari suatu instrumen mewakili apa yang seharusnya diukur, dan

reliabilitas berarti seseorang memperoleh skor yang konsisten. Apabila suatu tes tidak reliabel, maka apabila dilakukan tes kembali hasilnya jelas akan berbeda (Sukardi, 2012).

Menurut Azwar (2016), terdapat tiga estimasi yang dapat digunakan dalam perhitungan reliabilitas, yaitu *tes-retest*, *alternate-forms* dan *internal consistency*. *Internal consistency* mengandalkan data hasil uji coba yang telah dilakukan sebelumnya. Terdapat berbagai teknik dalam *internal consistency*, antara lain: formula Spearman Brown, KR 20, KR 21, formula Kuder Richardson, formula Alpha, formula Rulon dan sebagainya. Pada bab ini, khusus akan memberikan penjelasan uji coba dengan menggunakan formula Spearman Brown.

Salah satu syarat utama penggunaan formula Spearman Brown adalah butir soal berjumlah genap. Pada hasil uji coba soal sesuai pada Tabel 3.1 selanjutnya dibuat menjadi butir ganjil dan butir genap sesuai Tabel 3.4. Pada Tabel 3.4, skor total butir ganjil diberi simbol X dan total butir genap diberi simbol Y.

Tabel 3. 4. Hasil pembagian butir ganjil dan butir genap

Siswa	Nomor Soal																								X	Y
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	Total	Total
Angga	1	1	1	1	0	1	1	1	1	1	0	1	0	1	1	1	0	1	1	1	0	0	1	0	7	11
Abdi	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	0	11	9
Asrii	1	1	1	1	0	1	1	0	1	1	1	0	0	0	1	0	0	1	1	1	1	1	1	0	9	7
Badi	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	0	0	1	0	1	1	1	1	0	9	8
Basri	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	12	9
Biyah	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	11	11
Cerika	1	1	1	1	0	0	1	0	1	1	1	1	0	1	1	0	1	0	1	1	1	1	1	0	10	7
Cendi	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	11	10
Capu	1	1	1	1	1	0	1	0	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	11	9
Delima	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	12	9
Delta	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	0	1	0	0	1	0	1	1	10	8
Donda	1	1	1	1	0	0	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	0	11	6
Engelina	1	1	1	1	0	0	1	1	0	1	1	0	1	1	0	1	0	1	1	1	1	0	1	0	9	7
Enre	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	11	10
Flor	1	1	1	1	1	1	0	1	1	1	1	0	0	1	0	0	1	0	1	0	1	1	1	0	8	9
Fredrika	1	1	1	1	1	0	1	1	0	0	1	0	1	0	0	0	0	1	0	0	1	1	0	0	7	5
Gustafo	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	10	11
Gindal	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	0	10	10
Halima	1	1	1	1	0	1	1	1	1	1	1	0	1	0	1	1	0	1	0	0	0	1	1	0	9	8
Husni	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	11	10
Proton	1	1	1	1	1	0	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	10	10
Prenda	1	1	1	1	1	0	0	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	0	11	8
Qris	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	12	9
Qiqy	1	1	1	1	0	0	1	0	1	1	0	1	0	1	1	0	0	1	0	1	0	1	0	1	7	7
Raul	1	1	1	1	1	0	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	0	12	6
Risma	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	0	1	1	1	1	1	1	1	8	12
Tata	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	0	1	0	1	1	1	1	1	0	11	7
Titii	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	12	10
Umar	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	11	11
Udin	1	1	1	1	1	0	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	12	7

Rumus formula Spearman Brown menurut Azwar (2013) adalah:

$$r_{xx} = \frac{2 (r_{xy})}{1+r_{xy}} \dots\dots\dots 2)$$

Dimana r_{xy} adalah korelasi *product moment*.

Dengan menggunakan rumus 1 untuk menentukan nilai korelasi product moment, maka diperoleh r_{xy} sebesar 0,08, sehingga diperoleh nilai formula Spearman Brown:

$$r_{xx} = \frac{2 (0.08)}{1+0.08}$$

$$r_{xx} = \frac{0.16}{1.08}$$

$$r_{xx} = 0,1481$$

Sama halnya dengan validitas, nilai dari reliabilitas harus disesuaikan dengan koefisien reliabilitas yang diperoleh. Klasifikasi reliabilitas menurut Istiyono (2018) sesuai Tabel 3.5.

Tabel 3. 5. Interpretasi koefisien reliabilitas

No	Koefisien reliabilitas	Tingkat reliabilitas
1	0,00 – 0,20	Sangat rendah
2	0,20 – 0,40	Rendah
3	0,40 – 0,60	Cukup
4	0,60 – 0,80	Tinggi
5	0,80 – 1,00	Sangat tinggi

Berdasarkan nilai koefisien reliabilitas melalui formula Spearman Brown sebesar 0,1481 dengan melihat Tabel 3.5, maka soal yang dikembangkan mempunyai tingkat reliabilitas yang sangat rendah.

Nilai koefisien reliabilitas melalui formula Spearman Brown juga dapat ditentukan dengan menggunakan SPSS dan R Program. Khusus untuk SPSS, cara menentukan nilai reliabilitas adalah sama dengan cara penentuan *product moment* terlebih dahulu kemudian dimasukkan dalam rumus 2, artinya tidak ada cara langsung melalui SPSS.

Pada R program, nilai koefisien reliabilitas melalui formula Spearman Brown dapat langsung ditentukan dengan menggunakan sintaks:

```
#Spearman Brown

setengah_tes_X <- c(7, 11, 9, 9, 12, 11, 10,
11, 11, 12, 10, 11, 9, 11, 8,
7, 10, 10, 9, 11, 10, 11, 12,
7, 12, 8, 11, 12, 11, 12)
setengah_tes_Y <- c(11, 9, 7, 8, 9, 11, 7,
10, 9, 9, 8, 6, 7, 10, 9,
5, 11, 10, 8, 10, 10, 8, 9, 7,
6, 12, 7, 10, 11, 7)

# Hitung koefisien korelasi antara dua setengah tes
r_AB <- cor(setengah_tes_X, setengah_tes_Y)

# Gunakan rumus Spearman-Brown
reliabilitas_prediksi <- (2 * r_AB) / (1 + r_AB)

# Tampilkan hasil
print(reliabilitas_prediksi)
```

Setelah memasukkan sintaks dan angka butir ganjil (X) dan angka butir genap (Y), selanjutnya *run* pada R Program, sehingga akan diperoleh hasil sesuai Gambar 3.3.

```

2032 #Spearman Brown
2033 # Contoh data setengah tes X dan Y
2034 setengah_tes_X <- c(7, 11, 9, 9, 12, 11, 10, 11, 11, 12, 10, 11, 9, 11, 8,
2035 7, 10, 10, 9, 11, 10, 11, 12, 7, 12, 8, 11, 12, 11, 12)
2036 setengah_tes_Y <- c(11, 9, 7, 8, 9, 11, 7, 10, 9, 9, 8, 6, 7, 10, 9,
2037 5, 11, 10, 8, 10, 10, 8, 9, 7, 6, 12, 7, 10, 11, 7)
2038
2039 # Hitung koefisien korelasi antara dua setengah tes
2040 r_AB <- cor(setengah_tes_X, setengah_tes_Y)
2041
2042 # Gunakan rumus Spearman-Brown
2043 reliabilitas_prediksi <- (2 * r_AB) / (1 + r_AB)
2044
2045 # Tampilkan hasil
2046 print(reliabilitas_prediksi)
2047

```

Console Output:

```

R 4.3.1 ~:/~
> 7, 12, 8, 11, 12, 11, 12)
> setengah_tes_Y <- c(11, 9, 7, 8, 9, 11, 7, 10,
9, 9, 8, 6, 7, 10,
+ 5, 11, 10, 8, 10, 10, 8, 9,
7, 6, 12, 7, 10, 11, 7)
> # Hitung koefisien korelasi antara dua setengah tes
> r_AB <- cor(setengah_tes_X, setengah_tes_Y)
> # Gunakan rumus Spearman-Brown
> reliabilitas_prediksi <- (2 * r_AB) / (1 + r_AB)
> # Tampilkan hasil
> print(reliabilitas_prediksi)
[1] 0.1490385

```

Environment Values:

r_AB	0.0805195165201946
reliabilitas_pr...	0.149038523208738
setengah_tes_X	num [1:30] 7 11 9 9 12 11 10 11 11 12 ...
setengah_tes_Y	num [1:30] 11 9 7 8 9 11 7 10 9 9 ...

Gambar 3. 3. Hasil formula Spearman Brown dengan R Program

Hasil yang diperoleh menggunakan R sesuai Gambar 3.3 diperoleh nilai r_{xx} sebesar 0,1490385. Nilai yang diperoleh ini hampir sama dengan nilai yang diperoleh dengan perhitungan manual yaitu 0,148. Adapun perbedaannya disebabkan karena pembulatan angka dibelakang koma. Hasil yang sama ini menunjukkan bahwa soal yang dikembangkan ini memiliki tingkat reliabilitas yang sangat rendah. Apabila mendapatkan hasil seperti ini, maka soal yang dikembangkan sebaiknya ditinjau ulang sebelum digunakan ke skala yang lebih luas.

3.4 Validitas dan Reliabilitas Soal Secara Modern

Secara garis besar terdapat dua macam analisis tes yang dapat digunakan dalam proses pengukuran, yaitu *Classical Test Theory*/CTT) dan teori respon butir (*Item Response Theory*/IRT atau lebih dikenal dengan teori modern. Teori klasik mempunyai beberapa keterbatasan dan kelemahan sehingga memunculkan

teori respon butir. Model probabilistik merupakan model yang digunakan pada teori respon butir (Mardapi, 2018a). Beberapa kelemahan teori klasik menurut Mansyur, Rasyid and Suratno (2015), statistik butir tes yang digunakan sangat bergantung pada karakteristik peserta tes, butir soal yang diujikan mempengaruhi estimasi kemampuan peserta, estimasi skor kesalahan berlaku untuk semua peserta tes, tidak ada informasi respons untuk setiap butir soal peserta ujian serta asumsi paralel susah diterapkan untuk memenuhi estimasi keterandalan alat tes.

Hambleton (2006), menyatakan bahwa pada prinsipnya teori respon butir menyediakan cara untuk membangun tes yang konten dan statistik optimal untuk peserta ujian, dan memberikan dasar langsung untuk membandingkan peserta ujian satu dengan yang lain dan dengan standar kinerja, meskipun peserta ujian mungkin telah mengambil tes yang sangat berbeda dalam kesulitan dan panjangnya.

Paek and Cole (2020) menyatakan bahwa IRT adalah model matematis yang menghubungkan sifat laten atau skor kemampuan peserta tes dengan kemungkinan merespons dalam kategori respons tertentu dari suatu item. Model IRT populer untuk tanggapan item yang dinilai secara dikotomis diberi nama sesuai dengan jumlah parameter yang memodelkan karakteristik item dan bentuk fungsinya. Jika satu karakteristik item dimodelkan (misalnya, kesulitan item), model IRT disebut model 1-parameter. Jika dua karakteristik item dimodelkan (misalnya, kesulitan item dan diskriminasi item), model IRT disebut model 2-parameter. Ketika tiga karakteristik item dimodelkan (misalnya, kesulitan item, diskriminasi item, dan tebakan semu item, yang umum untuk tes item pilihan ganda), model IRT disebut model 3-parameter.

Terdapat item selain dikotomis yang sering digunakan dalam penilaian, yaitu item yang mempunyai lebih dari dua kategori. Item yang memiliki lebih dari dua kategori disebut item sebagai

politomus. Model IRT yang sering digunakan untuk item politomus, antara lain GRM, PCM, GPCM, dan PCM 1 PL.

Pada IRT, istilah valid jarang digunakan tetapi lebih cenderung menggunakan istilah fit dan tidak fit, hal ini dapat dilihat dari hasil penelitian Suparman, Rohaeti and Wening (2022), dimana salah satu hasilnya seperti Tabel 3.6.

Tabel 3. 6. Hasil penelitian penggunaan fit dan tidak fit

Category Name	Index name	Admission Category	Analysis Results	Information
Absolute fit	Chi-Square	P-value > .05	.051	Fit
	RMSEA	RMSEA < .08	.067	Fit
	GFI	GFI > .90	.91	Fit
Incremental fit	AGFI	AGFI > .90	.86	Not fit
	CFI	CFI > .90	.96	Fit
	IFI	IFI > .90	.96	Fit
	NFI	NFI > .90	.93	Fit
Parsimonious fit	Chisq/df	Chi-Suare/ df < 3.0	2,12	Fit

Reliabilitas pada teori tes modern dapat dilakukan dengan melihat fungsi informasi, dimana fungsi informasi ada dua macam, yaitu fungsi informasi tes dan fungsi informasi butir. Fungsi informasi dapat ditentukan melalui model logistic yang terdiri atas 1PL, 2PL dan 3PL. Perbedaan reliabilitas pada teori klasik dan teori modern adalah pada tes modern tidak memerlukan tes paralel untuk menghitung koefisien reliabilitas tetapi hanya menggunakan fungsi informasi (Istiyono, 2018). Menurut Hambleton, Swaminathan and Rogers (1991) fungsi informasi item mempunyai peran penting dalam pengembangan tes dan evaluasi item karena menampilkan kontribusi item terhadap estimasi kemampuan.

DAFTAR PUSTAKA

- Arikunto, S. (2018) *Dasar-dasar evaluasi pendidikan*. Jakarta: Bumi Aksara.
- Azwar, S. (2013) *Reliabilitas dan validitas*. Keenam. Yogyakarta: Pustaka Pelajar.
- Azwar, S. (2016) *Tes prestasi*. Kedua. Yogyakarta: Pustaka Pelajar.
- Cohen, L., Manion, L. and Morrison, K. (2018) *Research methods in education*. Eighth. New York: Routledge.
- Finch, W. H. and French, B. F. (2019) *Educational and psychological measurement*. First, *Educational and Psychological Measurement*. First. New York: Routledge. doi: 10.4324/9781315650951.
- Hambleton, R. K. (2006) 'Psychometric Models, Test Designs and Item Types for the Next Generation of Educational and Psychological Tests', in Bartram, D. and Hambleton, R. K. (eds) *Computer Based Testing and the Internet*. Chichester: John Wiley & Sons Ltd, pp. 77–90.
- Hambleton, R. K., Swaminathan, H. and Rogers, H. J. (1991) *Fundamentals of item response theory library*. London: Sage Publications.
- Istiyono, E. (2018) *Pengembangan instrumen penilaian dan analisis hasil belajar fisika*. Pertama. Yogyakarta: UNY Press.
- Mansyur, Rasyid, H. and Suratno (2015) *Asesmen pembelajaran di sekolah*. Yogyakarta: Pustaka Pelajar.
- Mardapi, D. (2018a) *Pengukuran, penilaian dan evaluasi pendidikan*. Yogyakarta: Parama Publishing.
- Mardapi, D. (2018b) *Teknik penyusunan instrumen tes dan nontes*. Yogyakarta: Parama Publishing.

- Paek, I. and Cole, K. (2020) *Using R for item response theory model applications*. First. New York: Routledge.
- Schaughency, E. *et al.* (2012) 'Classical test theory and higher education: Five questions', in Secolsky, C. and Denison, D. B. (eds) *Handbook on Measurement, Assessment, and Evaluation in Higher Education*. First. New York: Routledge, pp. 117–131.
- Sukardi (2012) *Evaluasi pendidikan prinsip & operasionalnya*. Jakarta: Bumi Aksara.
- Sumintono, B. and Widhiarso, W. (2015) *Aplikasi pemodelan rasch pada assessment pendidikan*. Cimahi: Trim Komunikata.
- Suparman, A. R., Rohaeti, E. and Wening, S. (2022) 'Development of attitude assessment instruments towards socioscientific issues in chemistry learning', *European Journal of Educational Research*, 11(4), pp. 1947–1958. doi: 10.12973/eujer.11.4.1947.
- Thorndike, R. M. and Chris, T. T. (2014) *Measurement and evaluation in psychology and education*. Eighth. Harlow: Pearson Education Limited.

BAB 4

REVISI BUTIR SOAL DAN MENAFSIRKAN HASIL TES

Oleh Suwarno

4.1 Pendahuluan

Salah satu bagian penting dalam kegiatan pembelajaran adalah penilaian. Penilaian dilakukan untuk memperoleh informasi mengenai pencapaian tujuan pembelajaran yang sudah ditentukan. Penilaian merupakan metode dokumentasi dan penyusunan informasi sebagai upaya untuk mengetahui keberhasilan proses belajar mengajar.

Evaluasi pendidikan adalah proses yang melibatkan penciptaan, penerapan, dan analisis instrumen pengukuran pendidikan. Fungsi utama alat ukur pendidikan adalah memberikan informasi yang menjadi dasar pengambilan keputusan yang tepat, bila alat tersebut dikembangkan sebagai alat untuk menyimpulkan kemampuan suatu obyek.

Dalam penggunaan materi pendidikan yang berskala besar dan berisiko, seperti ujian masuk, perlu diketahui indikator teknis yang menentukan kualitas materi yang digunakan. Tes skala besar adalah tes yang dirancang untuk dilaksanakan di beberapa kelompok sekolah. Karena dimensinya dan dampak sosialnya yang besar, desainnya harus memenuhi standar kualitas yang ketat (Aiken, 1996).

Penilaian pembelajaran siswa merupakan berbagai proses untuk memperoleh informasi pembelajaran dan pengambilan keputusan mengenai keterampilan dan hasil belajar siswa.

Penilaian hasil belajar siswa merupakan kegiatan yang berkaitan dengan keterampilan guru dalam proses pembelajaran atau pencapaian hasil belajar siswa (Miller, Linn, dan Gronlund, 2012). Penilaian merupakan metode untuk menentukan nilai sesuatu berdasarkan standar dan memperoleh pengetahuan dan informasi tentang kemampuan siswa dan tingkat interpretasinya, termasuk informasi dan data kognitif, psikomotorik, atau emosional. Perlunya memperbaiki apa yang telah diperbaiki dalam proses pembelajaran.

Evaluasi adalah proses penentuan kualitas, nilai, dan kepentingan berdasarkan aspek dan kriteria tertentu dalam konteks pengambilan keputusan. Alat yang digunakan mempunyai dampak yang sangat besar terhadap evaluasi hasil dan kualitas pembelajaran serta penilaian. Instrumen adalah suatu perangkat untuk mengukur kegiatan penilaian pembelajaran. Salah satunya adalah tes (Zainal Arifin, 2011)

Tes yang digunakan di sekolah biasanya untuk menilai kemampuan dan prestasi siswa dalam beberapa bidang kognitif seperti pengetahuan, keterampilan, analisis, sintesis, dan evaluasi. Tes adalah alat atau instrumen yang digunakan untuk mengumpulkan informasi tentang suatu ranah kognitif (Purwanto, 2013). Hasil respon siswa terhadap suatu pertanyaan yang mengharuskan siswa tersebut menjawab atau menyelesaikan suatu tugas.

Tes yang dilakukan sebaiknya juga dilakukan dalam jumlah yang lebih banyak. Hasil perangkat tes yang telah teruji validitas dan reliabilitasnya serta dapat digunakan sebagai alat ukur yang berkualitas dan efektif sesuai dengan kemampuan siswa. Tes yang efektif adalah tes yang secara akurat mengukur situasi yang hendak diukur. Sebaliknya jika suatu tes mengukur suatu kondisi yang tidak dapat diukur secara akurat oleh tes tersebut, maka tes tersebut dianggap tidak valid. Tes yang baik bertanggung jawab dalam hal kelayakan, validitas, reliabilitas, interpretabilitas,

kegunaan, dan validitas unsur soal (seperti tingkat kesukaran dan daya pembedaan soal yang sesuai) (Febriani dan Saksono, 2016).

Reliabilitas berhubungan dengan keakuratan suatu perangkat atau ketepatan pengukuran yang diperoleh dari suatu tes (Thorndike dan Hagen, 1977). Oleh karena itu, alat ukur hasil belajar yang baik harus memenuhi syarat validitas dan reliabilitas. Hasil pengukuran yang akurat diperoleh melalui percobaan yang memenuhi persyaratan peralatan pengukuran yang baik.

Soal yang baik adalah soal yang tidak terlalu sulit, tidak terlalu mudah, dan dapat dibuat sulit. Sebab, tingkat kesulitan soal berkorelasi dengan diskriminasi. Soal yang tingkat kesulitannya paling tinggi akan mempunyai tingkat diskriminasi yang paling rendah, dan soal yang paling mudah tidak akan membedakan. Oleh karena itu, yang terbaik adalah menjaga tingkat kesulitan pertanyaan dalam rentang yang dapat dibedakan.

Analisis pertanyaan dapat digunakan baik secara kualitatif maupun kuantitatif, keduanya memiliki kelemahan dan keunggulan. Beberapa teknik yang digunakan untuk analisis kualitatif: teknik moderasi dan teknik panel. Dalam metode ini, setiap pertanyaan didiskusikan dengan banyak pakar, termasuk guru materi pelajaran, pakar perangkat keras, penerjemah, dan pengembang perangkat lunak.

Teknik analisis kuantitatif adalah penggunaan data empiris untuk menyelidiki unsur-unsur suatu pertanyaan. Dimana informasi empiris ini diperoleh dari soal-soal yang diujikan. Ada dua analisis kuantitatif yaitu analisis klasik dan analisis modern. Analisis klasik mempertimbangkan soal-soal ujian berdasarkan respon siswa. Aspek yang sangat penting dan dipertimbangkan secara klasik ini dapat dikenali dari validitas, reliabilitas, kesukaran, daya pembeda butir soal, dan fungsi pengalih perhatian pada setiap pilihan jawaban.

Dalam penelitian kuantitatif terhadap soal, selain validitas dan reliabilitas, juga dilakukan penelitian terhadap tiga item pertanyaan yaitu analisis kesukaran, daya pembeda, dan faktor pengganggu/pengganggu berdasarkan data empiris pada setiap item pertanyaan.

4.2 Validitas

Validitas merupakan keakuratan penafsiran yang didapat dari kegiatan penilaian (Gronlund, 2009). Lebih spesifiknya, validitas yang mengacu pada sampai seberapa jauh suatu perangkat alat ukur menjalankan fungsi pengukurannya secara akurat dan tepat (Azwar, 2010). Namun secara spesifik, validitas suatu instrumen tes dapat diartikan sebagai kemampuan tes untuk mengukur apa yang ingin diukur (Allen & Yen, 1979).

Jenis validitas ada tiga yaitu: isi, konstruk, dan kriteria (Allen & Yen). Selanjutnya menurut (Azwar, 2011), ada dua jenis validitas isi: validitas muka dan validitas logis. Validitas isi adalah sejauh mana suatu perangkat dapat memberikan gambaran terhadap kemampuan yang dinilai secara menyeluruh. (Azwar, 2011). Hal ini berarti mengambil bentuk analisis rasional terhadap area yang diukur.

Validitas konstruk berkaitan dengan sejauh mana pengujian atau pengukuran Anda secara akurat menilai apa yang seharusnya dilakukan. Dalam penelitian, penting untuk mengoperasionalkan konstruksi menjadi karakteristik yang konkret dan terukur berdasarkan gagasan Anda tentang konstruksi dan dimensinya.

Perjelas cara anda mendefinisikan konstruksi dan bagaimana dimensi berhubungan satu sama lain sebelum anda mengumpulkan atau menganalisis data. Hal ini membantu Anda memastikan bahwa metode pengukuran apa pun yang Anda gunakan secara akurat menilai konstruksi spesifik yang Anda selidiki secara keseluruhan dan membantu menghindari bias dan

kesalahan seperti bias variabel yang dihilangkan atau bias informasi.

Contoh: Buatlah ukuran Anda mengembangkan kuesioner sederhana untuk menilai kecemasan sosial pada siswa. Anda membuat pertanyaan untuk mengukur konstruksi kecemasan sosial Anda: 1) Seberapa sering Anda menghindari memasuki ruangan ketika semua orang sudah duduk? 2) Apakah orang lain cenderung menggambarkan Anda sebagai orang yang pendiam? 3) Saat berbicara dengan kenalan baru, seberapa sering Anda khawatir akan mengatakan sesuatu yang bodoh?. 4) Sejauh mana Anda takut memberikan ceramah di depan audiens?. 5) Seberapa sering Anda menghindari kontak mata dengan orang lain?. 6) Apakah Anda lebih suka mempunyai teman dekat dalam jumlah sedikit dibandingkan dengan teman dalam jumlah besar?

Saat merancang atau mengevaluasi suatu pengukuran, penting untuk mempertimbangkan apakah pengukuran tersebut benar-benar menargetkan konstruksi yang diinginkan atau apakah pengukuran tersebut menilai konstruksi yang terpisah namun terkait.

Sangat penting untuk membedakan konstruksi Anda dari konstruksi terkait dan memastikan bahwa setiap bagian dari teknik pengukuran Anda hanya terfokus pada konstruksi spesifik Anda.

Contoh: Mengevaluasi ukuran Anda memeriksa kuesioner Anda dengan memikirkan beberapa pertanyaan: 1) Apakah kuesioner Anda hanya mengukur kecemasan sosial? 2) Apakah semua aspek kecemasan sosial tercakup dalam pertanyaan? 3) Apakah pertanyaan Anda menghindari pengukuran konstruksi relevan lainnya seperti rasa malu atau introversi?

Beberapa pertanyaan Anda menargetkan rasa malu dan introversi serta kecemasan sosial. Ini berarti kuesioner Anda terlalu luas dan perlu dipersempit agar hanya fokus pada kecemasan sosial.

Validitas kriteria (atau validitas terkait kriteria) mengukur seberapa baik suatu ukuran memprediksi hasil dari ukuran lainnya. Suatu tes memiliki validitas jenis ini jika berguna untuk memprediksi kinerja atau perilaku dalam situasi lain (masa lalu, sekarang, atau masa depan). Misalnya: Pelamar kerja mengikuti tes kinerja selama proses wawancara. Jika tes ini secara akurat memprediksi seberapa baik kinerja karyawan dalam pekerjaannya, tes tersebut dikatakan memiliki validitas kriteria.

Ukuran pertama terkadang disebut variabel prediktor atau estimator. Ukuran kedua disebut variabel kriteria sepanjang ukuran tersebut diketahui merupakan alat yang valid untuk memprediksi hasil. Validitas kriteria, disebut juga validitas prediktif, yaitu kemampuannya memprediksi keberhasilan siswa di masa depan. (Allen & Yen 1979; Azwar, 2011).

4.3 Reliabilitas

Indeks reliabilitas merupakan ukuran konsistensi hasil pengujian dari satu pengukuran ke pengukuran lainnya. Reliabilitas merujuk pada konsistensi atau stabilitas hasil penilaian (Reynold, 2006). Singkatnya, reliabilitas adalah stabilitas (Cohen, 2007). Jika suatu pengujian sangat konsisten, maka pengujian tersebut akurat, dapat direproduksi, dan dapat digeneralisasikan pada pengujian dan peralatan yang sama (Ebel & Frisbie, 1991). Reliabilitas adalah derajat konsistensi antara dua pengujian terhadap pada sasaran yang sama. (Mehrens & Lehmann, 1973). Arti ini dapat digambarkan pada seseorang yang mengukur tinggi badan dan hasilnya tidak berubah meskipun menggunakan alat ukur atau skala yang berbeda. Dalam dunia pendidikan, suatu hal dikatakan dapat dipercaya apabila kinerja atau kemampuan siswa dinilai. Reliabilitas ini berarti bahwa hasil pengukuran mengandung informasi yang sama terlepas dari apakah pengujinya berbeda, korektornya berbeda, atau sifat soalnya berbeda.

Suatu tes dianggap reliabel jika nilai pengamatan mempunyai hubungan yang mendekati dengan nilai sebenarnya (Allen & Yen, 1979). Lebih lanjut bahwa reliabilitas adalah koefisien korelasi antara dua nilai pengamatan yang diperoleh dari pengukuran dari pengujian paralel. Dari pernyataan tersebut kita dapat menyimpulkan bahwa perangkat tes dikatakan reliabel jika hasil pengukurannya mendekati keadaan peserta tes yang sebenarnya.

Di bidang pendidikan karakter atau sifat yang diukur tidak dapat diukur secara langsung. Karakter atau sifat ini masih abstrak. Hal ini menyulitkan perolehan instrumen yang stabil untuk mengukur karakteristik manusia (Mehrens & Lehmann, 1973).

Reliabilitas juga mengacu pada konsistensi hasil yang diperoleh ketika peserta tes yang sama mengerjakan tes yang sama pada waktu yang berbeda. Menurut (Saifudin Azwar, 2015), suatu alat tes mempunyai indeks reliabilitas yang baik jika nilai minimum indeks reliabilitasnya adalah 0,7.

Mengacu pada uraian di muka, maka membuat alat ukur di bidang pendidikan hendaknya diusahakan dengan hati-hati dan mematuhi ketentuan yang ditetapkan oleh ahli pengukuran di bidang pendidikan. Reliabilitas suatu alat ukur dapat dihitung secara statistik dalam bentuk indeks reliabilitas. Nilai ini biasa disebut dengan tingkat reliabilitas. Skor reliabilitas suatu tes dapat diketahui dengan rumus sebagai berikut:

$$R_{xy} = \frac{S_{xy}}{S_x S_y} \quad (1)$$

Dimana

S_{xy} = Kovarian antar tes

S_x dan S_y = Deviasi standar tes

Untuk mengetahui tingkat reliabilitas suatu instrumen tes dapat digunakan teknik korelasi dengan skor korelasi berkisar antara -1 sampai +1. Tanda minus dan plus menunjukkan arah hubungan tertentu. Korelasi angka 1 atau korelasi sempurna jarang sekali dicapai dalam dunia pendidikan. Secara umum nilai korelasi selalu kurang dari 1. Jika angka korelasinya lebih besar dari 0,60 dan kurang dari 1, maka instrumen tersebut berkorelasi tinggi atau reliabel. Sebaliknya jika angka korelasinya kurang dari atau sama dengan 0,50 maka instrumen tersebut mempunyai korelasi yang rendah atau reliabilitas yang rendah (Basuki & Haryanto, 2014).

4.3 Tingkat Kesukaran

Kesulitan dalam teori tes klasik adalah persentase jawaban benar (Djemari Mardapi, 2012). Secara umum, indeks kesulitan suatu tugas harus berada dalam kisaran 0,3 hingga 0,7 (Allen dan yen, 1979). Tingkat kesulitan menunjukkan berapa persentase peserta tes yang menjawab benar. Kesukaran merupakan persentase orang yang menjawab benar, dan besar kecilnya indeks kesukaran berkisar antara 0,00 hingga 1,00 (Mardapi, 2017). Semakin rendah indeks kesukarannya, maka semakin sulit soal tersebut. Sebaliknya, semakin tinggi indeks kesukarannya, maka semakin mudah pula soalnya. Pertanyaan yang baik adalah pertanyaan yang tidak terlalu sederhana dan tidak terlalu sulit (Amalia & Widayati, 2012).

Formula untuk menghitung indeks tingkat kesukaran butir soal pilihan ganda adalah sebagai berikut.

$$p = \frac{\text{Jumlah siswa yang mnjawab benar}}{\text{jumlah siswa yang mengikuti tes}} \quad (2)$$

Formula untuk menghitung indeks tingkat kesukaran butir soal uraian adalah sebagai berikut.

$$p = \frac{\text{Rata rata skor} - \text{skor minimem}}{\text{Skor maksimem} - \text{skor minimem}} \quad (3)$$

Tingkat kesukaran soal dapat dikategorikan “sulit”, “sedang”, dan “mudah” (Sumintono, 2015). Kesukaran digolongkan “mudah” jika indeks kesukaran mendekati -2,00, “sedang” jika indeks kesukaran antara -1,00 dan +1,00, dan “sulit” jika indeks kesukaran mendekati +2,00. Tingkat kesulitan soal yang baik adalah jika tingkat kesulitannya antara 0,30 sampai 0,70. Apabila indeks kesukaran butir soal antara 0,10 hingga 0,29, atau antara 0,71 hingga 0,90, maka butir angket tersebut harus direvisi (Mardapi, 2017; Azwar, 2016 & Kartowagiran, 2012). Selain itu, jika indeks kesulitan suatu item kurang dari 0,10 atau lebih besar dari 0,90, item tersebut akan ditolak atau dibatalkan.

Kesulitan soal tes memiliki dua kegunaan: untuk guru dan untuk tes dan pengajaran (Nitko, 2011: 303-305). Kegunaan bagi guru antara lain: a) Memberikan informasi kepada siswa sebagai pengenalan konsep pembelajaran ulang dan tentang hasil belajar. b) Periksa fokus silabus dan periksa apakah pertanyaannya bias. Kegunaannya dalam pengujian dan pendidikan meliputi: a) Memperkenalkan konsep-konsep yang memerlukan pendidikan ulang. b) Mengidentifikasi kelebihan dan kelemahan kurikulum. c) Memberikan masukan kepada siswa. d) Pertimbangkan kemungkinan pertanyaan yang bias.

Indeks kesukaran dapat diklasifikasikan sebagai berikut:

Tabel 4. 1. Kategori Tingkat Kesukaran

Besarnya koefisien	Kelompok (kategori)
$P < 0,3$	Sukar
$0,3 \leq p \leq 0,7$	Sedang
$p > 0,7$	Mudah

Sumber: Surapranata, 2009.

Tingkat Kesukaran soal adalah persentase atau perbandingan jumlah siswa yang bisa menjawab pertanyaan dengan betul. Tingkat kesulitan berkisar dari 0,00 hingga 1,00. Semakin tinggi tingkat kesulitan yang ditentukan oleh hasil perhitungan, maka semakin mudah soal tersebut dan semakin banyak pula kebutuhan untuk memperbaiki soal tersebut. Jika $p = 0,00$ pada suatu soal berarti tidak ada siswa yang menjawab benar, dan jika $p = 1,00$ berarti semua siswa menjawab benar. Jika tingkat kesukaran optimal 0,50 maka terdapat keberagaman antar peserta tes. Artinya, sejumlah lima puluh persen siswa menjawab dengan betul dan sejumlah lima puluh persen siswa menjawab dengan salah. Namun tidak semua soal diharuskan mempunyai tingkat kesulitan 0,50. Mungkinkah rata-rata seluruh soal tetap 0,50 meskipun nilai soalnya kurang dari atau lebih besar dari 0,50?

4.4 Daya pembeda

Indeks Kekuatan Diferensial menunjukkan kemampuan suatu soal tes dalam membedakan peserta tes berkemampuan tinggi dan rendah. Keunikan soal adalah kemampuan suatu soal untuk membedakan siswa pintar dengan siswa tidak pintar (Solichin (2017)). Ciri-ciri pertanyaan adalah kemampuan unsur-unsur pertanyaan untuk membedakan antara siswa yang sudah menguasai isi yang ditanyakan dan siswa yang belum/belum menguasai/belum menguasai isi yang ditanyakan.

Rumus untuk menghitung daya pembeda butir soal pilihan ganda adalah sebagai berikut

$$DP = \frac{JB_A - JB_b}{n} \quad (4)$$

Dimana:

DP = daya pembeda

JB_A = banyak siswa kelompok atas yang memberikan jawaban dengan betul

J_{BB} = banyak siswa kelompok bawah yang memberikan jawaban dengan betul

Rumus untuk menghitung daya pembeda butir soal butir soal uraian adalah sebagai berikut

$$DP = \frac{\bar{X} \text{ kelompok atas} - \bar{X} \text{ kelompok bawah}}{\text{skor maksimum soal}} \quad (5)$$

Dimana:

DP = Daya Pembeda

Daya Pembeda bisa juga menggunakan rumus

$$d = (NiT / NT) - (NiR / NR) \quad (6)$$

dimana:

d = daya pembeda

NiT = banyak siswa kelompok atas yang memberikan jawaban dengan betul

NT = banyak siswa kelompok atas

NiR = banyak siswa kelompok bawah yang memberikan jawaban dengan betul

NR = banyak siswa kelompok bawah

(Azwar, 2016)

Keunggulan daya pembeda item pertanyaan adalah sebagai berikut. a) Meningkatkan kualitas setiap item pertanyaan dengan menggunakan data empiris. Anda dapat menggunakan indeks diskriminasi untuk menentukan apakah setiap item pertanyaan bagus, dimodifikasi, atau ditolak. b) Menguji sejauh mana setiap item pertanyaan dapat mengidentifikasi/membedakan kemampuan siswa, terutama yang memahami atau tidak memahami materi yang diajarkan guru.

Jika angket tidak dapat membedakan kemampuan dua siswa, maka angket akan terlihat seperti ini: a) Kunci jawaban soal salah. b) Butir soal mempunyai dua atau lebih kunci jawaban yang benar. c) Kemampuan yang diukur tidak jelas. d) Alat pengalih perhatian tidak berfungsi. e) Banyak siswa yang menduga materi yang dibutuhkan terlalu sulit. f) Sebagian besar siswa yang memahami apa yang ditanyakan berpendapat bahwa pertanyaan mengandung informasi yang salah.

Indeks kekuatan masing-masing elemen juga biasanya dinyatakan secara proporsional. Semakin tinggi indeks keunikan suatu soal, maka pertanyaan tersebut semakin baik dalam membedakan antara siswa yang memahami isi dan yang tidak. Indeks selektivitas berkisar dari $-1,00$ hingga $+1,00$. Semakin membedakan suatu pertanyaan, semakin kuat/baik pertanyaan tersebut. Apabila daya pembeda negatif, berarti jumlah siswa yang berada di golongan bawah lebih banyak menjawab benar dibandingkan dengan jumlah siswa yang berada di golongan atas.

Proses penghitungan gaya diferensial dilakukan ketika proses penghitungan gaya diferensial kelompok kecil dan kelompok besar dibedakan. Pertama, tentukan sampel mana yang diambil kelompok atas dan sampel mana yang diambil kelompok bawah. Namun sebelumnya data tersebut dipilah berdasarkan siswa yang berprestasi baik dan siswa yang berprestasi kurang. Kemudian, pada kelompok kecil, kelompok cerdas dan kelompok kecerdasan rendah diambil sampelnya dari 50% populasi, dan pada kelompok besar, kelompok cerdas dan kelompok kecerdasan rendah masing-masing diambil sampelnya dari 27% populasi, dan kelompok cerdas dan kelompok cerdas diambil sampelnya masing-masing sebanyak 27% dari populasi. kelompok perantara Itu akan menjadi 54% populasi mengambil tindakan.

Tabel 4. 2. Kriteria Pemilihan Soal Berdasarkan Daya Pembeda

Kriteria	Koefisien	Keputusan
Daya Pembeda	$> 0,30$	Diterima
	$0,10 \text{ s.d } 0,29$	Direvisi
	$< 0,10$	Ditolak

Sumber: Surapranata, 2009

4.5 Analisis Distraktor/pengecoh

Tes pilihan ganda disusun menurut elemen masalah yang disajikan (atau pertanyaan utama) dan pilihan (atau kumpulan alternatif jawaban). Pilihannya sendiri terbagi menjadi dua bagian: Jawaban yang benar dan jawaban pengecoh. Artinya di antara jawaban-jawaban yang diberikan, disitu Cuma ada satu jawaban yang benar disebut dengan kunci jawaban, dan kemungkinan alternatif jawaban yang lain (alternatif yang salah) disebut pengecoh.

Orang yang mengajukan pertanyaan pilihan ganda mungkin tidak mudah teralihkan perhatiannya. Rendahnya jumlah gaya pembeda seringkali disebabkan oleh berkurangnya fungsi pengecoh (Surapranata, 2009).

Tujuan dari analisis pengecoh adalah untuk mengetahui jumlah siswa kelompok atas yang menjawab soal yang bersangkutan dengan benar dan berapa jumlah siswa kelompok bawah yang menjawab soal yang bersangkutan dengan benar. Maka, ada dua pertanyaan yang muncul dalam kasus ini. Misalnya peserta tes tidak memilih pengecoh. Artinya distraktor tidak berfungsi. Jika suatu pengecoh dengan jelas menunjukkan bahwa tes tersebut salah, tetapi siswa tidak ada yang memilih alternatif jawaban tersebut maka alternatif tersebut harus diubah.

Permasalahan selanjutnya: Apakah pengalih perhatian yang dirancang dengan menarik lebih mungkin dipilih oleh siswa dari kelompok yang lebih rendah dibandingkan oleh siswa dari

kelompok yang lebih tinggi? Jika demikian, itu adalah pengalih perhatian yang bagus. Pengalih berfungsi untuk mengidentifikasi peserta tes yang berkinerja tinggi. Distraktor dikatakan bekerja lebih efektif ketika mereka sering dipilih oleh peserta tes dari kelompok yang lebih rendah. Sebaliknya distraktor/distractor tidak akan berfungsi dengan baik jika sering dipilih oleh kelompok yang lebih unggul.

4.6 Contoh Analisis Soal

Di tampilan variabel pada Editor SPSS, tentukan variabel untuk setiap item pertanyaan dari item 1 hingga item 10, dan terakhir skornya. Pada kolom Nama, masukkan dan evaluasi Elemen 1, Elemen 2 hingga elemen terakhir. Ubah angka pada kolom desimal menjadi nol dan abaikan kolom lainnya.

Setelah variabel terisi, klik data view dan masukkan data yang telah disiapkan diawal tadi seperti contoh berikut ini.

4.6.1. Uji Validitas

Analyze --> Correlate --> Bivariate

Kotak "Variables" diisi

*Soal1 Soal2 Soal3 Soal4 Soal5 Soal6 Soal7 Soal8 Soal9 Soal10 Soal11
Soal12 Soal13 Soal 14 Soal 15 Jumlah*

Klik Pearson, Klik Two-tailed, Klik Flag significant correlation

Klik OK

Analisis: Hasil pada kolom "Jumlah" jika nilai sig < 0,05 menyatakan butir soal tersebut Valid. Contoh untuk soal nomor 1, nilai sig = 0,048, dimana nilai tersebut < 0,05, maka soal nomor 1 dinyatakan valid.

4.6.2. Uji Reliabilitas

Untuk menghitung reliabilitas tes

Analyze→Scale → Reliability analysis

Pada kotak “item” diisi

Soal1, soal2, soal3, ds.t. sampai dengan soal15

Pada kolom model, Klik alpha, Klik “list item label”,

dan Klik kstatistics, klik item, klik Continue

Kemudian Klik OK

Analisis: Berdasarkan nilai Cronbach's Alpha pada tabel Reliability Statistics diperoleh nilai 0,492 dimana nilai ini ditafsirkan dengan kriteria

Mengacu pada Cronbach's Alpha, reliabilitas didapat nilai 0,492 maka nilai tersebut berarti

Tingkat reliabilitas	Kategori
0,800 s.d. 1,000	sangat tinggi
0,600 s.d. 0,799	Tinggi
0,400 s.d. 0,599	cukup
0,200 s.d. 0,399	rendah
0,000 s.d. 0,199	sangat rendah

maka bisa diambil kesimpulan bahwa perangkat soal ini mempunyai tingkat reliabilitas yang cukup

4.6.3. Tingkat Kesukaran

Analyze→Descriptive Statistics → Frequencies

Pada kotak “variable” diisi

Soal1, soal2, soal3, ds.t. sampai dengan soal15

Klik statistics, Klik juga Mean, dan Klik Continue

Kemudian Klik OK

Analisis : dari hasil yang ditunjukkan nilai MEAN pada tabel statistics ditafsirkan pada rentang tingkat kesukaran, yaitu

0,00 - 0,20 = Sukar

0,21 - 0,70 = Sedang

0,71 - 1,00 = Mudah

Maka untuk soal nomor satu, diperoleh nilai 0,40 yang berarti tingkat kesukaran soal nomor satu adalah SEDANG.

4.6.4. Daya Pembeda

Untuk menentukan daya pembeda, maka nilai perhitungan yang digunakan adalah r_{hitung} pada SPSS yang dibandingkan dengan kriteria :

- 0.40 – 1.00 = Soal Baik
- 0.30 – 0.39 = Soal diterima dan diperbaiki
- 0.20 – 0.29 = Soal Diperbaiki
- 0.00 – 0.19 = Soal ditolak

R hitung dapat dilihat dari nilai pearson correlation pada uji validitas.

Maka hasil yang diperoleh untuk soal nomor satu pada kolom jumlah adalah 0,636 yang berarti SOAL BAIK.

4.7 Penutup

Teknik analisis data dilakukan dengan menganalisis tingkat kesukaran, kekhasan, dan fungsi pengecoh soal tes. Metode analisis data menggunakan pendekatan kuantitatif berdasarkan teori tes klasik. Program SPSS digunakan untuk menghitung indeks setiap parameter karakteristik masing masing soal dan hasilnya dianalisis untuk mengetahui kualitas masing masing butir soal. Kriteria penentuan ciri-ciri pertanyaan yang baik dalam analisis data ini adalah sebagai berikut:

Tabel 4. 3. Kriteria Keputusan

Karakteristik Butir Soal	Indeks/Koefisien	Keputusan
Tingkat kesukaran soal	0.30 - 0.70	Diterima
	0.10 - 0.29	Direvisi
	0.71 - 0.90	Direvisi
	Kurang dari 0.10	Ditolak
	Lebih dari 0.90	Ditolak

Karakteristik Butir Soal	Indeks/Koefisien	Keputusan
Daya Pembeda soal	≥ 0.30 $0.10 - 0.29$ ≤ 0.10	Ditrima Direvisi Ditolak
Keefektifan pengecoh	≥ 0.05 ≤ 0.05	Diterima direvisi

Sumber: Mardapi (2017), Azwar (2016), Kartowagiran (2012)

DAFTAR PUSTAKA

- Allen, M. J & Yen, W. M. (1979). *Introduction to measurement theory*. Monterey: Brooks: Cole Publishing Company.
- Anastasi, Anne, dan Urbina, Susana. (1997). *Psychological testing*. Seventh edition. Upper saddle River, Nj: Prentice Hall Inc.
- Arifin Zainal. 2012. *Evaluasi Pembelajaran*. Jakarta: Direktorat Jenderal Pendidikan Islam Kementrian Agama RI.
- Arifin, Zainal. (2012). *Evaluasi Pembelajaran*. Bandung: Remaja Rosdakarya.
- Arikunto Suharsimi. 2001. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: PT Bumi Aksara.
- Asrul, dkk. 2014. *Evaluasi Pembelajaran*. Medan: Cipta pustaka Media.
- Azwar, S. (2015). *Reliabilitas dan validitas*. Yogyakarta: Pustaka Pelajar.
- Azwar, S. (2016). *Tes prestasi: Fungsi pengembangan pengukuran prestasi belajar*. Yogyakarta: Pustaka Pelajar.
- Danial, D., Nurjannah, N., & Mirna, M. 2019. *Evaluation of The Learning*.
- DeMars, C. (2010). *Item response theory understanding statistics measurement*. New York: Oxford University Press, Inc.,
- Djemari Mardapi. 2008. *Teknik Penyusunan Instrumen Tes Dan Non Tes*. Yogyakarta : Mitra Cendekia, 88.
- Elis Ratnawulan. 2014. *Evaluasi Pembelajaran*. Bandung: Penerbit pustaka.

- Febriani, Irene Melinda, and Lutfi Saksono. 2016. *Analisis Butir Soal Ujian Akhir Semester (UAS) Bahasa Jerman Kelas X MIA 6 SMA Negeri 1 Maospati Tahun Pelajaran 2015/2016*. Universitas Negeri Surabaya: Jurnal Laterne, Volume V Nomor 02.
- Fitriani. 2019. Kemampuan Pemecahan Masalah Siswa SMP. *JTMT: Journal Tadris Matematika*, 01(01), 25–30. <http://journal.iaimsinjai.ac.id/index.php/Jtm/article/view/393>.
- Hamalik, Oemar. 2010. *Manajemen Pengembangan Kurikulum*. Bandung : PT Remaja Rosdakarya.
- Hanifah, Nurdinah. (2019). *Pengembangan instrumen penilaian Higher Order Thinking Skill (HOTS) di sekolah dasar*. Conference Series Journal, 1(1), 2.
- Ida Farida. 2017. *Evaluasi Pembelajaran*. Bandung:PT Remaja Rosdakarya.
- Imania, Kuntum An Nisa, dkk. 2019. *Rancangan Pengembangan Instrumen Penilaian Pembelajaran Berbasis Daring*. Jurnal PETIK, 5(1), 2.
- Kunanddar. 2014. *Penilaian Autentik Penilaian Hasil Belajar Peserta Didik Berdasarkan Kurikulum 2013*. Jakarta : Raja Grafindo Persada.
- Kusairi dan Suprananto. (2012). *Pengukuran dan Penilaian Pendidikan*. Yogyakarta: Graha Ilmu. Mardapi, D. (2008). *Teknik Penyusunan Instrumen Tes dan Non Tes*. Yogyakarta: Mitra Cendekia. Purwanto. (2013). *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Belajar.
- Kussaeri, K. (2014). *Acuan dan Teknik proses dan Hasil Belajar dalam Kurikulum 2013*. Yogyakarta: Ar-Ruqq Media.
- Mardapi, Djemari. (2012). *Pengukuran, penilaian & evaluasi pendidikan*. Yogyakarta: Nuha Medika.

- Muhammad Afandi.2013. *Evaluasi Pembelajaran Sekolah Dasar*. Semarang:UNISSULA Press
- Nurfillaili, Ulfa, dkk. (2016). *Pengembangan Instrumen Tes Hasil Belajar Kognitif Mata Pelajaran Fisika Pada Pokok Bahasan Usaha Dan Energi Sma Negeri Khusus Jeneponto Kelas Xi Semester I*. Jurnal Pendidikan Fisika, 4(2), 83-84
- Program of Mathematics Study Program at Islamic Institute Of Muhammadiyah Sinjai. Matematika Dan Pembelajaran*, 7(1), 65. <https://doi.org/10.33477/mp.v7i1.1046>
- Purnomosidi. 2017. *Kebersamaan (Buku Tematik Terpadu Kurikulum 2013)*. Jakarta : Kementerian Pendidikan dan Kebudayaan.
- Solichin, M. (2017). Analisis Daya Beda Soal, Taraf Kesukaran, Validitas Butir Tes, Interpretasi Hasil Tes, dan Validitas Ramalan dalam Evaluasi Pendidikan. *Dirāsāt: Jurnal Manajemen & Pendidikan Islam*, Volume 2, Nomor 2, 192-213.
- Sugiyono. *Metode Penelitian Kuntitatif Kualitatif dan R&D*. Bandung: Alfabeta, 2011. 4 Thorndike, Robert L dan Hagen, Elizabeth P. (1977). *Measurement and evaluation in psychology and education*. 4th edition. New York: Jhon Wiley & Sons.
- Sukardi. 2009. *Evaluasi Pendidikan*. Jakarta: PT Bumi Aksara.
- Sumintono, B. & Widhiarso, W. (2015) *Apliksi Pemodelan Rasch pada Assessment Pendidikan*. Cimahi: Trim Komunikata
- Tim Pengembangan MKDP Kurikulum dan Pembelajaran. (2011).*Kurikulum dan Pembelajaran*. Jakarta Utara: PT Raja Grafindo Persada.
- UU RI Nomor 20 Pasal 1 tahun 2003: Sistem Pendidikan Nasional.

BAB 5

PENGEMBANGAN TES SIKAP DAN RUBRIK PENILAIANNYA

Oleh Cahya Amalia Chusna

5.1 Pengertian Tes Sikap

Sikap berawal dari perasaan yang mengarah pada kecenderungan seseorang untuk bertindak terhadap sesuatu atau objek. Sikap bisa dibentuk supaya perilaku yang diinginkan dapat terwujud. Menurut Permendikbud Nomor 23 Tahun 2016, tes atau penilaian sikap adalah cara yang dilakukan oleh guru untuk menghimpun data mengenai tingkah laku peserta didik. (Kemendikbud, 2016b). Ada perbedaan dalam karakteristik penilaian atau tes sikap jika dibandingkan dengan penilaian pengetahuan dan keterampilan. Perbedaan inilah yang menyebabkan perlunya pengembangan instrumen penilaian sikap peserta didik. Penilaian sikap dilakukan untuk mengetahui pencapaian dan pengembangan perilaku berdasarkan nilai-nilai sikap yang termuat di Kompetensi Dasar (KD) dan Kompetensi Inti (KI). Kompetensi Inti (KI) tersebut terdiri dari kompetensi sikap spiritual (KI.1) dan kompetensi sikap sosial (KI.2), yang terintegrasi dengan Kompetensi Dasar (KD) dari kompetensi pengetahuan (KD.3) dan kompetensi keterampilan (KD.4).

5.2 Lingkup Penilaian Sikap

Pendidik menilai hasil belajar dalam hal sikap spiritual dan sikap sosial dengan fokus pada target atau lingkup seperti menerima (*receiving*), merespon (*responding*), menilai (*valuing*), mengorganisasi (*organization*), dan karakterisasi (*characterization*). Berikut deskripsi masing-masing tingkatan sikap.

Tabel 5. 1. Tingkatan Sikap

Tingkatan Sikap	Deskripsi
menerima (<i>receiving</i>)	reaksi peserta didik terhadap hal-hal yang terjadi di sekitarnya, seperti masalah, gejala, dan kondisi, dan sebagainya dapat dilihat dari tingkat keterlibatan atau perasaan yang ditunjukkan. Contoh peserta didik menyadari pentingnya menerapkan sikap disiplin serta menghilangkan sifat malas.
merespon (<i>responding</i>)	kemampuan yang dimiliki peserta didik untuk terlibat aktif dan merespon fenomena tertentu. Contohnya munculnya keinginan dari dalam diri peserta didik untuk mempelajari lebih dalam tentang konsep disiplin.
menilai (<i>valuing</i>)	Kemampuan untuk mengevaluasi atau mengapresiasi terhadap suatu kegiatan ataupun objek. Contohnya dorongan dari dalam diri peserta didik untuk mengimplementasikan perilaku disiplin di berbagai lingkungan.

Tingkatan Sikap	Deskripsi
mengorganisasi (<i>organization</i>)	kemampuan mempertemukan berbagai perbedaan nilai yang menghasilkan nilai baru yang membawa manfaat. Contohnya peserta didik memberikan dukungan untuk menegakkan sikap disiplin di segala tempat dan waktu.
karakterisasi (<i>characterization</i>)	Kemampuan untuk mengatur nilai-nilai sistem yang mempengaruhi pola kepribadian dan perilaku peserta didik. Contohnya sikap disiplin dijadikan pedoman peserta didik untuk bertindak di lingkungan sekolah, rumah dan masyarakat.

Sumber: (Krathwohl, Bloom, dan Masia 1964)

5.3 Perumusan Indikator Penilaian Sikap

Perencanaan penilaian sikap diawali dengan mengidentifikasi sikap sikap spiritual dan sosial yang ada pada KI.1 (Menerima dan menjalankan ajaran agama yang dianutnya) dan KI.2 (Menunjukkan perilaku jujur, disiplin, tanggung jawab, santun, peduli, dan percaya diri dalam berinteraksi dengan keluarga, teman, dan guru) serta sikap yang diinginkan oleh sekolah yang telah disebutkan dalam program kurikulum 2013. (Kemendikbud, 2016a).

Tanda tercapainya KI.1 dan KI.2 karena adanya indikator yang terukur. Indikator tes atau penilaian sikap merupakan sesuatu (sikap atau perilaku) yang terlihat dari diri peserta didik dimana guru bisa menilai melalui pengamatan atau observasi. (Asmarawati, Riyadi, dan Sujadi, 2016). Tabel 5.2 berikut berisi contoh rumusan indikator penilaian sikap berdasarkan KI.1 dan KI.2

Tabel 5. 2. Contoh Rumusan Indikator Penilaian Sikap

Kompetensi Inti (KI)	Indikator
Sikap Spiritual	
Menerima dan menjalankan ajaran agama yang dianutnya	• memanjatkan doa sebelum dan setelah menjalankan aktivitas • Melaksanakan kegiatan keagamaan sesuai dengan keyakinan agama yang dipeluk • memulai dan mengakhiri kegiatan dengan ucapan salam • mengucapkan puji syukur kepada Tuhan YME atas nikmat yang diberikan • mengucapkan rasa syukur atas kemampuan untuk mengontrol diri sendiri • mengucapkan rasa syukur setelah sukses menyelesaikan tugas • menerima takdir dengan ikhlas setelah berusaha sebaik mungkin • menjaga keharmonisan hubungan antar manusia • Menyatakan rasa terima kasih kepada Tuhan YME sebagai warga negara Indonesia • Menghormati individu lain yang sedang melaksanakan praktik keagamaan sesuai agamanya.

Kompetensi Inti (KI)	Indikator
Sikap Sosial Menunjukkan perilaku jujur, disiplin, tanggung jawab, santun, peduli, dan percaya diri dalam berinteraksi dengan keluarga, teman, dan guru	
1. Jujur Adalah perilaku yang dapat diandalkan dalam perbuatan, perkataan ataupun pekerjaan.	• Tidak melakukan kecurangan saat mengerjakan soal ujian • Tidak menyalin karya orang lain • Menyampaikan emosi sejalan dengan perasaan yang dirasakan • Menyusun laporan sesuai dengan situasi dan data yang terkini di lapangan • Mengakui kesalahan yang sudah dilakukan dan mengakui kelemahan yang ada pada diri sendiri
2. Disiplin adalah suatu sikap yang mencerminkan perilaku yang teratur dan taat terhadap aturan dan peraturan.	• datang tepat waktu • menyelesaikan tugas sebelum batas waktu yang ditetapkan • melaksanakan peraturan dan ketentuan yang berlaku di sekolah
3. Tanggung jawab adalah perilaku seseorang dalam menjalankan tanggung jawab terhadap diri sendiri, lingkungan, masyarakat, negara, dan Tuhan	• Menuntaskan tugas secara mandiri dengan mutu yang tinggi. • Menghadapi konsekuensi dari tindakan yang telah dilakukan • Tidak memberikan kritikan atau tuduhan tanpa memiliki bukti yang kuat • Memberi pengakuan akan

Kompetensi Inti (KI)	Indikator
YME.	kesalahan yang telah dilakukan dan memohon maaf atas perbuatan tersebut. • Berkomitmen dengan janji yang telah disepakati • Tidak melimpahkan kesalahan yang telah diperbuat kepada orang lain • Melakukan sesuatu secara sukarela dan tanpa harus diminta sesuai dengan apa yang telah diucapkan.
4. Toleransi adalah perilaku yang mengapresiasi perbedaan latar belakang, perspektif, dan kepercayaan individu.	• Tidak mempengaruhi teman yang berbeda pendapat • Menyetujui perjanjian/ kesepakatan walaupun memiliki pandangan yang tidak sama • Sanggup menerima ketidaksempurnaan orang lain • Sanggup mengampuni kesalahan orang lain • Sanggup dan bersedia bekerjasama dengan siapa pun yang berasal dari berbagai latar belakang, memiliki pandangan yang berbeda, dan keyakinan yang beragam • Menghormati pilihan dan keyakinan orang lain tanpa memaksa mereka untuk menerima pandangan kita • Siap untuk menerima sesuatu yang belum pernah ditemui

Kompetensi Inti (KI)	Indikator
	sebelumnya
5. Gotong royong adalah kolaborasi antara individu dengan yang lainnya untuk mencapai suatu tujuan bersama dengan cara berbagi tanggung jawab dan memberikan bantuan secara tulus.	• Ikut berpartisipasi sepenuh hati dalam kegiatan bakti sosial untuk membersihkan ruang kelas atau seluruh area sekolah • Sanggup melaksanakan tugas sesuai dengan permintaan • Sanggup memberikan bantuan kepada sesama tanpa mengharapkan balasan • Terlibat dalam kegiatan kolaboratif • Mengutamakan kepentingan bersama daripada kepentingan diri sendiri • Menemukan solusi untuk mencapai kesepakatan dalam mengatasi perbedaan perspektif antara individu dan orang lain • Melakukan kolaborasi dengan berbagai pihak demi mencapai tujuan yang diinginkan
6. Santun atau sopan adalah tindakan positif untuk memperlihatkan perilaku yang baik di lingkungan sosial, baik melalui kata-kata maupun perilaku nyata	• Menunjukkan rasa hormat kepada orang yang tua • Berbicara dengan sopan, halus dan rendah hati. • Tidak meludah di sembarang tempat. • Menghormati orang yang sedang berbicara tanpa menginterupsi. • Menyampaikan rasa terima kasih atas bantuan atau pertolongan yang diberikan

Kompetensi Inti (KI)	Indikator
	• Menunjukkan sikap ramah dengan memberikan salam, senyuman, dan sapa • Menghormati privasi orang lain dengan menanyakan izin sebelum masuk ke ruangan • Menyayangi dan menghormati orang lain dengan cara sebagaimana kita menginginkan diri kita sendiri dihargai.
7. Percaya diri adalah keadaan psikologis atau mental yang memberikan keyakinan kuat dalam mengambil tindakan.	• Menyampaikan pendapat dan mengambil tindakan tanpa keraguan. • Mampu membuat keputusan dengan cepat dan tepat • memiliki tekad kuat • Berani dalam bertindak • Memiliki keberanian untuk presentasi di hadapan seluruh anggota kelas • Memiliki keberanian untuk menyampaikan pendapat, mengajukan pertanyaan, atau memberikan jawaban
8. Peduli adalah sikap dan perbuatan yang senantiasa berusaha untuk menghindari dan memperbaiki kesalahan serta kerusakan yang	• Menolong orang yang membutuhkan • Menahan diri untuk tidak melakukan kegiatan yang mengganggu maupun merugikan orang lain • Berpartisipasi dalam kegiatan sosial dengan tujuan membantu mereka yang sedang

Kompetensi Inti (KI)	Indikator
terjadi pada manusia, lingkungan, dan tata aturan.	membutuhkan bantuan • menjaga lingkungan sekolah • merawat tanaman di lingkungan sekolah

Sumber: (Kemendikbud, 2017b)

5.4 Teknik, Instrumen dan Rubrik Penilaian Sikap

Ada beberapa metode yang bisa digunakan untuk menilai sikap spiritual dan sosial peserta didik. Pertama adalah melakukan pengamatan langsung terhadap perilaku peserta didik atau observasi. Kedua, peserta didik melakukan evaluasi diri untuk melihat kemajuan dalam dirinya. Ketiga, saran dan pendapat dari teman sejawat juga merupakan sumber yang baik untuk mendapatkan umpan balik tentang perkembangan sikap peserta didik. Keempat menyimpan catatan dalam jurnal juga dapat membantu dalam melacak perkembangan dan perubahan sikap peserta didik. Alat yang digunakan untuk menilai sikap peserta didik berupa skala penilaian (*rating scale*) dan daftar *checklist* yang dilengkapi dengan rubrik penskoran. Data yang diperoleh dihitung menggunakan rumus kemudian disesuaikan dengan predikat yang tertera dalam Permendikbud Nomor 81A Tahun 2013 (Kemendikbud, 2014).

5.4.1 Teknik Penilaian Sikap

1. Obsevasi

Kunandar (2015) menyatakan bahwa observasi adalah teknik penilaian berulang menggunakan indra yang dilakukan secara langsung maupun tidak langsung dengan menggunakan pedoman atau lembar observasi yang mengandung sejumlah indikator perilaku atau elemen yang diamati. Observasi tidak

langsung dapat dilaksanakan melalui bantuan individu lain, seperti pendidik lain, wali murid, siswa, dan staf sekolah. Observasi langsung guru melakukan observasi secara mandiri tanpa bergantung pada bantuan orang lain.

Instrumen yang dipakai adalah lembar observasi yang kegunaannya untuk mempermudah guru (pendidik) menyusun laporan terkait hasil pengamatan sikap peserta didik yang mencakup indikator pada KI.1 dan KI.2. (Kemendikbud, 2017b)

Lembar observasi terdiri dari identitas peserta didik yang akan diobservasi, rubrik dan skala pengukuran. Rubrik yang disusun berisi petunjuk mengenai penggunaan skala penilaian atau daftar cek. Dalam hal ini, petunjuk penskoran mencakup metode memberikan nilai dan menghitung nilai akhir. Berikut beberapa hal yang harus diperhatikan agar kegiatan observasi lebih efektif dan efisien:

- Merencanakan tujuan dan indikator yang akan diobservasi.
- Membuat pedoman observasi dalam bentuk daftar *check list* atau skala penilaian.
- Mencatat hasil observasi segera mungkin.
- Membuat kesimpulan berdasarkan hasil observasi.

2. Penilaian Diri

Penilaian diri merupakan suatu metode dimana seseorang (peserta didik) menilai dirinya sendiri dengan mengenali baik buruknya sikapnya dalam bertindak. (Kemendikbud, 2017b).

Penilaian diri memiliki peran penting bersamaan dengan pergeseran fokus pembelajaran dari *teaching learning* (berpusat pada guru) menjadi *student learning* (berpusat pada peserta didik), yang didasarkan pada gagasan belajar mandiri. Tujuan penilaian diri ini untuk memberikan dukungan pada proses pembelajaran peserta didik (Kemendikbud, 2014).

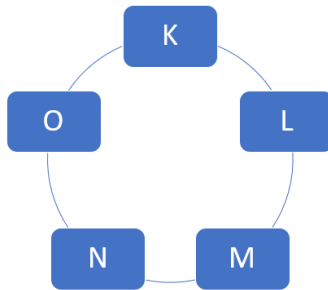
Instrumen yang dipakai adalah lembar penilaian diri yang dibuat dengan format yang jelas, sederhana, tidak ambigu, dan mudah diisi oleh peserta didik. Lembar penilaian diri dirancang untuk menunjukkan sikap peserta didik dalam situasi nyata, memiliki makna, dan membantu peserta didik mengenali kelebihan atau kekurangannya serta untuk menghilangkan kebiasaan peserta didik dalam menilai diri mereka dengan cara yang bersifat pribadi atau berdasarkan pendapat subjektif. Berikut adalah langkah-langkah penilaian diri peserta didik

- buatlah tujuan penilaian diri dengan jelas
- tentukan indikator yang akan diamati, sikap spiritual tercantum dalam KI.1 sedangkan sikap sosial tercantum dalam KI. 2 pada kurikulum 2013
- tentukan kriteria penilaiannya
- buatlah format penilaian, yang terdiri dari identitas peserta didik, petunjuk pengisian, daftar *checklist*, skala penilaian, pedoman penskoran dan predikat sikap.

3. Penilaian Teman Sejawat

Penilaian teman sejawat merupakan metode penilaian di mana seorang peserta didik (penilai) mengevaluasi sikap atau perilaku peserta didik lain. Tujuan penilaian ini adalah untuk mendorong peserta didik agar menjadi lebih objektif, memiliki rasa empati, menghargai keragaman, dan melakukan refleksi diri. (Kemendikbud 2017a).

Penilaian teman sejawat dilakukan saat peserta didik melakukan aktivitas di dalam atau di luar ruang kelas. Misalnya, selama kegiatan berkelompok, setiap peserta didik diminta untuk melihat dan mengevaluasi anggota kelompok yang lain. Sedangkan dalam kelompoknya sendiri, dia juga dievaluasi oleh dua temannya yang lain, seperti yang ditunjukkan pada gambar di bawah ini.



Gambar 5. 1. Diagram Penilaian Teman Sejawat

Sumber : (Kemendikbud 2017a)

Diagram pada Gambar 5.1 menunjukkan alur bagaimana teman sejawat saling mengevaluasi sikap dan perilaku satu sama lain.

- Peserta didik kode K melihat dan menilai peserta didik kode L dan peserta didik kode E. Peserta didik kode A juga mendapat penilaian dari peserta didik kode L dan O.
- Peserta didik kode L melihat dan menilai peserta didik kode K dan peserta didik kode M. Peserta didik kode L juga mendapat penilaian dari peserta didik kode K dan M.
- Peserta didik kode M melihat dan menilai peserta didik kode L dan peserta didik kode N. Peserta didik kode M juga juga mendapat penilaian dari peserta didik kode L dan N.
- Peserta didik kode N melihat dan menilai peserta didik kode M dan peserta didik kode O. Peserta didik kode N juga mendapat penilaian dari peserta didik kode M dan O.
- Peserta didik kode O melihat dan menilai peserta didik kode N dan peserta didik kode K. Peserta didik kode O juga mendapat penilaian dari peserta didik kode N dan K.

Alat atau instrumen yang digunakan dalam penilaian diri berupa lembar penilaian diri. Instrumen ini terdiri dari pernyataan-pernyataan yang menggambarkan sikap positif dengan pilihan jawaban "Ya" dan "Tidak" atau melalui penggunaan *Skala Likert*. Satu lembar penilaian diri dapat

digunakan untuk penilaian sikap spiritual dan sikap sosial secara bersamaan.

4. Jurnal

Jurnal mencakup serangkaian catatan yang dibuat oleh para pendidik dan/atau staff pendidikan mengenai sikap dan perilaku positif atau negatif, selama di dalam dan di luar pembelajaran (Kemendikbud, 2014). Beberapa hal yang perlu diperhatikan dalam membuat jurnal adalah:

- Jurnal atau catatan yang dibuat harus seobjektif mungkin.
- Pengamatan dilaksanakan dengan cara mencatat kejadian / peristiwa yang berkaitan dengan kompetensi inti yang diujikan.
- Mencatat jurnal sesegera mungkin.

5.4.2 Instrumen dan Rubrik Penilaian Sikap

1. Observasi

Instrumen Penilaian Sikap Spiritual (KI.1)

Satuan Pendidikan	: SMP Nusa Indah
Kelas/Semester	: VII/Dua
Tanggal Observasi	: 6 Desember 2023
Materi Pokok	: Tata Surya

A. Petunjuk Umum

Berilah tanda centang (√) pada kolom yang tersedia dengan kriteria sebagai berikut:

Indikator mengucap salam sebelum dan setelah mengikuti pembelajaran IPA materi Tata Surya

- a) Menyapa guru
- b) Mengucap salam
- c) Mencium tangan guru
- d) Bersikap tertib

Indikator berdoa sebelum dan sesudah melakukan mengikuti pembelajaran IPA materi Tata Surya

- a) Duduk dengan tenang
- b) Bersikap sopan
- c) Membaca doa dengan baik
- d) Sungguh-sungguh

Indikator mensyukuri berbagai bentuk dan keragaman di Kelas

- a) Berbicara sopan
- b) Tidak memilih-milih teman
- c) Menghormati guru
- d) Menghargai teman

B. Rubrik Penilaian

Skor			
1	2	3	4
Hanya memenuhi satu kriteria	Memenuhi dua kriteria	Memenuhi tiga kriteria	Memenuhi empat kriteria

C. Lembar Observasi

Nama	Aspek 1 (salam)				Aspek 2 (Berdoa)				Aspek 3 (Syukur)				Skor	Predikat
	4	3	2	1	4	3	2	1	4	3	2	1		

D. Skor Penilaian

$$skor = \frac{\text{skor yang diperoleh}}{\text{skor maksimal}} \times 4$$

E. Predikat

Berdasarkan Permendikbud Nomor 81A Tahun 2013 peserta didik memperoleh nilai dengan kriteria sebagai berikut

Sangat Baik (SB)	: $3,33 < \text{skor} \leq 4,00$
Baik (B)	: $2,33 < \text{skor} \leq 3,33$
Cukup (C)	: $1,33 < \text{skor} \leq 2,33$
Kurang (K)	: $\text{skor} \leq 1,33$

Instrumen Penilaian Sikap Sosial (KI.2)

Satuan Pendidikan : SMP Nusa Indah
Kelas/Semester : VII/Dua
Tanggal Observasi : 6 Desember 2023
Materi Pokok : Tata Surya

A. Petunjuk Umum

Berilah tanda centang (v) di kolom skor sesuai dengan sikap spiritual yang ditunjukkan oleh peserta didik, dengan kriteria di bawah ini

Indikator sikap toleransi terhadap keberagaman pendapat dalam proses pemecahan masalah materi Tata Surya

1. *Kurang baik* apabila sama sekali tidak memiliki sikap toleran terhadap keberagaman pendapat dalam proses pemecahan masalah materi tata surya.
2. *Cukup* apabila ada sedikit usaha untuk bersikap toleran terhadap keberagaman pendapat dalam proses pemecahan masalah materi tata surya.
3. *Baik* apabila menunjukkan sudah ada usaha untuk bersikap toleran terhadap keberagaman pendapat dalam proses pemecahan masalah materi tata surya tetapi masih belum konsisten.
4. *Sangat baik* apabila menunjukkan sudah ada usaha untuk bersikap toleran keberagaman pendapat dalam proses pemecahan masalah materi tata surya secara terus menerus dan konsisten.

Indikator sikap bekerjasama tim dalam memecahkan masalah materi Tata Surya.

1. *Kurang baik* apabila sama sekali tidak berusaha untuk bekerjasama tim dalam kegiatan memecahkan masalah materi tata surya.

2. *Cukup* apabila menunjukkan ada sedikit usaha untuk bekerjasama tim dalam kegiatan memecahkan masalah materi tata surya tetapi masih belum konsisten.
3. *Baik* apabila menunjukkan sudah ada usaha untuk bekerjasama tim dalam kegiatan memecahkan masalah materi tata surya tetapi masih belum konsisten.
4. *Sangat baik* apabila menunjukkan sudah ada usaha untuk bekerjasama tim dalam memecahkan masalah materi tata surya secara terus menerus dan konsisten.

Indikator sikap aktif dalam kegiatan berdiskusi dengan tim.

- Kurang baik* apabila sama sekali tidak ambil bagian dalam kegiatan diskusi tim materi tata surya.
- Cukup* apabila ada sedikit usaha untuk ambil bagian dalam kegiatan diskusi tim materi tata surya tetapi masih belum konsisten.
- Baik* apabila menunjukkan sudah ada usaha untuk ambil bagian dalam kegiatan diskusi tim materi tata surya tetapi masih belum konsisten.
- Sangat baik* apabila menunjukkan sudah ada usaha untuk selalu aktif dan konsisten dalam kegiatan diskusi tim materi tata surya.

B. Rubrik Penilaian

Skor			
1	2	3	4
Kurang baik	Cukup baik	Baik	Sangat baik

C. Lembar Observasi

[illegible]

D. Skor Penilaian

$$\text{skor} = \frac{\text{skor yang diperoleh}}{\text{skor maksimal}} \times 4$$

E. Predikat

Sangat Baik (SB)	: 3,33 < skor ≤ 4,00
Baik (B)	: 2,33 < skor ≤ 3,33
Cukup (C)	: 1,33 < skor ≤ 2,33
Kurang (K)	: skor ≤ 1,33

4. Penilaian Diri

Instrumen Penilaian Diri Sikap Spiritual dan Sosial

Nama :
Kelas/Semester :
Tanggal Penilaian :
Materi :

A. Petunjuk Pengisian

- Isilah identitas diri pada tempat yang disediakan
- Berilah penilaian terhadap sikapmu dengan memberikan tanda centang (√) pada kolom yang telah disediakan berdasarkan kriteria berikut:
Skor 1: saya **jarang** melakukan perilaku yang dinyatakan
Skor 2: saya **kadang-kadang** melakukan perilaku yang dinyatakan
Skor 3: saya **sering** melakukan perilaku yang dinyatakan
Skor 4: saya **selalu** melakukan perilaku yang dinyatakan
- Hitunglah skor yang diperoleh serta tentukan predikatnya

B. Lembar Penilaian Diri

Butir Nilai : Menghargai Perilaku Bertakwa kepada Tuhan YME

No	Pernyataan	4	3	2	1
1	Saya menjalankan ibadah sesuai dengan keyakinan agama yang saya peluk				
2	Saya mengucapkan salam dan berdoa				

No	Pernyataan	4	3	2	1
	setiap kali saya belajar				
3	Saya dengan tekun mengikuti proses pembelajaran				
4	Saya menjaga hubungan yang baik dengan semua makhluk hidup ciptaan Tuhan				
Total					
Predikat					

Butir Nilai : Mengembangkan Kejujuran

No	Pernyataan	4	3	2	1
1	Saya berusaha keras menyelesaikan ulangan tanpa melihat pekerjaan teman				
2	Saya menyelesaikan tugas secara mandiri, tanpa batasan dan tekanan				
3	Saya mengekspresikan perasaan saya dengan jujur tanpa ada tekanan				
4	Saya membuat laporan berdasarkan pada fakta				
Total					
Predikat					

Butir Nilai : Bertanggung Jawab

No	Pernyataan	4	3	2	1
1	Setiap kali ada jadwal Pelajaran IPA, saya selalu membawa buku pelajarannya.				
2	Saya dengan sepenuh hati melakukan semua tugas dan instruksi dari guru				
3	Saya mengumpulkan tugas sesuai dengan jadwal yang telah ditetapkan				
4	Saya bersungguh-sungguh setiap melakukan kegiatan praktikum IPA				
Total					
Predikat					

Butir Nilai : Kerjasama Kelompok

No	Pernyataan	4	3	2	1
1	Saya merasa gembira Ketika bekerja dalam tim				
2	Saya memberikan pendapat ketika bekerja dalam tim				
3	Saya menjelaskan kepada anggota kelompok mengenai kesulitan/hambatan yang sedang saya hadapi				
4	Saya turut serta mencari solusi dan bekerja sama dengan anggota kelompok untuk menyelesaikan masalah				
Total					
Predikat					

C. Skor Penilaian

$$skor = \frac{skor\ yang\ diperoleh}{skor\ maksimal} \times 4$$

D. Predikat

Sangat Baik (SB)	: 3,33 < skor ≤ 4,00
Baik (B)	: 2,33 < skor ≤ 3,33
Cukup (C)	: 1,33 < skor ≤ 2,33
Kurang (K)	: skor ≤ 1,33

5. Penilaian Teman Sejawat

Lembar Penilaian Teman Sejawat

Nama peserta didik yang dinilai :

Kelas :

Materi :

Tanggal :

A. Petunjuk Pengisian

- Isilah identitas teman sejawat yang akan dinilai
- Berilah tanda centang di kolom yang tersedia sesuai dengan kriteria di bawah ini:
4 = *selalu*, jika selalu sesuai dengan pernyataan
3 = *sering*, jika sering melakukan sesuai pernyataan dan kadang-kadang tidak melakukan
2 = *kadang-kadang*, jika kadang-kadang melakukan dan sering tidak melakukan
1 = *tidak pernah*, jika tidak pernah melakukan
- Hitunglah skor yang diperoleh serta tentukan predikatnya

B. Lembar Penilaian Teman Sejawat

No	Aspek Pengamatan	Skor			
		4	3	2	1
1	Teman saya tidak melakukan kecurangan saat mengerjakan tugas maupun ulangan				
2	Teman saya tidak meniru atau mencontoh karya yang dihasilkan orang lain				
3	Teman saya mengemukakan perasaan apa adanya				
4	Teman saya berkata apa adanya sesuai fakta				
Total					
Predikat					

C. Skor Penilaian

$$\text{skor} = \frac{\text{skor yang diperoleh}}{\text{skor maksimal}} \times 4$$

D. Predikat

Sangat Baik (SB)	: 3,33 < skor ≤ 4,00
Baik (B)	: 2,33 < skor ≤ 3,33
Cukup (C)	: 1,33 < skor ≤ 2,33
Kurang (K)	: skor ≤ 1,33

DAFTAR PUSTAKA

- Asmarawati, Endah, Riyadi, dan Imam Sujadi. 2016. "The process of integrating social and spiritual attitudes." *Jurnal Elektronik Pembelajaran Matematika* 4(1): 58–69.
<https://media.neliti.com/media/publications/120234-ID-proses-integrasi-sikap-sosial-dan-spirit.pdf>.
- Kemendikbud. 2014. *Permendikbud Nomor 104 Tahun 2014 Tentang Penilaian Hasil Belajar oleh Pendidik pada Pendidikan Dasar dan Pendidikan Menengah*. <http://pgsd.uad.ac.id/wp-content/uploads/lampiran-permendikbud-no-104-tahun-2014.pdf>.
- . 2016a. *Permendikbud No 24 Tahun 2016 Tentang Kompetensi Inti dan Kompetensi Dasar Pelajaran pada Kurikulum 2013 pada Pendidikan Dasar dan Pendidikan Menengah*.
- . 2016b. *Permendikbud Nomor 23 tahun 2016 Tentang Standar Penilaian Pendidikan*.
<http://arxiv.org/abs/1011.1669v0><http://dx.doi.org/10.1088/1751-8113/44/8/085201>.
- . 2017a. *Panduan Penilaian oleh Pendidik dan Satuan Pendidikan Sekolah Menengah Atas*. Jakarta: Direktorat Pembinaan SMA Ditjen Pendidikan Dasar dan Menengah.
- . 2017b. *Panduan Penilaian oleh Pendidik dan Satuan Pendidikan Sekolah Menengah Pertama*. Jakarta: Kementerian Pendidikan dan Kebudayaan Direktorat Jenderal Pendidikan Dasar dan Menengah Direktorat Pembinaan Sekolah Menengah Pertama.
<http://repositori.kemdikbud.go.id/18051/1/1>. Panduan Penilaian SMP - Cetakan Keempat 2017.pdf.
- Krathwohl, Bloom, dan Masia. 1964. *Taxonomy of Educational Objectives Book 2 Affective Domain*. New York: Longman Inc.
- Kunandar. 2015. *Penilaian Autentik (Penilaian Hasil Belajar Peserta Didik. Berdasarkan Kurikulum 2013)*. Jakarta: PT Raja Grafindo Persada.

BAB 6

PENILAIAN BERBASIS KELAS

Oleh Evi Ristiana

6.1 Pendahuluan

Penilaian Berbasis Kelas (PBK) merupakan pendekatan penilaian yang menempatkan pengajar sebagai penilai utama dalam menilai kemajuan dan capaian peserta didik. Proses PBK, guru menjadi sentral dalam mengumpulkan informasi mengenai hasil belajar murid, menganalisis, serta informasi feed back yang dapat membantu peserta didik terus berkembang.

Berikut karakteristik dari penilaian berbasis kelas:

1. Mengumpulkan data-data kemajuan peserta didik: data-data peserta dikumpulkan oleh guru dan informasi mengenai kemajuan peserta didik dari berbagai sumber, seperti tugas, ujian, proyek, observasi kelas, dan interaksi sehari-hari di kelas. Informasi yang dikumpulkan mencakup berbagai aspek, termasuk pemahaman konsep, keterampilan, sikap, dan partisipasi siswa.
2. Penggunaan Informasi untuk Peningkatan Pembelajaran: Hasil penilaian peserta didik digunakan untuk membimbing proses pembelajaran selanjutnya. Guru dapat menyesuaikan metode mengajar, kurikulum, atau memberikan dukungan kepada siswa yang belum baik hasil belajarnya. PBK bertujuan untuk lebih dari sekadar memberi nilai; ia lebih fokus pada memberikan umpan balik konstruktif yang dapat membantu siswa dalam pengembangan pribadi dan akademiknya.

3. Konsistensi dan Akurasi Penilaian: Guru berusaha untuk memberikan penilaian yang konsisten dan akurat terhadap kemajuan siswa. Standar penilaian yang jelas, termasuk standar kompetensi, kompetensi dasar, dan indikator hasil belajar, membantu memastikan konsistensi penilaian di seluruh kelas.
4. Deskripsi Kompetensi dan Kemajuan Siswa: PBK memberikan gambaran yang holistik tentang kompetensi, keterampilan, dan kemajuan siswa, bukan hanya berfokus pada nilai numerik. Guru dapat memberikan informasi lebih rinci tentang kekuatan dan kelemahan siswa, sehingga dapat merancang strategi pengajaran yang lebih efektif.
5. Pemberdayaan Siswa: PBK dapat mendorong keterlibatan siswa dalam pemantauan dan penilaian terhadap hasil belajarnya sendiri. Dengan pemahaman yang lebih baik tentang kemajuan mereka, siswa dapat mengambil peran aktif dalam pengelolaan pembelajaran pribadi mereka.

Evaluasi adalah suatu proses yang berkelanjutan dan dapat melibatkan revisi jika diperlukan. Evaluasi dalam konteks pendidikan memang mencakup penilaian sebagai salah satu komponennya. Penilaian berperan penting dalam membantu mengukur pencapaian siswa, memahami keefektifan metode pengajaran, dan menyediakan informasi yang berguna untuk perbaikan. Penilaian Berbasis Kelas (PBK) merupakan strategi guru untuk mengarahkan fokus pembelajaran di kelas. Konteks PBK, pengajar melakukan penilaian secara berkelanjutan selama proses pembelajaran berlangsung. Beberapa elemen yang terkait dengan PBK melibatkan pengumpulan data terus menerus. Proses PBK tidak hanya tentang peserta didik dinilai, juga feed back yang membangun kepada siswa. Umpan balik ini membantu peserta didik memahami kelebihan dan kekurangan peserta didik serta memberikan arah untuk perbaikan. Guru menggunakan data hasil penilaian untuk mengevaluasi efektivitas metode pengajaran. Jika ditemukan masalah atau kesenjangan dalam pemahaman siswa,

guru dapat menyesuaikan pendekatan pengajaran mereka. PBK dapat melibatkan siswa dalam pemantauan dan penilaian terhadap kemajuan mereka sendiri. Ini mendorong tanggung jawab siswa terhadap pembelajaran mereka sendiri. Evaluasi PBK mengungkapkan kekurangan atau masalah dalam pembelajaran, guru dapat merevisi rencana pembelajaran mereka. Tujuannya adalah meningkatkan efektivitas pengajaran dan hasil belajar siswa.

Penggunaan PBK, pendidik dapat memahami lebih baik kebutuhan individual siswa dan merespons dinamika kelas dengan lebih fleksibel. Ini adalah suatu pendekatan yang adaptif dan responsif terhadap perubahan dalam kebutuhan belajar siswa. Penilaian Berbasis Kelas memainkan peran penting dalam mendukung upaya peningkatan kualitas pendidikan dengan fokus pada pembelajaran yang efektif dan relevan bagi siswa.

Konsep penilaian yang dijelaskan oleh Schwartz dan kawan-kawan, sebagaimana dikutip oleh Oemar Hamalik, menekankan pada dua aspek utama: “pertama, bentuk penilaian sebagai usaha untuk memeriksa sejauh mana anak telah mengalami kemajuan belajar atau telah mencapai tujuan belajar; kedua, penilaian sebagai program yang memberikan pendapat dan menentukan arti atau faedah suatu pengalaman”.

Konsep penilaian ini menekankan pentingnya pengukuran dalam menilai kemajuan siswa dan keberhasilan proses pembelajaran, serta merinci peran penilaian kelas sebagai alat penting untuk pengambilan keputusan guru. Data hasil penilaian dianggap sebagai informasi kritis untuk membantu merancang dan mengelola pembelajaran dengan baik.

6.2. Tujuan Penilaian Berbasis Kelas

Empat tujuan dari penilaian berbasis kelas (PBK) yang dilakukan oleh guru atau dosen:

1. Penelusuran (*Keeping Track*): Tujuan pada pemantauan terus-menerus terhadap proses pembelajaran peserta didik untuk sesuai dengan rencana. Penilaian digunakan sebagai alat untuk menelusuri perkembangan siswa selama proses pembelajaran pada tahun pelajaran dan memberikan gambaran tentang capaian kompetensi.
2. Pengecekan (*Checking-up*): Pengecekan dilakukan untuk mengidentifikasi kekurangan proses belajar yang dirasakan oleh siswa selama proses pembelajaran. Guru menggunakan evaluasi untuk mengetahui kemampuan memahami murid terhadap kompetensi tertentu yang mungkin sulit untuk dikuasai.
3. Pencarian (*Finding-out*): Tujuan kegiatan pencarian untuk menemukan penyebab terjadinya kelemahan atau kesulitan dalam pembelajaran. Guru melakukan analisis hasil penilaian untuk mencari hal-hal yang dapat menjadi penghambat kegiatan belajar dan mencari solusi untuk membantu mereka.
4. Kegiatan menyimpulkan: Kegiatan menyimpulkan telah dijadwalkan dalam kurikulum. Penilaian di kelas bertujuan untuk mengetahui proses belajar murid, pemahaman murid selama proses belajar. Kesimpulan memiliki signifikansi yang penting dalam melaporkan hasil pembelajaran kepada berbagai pihak terkait, termasuk orang tua siswa, kepala sekolah.

Fungsi-fungsi penting dari penilaian berbasis Kelas (PBK) bagi murid dan pengajar:

1. Fasilitasi perkembangan pribadi murid: PBK membimbing siswa dalam mengubah perilaku mereka menuju kemajuan yang positif, memberikan umpan balik konstruktif, dan mendorong mereka mencapai potensi maksimal.

2. Memberikan Kepuasan kepada siswa: PBK bukan saja mengukur pencapaian siswa, tetapi juga memberikan pengakuan dan kepuasan kepada mereka atas usaha dan prestasi yang telah dicapai.
3. Evaluasi Efektivitas Metode Pengajaran: PBK membantu pengajar mengevaluasi keefektifan metode pengajaran, memberikan wawasan tentang strategi yang efektif, dan memungkinkan penyesuaian pendekatan instruksional.
4. Dasar Pertimbangan dan Keputusan Administratif: PBK menyediakan landasan bagi guru untuk membuat pertimbangan dan keputusan administratif terkait manajemen kelas, penempatan siswa, dan perencanaan pembelajaran
5. Identifikasi Kesulitan dan Kelemahan Siswa: PBK memberikan gambaran mendalam tentang kemajuan siswa, mengidentifikasi kesulitan belajar, dan menyoroti area yang memerlukan perhatian khusus untuk perbaikan.
6. Kontrol dan Evaluasi Kondisi Belajar: PBK berfungsi sebagai alat kontrol untuk menilai kesesuaian kondisi belajar di sekolah, menyediakan data untuk perbaikan dan peningkatan kondisi pembelajaran.
7. Pertimbangan untuk Perencanaan dan Pedoman Sekolah: PBK memberikan informasi yang mendukung perencanaan pembelajaran di tingkat kelas dan sebagai panduan bagi sekolah untuk mencapai standar pendidikan yang ditetapkan.

Dengan fungsi-fungsi ini, PBK tidak hanya berperan sebagai alat penilaian semata, melainkan juga sebagai sistem yang mendukung perkembangan siswa, evaluasi metode pengajaran, serta pengambilan keputusan di tingkat kelas dan sekolah.

Pedoman yang sebaiknya diikuti oleh pengajar dalam proses penilaian kelas dapat diuraikan sebagai berikut:

1. Integrasi Penilaian dengan Kegiatan Belajar Mengajar: Pengajar menganggap penilaian sebagai bagian yang tidak terpisah dari kegiatan belajar . Pastikan bahwa penilaian terpadu dengan proses pengajaran, menciptakan gambaran menyeluruh tentang pencapaian siswa.
2. Penguatan Penilaian sebagai Refleksi Diri: Guru perlu mengembangkan strategi untuk mendorong siswa melihat penilaian sebagai refleksi diri. Bantu siswa memahami dan mengidentifikasi kekuatan serta kelemahan mereka melalui proses penilaian.
3. Penerapan Berbagai Strategi Penilaian: Pengajar disarankan menggunakan beragam pendekatan penilaian dalam proses belajar mengajar, dapat memberikan pandangan yang lebih holistik tentang hasil belajar siswa.
4. Perhatian terhadap siswa berkebutuhan khusus: Guru perlu memperhatikan serta mempertimbangkan kebutuhan khusus peserta didik. Penyesuaian strategi penilaian membantu mendukung keberhasilan setiap siswa, termasuk mereka yang memiliki kebutuhan khusus.
5. Penggunaan Sistem Pencatatan yang Beragam: Penting untuk mengembangkan sistem pencatatan yang bervariasi untuk mengamati kegiatan belajar siswa. Sistem pencatatan yang beragam dapat membantu guru merekam dan memahami kemajuan siswa secara lebih komprehensif.
6. Variasi Cara dan Alat Penilaian: Guru sebaiknya menggunakan berbagai cara dan alat penilaian. Dengan variasi ini, guru dapat menilai berbagai aspek hasil belajar, termasuk pengetahuan, keterampilan, sikap, dan penerapan konsep dalam konteks nyata.
7. Pendidikan dan Peningkatan Mutu Proses Pembelajaran: PBK bukan hanya untuk menilai, melainkan juga untuk mendidik dan meningkatkan mutu proses pembelajaran. Guru perlu

terus berupaya meningkatkan efektivitas metode pengajaran dan memberikan umpan balik konstruktif kepada siswa.

Pedoman ini, mengarahkan pengajar agar memastikan bahwa PBK tidak hanya berfokus pada pemberian nilai, tetapi juga berkontribusi pada pembelajaran siswa dan perbaikan proses pembelajaran secara keseluruhan.

6.3. Prinsip-prinsip PBK

Prinsip-prinsip yang sangat relevan dalam Penilaian Berbasis Kelas (PBK). Berikut adalah penjelasan lebih lanjut tentang setiap prinsip:

1. Validitas:

Validitas mencerminkan seberapa tepat penilaian mengukur sesuatu yang seharusnya diukur. Penilaian Berbasis Kelas (PBK), harus dapat menggunakan alat yang sesuai kompetensi murid yang diharapkan. Validitas tercapai ketika instrumen penilaian mampu secara akurat mengukur keterampilan atau pengetahuan yang diinginkan.

2. Reliabilitas

Reliabilitas mencerminkan tingkat konsistensi hasil penilaian. Dalam Penilaian Berbasis Kelas, konsistensi dalam memberikan nilai atau skor menjadi hal kritis untuk memastikan bahwa hasil penilaian dapat diandalkan dan memberikan gambaran yang tepat tentang kemampuan siswa.

3. Menyeluruh

Prinsip ini menekankan bahwa penilaian harus menjangkau pengukuran yang terkait dengan setiap kompetensi dasar. PBK memberikan gambaran komprehensif tentang pencapaian siswa, mencakup alternatif dan alat

penilaian untuk mengevaluasi berbagai kemampuan dasar murid

4. Berkesinambungan

Penilaian dilakukan terencana, bertahap, dan berkelanjutan selama proses pembelajaran berlangsung. Penilaian dilakukan berkesinambungan, guru dapat memantau perkembangan siswa dari waktu ke waktu dan memberikan umpan balik yang relevan.

5. Objektif

Objektivitas menekankan pada keadilan dan keterencanaan dalam penilaian. Penting untuk memastikan bahwa penilaian hasilnya dapat dianggap obyektif.

6. Mendidik:

Penilaian tidak hanya tentang memberikan nilai, tetapi juga memberikan edukasi. Kegiatan penilaian meningkatkan motivasi murid untuk belajar. meningkatkan metode pembelajaran untuk pengajar, serta membimbing murid menuju pertumbuhan dan perkembangan yang optimal. Memahami dan menerapkan prinsip-prinsip ini, PBKeningkatkan kualitas pembelajaran di kelas.

6.3 Teknik Penilaian Berbasis Kelas

Terdapat berbagai pendekatan/strategi digunakan mengenai perkembangan belajar murid, baik itu terkait dengan proses belajar maupun hasil. Prinsip dasar dari teknik-teknik pengumpulan informasi ini adalah menilai perkembangan kemampuan kognitif, psikomotori dan afektif murid sesuai kompetensi yang harus mereka capai. Penilaian terhadap kompetensi dasar ini dilakukan dengan merujuk pada indikator pencapaian kompetensi yang mencakup satu ranah atau lebih. Dengan menggunakan indikator-

indikator ini, dapat ditetapkan metode penilaian yang sesuai. Oleh karena itu, terdapat tujuh teknik yang dapat digunakan untuk mengumpulkan data, sebagai berikut:

1. Penilaian Unjuk Kerja

Penilaian tindakan atau uji praktik adalah metode yang efektif untuk mengumpulkan berbagai informasi tentang perilaku yang diharapkan muncul dalam siswa (keterampilan). Penilaian perbuatan melibatkan observasi terhadap kegiatan peserta didik saat mereka melakukan suatu tindakan. Jenis penilaian ini sangat cocok digunakan untuk mengevaluasi pencapaian kompetensi yang mengharuskan siswa menunjukkan unjuk kerja. Sebaliknya, penilaian kinerja (produk) adalah penilaian terhadap kemampuan siswa dalam mengawasi proses dari manfaat dan pemanfaatan atau penggunaan bahan untuk menghasilkan suatu hasil, pekerjaan praktis, atau kualitas estetik dari sesuatu yang mereka hasilkan. Sebagai contoh, karya seni seperti menggambar, melukis, atau kerajinan tangan. Aspek-aspek yang dinilai dalam menilai hasil karya ini mencakup kemampuan untuk bereksplorasi dan mengembangkan gagasan desain, memilih bahan yang sesuai, menggunakan alat dengan kreativitas dan inovasi, serta memilih bentuk dan gaya dalam karya seni. Fasilitas-fasilitas produksi melibatkan persiapan, produksi, dan refleksi.

2. Penilaian Sikap

Penilaian sikap dapat dilaksanakan dalam konteks berbagai objek sikap, seperti pandangan terhadap mata pelajaran, hubungan dengan guru, respons terhadap proses pembelajaran, dan sikap terhadap materi pelajaran. Selain itu, sikap juga dapat dikaitkan dengan pembentukan nilai-nilai yang diinginkan dalam diri peserta didik melalui suatu kerangka kerja khusus. Ada variasi metode yang dapat diterapkan untuk mengukur sikap, termasuk pengamatan

perilaku, penggunaan pertanyaan langsung, analisis laporan pribadi, dan penerapan skala sikap.

3. Penilaian Tertulis

Penilaian Tertulis adalah bentuk uji di mana pertanyaan dan jawaban disampaikan kepada peserta didik dalam bentuk tulisan. Peserta didik memiliki opsi untuk memberikan tanggapan tidak hanya melalui penulisan jawaban, tetapi juga melalui tindakan seperti memberikan tanda, mewarnai, menggambar, dan sebagainya. Dalam merancang uji tertulis, beberapa aspek perlu diperhatikan agar pengembangannya dapat mencerminkan atau mengukur penguasaan kompetensi secara akurat. Salah satu hal yang perlu diperhatikan adalah kemampuan uji untuk mengevaluasi hasil pembelajaran sesuai dengan tujuan pembelajaran, di mana pertanyaan uji menjadi contoh representatif dari materi yang telah diajarkan.

4. Penilaian Proyek

Penilaian Proyek adalah evaluasi terhadap tugas yang harus diselesaikan dalam periode tertentu. Proses ini mencakup investigasi dari tahap perencanaan, pengumpulan data, pengorganisasian, pengolahan, hingga presentasi informasi. Penilaian proyek digunakan untuk mengukur pemahaman dan pengetahuan dalam bidang tertentu, kemampuan peserta didik dalam menerapkan pengetahuan tersebut, terutama dalam penyelidikan khusus, serta keterampilan peserta didik dalam menyajikan informasi tentang subjek tersebut.

Proses penilaian proyek melibatkan pendekatan sistematis yang melibatkan pengumpulan informasi (dalam bentuk angka atau deskripsi verbal) dan analisis informasi untuk mendukung pengambilan keputusan. Karakteristik penilaian kinerja mencakup identifikasi semua aspek yang

relevan, dokumentasi seluruh kebutuhan yang ditunjukkan, dan fokus pada kemampuan yang akan dievaluasi. Model-model penilaian proyek mencakup skala penilaian, checklist, dan kombinasi dari skala penilaian dan checklist sebagai contoh. **Penilaian Produk**

Penilaian Produk adalah evaluasi terhadap keahlian dalam menciptakan suatu produk dan kualitas produk tersebut. Penilaian terhadap produk tidak hanya didasarkan pada hasil akhirnya, melainkan juga memperhatikan seluruh proses pembuatannya. Penilaian produk mencakup evaluasi terhadap kemampuan peserta didik dalam menciptakan berbagai produk dalam bidang teknologi dan seni, seperti makanan, pakaian, karya seni (patung, lukisan, gambar), serta barang-barang yang terbuat dari bahan seperti kayu, plastik, keramik, dan logam.

5. Penilaian Hasil Akhir Kegiatan

Penilaian dilakukan dalam bentuk portofolio, berasal dari istilah bahasa Inggris yang merujuk pada dokumen atau kumpulan surat-surat. Portofolio ini merupakan kumpulan hasil karya siswa yang mencerminkan pencapaian belajarnya. Selain memberikan informasi mengenai kemampuan dan pemahaman siswa, serta menggambarkan sikap dan minat siswa terhadap materi pelajaran, portofolio juga mampu menunjukkan perkembangan atau peningkatan yang diperoleh siswa selama proses pembelajaran. Beberapa karakteristik penilaian portofolio mencakup pengumpulan data melalui karya siswa, evaluasi yang dilakukan secara berkelanjutan, dan kemampuannya untuk memperlihatkan perkembangan peserta didik.

Penilaian portofolio adalah bentuk evaluasi berkelanjutan yang didasarkan pada koleksi informasi yang mencerminkan perkembangan kemampuan peserta didik dalam suatu periode tertentu. Dalam konteks kelas, penilaian

portofolio dilakukan terhadap rangkaian karya siswa yang terstruktur dengan baik, diambil selama periode pembelajaran tertentu. Guru dan siswa memanfaatkan portofolio ini untuk memonitor kemajuan pengetahuan, keterampilan, dan kesiapan siswa dalam suatu mata pelajaran.

6. Penilaian Diri

Penilaian diri merupakan suatu pendekatan evaluasi di mana individu yang sedang dinilai diminta untuk menilai dirinya sendiri terkait status, proses, dan tingkat pencapaian kompetensi yang telah dikuasainya dalam suatu mata pelajaran tertentu. Teknik penilaian diri dapat digunakan dalam berbagai aspek penilaian yang mencakup kompetensi kognitif, afektif, dan psikomotor.

Keunggulan penilaian berbasis kelas adalah:

1. Pengumpulan informasi kemajuan belajar baik formal maupun non formal diadakan secara bersamaan secara menyenangkan, serta senantiasa memungkinkan kesempatan bagi murid untuk menunjukkan apa yang diketahui, dipahami dan mampu dikerjakan peserta didik.
2. Pencapaian hasil belajar peserta didik tidak dibandingkan dengan prestasi kelompok (norm reference asesment), tetapi dibandingkan dengan kemampuan sebelumnya kriteria pencapaian kompetensi, standar pencapaian dan level pencapaian nasional, dalam rangka membantu pebelajar mencapai apa yang ingin dicapai bukan untuk menghakiminya.
3. Pengumpulan informasi menggunakan berbagai cara, agar kemajuan belajar peserta didik dapat terdeteksi secara lengkap.
4. Peserta didik perlu dituntut agar dapat mengeksplorasi dan memotivasi diri untuk mengarahkan semua potensi dalam menanggapi, mengatasi semua masalah yang dihadapi dengan caranya sendiri, bukan sekedar melatih siswa memilih jawaban

yang tersedia 5. Untuk menentukan ada tidaknya kemajuan belajar dan perlu tidaknya bantuan secara berencana, bertahap dan berkesinambungan, berdasarkan fakta dan bukti yang cukup akurat.

6.4 Bentuk Penilaian Berbasis Kelas

Penilaian berbasis kelas dapat dibagi menjadi beberapa jenis.

Pertama bisa dilakukan dengan kuis, yang digunakan untuk menanyakan konsep dasar dari pelajaran sebelumnya dalam bentuk isian singkat dan dilakukan sebelum pelajaran. Kuis dilaksanakan untuk mengetahui murid memahami pelajaran yang diterima dan memahami hubungan antara pelajaran yang akan datang dan pelajaran sebelumnya (apresiasi).

Kedua dengan ujian lisan di kelas; digunakan untuk menunjukkan penguasaan peserta didik tentang pemahaman fakta, ide, prinsip, dan teknik yang berkaitan dengan subjek yang dipelajari. Dengan demikian, diharapkan siswa memiliki bangunan keilmuan yang kukuh dan landasan yang kuat untuk mempelajari topik terkait.

Ketiga dalam bentuk ulangan harian; melakukan secara berkala pada akhir pembelajaran untuk menunjukkan penguasaan kognitif murid dan mengevaluasi keberhasilan pengguna sumber daya pembelajaran. Keempat dalam bentuk tugas individu; dilakukan secara periode untuk diselesaikan oleh peserta didik dan dapat berupa tugas dikelas dan di rumah. Tugas individu dipakai untuk mengungkapkan kemampuan teoritis dan praktis penguasaan hasil penilaian dalam penggunaan media, metode, strategi, dan prosedur tertentu.

Kelima, melalui tugas kelompok; digunakan untuk mengevaluasi kemampuan kerja kelompok dalam mencari solusi masalah, sekaligus membangun sikap kebersamaan pada peserta didik. Tugas kelompok ini akan lebih efektif jika difokuskan pada

penyelesaian aspek-aspek yang bersifat empiris dan kasuistik. Bila memungkinkan, peserta didik dalam kelompok dapat diminta untuk melakukan observasi langsung atau merancang proyek dengan menggunakan data dan informasi lapangan.

Keenam, melalui ulangan persemester; digunakan untuk menilai penguasaan kompetensi pada akhir periode semester. Kompetensi dievaluasi berdasarkan kisi-kisi yang mencerminkan kompetensi dasar, hasil pembelajaran, dan indikator pencapaian hasil belajar yang telah dikembangkan selama semester tersebut.

Ketujuh, melalui ulangan kenaikan kelas; digunakan untuk menilai tingkat kemahiran peserta didik dalam menguasai materi pada bidang studi tertentu selama satu tahun ajaran. Pemilihan kompetensi ujian seharusnya merujuk pada kompetensi dasar yang berkelanjutan, memiliki nilai aplikatif, atau relevan untuk pembelajaran pada bidang lain yang terkait.

Kedelapan, melalui responsi atau ujian praktek; digunakan pada mata pelajaran yang melibatkan aspek praktik, seperti untuk menilai penguasaan akhir peserta didik dari segi psikomotorik, afektif, dan kognitifnya.

Penilaian berbasis kelas mengikuti prinsip berkelanjutan dan komprehensif yang bertujuan untuk memberdayakan peserta didik agar dapat belajar mandiri, bekerja sama, dan melakukan penilaian diri. Oleh karena itu, proses penilaian dilakukan sebagai bagian integral dari Penilaian Berbasis Kelas (PBK).

DAFTAR PUSTAKA

- Arifin Zainal. 2012. *Evaluasi Pembelajaran*. Jakarta: Direktorat Jenderal Pendidikan Islam Kementerian Agama RI.
- Arikunto Suharsimi. 2009. *Dasar-Dasar Evaluasi Pendidikan* (Edisi Revisi). Jakarta: PT Bumi Aksara.
- Asrul, dkk. 2014. *Evaluasi Pembelajaran*. Medan: Cipta pustaka Media.
- Daryanto, 2001, *Evaluasi Pendidikan*, Jakarta: PT. Rineka Cipta.
- Hamalik, Oemar. 2009. *Psikologi Belajar Dan Mengajar*, Bandung, Sinar Baru Algensindo.
- Hamalik, Oemar. 2010. *Manajemen Pengembangan Kurikulum*. Bandung : PT Remaja Rosdakarya.
- Mulyasa, 2006. *Kurikulum Yang Disempurnakan (Pengembangan Standar Kompetensi dan Kompetensi Dasar)*, , Bandung: PT Remaja Rosdakarya.
- Mulyasa, 2007. *Kurikulum Tingkat Satuan Pendidikan*, , Bandung: Remaja Rosdakarya.
- Sugiono, 2007. *Pengantar Evaluasi Pendidikan*, Jakarta: PT Raja Grafindo Persada.
- Sukiman, 2012. *Pengembangan Sistem Evaluasi*. Yogyakarta, Insan Madani.
- Ristiana. 2018. *Pemberian Umpan Balik Pada Pembelajaran*. Jurnal Riset Pendidikan Dasar. Hal 65-71, Vol. 1 April. ISSN 2615-1766

BAB 7

PENILAIAN KINERJA

Oleh Yuli Tamar Filindity

7.1 Penilaian Kinerja

Penilaian pembelajaran merupakan bagian evaluasi terhadap proses pembelajaran dan hasil belajar. Penilaian proses pembelajaran bertujuan untuk mengungkapkan kemampuan dan sikap siswa terhadap partisipasi dalam proses pembelajaran. Penilaian dalam pendidikan adalah proses pengumpulan, pengolahan, dan interpretasi informasi tentang hasil belajar siswa sebagai dasar pengambilan keputusan.

Penilaian adalah suatu proses sistematis dalam mengumpulkan, menganalisis, dan menyajikan informasi dengan cara yang tepat dan efektif untuk menilai keberhasilan siswa. Tujuan penilaian antara lain: menilai kekuatan individu melalui kegiatan tertentu, mengidentifikasi kebutuhan pembelajaran, mendukung dan mendorong siswa, mendukung dan mendorong guru untuk meningkatkan pengajaran, mengidentifikasi strategi pembelajaran, memberikan akuntabilitas dan meningkatkan kualitas pengajaran. Alat evaluasi yang baik adalah alat yang dapat mengukur keberhasilan suatu pelatihan secara akurat.

Penilaian hasil belajar adalah serangkaian strategi yang menunjukkan hasil pengetahuan, keterampilan, dan kinerja siswa melalui kegiatan yang bermakna dan menarik. Penilaian kinerja dilakukan untuk mengumpulkan data mengenai kemampuan siswa dalam melakukan tugas-tugas yang berkaitan dengan mata pelajaran yang dipelajari. Penilaian kinerja ini mengharuskan siswa

untuk mempraktikkan dan menggunakan pengetahuan yang diperoleh dalam banyak situasi sesuai standar yang dipersyaratkan. Penilaian membantu guru mengidentifikasi masalah dan kesulitan yang dihadapi anak-anak (misalnya, beberapa anak kesulitan dengan keterampilan komunikasi) dan dapat merencanakan serta mengembangkan program pembelajaran yang tepat untuk merangsang kemampuan tersebut.

Penilaian kinerja, sering disebut penilaian autentik, adalah metode evaluasi multi-segi yang diperoleh melalui proses menilai tertulis, menilai perilaku, dan pekerjaan rumah. Penilaian kinerja merupakan jenis penilaian yang mengharuskan siswa untuk mempraktikkan dan menerapkan pengetahuan/konsep yang ajarkan dalam kondisi dan situasi sesuai standar yang dipersyaratkan. Penilaian kinerja mengharuskan siswa untuk membuat suatu produk dan/atau mendemonstrasikan satu atau dua proses, dan menilai kualitas pekerjaan siswa menggunakan kriteria yang ditetapkan dengan jelas. Misalnya siswa melakukan kegiatan pengembangan pengetahuan seperti menulis laporan atau berlatih mengukur suhu suatu zat di laboratorium.

7.2 Bentuk Penilaian Kinerja

Ada dua aspek utama dalam evaluasi kerja: penggunaan energi atau input dalam suatu pekerjaan untuk menghasilkan suatu produk (proses) dan produksi suatu produk dari pekerjaan yang diperlukan. Evaluasi kinerja dapat mengevaluasi proses, produk, atau keduanya (kinerja dan produk) berdasarkan dua aspek dasar tersebut. Penilaian kinerja dilakukan dengan mengamati siswa ketika melakukan atau menciptakan kegiatan berdasarkan tujuan pembelajaran, dengan mengamati keluaran/hasil kinerja, atau kedua-duanya. Kompetensi yang ditunjukkan peserta didik merupakan bidang yang dinilai. Evaluasi keterampilan didasarkan pada kualitas kinerja siswa terhadap tujuan yang telah ditetapkan.

Proses evaluasi dilakukan mulai dari perencanaan dan pelaksanaan tugas hingga pencapaian hasil akhir.

Jenis penilaian kinerja yang akan diberlakukan kepada peserta didik dapat berupa penilaian praktek, penilaian produk, dan penilaian proyek.

1. Penilaian Praktek

Menilai praktikum dilakukan melalui observasi saat siswa mendemonstrasikan atau mempraktikkan aktivitas sesuai dengan tujuan kompetensi. Selama penilaian praktek, guru dapat mengevaluasi aspek kognitif (pengetahuan), aspek psikomotor (keterampilan), dan aspek afektif (sikap siswa). Penilaian praktek dapat dilakukan pada semua mata pelajaran seperti terlihat pada tabel dibawah ini :

Tabel 7. 1. Tabel Penilaian praktek

No.	Kelompok Mata Pelajaran	Contoh Penilaian Praktek
1	IPA	• Percobaan di laboratorium • Presentase hasil percobaan/proyek
2	Bahasa Indonesia	• Berpidato • Membaca puisi
3	Matematika	• Menggunakan konsep matematika untuk mengukur tinggi pohon • Menciptakan jaringan spasial

Aspek-aspek yang dinilai pada saat penilaian sebenarnya akan dimasukkan dalam formulir penilaian. Rubrik tersebut memuat kriteria-kriteria yang berkaitan dengan langkah-langkah yang dilakukan ketika melakukan suatu kegiatan. Langkah-langkahnya dilakukan secara

berurutan, mudah dipahami, jelas, mudah diamati dan terukur.

Penilaian yang dilakukan guru juga selalu menitikberatkan pada penilaian kelompok, bukan penilaian individu. Hal ini disebabkan karena keterbatasan guru dalam memperhatikan prestasi individu siswa dan terbatasnya waktu dalam menyelesaikan latihan yaitu 2x40 menit untuk 1 kali pertemuan dan seminggu sekali, menyebabkan guru kesulitan untuk melakukan evaluasi kinerja individu. Dari beberapa permasalahan yang ditemui dalam praktek, perlu dikembangkan suatu alat penilaian kinerja yang dapat mengevaluasi kinerja siswa secara detail dan alternatif teknik penilaian yang dapat mengevaluasi kinerja siswa dalam kegiatan praktek dengan lebih realistis

2. Penilaian Produk.

Penilaian produk didasarkan pada kualitas teknis dan estetika karya atau produk yang diciptakan siswa. Hasil karya siswa dapat berupa kain, kertas, logam, kayu, plastik, tanah liat. Seni seperti melukis, menggambar, patung dan menulis; dan laporan penelitian/hasil ilmiah. Penilaian produk dapat dilakukan pada mata pelajaran seperti pada Tabel 7.1.

Tabel 7. 2. Penilaian produk

No.	Kelompok Mata Pelajaran	Contoh Penilaian Praktek
1	IPA	• Membuat laporan praktikum • Membuat produk bioteknologi
2	Bahasa Indonesia	• Membuat resensi • Membaca karya sastra
3	IPS	• Membaca peta • Membuat laporan keuangan

Aspek-aspek yang dievaluasi dalam evaluasi produk dimasukkan dalam lembar evaluasi. Bentuk evaluasi mencakup kriteria yang berkaitan dengan kualitas teknis dan estetika produk. Kriteria tersebut harus lengkap, jelas, mudah diamati dan diukur.

3. Penilaian Proyek

Penilaian proyek merupakan penilaian yang akan dilaksanakan dalam suatu periode tertentu. Kegiatannya meliputi perencanaan, pengumpulan data, analisis data, penyajian data, dan pelaporan. Waktu penyelesaiannya tergantung tingkat kesulitan tugas (misal 1 minggu, 2 minggu, 1 bulan, atau 1 minggu).

Memulai sebuah proyek memerlukan data primer, data sekunder, kolaborasi dengan berbagai kelompok, dan kemampuan untuk mengevaluasi hasil. Oleh karena itu, penilaian mata pelajaran dapat dilakukan secara bersama-sama untuk semua mata pelajaran atau untuk setiap mata pelajaran pada semua jenjang pendidikan.

Penilaian proyek berfokus pada perencanaan, pelaksanaan, dan hasil proyek. Berkaitan dengan hal tersebut, guru perlu melakukan serangkaian kegiatan antara lain menyiapkan desain dan alat penilaian, mengumpulkan data, menganalisis data, dan menyiapkan laporan. Penilaian proyek dapat menggunakan alat pengujian, skala penilaian sampai dengan rubrik. laporannya dapat disajikan dalam bentuk poster maupun dalam bentuk tertulis. Produk akhir suatu proyek mungkin memerlukan penilai khusus. Tujuan dari menilai produk suatu proyek adalah untuk mengevaluasi kualitas dan tampilan hasil akhir secara komprehensif dan analitis.

Sebuah proyek dapat dievaluasi pada setiap tahap, termasuk persiapan (perencanaan), pelaksanaan dan pelaporan. Hasil pembelajaran yang dapat dinilai pada bagian ini adalah:

- a. Tahap persiapan (perencanaan, mengatur proyek, data awal)
- b. Tahap implementasi (kerja tim, kerja individu, identifikasi, analisis)
- c. Tahap pelaporan (interpretasi data, laporan, presentase hasil)

Dalam pembelajaran di kelas, pendidik dapat melakukan penilaian selama pembelajaran sebagai cara untuk memantau keterampilan perencanaan, penyelidikan, dan analisis, interpretasi dan penyampaian laporan dari proyek. Saat menggunakan proyek dalam penilaian sumatif, fokusnya sering kali tertuju pada produk. Aspek-aspek yang dievaluasi dalam evaluasi proyek dimasukkan dalam lembar evaluasi.

Siswa dapat berbagi pengetahuan pengalaman belajar tentang suatu materi tau tema, menyampaikan pertanyaan, dan mempelajarinya melalui hasil baca dan wawancara. Kegiatan ini dapat digunakan untuk menilai kemampuan siswa secara individu maupun dalam tim kerja. guru juga menggunakan hasil akhir proyek/ produk untuk menilai kemampuan peserta didik dengan mengungkapkan hasilnya dalam bentuk presentasi/ unjuk kerja (evaluasi langsung) dan dalam bentuk laporan (laporan kinerja).

7.3 Pengembangan Penilaian Kinerja

1. Rubrik Penilaian

Rubrik merupakan pedoman untuk yang membantu pendidik/guru menilai mutu hasil pekerjaan siswa secara berkala dan dapat dipertanggungjawabkan terhadap kualitas pekerjaan siswa. Rubrik juga dapat digunakan sebagai *feedback* terhadap hasil kinerja siswa. Penggunaan rubrik sebagai alat penilaian dapat membantu guru mengevaluasi pekerjaan siswa secara konsisten dan obyektif, meminimalisir waktu penilaian dalam jangka pendek dan jangka panjang, serta memberikan umpan balik yang efektif, tepat waktu dan mendorong proses belajar secara berlanjut.

Rubrik merupakan pedoman yang digunakan untuk mengevaluasi hasil belajar siswa. Hasil perolehan yang berasal dari rubrik dapat digunakan untuk meningkatkan kualitas pembelajaran. Selain itu, hasil atau nilai yang diperoleh dari rubrik dipergunakan untuk mengevaluasi perolehan keterampilan belajar.

Menggunakan rubrik sebagai alat penilaian dapat membantu guru dalam beberapa hal berikut. Menilai kinerja siswa secara teratur dan bermakna. Meminimalisir waktu untuk penilaian secara efektif dan efisien dalam jangka waktunya yang ditetapkan. Memberikan *feedback* yang efektif dan tepat waktu serta memotivasi siswa untuk terus belajar. Grade penilaian terdiri dari sejumlah bagian yang sesuai dengan karakteristik tugas yang dilakukan. Grade penilaian terdiri dari sejumlah bagian yang sesuai dengan karakteristik tugas yang dilakukan. Format pertama terdiri dari tiga bagian yaitu aspek, kriteria dan tingkat kinerja. Aspek adalah komponen, ruang lingkup atau aspek yang akan dinilai, misalnya pada saat kita menilai laporan praktek pada mata pelajaran IPA, maka aspek yang dinilai meliputi pendahuluan, teori penelitian teoritis, metodologi, hasil

observasi dan pembahasan. Kriteria merupakan uraian atau penjelasan yang mencerminkan hubungan antar aspek dan tingkat suatu skala kinerja.

Nilai skala kinerja menunjukkan tingkat ketercapaian kinerja siswa, yang dapat ditulis dalam bentuk angka-angka, seperti 0, 1, 2, dst, 1, 2, 3, dst, kurang atau, cukup, baik, dst. atau indikator tingkat kinerja lainnya. Format yang pertama dapat ditunjukkan dibawah ini

Tabel 7. 3. Contoh format rubrik

No.	Komponen Penilaian	Skala				Perolehan
		1	2	3	4	
1	Komponen 1	Kriteria	Kriteria	Kriteria	Kriteria	
2	Komponen 2	Kriteria	Kriteria	Kriteria	Kriteria	
3		...	...	...	...	...
4		...	...	...	...	...

Menggunakan kata kerja operasional untuk mendeskripsikan karakteristik/kriteria untuk menilai kemampuan, bukan pengetahuan. Oleh karena itu, nilai bukanlah tujuan utama menilai keterampilan komposisi, termasuk rubrik yang dapat disajikan secara deskriptif. Untuk penilaian sumatif, judul suatu sifat harus diukur secara jelas dan tidak ditafsir secara berulang kali, sehingga penghargaan terhadap suatu sifat/kriteria menunjukkan keunggulan atau kriteria lain dapat dipenuhi.

Rubrik dalam mengevaluasi setiap peserta didik secara personal berbeda dengan rubrik evaluasi untuk tim. Rubrik terhadap tim memperhitungkan kualitas kerjasama dan kualitas kontribusi setiap anggota tim. Dengan demikian,

anggota tim yang kontribusinya berkualitas tinggi akan memperoleh hasil yang lebih baik dibandingkan anggota tim lainnya. Langkah-langkah dalam menyusun suatu rubrik, antara lain:

- a. Identifikasi komponen-komponen utama yang mempengaruhi hasil.
- b. Mengidentifikasi komponen-komponen utama sehingga memperoleh hasil yang optimal.
- c. Gunakan beberapa komponen saja sehingga siswa dapat mengamati semuanya saat mereka menyelesaikan tugas.
- d. Menjelaskan setiap aspek proses menjadi referensi yang lebih nyata.
- e. Menentukan tingkat kemahiran pada tingkat kemahiran yang berbeda untuk mengalokasikan poin/poin.
Pada dasarnya tingkat kemampuan berbentuk skala numerik (0, 1, 2, 3...) yang menunjukkan tingkat kualitas pencapaian di setiap aspek/komponen, misalnya: (0) Tidak ditemukan aktivitas, tidak ada kategori atau kriteria bertemu. (1) Buruk, (2) cukup, (3) baik, dan (4) sangat baik. Tingkat kemampuan juga dapat diisi dengan indikator berdasarkan kata kunci dalam mengukur berbagai faktor.
- f. Penentuan bobot masing-masing aspek atau kelompok aspek yang dinilai tergantung pada tahapan proses kerja, kompleksitas dan urgensi aspeknya yang bersangkutan. Pembobotan masing-masing aspek atau kelompok aspek menjadi penilaian kelompok penilai/ tim pengajar.
Jika tidak dilakukan penentuan pembobotan, maka penilaian siswa diperoleh dengan rumus berikut:

$$Nilai = \frac{\text{Jumlah Skor Perolehan}}{\text{Jumlah Skor Maksimum}} \times 100$$

- g. Periksa dan bandingkan hasil kerja seluruh aspek dengan apa yang telah dilakukan orang lain di lapangan sebelumnya (bila ada pembandingan) untuk menjamin validitas dan reliabilitas rubrik.

Manfaat membuat rubrik adalah sebagai berikut :

- Siswa memperoleh informasi mengenai aspek-aspek yang akan dievaluasi dan memberikan pembobotan pada masing-masing aspek dan kelompok yang sesuai.
- Siswa didorong untuk mencapai hasil maksimalnya.
- Memberikan umpan balik kepada siswa untuk mengevaluasi hasil keterampilannya (self-assessment).
- Memberikan bimbingan kepada pendidik dalam melakukan observasi agar dapat berfokus pada aspek kinerja dan produk yang dievaluasi. Jadilah pembimbing bagi siswa Anda ketika merancang bahan-bahan yang dibutuhkan untuk suatu prosedur kerja dan melaksanakan tugas.

Sebagaimana diketahui Jenis penilaian kinerja siswa meliputi evaluasi kinerja, evaluasi produk, dan evaluasi proyek, maka berikut diberi contoh penialain kinerja yaitu :

A. Penilaian Kinerja Praktek untuk SD

Mapel IPA

Kelas

VI

Kurikulum	2013
Kompetensi Dasar	4.4. Melakukan percobaan rangkaian listrik sederhana secara Seri dan Paralel
Materi	Rangkaian Listrik Seri dan Paralel
Indikator Soal	Siswa dapat : • Melakukan pengujian rangkaian listrik sederhana secara seri dan secara paralel • Merancang rangkaian listrik seri dan paralel

Tugas menilai Praktikum

Alat dan bahan :

1. Empat buah lampu senter berlabel A,B, C dan D
2. Kabel yang cukup
3. 2 baterai
4. Gunting
5. Selotip

Cara Kerja :

1. Membuat rangkaian seri dengan menggunakan lampu berlabel A dan B, kabel serta lampu sambil amati ada atau munculnya lampu atau tidak.
2. Membuat rangkaian paralel dengan menggunakan lampu berlabel C dan D, kabel, dan baterai dan sambil amati ada atau munculnya lampu atau tidak.
3. Catatlah pengamatanmu pada tabel berikut.

Jenis Rangkaian	Lampu	Nyala lampu
Rangkaian Seri	Lampu A	
	Lampu B	
Rangkaian Paralel	Lampu C	
	Lampu D	

Rubrik Penilaian Praktek

No.	Aspek / Kriteria yang dinilai	Skor
1	Alat dan bahan • 4 senter berlabel A, B, C dan D, beberapa kabel, 2 baterai, cutter dan selotip. • Tidak membawa peralatan yang diperlukan • Tidak semua peralatan dibawa atau semua peralatan tidak memenuhi persyaratan.	1 - 3 3 2 1
2	Merangkai rangkaian seri/ paralel • Tepat dalam menyusun rangkaian secara seri pada lampu, kabel mauun baterai. • Tidak tepat dalam merangkai lampu, kabel, atau baterai dalam menyusun rangkaian seri • Tidak melakukan percobaan membuat rangkaian seri	1 – 3 3 2 1
3	Mengisi Tabel Pengamatan • Diisi berdasarkan nyala lampu yang dihasilkan • Terdapat kesalahan dalam mengisi data pengamatan • Tidak mengisi data pengamatan	1 – 3 3 2 1

B. Penilaian Kinerja Proyek untuk SMA

Mapel Kimia

Kelas

XI

Kurikulum	2013
Kompetensi Dasar	4.13. menyimpulkan hasil analisis data percobaan titrasi asam-basa
Materi	Titrasi Asam-basa
Indikator Soal	Peserta didik dapat : merancang, mengimplementasikan dan melaporkan proyek yang melibatkan analisis kandungan boraks dalam ahan makanan menggunakan titrasi asam basa.

Tugas Penilaian Proyek :

Siswa ditugaskan dalam kerja kelompok untuk menjawab pertanyaan-pertanyaan berikut:

1. Dapatkah boraks diukur dan dideteksi dalam bahan makanan?
2. Apakah kandungan boraks dalam makanan dapat ditentukan dengan titrasi asam-basa?
3. Adakah hubungan data eksperimen dengan prinsip dasar titrasi asam basa?

Cara Kerja:

Proyek ini diselesaikan dalam 4 tahap:

Tahap 1.

1. Proyek ini selesai pada waktu 4 minggu, rincian kerjanya sebagai berikut
2. Bagilah kelas menjadi kelompok-kelompok terdiri dari 5 atau 6 orang.

3. Minta setiap anggota kelompok untuk menemukan artikel tentang analisis kualitatif dan kuantitatif (menentukan pH) boraks dalam makanan.
4. Setiap kelompok membawa satu sampel makanan yang dapat diuji kualitasnya selama produksi menggunakan boraks pada pertemuan berikutnya.

Tahap 2.

1. Pengujian pertama menggunakan kunyit dan reagen kimia yang disediakan untuk menganalisis kualitas boax pada makanan yang dijual di daerah tersebut.
2. Tes kedua adalah melakukan analisis kuantitatif reagen.
3. Ukur kandungan boraks dalam sampel, dan hasilnya positif boraks.
4. Anggota tim menuliskan pengamatannya.
5. Mengolah dan menganalisis data yang diperoleh.
6. Meringkas data yang diperoleh dari membaca informasi titrasi asam basa dari buku, perpustakaan, atau sumber lain pada saat pembahasan laporan.

Tahap 3

Menyusun dan mengumpulkan laporan

Rubrik Penilaian Proyek

No.	Aspek / Kriteria yang dinilai	Skor
1	• Persiapan Sampel ○ Menunjukkan sampel yang sesuai ○ Menunjukkan sampel yang tidak sesuai ○ Tidak menunjukkan sampel	1 - 3 3 2 1
	• Praktikum Persiapan alat dan bahan ○ Mempersiapkan seluruh alat dan bahan yang diperlukan dengan benar tanpa bantuan	1 - 3 3

No.	Aspek / Kriteria yang dinilai	Skor
2	kelompok lain	
	○ Mempersiapkan seluruh alat serta bahan yang diperlukan dengan benar dengan bantuan kelompok lain.	2
	○ Tidak mempersiapkan alat dan bahan	1
	<i>Kebersihan Peralatan</i>	1 - 3
	○ Semua peralatan dibersihkan sebelum dan sesudah digunakan	3
	○ Hanya peralatan tertentu yang dibersihkan sebelum dan sesudah digunakan.	2
	○ Alat tidak dibersihkan	1
	<i>Ketelitian Pengamatan</i>	1 - 3
	○ Mengamati secara teliti terhadap seluruh kriteria: menimbang, membaca miniskus, mengukur volume, perubahan warna	3
	○ Mengamati dengan teliti terhadap 3 kriteria: menimbang, membaca miniskus, mengukur volume, perubahan warna	2
	○ Mengamati dengan teliti terhadap 1-2 kriteria: menimbang, membaca miniskus, mengukur volume, perubahan warna	1
3	Mencatat Data Observasi <i>Hasil Observasi</i>	1 - 3
	○ Hasil pengamatan ditulis secara lengkap dan jelas dengan foto dan video.	3
	○ Hanya tulis atau difoto atau divideo dengan lengkap	2
	○ Hasil pengamatan tidak dicatat/ difoto/ divideo	1

1. Pelaksanaan Penilaian Kinerja

Evaluasi prestasi kerja dilakukan dengan menggunakan metode observasi individu dan kelompok tergantung tugas yang diberikan. Metode penilaiannya komprehensif dan analitis. Metode keseluruhan digunakan jika penilai hanya memberikan satu jenis nilai berdasarkan kesan keseluruhannya terhadap kinerja siswa. Dalam metode analisis, skor diberikan pada berbagai aspek kinerja berdasarkan kriteria yang disepakati. Metode analisis dapat menggunakan alat berupa daftar periksa atau skala penilaian.

Hasil penilaian kinerja dapat disajikan dalam bentuk penilaian formatif maupun penilaian sumatif. Jika penilaian diberikan untuk tujuan formatif maka perlu dilakukan lebih banyak proses umpan balik agar siswa menyadari kesenjangan/kelemahan hasil belajarnya. Pemberian umpan balik secara langsung diberikan pada saat pelaksanaan tugas berdasarkan hasil observasi pendidik. Jika penilaian kinerja dilakukan untuk melihat (ringkasan) kinerja yang telah dikuasai siswa, maka menilai umpan balik siswa bukanlah tujuan utama dan skor yang dicapai siswa merupakan hasil pencapaian Keterampilan yang dinilai dalam penilaian kinerja. Oleh karena itu, dalam penilaian sumatif, setiap matrik untuk setiap aspek atau kriteria harus mencerminkan apakah sebagian atau seluruh kriteria terpenuhi. Oleh karena itu, data dalam rubrik harus mempunyai kriteria yang jelas, terukur dan tidak dapat ditafsirkan.

Kesulitan utama dalam mengevaluasi kinerja terletak pada penilaian. Setidaknya ada 3 kesalahan penilaian kinerja, yaitu:

1. Skala penilaian

Skala penilaian/petunjuk penilaian tidak jelas sehingga menyulitkan penilai dalam menggunakannya.

2. Cara Kerja

Prosedur yang digunakan untuk menilai kinerja karena itu dapat mempengaruhi hasil penilaian.

3. Pemberian skor

Penskor cenderung mengalami kesulitan menghilangkan masalah-masalah individual atau pribadi yang sering di kenal penilaian subjektif.

Dalam menilai hasil belajar siswa, besar kemungkinan penilai melakukan “kesalahan besar”, yaitu cenderung memberikan nilai yang tinggi, padahal hasil belajar siswa ternyata kurang baik.

DAFTAR PUSTAKA

- Clauser, B. E. 2000. Recurent Issues and Recent Advances in Scoring Performance Assessment, *Applied Psychological Measurement*, 24, 4, (pp. 310- 324). Sage Publications, Inc.
- Forster, Margaret. & Masters, Geoff. 1996. Portfolios Assessment Resource Kit. The Australian Council for Educational Research Ltd.
- Hargreaves, A. Earl, Lorna, M. and Schmidt, M. 2002. Perspective on Reform. *American Educational Research Journal*, Sprinf 2002, Vol.39. No.1.
- Mueller, J. 2006. Authentic Assessment. North Central College
- Nitko, A.J & Brookhart, S.M. 2016. Educational Assessment of Student. USA:Pearson
- Wulan, A.R. 2018. Menggunakan Asesmen Kinerja Untuk Pembelajaran Sains dan Penelitian. Bandung: UPI Press.
- Setiadi, H. 2006. Penilaian Kinerja. Jakarta: Pusat Penilaian Pendidikan Balitbang Depdiknas.
- Stiggins, R.J. 1994. Student-Centered Classroom Assessment. New York: MacMillan Publishing Company.

BIODATA PENULIS



Dr. Tarzan Purnomo, M.Si.

Dosen Program Studi Biologi
Fakultas MIPA Universitas Negeri Surabaya

Dr. Tarzan Purnomo, M.Si., lahir di Mojokerto, Jawa Timur tanggal 5 Mei 1965. Penulis adalah dosen tetap pada Program Studi Biologi, Fakultas MIPA, Universitas Negeri Surabaya sejak 1992. Menyelesaikan pendidikan S1 pada Jurusan Pendidikan Biologi FPMIPA IKIP Surabaya (1991), S2 Program Studi Biologi, Program Pascasarjana Universitas Gadjah Mada Yogyakarta (1998), dan Program Doktor (S3) Lingkungan Pesisir dan Lautan, Program Pascasarjana Universitas Brawijaya Malang (2011). Mengikuti *Postdoctoral Programe* Marine Ecology di Victoria University of Wellington, New Zealand (2013). Penulis menekuni bidang Ekologi, Ekologi Perairan, Ekologi Laut, Biologi Perairan, Ilmu Lingkungan, Konservasi Sumber Daya Alam, Manajemen Ekosistem Perairan, serta Ekologi Pesisir dan Kelautan.

BIODATA PENULIS



Dyah Ayu Fitriana, M.Pd.

Email: dyahayufitriana@gmail.com

Dosen Program Studi Pendidikan Guru Madrasah Ibtidaiyah
Fakultas FITK Institut Agama Islam Khozinatul Ulum Blora

Penulis lahir di Blora tanggal 3 Maret 1995. Penulis adalah dosen tetap pada Program Studi Pendidikan Guru Madrasah Ibtidaiyah Fakultas FITK, Institut Agama Islam Khozinatul Ulum Blora. Menyelesaikan pendidikan S1 pada Jurusan Pendidikan Matematika Universitas Muhammadiyah Surakarta dan melanjutkan S2 pada Jurusan Penelitian dan Evaluasi Pendidikan Universitas Negeri Yogyakarta.

BIODATA PENULIS



Achmad Rante Suparman

Staf Dosen Pendidikan Kimia Universitas Papua

Penulis lahir di Makale tanggal 6 September 1986. Penulis merupakan dosen tetap pada Program Studi Pendidikan Kimia FKIP Universitas Papua. Menyelesaikan pendidikan S1 pada tahun 2009 pada Prodi Pendidikan Kimia Universitas Negeri Makassar dan pada tahun 2013 menyelesaikan pendidikan S2 pada Prodi Pendidikan Kimia Universitas Negeri Makassar. Pada tahun 2019 penulis melanjutkan pendidikan Doktor (S3) pada Prodi Penelitian dan Evaluasi Pendidikan (PEP) Sekolah Pascasarjana Universitas Negeri Yogyakarta melalui Beasiswa Unggulan Dosen Indonesia (BUDI) Lembaga Pengelola Dana Pendidikan (LPDP) dan mendapat gelar Doktor pada tahun 2023.

BIODATA PENULIS



Dr. Suwarno, M.Pd

Dosen Program Studi Pendidikan Ekonomi
Fakultas Keguruan dan Ilmu Pendidikan
Universitas PGRI Ronggolawe Tuban

Penulis lahir di Klaten, Jawa Tengah tanggal 14 Agustus 1964. Penulis adalah dosen tetap pada Program Studi Pendidikan Ekonomi, Fakultas Keguruan dan Ilmu Pendidikan Universitas PGRI Ronggolawe Tuban, Jawa Timur.

Menyelesaikan Pendidikan S1 pada pendidikan Akuntansi, Universitas Negeri Yogyakarta, melanjutkan pendidikan S2, pada jurusan Penelitian dan Evaluasi Pendidikan di Pasca Sarjana, Universitas Negeri Yogyakarta dan Pendidikan S3 pada jurusan Penelitian dan Evaluasi Pendidikan di Pasca Sarjana, Universitas Negeri Yogyakarta

BIODATA PENULIS



Cahya Amalia Chusna, M.Pd.

Dosen Program Studi Pendidikan Guru Madrasah Ibtidaiyah
Fakultas Ilmu Tarbiyah dan Keguruan
Institut Agama Islam Khozinatul Ulum Blora

Penulis lahir di Blora tanggal 14 November 1995. Penulis adalah dosen tetap pada Program Studi Pendidikan Guru Madrasah Ibtidaiyah, Fakultas Ilmu Tarbiyah dan Keguruan, Institut Agama Islam Khozinatul Ulum Blora. Menyelesaikan pendidikan S1 pada Jurusan Pendidikan Matematika Universitas Negeri Semarang dan melanjutkan S2 pada Pendidikan Matematika Universitas Negeri Semarang

BIODATA PENULIS



Dr. Evi Ristiana, S.Pd., M.Pd.

Dosen Program Studi Pendidikan Biologi
Fakultas Matematika dan IPA Universitas Negeri Makassar

Penulis lahir di Bantaeng, salah satu kabupaten di Sulawesi Selatan, pada tanggal 2 Februari 1985. Penulis adalah dosen tetap pada Program Studi Pendidikan Biologi FMIPA UNM. Menyelesaikan pendidikan S1 Pendidikan Biologi pada tahun 2007 di Universitas Negeri Makassar. S2 Pendidikan Biologi di Universitas Negeri Malang Pada tahun 2008 dan selesai tahun 2010 Penulis menekuni bidang Menulis. Melanjutkan studi S3 Pendidikan Biologi di Universitas Negeri Malang selesai tahun 2017.

Penulis pernah aktif mengajar di prodi PGSD dan Magister Pendidikan Dasar Universitas Muhammadiyah Makassar tahun 2016 sampai 2021. Penulis termasuk penyusun borang akreditasi pertama Magister Pendidikan Dasar di Unismuh Makassar pada tahun 2018. Tulisan dalam buku ini merupakan hasil karya penulis untuk tidak melupakan sejarah penulis yang pernah mengabdikan sebagai dosen Unismuh Makassar. Semoga karya ini bermanfaat untuk semua.

BIODATA PENULIS



Yuli Tamar Filindity

Dosen Program Studi Pendidikan Kimia

Fakultas Keguruan dan Ilmu Pendidikan -Universitas Pattimura

Penulis lahir di Tehoru, tanggal 2 Juli 1980. Penulis adalah dosen tetap pada Program Studi Dosen Program Studi Pendidikan Kimia Fakultas Keguruan dan Ilmu Pendidikan -Universitas Pattimura. Menyelesaikan pendidikan S1 pada Program Studi Pendidikan Kimia, FKIP Unpatti, dan melanjutkan S2 pada Program Pascasarjana Program Studi Pendidikan Sains Universitas Negeri Yogyakarta. Penulis lebih menekuni bidang Pendidikan Sains