

A. Jenis Alat Penilaian

Secara garis besar alat penilaian digolongkan menjadi 2 jenis, yaitu non-tes dan tes. Seringkali kedua jenis alat penilaian tersebut dinamakan teknik penilaian.

1. Teknik Non-Tes

Teknik non-tes biasanya digunakan untuk mengevaluasi bidang afektif atau psikomotorik. Hal ini bisa dilakukan dengan cara seperti berikut ini.

a) Angket (*questionnaire*)

Angket adalah sebuah daftar pertanyaan atau pernyataan yang harus dijawab oleh responden. Angket berfungsi sebagai alat pengumpul data. Data tersebut berupa keadaan atau data diri, pengalaman, pengetahuan, sikap, pendapat mengenai suatu hal.

b) Wawancara (*interview*)

Wawancara merupakan teknik non-tes secara lisan. Pertanyaan yang diungkapkan umumnya menyangkut segi-segi sikap dan kepribadian siswa dalam proses belajarnya. Teknik ini dilakukan secara langsung dan dimaksudkan untuk memperoleh bahan-bahan penilaian bagi siswa.

c) Observasi (*pengamatan*)

Observasi adalah suatu teknik penilaian non-tes yang menginventarisasikan data tentang sikap dan kepribadian siswa dalam kegiatan belajarnya. Observasi dilakukan dengan mengamati kegiatan dan perilaku siswa secara langsung. Data yang diperoleh dijadikan bahan penilaian.

d) Inventori (*inventory*)

Inventori pada hakekatnya tidak banyak berbeda dengan angket. Inventori mengandung sejumlah pertanyaan yang tersusun dalam rangka mengetahui tentang sikap, pendapat dan perasaan siswa terhadap kegiatan pembelajaran. Data sebagai informasi umumnya telah disediakan dalam bentuk pilihan ganda, yang harus dipilih oleh siswa.

e) Daftar cek (*Checklist*) dan Daftar Bertingkat (*Rating Scale*)

Bila kita melakukan tes secara tertulis dan secara lisan, maka kita hanya mengukur kemampuan siswa dalam daerah kognitif saja. Sistem tes tertulis (*pencil and paper test*) seperti itu tidaklah mungkin dapat mengungkapkan kemampuan siswa dalam hal keterampilan. Perubahan tingkah laku dalam hal sikap, minat, kebiasaan, dan penyesuaian diri perlu mendapat perhatian yang tak dapat diungkapkan hanya dengan tes lisan dan tulisan. Oleh karena itu perlu tes lain, yaitu tes perbuatan. Yang dimaksud dengan daftar cek adalah sederetan pertanyaan atau pernyataan yang dijawab oleh responden dengan membubuhkan tanda cek (\checkmark) pada tempat yang telah disediakan. Sedangkan skala bertingkat adalah sejenis daftar cek dengan kemungkinan jawaban terurut menurut tingkatan atau hierarki.

Contoh:

Tabel 7.1 Contoh Daftar Skala Bertingkat

Daftar skala bertingkat untuk kegiatan mengukur tinggi pohon dengan menggunakan klinometer.

Nama Siswa : _____
Kelas : _____

Kriteria Evaluasi	SB	B	LC	C	K	SK
Jenis Kegiatan						
Kecermatan melakukan pengukuran jarak						
Kecermatan mengukur sudut elevasi						
Kecermatan menggambar sketsa hasil pengamatan						
Ketepatan menggunakan konsep matematika						
Kecermatan melakukan perhitungan						

Keterangan: SB = sangat baik
 B = baik
 LC = lebih dari cukup
 C = cukup
 K = kurang
 SK = sangat kurang

2. Teknik Tes

Tes adalah suatu alat atau prosedur yang sistematis dan obyektif untuk memperoleh data atau keterangan tentang seseorang, dengan cara tepat dan cepat.

Tes matematika adalah alat pengumpul informasi tentang hasil belajar matematika.

Teknik tes atau cara melaksanakan tes dapat digolongkan ke dalam 3 cara, yaitu

- i) Tes Tertulis,
- ii) Tes Lisan, dan
- iii) Tes Perbuatan

Ketiga macam teknik tes tersebut perbedaannya dititikberatkan pada segi cara menjawabnya, bukan dari cara menyajikan atau memberikan tes itu. Jadi orientasinya adalah tes, bukan instrumen tes atau tester.

Dalam tes tertulis testi menjawab tes tersebut secara tertulis pada lembar jawaban. Tes tertulis sangat bermanfaat untuk mengetahui kemahiran testi dalam teknik menulis yang benar, menyusun kalimat menurut kaidah bahasa yang baik

dan benar secara efisien, mengungkapkan buah pikiran melalui bahasa tulisan dengan kata-kata sendiri.

Dalam tes lisan, jawaban yang diberikan oleh testi dalam bentuk ungkapan lisan. Instrumen yang digunakan bisa saja disajikan dalam bentuk tulisan bisa pula dalam bentuk lisan. Pada umumnya tes lisan berbentuk tanya jawab langsung secara lisan antara tester dengan testi. Tes lisan ini sangat berguna bagi siswa untuk melatih diri dalam mengungkapkan pendapat atau buah pikirannya secara lisan dan mengembangkan kemampuan berbicara.

Tes perbuatan menuntut testi untuk melakukan perbuatan tertentu, tidak cukup hanya dengan menjawab tes tersebut. Tes perbuatan diberikan dalam bentuk tugas atau latihan yang harus diselesaikan secara individual atau kelompok. Dalam pembelajaran matematika, tes perbuatan bisa berupa memperagakan apakah suatu bangun datar merupakan jaring-jaring kubus atau bukan, menggambarkan suatu bangun ruang dan menunjukkan semua bidang diagonal serta diagonal bidang, membuat lukisan dengan menggunakan jangka, mistar, dan busur derajat, dan sebagainya.

B. Tipe dan Bentuk Tes

Tes tertulis menurut tipenya dikelompokkan menjadi 2 macam, yaitu tes tipe uraian dan tes tipe objektif.

1. Tes Tipe Uraian

Bentuk soal tes tipe uraian menuntut siswa untuk menyusun jawaban secara terurai. Jawaban tidak cukup hanya dengan satu atau dua kata saja, tetapi memerlukan uraian yang lengkap dan jelas. Selain harus menguasai materi tes, siswa dituntut untuk bisa mengungkapkannya dalam bahasa tulisan dengan baik.

Penyajian soal tipe uraian mempunyai beberapa kelebihan, yaitu:

- a) Pembuatan soal bentuk uraian relatif lebih mudah dan bisa dibuat dalam waktu yang tidak terlalu lama. Hal ini disebabkan karena jumlah soalnya tidak terlalu banyak.
- b) Karena dalam menjawab soal bentuk uraian siswa dituntut untuk menjawabnya secara rinci, maka proses berpikir, ketelitian, sistematika penyusunan dapat dievaluasi.
- c) Proses pengerjaan tes akan menimbulkan kreativitas dan aktivitas positif siswa, karena tes tersebut menuntut siswa agar berpikir secara sistematis, menyampaikan pendapat dan argumentasi, mengaitkan fakta-fakta yang relevan.

Di samping kelebihan yang dimiliki soal bentuk uraian, ia tidak luput dari kelemahan. Kelemahan soal bentuk uraian antara lain:

- a) Ruang lingkup materi yang disajikan dalam bentuk uraian kurang menyeluruh. Hal ini disebabkan waktu yang diperlukan untuk menyelesaikan setiap butir soal cukup banyak, sehingga butir soal yang disajikan sedikit.
- b) Soal tipe uraian, dalam pemeriksaan dan pemberian nilai akhir seringkali dipengaruhi faktor subyektifitas dari pemeriksa, sehingga nilai akhir ada kemungkinan bias, kurang mencerminkan kemampuan sebenarnya.
- c) Pemeriksaan jawaban soal bentuk uraian ini tidak bisa dilakukan oleh sembarang orang, tetapi harus diperiksa oleh orang yang benar-benar ahli dalam bidangnya. Bila pemeriksa kurang mengetahui pokok persoalan yang diujikan, akan mengakibatkan hasil pemeriksaan dapat merugikan siswa.

- d) Periksa jawaban tes bentuk uraian cukup rumit sehingga memerlukan waktu yang cukup banyak. Tiap siswa akan memberikan uraian yang berlainan dan bermacam-macam, apalagi jika persoalannya divergen.

2. Tes Tipe Objektif

Istilah objektif adalah tidak adanya faktor lain yang mempengaruhi proses pemeriksaan, jadi benar-benar murni hasil pekerjaan siswa.

Istilah lain dari tes tipe objektif adalah tes dengan jawaban singkat (*short answer test*). Tes tipe objektif dewasa ini seringkali digunakan karena memiliki banyak kelebihan.

Kelebihan atau keunggulan disajikan soal tipe objektif antara lain:

- a) Proses dan hasil pemeriksaan bersifat objektif sehingga hasilnya sesuai dengan kenyataan sebenarnya.
- b) Ruang lingkup materi yang diujikan lebih menyeluruh sehingga cukup representatif mewakili materi yang telah dipelajari siswa.
- c) Pemeriksaan tidak harus oleh penyusun soal itu, namun dapat dilakukan oleh orang lain asalkan sudah memahami patokan dalam hal pedoman pemeriksaannya. Komputer pun dapat dimanfaatkan untuk proses pemeriksaan dan pengolahan nilainya.
- d) Jawaban yang benar sudah tertentu dan pasti. Siswa yakin akan jawabannya apabila ia sudah mempelajari bahan yang diujikan, sehingga tidak timbul keraguan pada dirinya dalam hal menentukan jawaban yang tepat.
- e) Pemeriksaan dapat dilakukan dengan mudah dan cepat.
- f) Ketidakmampuan testi dalam bagian-bagian tertentu pada sebuah konsep atau topik lebih mudah dikenali secara langsung dari jawaban butir soal yang salah.

Tes tipe objektif memiliki kelemahan-kelemahan, diantaranya adalah:

- a) Proses berpikir siswa tidak dapat dievaluasi. Kita tidak tahu apakah siswa bekerja dengan benar, sesuai dengan yang diperintahkan.
- b) Kesempatan testi untuk menerka-nerka cukup besar, sehingga siswa yang tidak belajar pun atau tidak menguasai materi dengan baik, mungkin saja dapat menjawab dengan tepat.
- c) Tes tipe objektif kurang mampu memberikan gambaran sampai sejauhmana daya analisis siswa dan mengemukakan pikiran serta gagasannya.
- d) Pembuatan tes tipe objektif bukan saja sulit namun membutuhkan waktu yang tidak sedikit pula.

Menurut bentuknya tes tipe objektif terdiri dari 2 macam, yaitu:

- 1) Bentuk Benar-Salah (*True-False*)

Tes bentuk Benar-Salah soalnya disajikan dalam bentuk pernyataan. Pernyataan tersebut mengandung nilai kebenaran Benar (B) atau Salah (S), tetapi tidak keduanya sekaligus.

- 2) Bentuk Pilihan Ganda (*Multiple Choice*)

Soal tipe objektif bentuk pilihan ganda pada saat ini mendapat perhatian dan seringkali digunakan dalam penilaian pendidikan. Hal ini dilakukan mengingat bentuk ini lebih banyak keunggulannya jika dibandingkan dengan soal tipe obyektif yang lainnya.

Untuk mempertinggi ketetapan penilaian perlu diperhatikan kaidah-kaidah sebagai berikut.

1. Pokok soal (stem) yang merupakan permasalahan harus dirumuskan secara jelas.
2. Perumusan pokok soal dan alternatif jawaban hendaknya merupakan pernyataan yang diperlukan saja.

3. Untuk setiap soal hanya ada satu jawaban yang benar atau yang paling benar.
4. Pada pokok soal sedapat mungkin dicegah perumusan pernyataan yang bersifat negatif.
5. Alternatif jawaban (option) sebaiknya logis dan pengecoh harus berfungsi.
6. Diusahakan agar tidak ada “petunjuk” untuk jawaban yang benar.
7. Diusahakan untuk mencegah penggunaan option terakhir yang berbunyi “Semua alternatif jawaban di atas benar” atau “Semua pilihan jawaban di atas salah”.
8. Diusahakan agar alternatif jawaban homogen, baik dari segi isi atau materi maupun panjang-pendeknya kalimat (pernyataan).
9. Apabila alternatif jawaban berbentuk angka (bilangan), susunlah secara berurutan mulai angka terkecil di atas dan yang terbesar di bawah.
10. Di dalam pokok soal diusahakan tidak menggunakan ungkapan atau kata-kata yang bersifat tidak tentu, seperti kebanyakan, seringkali, kadang-kadang, dan sejenisnya.
11. Diusahakan agar jawaban butir soal yang satu tidak tergantung dari jawaban butir soal yang lain.
12. Dalam merakit soal harus diusahakan agar jawaban yang benar (kunci jawaban) tidak terpola, tetapi letaknya tersebar dan ditentukan secara random.
13. Penulisan alternatif jawaban diusahakan tersusun dari atas ke bawah, tidak ke samping.

C. Kualitas Alat Penilaian

Ruseffendi (1980) menyatakan bahwa dalam suatu kelompok individu (siswa) yang tidak dipilih secara khusus memiliki karakteristik tertentu yang frekuensinya berdistribusi normal. Begitu pula kepandaiannya dalam suatu mata pelajaran tertentu. Dengan demikian suatu alat penilaian yang baik akan mencerminkan kemampuan sebenarnya dari testi yang dinilai dan bisa membedakan siswa yang pandai (di atas rata-rata), siswa yang kemampuannya sedang (pada kelompok rata-rata), dan siswa yang kemampuannya kurang (di bawah rata-rata), sehingga penyebaran skor atau nilai tersebut berdistribusi normal.

Untuk mendapatkan hasil penilaian yang baik tentunya diperlukan alat penilaian yang kualitasnya baik pula di samping faktor lain yang dapat mempengaruhinya. Misalnya pelaksanaan penilaian (pengawasan), kondisi tester (pembuat dan pemeriksa hasil tes), dan keadaan lingkungan. Untuk mendapatkan hasil penilaian yang kualitasnya baik perlu diperhatikan beberapa kriteria yang harus dipenuhi. Alat penilaian yang baik dapat ditinjau dari hal-hal berikut ini:

- a) Validitas,
- b) Reliabilitas,
- c) Objektivitas,
- d) Praktikabilitas,
- e) Derajat Kesukaran,
- f) Daya Pembeda,
- g) Efektivitas Option, dan
- h) Efisiensi.

1. Validitas

Suatu alat penilaian disebut valid (absah atau sah) apabila alat tersebut mampu menilai apa yang seharusnya dinilai. Dengan demikian suatu alat penilaian disebut valid jika ia dapat menilai dengan tepat sesuatu yang dinilai itu.

Alat evaluasi yang valid untuk suatu tujuan tertentu belum tentu valid untuk tujuan (karakteristik) lain. Dengan kata lain validitas (keabsahan, ketepatan) dari suatu alat evaluasi harus ditinjau dari karakteristik tertentu. Oleh karena itu untuk menentukan validitas suatu alat evaluasi hendaklah dilihat dari berbagai aspek, diantaranya validitas isi, validitas muka (luar), validitas konstruksi (psikologis), validitas ramal, dan validitas banding.

Cara menentukan tingkat (indeks) validitas kriterium ialah dengan menghitung koefisien korelasi antara alat penilaian yang akan diketahui validitasnya dengan alat ukur lain yang telah dilaksanakan dan diasumsikan telah memiliki validitas yang tinggi (baik), sehingga hasil penilaian yang digunakan sebagai kriterium itu telah mencerminkan kemampuan siswa sebenarnya. Makin tinggi koefisien korelasinya makin tinggi pula validitas alat ukur tadi.

Banyak faktor yang dapat mempengaruhi hasil suatu evaluasi sehingga menjadi bias, menyimpang dari keadaan yang sebenarnya untuk suatu penggunaan yang dimaksudkan. Beberapa diantaranya adalah berasal dari dalam alat penilaian itu sendiri. Pemeriksaan secara cermat tentang pokok-pokok uji (materi) akan menunjukkan apakah alat penilaian itu dapat mengukur isi materi pelajaran dan fungsi-fungsi intelektual yang ingin diketahui.

Dalam hubungannya dengan kegiatan belajar mengajar matematika, faktor-faktor ini akan dapat mengurangi fungsi pokok uji sesuai dengan yang diharapkan sehingga bisa merendahkan validitas alat evaluasi tersebut.

- a) Petunjuk yang tidak jelas
- b) Perbendaharaan kata dan struktur kalimat yang sukar
- c) Penyusunan soal yang kurang baik
- d) Kekaburan
- e) Derajat kesukaran soal yang tidak cocok
- f) Materi tes tidak representatif

- g) Pengaturan soal yang kurang tepat
- h) Pola jawaban yang dapat diidentifikasi

Jika dirinci lebih lanjut, faktor-faktor yang dapat mempengaruhi kadar validitas alat evaluasi tersebut akan lebih banyak lagi. Prinsip yang penting agar kita terhindar dari pembuatan soal (matematika) yang validitasnya rendah, secara teoritik pahami konsep-konsep yang berkenaan dengan pengertian dan macam-macam validitas, secara praktek cara pembuatan soal tipe subyektif dan tipe obyektif harus pula dipahami

2. *Reliabilitas*

Suatu mistar terbuat dari kayu dipakai untuk mengukur panjang sebuah pensil, dilihat pada skalanya menunjukkan angka 30 cm. Mistar tersebut digunakan oleh orang lain untuk mengukur panjang pensil yang sama dan hasilnya tetap, yaitu 30 cm. Begitu pula jika pengukuran itu dilakukan pada waktu dan tempat yang berlainan, hasilnya tetap sama. Untuk kondisi tersebut, dikatakan bahwa mistar tersebut sebagai alat ukur panjang yang reliabel.

Dari contoh di atas, reliabilitas suatu alat ukur atau alat penilaian dimaksudkan sebagai suatu alat yang memberikan hasil yang tetap sama (konsisten, ajeg). Hasil pengukuran itu harus tetap sama (relatif sama) jika pengukurannya diberikan pada subjek yang sama meskipun dilakukan pada orang yang berbeda, waktu yang berbeda, dan tempat yang berbeda pula.

Berkenaan dengan penilaian, suatu alat penilaian (tes dan non tes) disebut reliabel jika hasil penilaian tersebut relatif tetap jika digunakan untuk subjek yang sama. Istilah relatif tetap di sini dimaksudkan tidak tepat sama, tetapi mengalami perubahan yang tak berarti (tidak signifikan) dan bisa diabaikan.

Andaikan suatu alat tes matematika diberikan kepada sekelompok siswa, hasil tes tersebut untuk setiap siswa relatif tetap (jika ada perubahan tidak mencolok)

sehingga rata-rata hitungnya (rerata, mean) tidak berbeda signifikan, untuk alat tersebut bisa dikatakan reliabel. Tetapi jika terjadi banyak siswa yang awalnya mendapat skor tinggi, pada tes berikutnya (dengan alat yang sama) mendapat skor rendah, atau sebaliknya, dikatakan bahwa alat ukur tersebut tidak reliabel.

Untuk mengestimasi reliabilitas suatu alat penilaian (tes dan non tes) ada tiga cara yang paling banyak dipergunakan, yaitu: Tes Tunggal (single test), Tes Ulang (test retest), dan Tes Ekuivalen (alternate test).

- a) Tes Tunggal adalah tes yang terdiri dari satu perangkat (satu set) yang dikenakan terhadap sekelompok subyek dalam satu kali pelaksanaan. Dengan demikian hasil evaluasi ini hanya terdapat satu kelompok data berupa skor hasil evaluasi itu. Dari kelompok data ini ditentukan reliabilitas alat evaluasi tersebut. Reliabilitas yang didasarkan atas tes tunggal dinamakan “internal consistency reliability”.
- b) Tes Ulang adalah tes yang terdiri dari seperangkat tes yang dikenakan terhadap sekelompok subyek dua kali. Reliabilitasnya dihitung dengan cara mengkorelasikan hasil evaluasi pertama dengan yang kedua.
- c) Tes Ekuivalen adalah tes yang terdiri dari dua perangkat dimana soal-soal pada perangkat pertama ekuivalen dengan soal-soal pada perangkat kedua. Pengertian ekuivalen di sini adalah soal-soal yang memuat konsep yang sama, tetapi soal tersebut tidak persis sama. Selain memuat konsep yang sama, tingkat kesukarannya pun harus sama. Untuk menentukan reliabilitasnya dihitung dengan cara mengkorelasikan hasil tes untuk soal perangkat pertama dengan hasil tes dari perangkat kedua.

Secara rinci, faktor-faktor yang dapat mempengaruhi reliabilitas alat penilaian adalah sebagai berikut ini.

a) Panjang Tes

Pada umumnya makin panjang tes (butir soal makin banyak), makin tinggi pula reliabilitasnya. Hal ini disebabkan karena tes yang butir soalnya lebih banyak akan memuat cukup banyak kemampuan kognitif siswa yang dapat diungkapkan.

b) Kondisi Testi

Suatu tes yang dicobakan kepada kelompok testi yang beraneka ragam kemampuannya akan menghasilkan skor yang heterogen, sehingga varians skor yang diperoleh akan besar. Hal ini akan mempengaruhi nilai koefisien reliabilitas sehingga menjadi lebih tinggi. Begitu pula jika peserta tes lebih banyak akan mengakibatkan keaneka-ragaman yang lebih bervariasi yang berakibat langsung pada penyebaran skor.

c) Kesukaran Tes

Materi tes yang terlalu mudah atau terlalu sulit cenderung akan merendahkan reliabilitas. Hal ini disebabkan karena skor yang diperoleh siswa untuk soal yang terlalu sulit atau terlalu mudah berkelompok pada skor tinggi atau skor rendah, jadi sebaran skornya terbatas.

d) Pelaksanaan Tes

Faktor yang bersifat administratif dalam melaksanakan tes akan mempengaruhi hasil tes, sehingga secara langsung akan mempengaruhi pula derajat reliabilitas tes tersebut. Petunjuk mengerjakan tes yang jelas akan memperlancar testi dalam mengerjakan tes tersebut sehingga ia dapat berpikir tenang dalam mencurahkan segala kemampuannya. Suasana yang tertib dan aman, tetapi tidak terlalu ketat, juga akan mempengaruhi ketenangan siswa dalam berpikir.

Dari uraian di muka mengenai reliabilitas, tampak bahwa reliabilitas ini berhubungan dengan masalah kepercayaan atau keterandalan. Suatu tes yang memiliki reliabilitas yang tinggi jika ia dapat memberikan hasil yang relatif tetap. Kualitas alat penilaian yang berkenaan dengan validitas dan reliabilitas tersebut penting. Akan tetapi validitas lebih penting karena validitas mempengaruhi derajat reliabilitas, sebaliknya belum tentu. Dengan demikian reliabilitas alat penilaian itu sifatnya perlu sebagai akibat dari validitas.

3. *Daya Pembeda (Discriminating Power)*

Pengertian Daya Pembeda (DP) dari sebuah butir soal menyatakan seberapa jauh kemungkinan butir soal tersebut mampu membedakan antara testi yang mengetahui jawabannya dengan benar dengan testi yang tidak dapat menjawab soal tersebut (atau testi yang menjawab salah). Dengan perkataan lain daya pembeda suatu butir soal adalah kemampuan butir soal itu untuk membedakan antara testi (siswa) yang pandai atau berkemampuan tinggi dengan siswa yang bodoh. Pengertian tersebut didasarkan pada asumsi Galton bahwa suatu perangkat alat tes yang baik harus bisa membedakan antara siswa yang pandai, rata-rata, dan yang bodoh karena dalam suatu kelas biasanya terdiri dari ketiga kelompok tersebut.

4. *Indeks Kesukaran (Difficulty index)*

Sejalan dengan asumsi Galton mengenai kemampuan tertentu (karakteristik), dalam hal ini kemampuan matematika, dari sekelompok siswa yang dipilih random (acak) akan berdistribusi normal, maka hasil penilaian dari suatu perangkat tes yang baik akan menghasilkan skor atau nilai yang membentuk distribusi normal. Hal ini mempunyai implikasi bahwa soal yang baik akan menghasilkan skor yang berdistribusi normal pula, sehingga sejalan dengan distribusi yang telah diuraikan pada pembicaraan mengenai daya pembeda.

Jika soal tersebut terlalu sukar, maka frekuensi distribusi yang paling banyak terletak pada skor yang rendah sebagian besar mendapat nilai yang jelek.

Jika soal seperti ini seringkali diberikan akan mengakibatkan siswa menjadi putus asa. Hal ini bukan berarti soal tersebut tidak boleh sukar, karena jika sewaktu-waktu atau dalam seperangkat diberikan beberapa soal yang sukar akan melatih siswa untuk berpikir lebih tinggi.

Sebaliknya jika soal yang diberikan terlalu mudah, maka frekuensi distribusi yang paling banyak berada pada skor yang tinggi, karena sebagian besar siswa mendapat nilai baik.

5. *Efektifitas Option*

Kata lain dari option adalah alternatif jawaban atau kemungkinan jawaban yang harus dipilih. Dengan demikian arti dari kata option adalah kemungkinan jawaban yang disediakan pada butir soal (tes) tipe objektif bentuk pilihan ganda atau memasangkan untuk dipilih oleh peserta tes, sesuai dengan petunjuk yang diberikan.

Suatu option disebut efektif jika memenuhi fungsinya atau tujuan yang disajikan option tersebut tercapai. Hal ini berarti bahwa setiap option yang disajikan masing-masing mempunyai kemungkinan yang sama untuk dipilih, jika testi menjawab soal itu dengan menerka-nerka (spekulasi).

Option yang merupakan jawaban yang benar disebut option kunci (key option), sedangkan option lainnya disebut option pengecoh (distractor option).

Agar suatu option yang disajikan efektif harus diusahakan homogen (serupa), baik dari isi (materi), notasi, maupun panjang pendeknya kalimat pada option tersebut. Berdasarkan distribusi pilihan pada setiap option untuk siswa kelompok

atas dan kelompok bawah, dapat ditentukan option yang berfungsi efektif dan yang tidak. Kriteria option yang berfungsi secara efektif adalah:

a) Untuk Option Kunci

- i) Jumlah pemilih kelompok atas harus lebih banyak daripada jumlah pemilih kelompok bawah, yaitu siswa yang pandai lebih banyak yang menjawab benar daripada siswa yang bodoh.
- ii) Jumlah pemilih kelompok atas dan kelompok bawah lebih dari 0,25 tetapi tidak lebih dari 0,75 dari seluruh siswa pada kelompok atas dan kelompok bawah. Jika jumlah tersebut kurang dari 0,25 berarti sebagian besar testi kelompok atas dan kelompok bawah menjawab salah untuk soal tersebut. Soal itu dikategorikan sukar atau terlalu sukar. Sebaliknya jika jumlah tersebut lebih dari 0,75 soal itu termasuk kategori mudah atau terlalu mudah.

b) Untuk Option Pengecoh

- i) Jumlah pemilih kelompok atas lebih sedikit (kurang) daripada jumlah pemilih kelompok bawah. Hal ini berarti untuk jawaban yang salah siswa yang bodoh lebih banyak yang memilih daripada siswa yang pandai. Idealnya siswa pandai tidak memilih jawaban yang salah dan siswa yang bodoh memilihnya.
- ii) Jumlah pemilih kelompok atas dan kelompok bawah minimal sebanyak 0,25 dari seperdua jumlah option pengecoh kali jumlah kelompok atas dan kelompok bawah. Dirumuskan dalam formula matematika menjadi

$$JP_A + JP_B \geq 0,25 \times \frac{1}{2} \times \frac{1}{n-1} \times JS_A + JS_B$$

dengan JP_A = jumlah pemilih kelompok atas
 JP_B = jumlah pemilih kelompok bawah
 n = banyak option pengecoh
 JS_A = jumlah subyek pada kelompok atas, dan
 JS_B = jumlah subyek pada kelompok bawah

Ada pakar lain yang mengemukakan bahwa rumus di atas terlalu menyulitkan, ia berpendapat bahwa selain jumlah pemilih kelompok bawah harus lebih banyak daripada pemilih kelompok atas untuk option pengecoh, option pengecoh itu harus dipilih minimum oleh 5% peserta tes pada kedua kelompok siswa.

- iii) Jika peserta tes mengabaikan semua option (tidak memilih) disebut omit. Option disebut efektif jika omit ini jumlahnya tidak lebih dari 10% jumlah siswa pada kelompok atas dan kelompok bawah.

Agar uraian diatas lebih dipahami, perhatikan contoh berikut ini. Misalkan sebuah butir soal bentuk pilihan ganda dengan 5 option dijawab oleh testi kelompok atas dan kelompok bawah seperti tampak pada tabel di bawah ini.

Tabel 7.2 Sebaran Pemilih pada Suatu Butir Soal
Sebaran Pemilih pada Suatu Butir Soal

Soal No. x	Option					Omti (o)
Kelompok	a	b	c	d	e	
Atas	2	6	0	7	4	1
Bawah	9	3	0	2	4	2

Berdasarkan pada tabel di atas akan diuji efektifitas setiap option (termasuk omit) sebagai berikut.

- i) Untuk option (a) sebagai pengecoh berfungsi efektif, sebab
- Jumlah pemilih kelompok atas kurang dari jumlah pemilih kelompok bawah,
 - Jumlah pemilih kelompok atas dan kelompok bawah $2 + 9 = 11$, sedangkan nilai

$$0,25 \times \frac{1}{2 \times 3} \times 40 = 1,25$$

JPA = jumlah pemilih kelompok atas dan 5% dari 40 = 2,00 sehingga $11 > 2,00$

- ii) Untuk option (b) sebagai pengecoh tidak efektif, sebab jumlah pemilih kelompok atas lebih banyak daripada jumlah pemilih kelompok bawah. Salah satu syarat tidak dipenuhi, syarat lainnya tidak perlu diperiksa.
- iii) Untuk option (c) dan (e) tidak efektif karena jumlah pemilih kelompok atas dan kelompok kelompok bawah sama.
- iv) Untuk option (d) sebagai kunci jawaban, jumlah pemilih kelompok atas lebih banyak daripada jumlah pemilih kelompok bawah dan jumlah pemilih kedua kelompok itu sebanyak

$$\frac{7+2}{40} \times 100\% = 1,225$$

Nilai tersebut kurang dari 0,25. jadi option (d) kurang efektif. Soal tersebut terlalu sukar.

- v) Untuk omit masih dibawah toleransi karena jumlahnya tidak lebih dari 4 (10% dari 40).

Perlu ditambahkan bahwa option (c) kemungkinan tidak homogen sehingga tampak tampak sekali bahwa option itu salah. Siswa pada kedua kelompok tidak ada yang memilihnya.

Uraian mengenai daya pembeda, derajat kesukaran, dan efektifitas option dimuka lebih dikenal dengan istilah Analisis Butir Soal atau Analisis Item. Analisis butir soal bertujuan untuk mengidentifikasi soal-soal yang baik, kurang baik (cukup), atau jelek (buruk) sehingga hasilnya dapat diperoleh informasi tentang kualitas soal yang kita buat. Hal ini diperlukan unutm mengadakan perbaikan seperlunya, minimal kita (sebagai guru) dapat menginstropeksi diri terhadap kemampuan kita dalam membuat alat evaluasi. Dengan pengalaman dan pengetahuan yang terus digali dan dimanfaatkan dalam pelaksanaan kegiatan belajar mengajar disekolah diharapkan soal yang kita buat akan semakin baik

kualitasnya. Dengan perkataan lain, jika hasil evaluasi yang kita laksanakan kurang baik, salah satu cara introspeksi adalah dengan mengkaji kembali kualitas tes yang dibuat. Sebaliknya jika hasil evaluasi baik atau semuanya baik, jangan berbangga dulu. Periksalah kembali soal yang dibuat, mungkin saja soalnya terlalu mudah.

6. *Objektivitas*

Dalam pemeriksaan hasil tes, faktor subjektif pemeriksaan biasanya berperan, apalagi bila tes itu berbentuk uraian. Meskipun demikian kita dapat mengurangi kelemahan ini sedapat mungkin, seperti yang telah diuraikan pada bagian sebelumnya.

Sebuah tes hendaknya bersifat objektif. Hal ini maksudnya adalah hasil dari tes tersebut harus selalu sama, meskipun diperiksa oleh orang yang berlainan. Tentu saja, agar harapan tersebut terpenuhi tes yang kita buat harus mempunyai jawaban yang jelas, tidak kabur, jawabannya tentu dan tidak terlalu memberikan jawaban yang beraneka ragam. Meski divergen jawaban siswa akan berakibat penilaian kurang objektif. Itulah sebabnya mengapa pada pemeriksaan tes bentuk uraian hendaknya terlebih dahulu dibuat kunci jawabannya atau paling tidak pokok-pokok jawabannya. Langkah ini bukan saja sebagai pengecekan ketepatan kalimat dalam soal, banyaknya waktu penyelesaian yang dibutuhkan, tetapi juga sekaligus memperkirakan sampai sejauh mana lingkup batasan jawaban siswa yang akan muncul.

Besarnya skor yang diberikan kepada testi menunjukkan sampai sejauh mana tingkat penguasaan materi yang telah dimiliki siswa tersebut. Gambaran yang dinyatakan dengan skor ini hendaknya bersifat seobjektif mungkin. Karena itu tes yang memberikan nilai (skor) tersebut harus objektif dan benar-benar mengevaluasi kemampuan siswa secara tepat.

Jika dikaitkan dengan reliabilitas, objektivitas memberi tekanan pada ketetapan sistem pemberian skor sedangkan reliabilitas memberikan penekanan pada ketetapan

hasil. Jelas bahwa keduanya saling mempengaruhi atau saling ketergantungan, sistem pemberian skor mempunyai dampak terhadap ketetapan hasil.

Ada dua faktor yang penting yang bisa mempengaruhi derajat objektivitas tersebut di atas, yaitu :

a) Tipe Tes

Tes dengan tipe uraian akan lebih banyak mengurangi objektivitas daripada tes tipe objektif. Soal bentuk uraian memerlukan proses pemeriksaan yang cukup memakan waktu, karena jawabannya bisa panjang lebar sehingga setiap proses (langkah) pengerjaan testi diperiksa. Disinilah seringkali muncul faktor yang mengurangi objektivitas tersebut. Lain halnya untuk soal tipe objektif yang jawabannya pasti.

b) Penilaian/Pemeriksa

Faktor-faktor yang dapat mempengaruhi objektivitas dari penilai antara lain kesan penilai terhadap pribadi siswa, tulisan, bahasa, kerapian pekerjaan. Selain daripada itu kondisi penilai dan lingkungan bisa juga mempengaruhi.

7. Praktikabilitas

Tes yang baik harus bersifat praktis, dalam arti mudah dilaksanakan dan efisien dari segi biaya dan tenaga. Dalam penyusunan tes hendaknya biaya yang diperlukan tidak terlampaui tinggi, namun masih memenuhi persyaratan sebuah tes yang baik. Sebuah tes juga disebut praktis bila pemeriksaannya mudah dan dapat dianalisis dalam waktu relatif singkat.

D. Penilaian Hasil Belajar

Evaluasi hasil belajar matematika pada umumnya dilaksanakan dalam bentuk tes tertulis. Dalam penyusunan tes ini langkah utama yang harus ditempuh yaitu:

1. *Menyusun Kerangka*

Suatu tes untuk mengevaluasi hasil belajar disebut baik jika materi yang terkandung dalam butir-butir tes tersebut dapat mewakili seluruh materi yang telah dipelajari siswa. Jika materi tes yang disajikan hanya memuat sebagian kecil dari materi yang telah dipelajari siswa, atau sebaliknya memuat materi yang belum dipelajari (diajarkan guru), maka tes tersebut kurang baik.

Pada kondisi pertama akan muncul unsur spekulatif. Artinya jika siswa kebetulan mempelajari atau mendalami materi yang sesuai dengan materi soal yang disajikan, ia akan mendapat hasil yang baik. Padahal ia tidak menguasai bagian materi yang lainnya. Sebaliknya jika siswa secara tidak kebetulan mempelajari suatu materi dan tidak tersaji dalam soal, ia akan mendapat hasil yang jelek. Pada kondisi kedua, kemungkinan besar hasil penilaian akan jelek sebab siswa belum mempelajarinya apalagi memahaminya.

Untuk menghindari kedua hal tersebut di atas mendapatkan suatu perangkat tes yang representatif seyogyanya dilakukan analisis rasional. Artinya dengan melaksanakan analisis berdasarkan pikiran logis tentang materi-materi yang akan diteskan, tujuan pembelajaran, tipe dan bentuk tes, dan jenjang kognitif yang akan dicapai.

Analisis rasional tersebut dituangkan dalam bentuk “blue printe” atau “lay out” atau “kisi-kisi” yang berisi pokok-pokok uji yang akan disajikan dalam tes. Kisi-kisi adalah suatu acuan berbentuk rangka mengenai alokasi bahan, tipe bentuk tes, aspek intelektual, taraf kesukaran, jumlah soal dan persentasenya.

Kisi-kisi berfungsi seperti halnya sebuah peta tentang penyebaran butir soal, sehingga bahan, bentuk soal, aspek intelektual, taraf kesukaran, jumlah soal dan persentasenya dapat tersebar secara merata. Kisi-kisi yang baik akan dapat

menentukan keberhasilan belajar testi dengan tepat, sehingga hasil penilaian bisa mencerminkan kemampuan siswa sebenarnya.

2. Menulis Butir Soal

Setelah melengkapi kisi-kisi yang menggambarkan keseluruhan isi yang akan dibuat, pembuat tes melengkapinya dengan format penulisan soal seperti tampak di bawah ini.

Tabel 7.3 Format Penulisan Soal

FORMAT PENULISAN SOAL

BIDANG STUDI :
 JENJANG SEKOLAH :
 PERKIRAAN WAKTU :

No	No. Pokok Bahasan Sub Pokok Bahasan	Bahan Kelas	Jenjang Kognitif	TP	Nomor Soal	Soal	Tingkat Kesukaran	Bentuk Soal	Kunci

KETERANGAN: JENJANG KOGNITIF TINGKAT KESUKARAN BENTUK SOAL

C1 = Pengetahuan	Md = Mudah	A = Melengkapi pilihan
C2 = Pemahaman	Sd = Sedang	B = Hubungan antar hal
C3 = Penerapan	Sk = Sukar	C = Tinjauan kasus
C4 = Analisis		D = Analisis pilihan ganda
C5 = Sistesis		(pilihan ganda kompleks)
C6 = Evalasi		E = Membaca diagram/grafik

Format penulisan soal tersebut diatas berguna untuk menyesuaikan butir soal yang dibuat dengan rumusan Tujuan Pembelajaran (TP) telah disusun sebelumnya,

jenjang kognitif yang akan dicapai, tingkat kesukaran, bentuk soal, dan aspek lainnya. Karena pengisian format ini merupakan penjabaran dari kisi-kisi yang dilengkapi dengan rumus tujuan pembelajaran dan soal secara lengkap (dengan kunci jawabannya), maka diperkirakan cara pengisiannya tidak akan sulit.

E. Penilaian Sikap

Telah diutarakan di muka bahwa pendidikan, termasuk di dalamnya pengajaran, selain daripada daerah kognitif dan psikomotorik adalah daerah afektif. Di antaranya adalah yang berkenaan dengan sikap (attitude) sebagai manifestasi dari minat, motivasi, perasaan, dan semacamnya. Oleh karena itu evaluasi untuk bidang studi matematika tidak hanya terbatas pada bidang afektif khususnya mengenai sikap siswa.

Pembentukan daerah afektif sebagai hasil belajar matematika relatif lebih lambat daripada pembentukan daerah kognitif dan psikomotorik, karena perubahan daerah afektif (baca sikap) memerlukan waktu yang lebih lama dan merupakan akibat dari pembentukan pada daerah kognitif dan psikomotorik. Gagne (1974) menyebutkan bahwa daerah afektif ini sebagai obyek matematika yang sifatnya tidak langsung, sedangkan daerah kognitif dan psikomotorik sebagai obyek langsung, yang dapat secara langsung dimiliki dalam diri siswa setelah kegiatan belajar mengajar berlangsung.

Dalam pelajaran matematika seringkali pembentukan sikap seseorang terhadap matematika sebagai akibat dari pembentukan daerah kognitifnya, meskipun kadang-kadang terjadi sebaliknya. Misalnya seorang siswa yang seringkali merasa mampu untuk mengerjakan soal-soal matematika, ia menjadi senang bahkan mengharapkan lebih banyak lagi belajar matematika. Sebaliknya jika ia sering tidak mampu akan mengakibatkan rasa segan atau bahkan menakutinya. Seorang guru matematika yang mengajar dengan baik sehingga mudah dimengerti oleh siswa, ramah kepada setiap siswa, tepat waktu jika ia

mengajar, memberikan nilai secara objektif, bisa membentuk sikap siswa menjadi positif terhadap matematika. Sebaliknya jika guru tersebut kurang jelas menerangkannya berperilaku tidak ramah, sering terlambat atau bolos dalam mengajar, memberikan nilai tidak objektif, bisa menimbulkan sikap siswa yang negatif terhadap matematika.

Dari contoh-contoh di atas tampak bahwa pembentuk sikap seseorang terhadap matematika memerlukan proses yang cukup panjang, sebagai akumulasi dari pengalaman-pengalaman dalam belajar, melalui proses kognitif dan psikomotorik. Dari kondisi ini pelaksanaan penilaian yang berkenaan dengan sikap seseorang terhadap matematika, tidak dapat dilakukan setiap saat seperti halnya penilaian untuk daerah kognitif dan psikomotorik. Oleh karena itu jika kita akan melakukan penilaian terhadap sikap ini sebaiknya dilakukan pada akhir semester atau pada akhir tahun ajaran. Dalam selang waktu tersebut akumulasi pengalaman belajar siswa dalam matematika sudah cukup banyak, sehingga dapat diharapkan cukup banyak pula komponen sikap siswa yang dapat dievaluasi.

Dengan melaksanakan penilaian sikap terhadap matematika, ada beberapa hal yang bisa diperoleh guru, antara lain:

- 1) memperoleh balikan (feed back) sebagai dasar untuk memperbaiki proses belajar mengajar dan program pengajaran remedial;
- 2) memperbaiki perilaku diri sendiri (guru) maupun siswa;
- 3) memperbaiki atau menambah fasilitas belajar yang masih kurang; dan
- 4) mengetahui latar belakang kehidupan siswa yang berkenaan dengan aktivitas belajarnya.

Pengertian sikap itu sendiri berkenaan dengan perasaan (kata hati) dan manifestasinya berupa perilaku yang bersifat positif (favorable) atau negatif (unfavorable) terhadap objek tertentu. Objek tersebut bisa diri sendiri, orang lain, kegiatan, keadaan,

lingkungan, dan lain sebagainya. Thurstone mendefinisikan sikap sebagai derajat perasaan positif atau negatif terhadap suatu objek yang bersifat psikologis. Sikap positif bisa diartikan sebagai menyukai, menyenangkan, menunjang, atau memihak terhadap objek tadi. Sedangkan sikap negatif bisa diartikan sebaliknya.

Pada bagian terdahulu telah dikemukakan bahwa untuk mengevaluasi sikap dapat dilakukan dengan wawancara, observasi, dan angket dengan segala kekurangan dan kelebihanannya. Berikut ini akan dibicarakan secara khusus mengenai angket skala sikap yang dikemukakan oleh Likert, Thurstone, Guttman, dan Diferensial Semantik.

Beberapa kriteria yang harus diperhatikan dalam menyusun pernyataan untuk mengungkapkan sikap seseorang terhadap suatu objek, sehingga bisa memperoleh jawaban yang benar-benar diharapkan, setiap pernyataan yang diajukan hendaknya tidak meragukan dan menimbulkan tafsiran lain. Edwards (1969) mengemukakan kriteria tersebut adalah:

- 1) Arahkanlah pernyataan yang dikemukakan yang berhubungan dengan kondisi aktual, tidak mengungkapkan hal-hal yang sifatnya sudah basi;
- 2) Hindarkan pernyataan yang merupakan suatu fakta atau dapat dianggap fakta sehingga tidak mengungkapkan sikap;
- 3) Pernyataan hendaknya tidak mengandung arti yang bermacam-macam, sehingga membingungkan;
- 4) Pernyataan hendaknya relevan dengan aspek psikologi yang akan dievaluasi;
- 5) Hindarkan pernyataan yang akan disetujui oleh semua responden, atau sebaliknya;
- 6) Pilihlah pernyataan-pernyataan yang dianggap mewakili isi keseluruhan skala sikap yang akan dievaluasi;
- 7) Susunlah pernyataan dalam bahasa yang sederhana, jelas, dan langsung pada pokok masalah;
- 8) Pernyataan yang disajikan usahakan singkat;

- 9) Tiap pernyataan diusahakan hanya memuat satu pokok pikiran;
- 10) Hindarkanlah kata-kata yang sifatnya universal, seperti semua, setiap, selalu, tak satu pun, atau tak pernah sehingga tidak menimbulkan keraguan;
- 11) Hati-hati menuliskan pernyataan dengan menggunakan kata “hanya”, dan sejenisnya;
- 12) Usahakan menyusun pernyataan dengan kalimat tunggal;
- 13) Gunakanlah kata-kata yang sudah banyak dikenal (familiar) agar tidak menyesatkan; dan
- 14) Hindarkan menggunakan pernyataan yang mengandung kata negatif lebih dari satu kali.

Keabsahan (validitas) suatu skala sikap biasanya dilakukan dengan cara mempertimbangkan (judgement) nya secara teoritik. Untuk validitas isi, bisa diperiksa apakah pernyataan-pernyataan yang disajikan telah memiliki komponen-komponen sikap yang akan dievaluasi. Untuk validitas tampilan (face validity), apakah butir-butir pernyataan telah cukup jelas, mudah dipahami, baik dari segi bahasa maupun dari segi teknis lainnya. Untuk validitas kontruksi psikologis, apakah seluruh pernyataan yang disajikan telah sesuai dengan tujuan penilaian sikap tersebut dan mampu mengungkapkannya.

1. Skala Likert

Dalam skala Likert, responden (subjek) diminta untuk membaca dengan seksama setiap pernyataan yang disajikan, kemudian ia diminta untuk menilai pernyataan-pernyataan itu. Penilaian terhadap pernyataan-pernyataan itu sifatnya subjektif, tergantung dari kondisi sikap masing-masing individu. Faktor dari luar yang bisa mempengaruhi diusahakan tidak ada.

Derajat penilaian siswa terhadap suatu pernyataan terbagi ke dalam 5 (lima) kategori yang tersusun secara bertingkat, mulai dari Sangat Tidak Setuju (STS), Tidak Setuju (TS), Netral (N), Setuju (S), dan Sangat Setuju (SS) atau bisa pula disusun sebaliknya.

Dalam menganalisis hasil angket, skala kualitatif di atas ditransfer ke dalam skala kuantitatif. Untuk pernyataan yang bersifat positif (favorable) kategori SS diberi skor tertinggi, makin menuju ke STS skor yang diberikan berangsur-angsur menurun. Sebaliknya untuk pernyataan yang bersifat negatif (unfavorable) untuk kategori SS diberi skor terendah, makin menuju ke STS skor yang diberikan berangsur-angsur makin tinggi.

Jika secara cermat kita teliti, biasanya setiap pernyataan yang disajikan dalam skala Likert ini, masing-masing memiliki kontribusi yang berlainan terhadap sikap individu tersebut. Sehingga sudah seharusnya pemberian bobot (skor) untuk setiap pernyataan berlainan pula. Pemberian skor untuk setiap pernyataan dengan memperhatikan hal tersebut di atas tidak sembarang bisa ditentukan, melainkan harus melalui uji coba terlebih dahulu. Bobot untuk setiap pernyataan tersebut sangat tergantung dari hasil uji coba yang dilakukan.

Pembicaraan mengenai bobot untuk setiap pernyataan untuk angket yang dibuat akan ditentukan secara kasar saja, dengan mengasumsikan bahwa setiap pernyataan yang disajikan memiliki kontribusi yang sama terhadap sikap individu secara keseluruhan. Hal ini dimaksudkan agar pembuatan angket skala sikap model Likert ini bisa mudah dipahami dan dilaksanakan.

Pembobotan yang paling sering dipakai dalam mentransfer skala kualitatif ke dalam skala kuantitatif adalah:

Untuk pernyataan favorable, jawaban

SS	diberi skor 5
S	diberi skor 4
N	diberi skor 3
TS	diberi skor 2
STS	diberi skor 1

Sebaliknya untuk pernyataan unfavourable, jawaban

SS	diberi skor 1
S	diberi skor 2
N	diberi skor 3
TS	diberi skor 4
STS	diberi skor 5

Misalkan untuk pernyataan favorable “Matematika melatih saya untuk berpikir kritis, sistematis, serta menumbuhkan sifat kreatif”. Responden yang memilih SS diberi skor 5, jika memilih S diberi skor 4, jika memilih N diberi skor 3, jika memilih TS diberi skor 2, jika memilih STS diberi skor 1. Untuk pernyataan unfavorable, misalkan pernyataan “Matematika adalah pelajaran yang membosankan, karena hanya membicarakan bilangan saja”. Responden yang memilih SS diberi skor 1, jika memilih S diberi skor 2, jika memilih N diberi skor 3, jika memilih TS diberi skor 4, jika memilih STS diberi skor 5. Hal ini logis, sebab pernyataan yang tidak mendukung (negatif), sikap yang negatif harus diberi skor tinggi karena sikap itu menyatakan sikap yang positif.

Hal yang perlu dicatat bahwa bobot (skor) untuk setiap pernyataan itu tidak disajikan dalam lembaran angket, tetapi hanya untuk keperluan pengolahan data saja. Jika disajikan kemungkinan unsur obyektivitas responden dapat dipengaruhi oleh adanya skor tersebut. Disamping itu, agar jawaban siswa tidak spekulatif, usahakan penyajian butir pernyataan yang favorable dan yang unfavorable tidak terpola.

Setelah angket terkumpul dan diolah dengan menggunakan cara seperti di atas, seorang subyek dapat digolongkan pada kelompok responden yang memiliki sikap positif atau bersikap negatif. Penggolongan dapat dilakukan dengan membandingkan

skor subjek dengan jumlah skor alternatif jawaban netral dari semua butir pernyataan. Jika skor subjek lebih besar daripada jumlah skor netral, maka subjek tersebut mempunyai sikap positif. Sebaliknya jika skor subjektif kurang dari jumlah skor netral maka subyek itu mempunyai sikap negatif. Proses lain bisa dilakukan dengan menghitung rerata skor subyek. Jika nilainya lebih besar daripada 3 (rerata skor untuk jawaban netral) ia bersikap positif. Sebaliknya jika reratanya kurang dari 3, ia bersikap negatif. Rerata skor subyek makin mendekati 5, sikap siswa makin positif. Sebaliknya jika mendekati 1, sikap siswa makin negatif.

Jika kita tidak menghendaki jawaban responden yang ragu-ragu (netral), dengan kata lain siswa dituntut untuk menjawab angket secara konsekuen, bisa saja alternatif jawaban yang disajikan menjadi 4 buah, tanpa alternatif N (netral). Dengan demikian pemberian skor untuk setiap pernyataan adalah 1 (STS), 2 (TS), 4 (S), 5 (SS) untuk pernyataan favorable, sebaliknya diberi skor 1 (SS), 2 (S), 4 (TS), 5 (STS) untuk pernyataan favorable.

2. *Skala Thurstone*

Selain skala Likert, skala lain yang banyak dipergunakan untuk mengungkapkan sikap individu adalah Skala Thurstone. Skala Thurstone memuat jumlah pernyataan yang harus dipilih oleh responden, yang masing-masing telah diberi skor (bobot) tertentu.

Pada skala Likert pembuat angket bisa saja mengasumsikan bahwa kontribusi setiap pernyataan terhadap sikap dari seorang individu sama, tetapi dalam skala Thurstone justru hal ini dipentingkan. Pernyataan yang kontribusinya terhadap sikap lebih tinggi diberi skor lebih besar, sebaliknya pernyataan yang kontribusinya lebih rendah diberi skor lebih kecil. Dengan demikian dalam skala ini pernyataan-pernyataan yang disajikan tidak dipilah ke dalam pernyataan yang favorable dan unfavorable.

Cara penentuan skor untuk setiap pernyataan yang disajikan dipertimbangkan oleh pembuat angket, atau (sebaiknya) meminta pertimbangan beberapa ahli agar lebih objektif. Hasil pertimbangan tersebut, berupa skor untuk masing-masing pernyataan, kemudian dicari reratanya. Seperti halnya pada skala Likert, skala Thurstone ini pun skor untuk setiap pernyataan itu tidak disajikan.

Pernyataan yang disajikan dengan menggunakan skala Thurstone ini biasanya dibuat sebanyak 9 (sembilan) butir. Responden diminta memilih 5 pernyataan yang paling sesuai dengan sikapnya. Skor yang diperoleh responden didapat dengan menghitung rerata skor dari pernyataan yang dipilih tadi. Interpretasi dari hasil angket bahwa responden bersikap positif atau bersikap negatif tergantung dari interval skor yang dipergunakan.

Berikut ini disajikan contoh angket yang disajikan dengan menggunakan model skala Thurstone.

Petunjuk: Pilihlah 5 (lima) buah pernyataan yang paling sesuai dengan sikap Anda terhadap pelajaran matematika, dengan cara membubuhkan tanda cek (\surd) di depan nomor pernyataan di dalam tanda kurung.

- () 1. Saya senang belajar matematika.
- () 2. Matematika adalah segalanya buat saya.
- () 3. Jika ada pelajaran kosong, saya lebih suka belajar matematika.
- () 4. Belajar matematika menumbuhkan sikap kritis dan kreatif.
- () 5. Saya merasa pasrah terhadap ketidak-berhasilan saya dalam matematika.
- () 6. Penguasaan matematika akan sangat membantu dalam mempelajari bidang studi lain.
- () 7. Saya selalu ingin meningkatkan pengetahuan dan kemampuan saya dalam matematika.

- () 8. Pelajaran matematika sangat menjemukan.
- () 9. Saya merasa terasing jika ada teman membicarakan matematika.

Misalkan pembuat angket menentukan bahwa skor yang akan dipakai untuk pernyataan yang kontribusinya paling tinggi adalah 9 dan untuk yang paling rendah diberi skor 1, sehingga skor tengahnya sama dengan 5. Hasil pertimbangannya, ia menyatakan bahwa pernyataan yang paling tinggi kontribusinya terhadap sikap positif untuk matematika adalah pernyataan nomor 2 sehingga ia memberi bobot skor 9. Agar hasil pertimbangan itu lebih objektif, ia meminta bantuan kepada teman seprofesinya yang dianggap mampu atau lebih mampu daripada dirinya sendiri. Misalkan ada 4 orang yang diminta pertimbangan itu, hasil pertimbangan untuk butir nomor 2 dari keempat orang itu masing-masing 8, 8, 9 dan 9. Dengan demikian skor untuk butir soal nomor 2 itu adalah

$$\frac{9+8+8+9+9}{5} = 8,6$$

Untuk butir nomor 8 pembuat angket memberi skor 2 karena ia menganggap kontribusinya rendah terhadap sikap siswa dalam matematika. Keempat teman lainnya masing-masing memberi skor 3, 4, 1, 2 sehingga skor untuk butir nomor 8 adalah

$$\frac{2+3+4+1+2}{5} = 2,4$$

Begitulah seterusnya cara pemberian skor untuk setiap butir pernyataan.

Misalkan skor untuk setiap butir soal, berturut-turut dari butir soal nomor 1 sampai dengan nomor 9 adalah sebagai berikut :

$$9,0; 8,6; 8,2; 7,6; 4,5; 6,0; 7,6; 2,4; 4,0; 5,3$$

Setelah angket diberikan kepada responden (siswa), misalkan untuk subjek A memilih butir-butir nomor 1, 4, 6, 7 dan 10. Rerata skor dari subyek A adalah

$$\frac{9,0 + 7,6 + 6,0 + 7,6 + 5,3}{5} = 7,1$$

Ini berarti sikap A terhadap matematika positif, karena skornya lebih daripada skor tengah (= 5).

Dibandingkan dengan skala Likert, skala Thurstone hanya menyajikan butir pernyataan yang sedikit sehingga aspek sikap yang bisa diungkapkan relatif sedikit pula. Namun demikian skala Thurstone mempunyai kelebihan pada ketajaman pernyataan untuk mengungkapkan sikap tersebut, sehingga lebih sedikit kemungkinan responden untuk menjawab dengan cara menebak. Untuk mengurangi kelemahan di atas, di samping cara pemberian skor yang cukup rumit, untuk setiap aspek mengenai sikap bisa dibuat satu set (10 butir) pernyataan. Misalkan dari segi materi matematika, pelaksanaan kegiatan belajar mengajar, sistem evaluasi, sarana dan prasarana, masing-masing 10 butir pernyataan sehingga seluruh aspek sikap terhadap matematika bisa terungkap.

3. *Skala Guttman*

Pada skala sikap model Guttman, tingkat ketajaman kontribusi pernyataan terhadap sikap yang akan diungkapkan lebih jelas lagi, sebab jawaban terhadap pernyataan pertama disusul (dilacak) oleh pernyataan kedua, dan pernyataan kedua disusul lagi oleh pernyataan ketiga, dan seterusnya. Pernyataan berikutnya merupakan pelacak tentang jawaban pada pernyataan sebelumnya. Jadi setiap pernyataan yang disajikan saling terkait, tidak saling lepas satu sama lain.

Karena sikap dari pernyataan yang disajikan dalam skala Guttman demikian, maka ruang lingkup masalah (aspek) yang berkenaan dengan sikap seseorang menjadi sempit. Seringkali hanya satu aspek saja yang bisa diungkapkan. Untuk menanggulangi kelemahan ini, bisa dibuat beberapa set pernyataan seperti halnya pada skala Thurstone.

Jawaban yang harus diberikan pada skala Guttman dengan cara membubuhkan tanda cek (√) pada kolom “YA” atau “TIDAK”.

Pernyataan-pernyataan dalam skala Guttman disusun sedemikianrupa hingga jika responden menjawab YA pada butir pernyataan nomor 1, 2, dan 3 kemudian menjawab TIDAK pada butir nomor 4, untuk butir berikutnya ia menjawab TIDAK. Jadi diharapkan responden menjawab YA pada butir-butir awal, sekali ia menjawab TIDAK pada suatu butir pernyataan maka ia akan menjawab TIDAK pada butir selanjutnya. Dengan demikian penilaian cukup dengan mempertimbangkan atau menghitung batas jawaban YA yang diberikan responden. Interpretasi untuk menentukan sikap responden bisa dilakukan dengan cara menghitung persentase banyaknya jawaban YA dari seluruh butir pernyataan yang disajikan.

Berikut ini disajikan suatu contoh instrumen (angket) dengan menggunakan skala Guttman.

Petunjuk: Bubuhkan tanda cek (√) pada tempat yang telah disediakan sesuai dengan sikap Anda terhadap pernyataan-pernyataan yang disajikan.

Tabel 7.4 Format Sikap terhadap Matematika

SIKAP ANDA TERHADAP PMBELAJARAN MATEMATIKA

NO	PERNYATAAN	YA	TIDAK
1	Pelajaran matematika sangat menyenangkan		
2	Saya selalu berusaha meningkatkan kemampuan saya dalam matematika		
3	Saya selalu berusaha untuk mengerjakan soal matematika dengan baik		
4	Saya selalu berusaha untuk melengkapi buku-buku matematika		
5	Saya merasa sedih jika tidak bisa mengerjakan soal matematika dengan baik		
6	Saya tidak merasa pasrah terhadap ketidakberhasilan saya dalam matematika		
7	Saya seringkali menambah pengetahuan matematika di luar jam pelajaran sekolah		

8	Saya merasa senang jika ada teman yang membicarakan matematika		
9	Matematika dirasakan ada manfaatnya dalam kehidupan sehari-hari		
10	Saya belajar matematika karena merupakan salah satu mata pelajaran yang wajib diikuti		

4. Skala Diferensial Semantik

Skala Diferensial Semantik mula-mula dikembangkan oleh Osgood dan kawan-kawan. Skala ini menuntut responden untuk memberikan penilaian tentang suatu obyek atau keadaan dengan memberikan tanda pada kontinum (selang) pernyataan yang ditulis ekstrimnya, yaitu ekstrim negatif dan ekstrim positif. Titik tengah kontinum itu sebagai titik netral (nol).

Untuk memberikan skor pada jawaban siswa, tempat-tempat tertentu pada kontinum itu diberi nilai, mulai dari nilai negatif menuju nilai positif, dari kiri ke kanan. Skala penilaian yang biasa dipergunakan adalah

$$-3, -2, -1, 0, 1, 2, 3$$

atau bisa juga ditulis sebaliknya.

Jika nilai rerata yang diperoleh seorang individu lebih besar daripada nol, maka ia mempunyai sikap positif, sebaliknya jika kurang daripada nol maka ia bersikap negatif.

Berikut ini disajikan contoh skala sikap model diferensial semantik tentang pengajaran matematika.

Tabel 7.5 Contoh Format Skala Diferensial Semantik

Petunjuk: Berilah tanda cek (✓) di atas tanda (-) sesuai dengan sikap anda terhadap pelajaran matematika.

PEMBELAJARAN MATEMATIKA DI KELAS

No	Pernyataan Kiri	Jawaban	Pernyataan Kanan
1	Sulit dimengeti	- - - - -	Mudah dimengerti
2	Tegang	- - - - -	Rileks
3	Ruwet	- - - - -	Sederhana
4	Kurang contoh	- - - - -	Contoh cukup banyak
5	Contoh kurang relevan	- - - - -	Contoh cukup relevan
6	Pemberian contoh tidak menambah pengertian	- - - - -	Pemberian contoh menambah pengertian
7	Pekerjaan rumah terlalu banyak	- - - - -	Pekerjaan rumah cukup
8	Evaluasi kurang objektif	- - - - -	Evaluasi cukup objektif

F. Penilaian Non-Tes dalam Matematika

Teknik penilaian non-tes dapat dikelompokkan atas pengamatan/observasi, wawancara, angket atau kuesioner, dan dokumen. Penilaian non-tes ini misalnya pengamatan dalam unjuk kerja (*performance*), pengamatan kegiatan dalam melaksanakan tugas (*proyek/project*), pengamatan terhadap hasil kerja dalam membuat suatu produk (*product*), dan pemeriksaan terhadap dokumen misalnya portofolio (*portfolio*).

Dalam pembelajaran matematika di SD, penilaian disamping dilakukan dengan menggunakan tes tertulis juga dapat dilakukan dengan tes non-tertulis dan non-tes. Namun demikian penggunaan kedua teknik penilaian yang terakhir ini perlu mempertimbangan kesesuaiannya dengan aspek-aspek yang terkandung dalam standar kompetensi yang harus dikuasai siswa. Berikut ini rambu-rambu penyusunan tes non-tertulis dan non-tes untuk mata pelajaran matematika SD.

1. *Penilaian Unjuk Kerja*

Penilaian unjuk kerja adalah penilaian yang dilakukan dengan cara mengamati kegiatan siswa dalam melakukan suatu kegiatan. Oleh karena itu dalam penilaian unjuk kerja diperlukan instrumen misalnya berupa lembar pengamatan atau lembar observasi. Penilaian unjuk kerja berguna untuk mengukur keterampilan atau sikap siswa melakukan kerja tertentu. Contoh unjuk kerja yang dapat diamati antara lain: kemampuan siswa dalam menggunakan atau mendemonstrasikan alat peraga matematika, memecahkan masalah kontekstual, dan melakukan penelitian matematika sederhana.

Instrumen penilaian unjuk kerja dapat terdiri dari lembar pengamatan saja, misalnya dalam kegiatan menggambar bangun datar dan memberi nama sudut, membagi sudut yang telah diketahui menjadi dua sama besar. Pada lembar pengamatan harus didefinisikan aspek yang dinilai berupa perilaku yang diharapkan muncul dari siswa selama proses unjuk kerja, aspek-aspek yang dinilai ini dipilih dengan mempertimbangkan kompetensi dasar yang harus dicapai siswa. Selain itu juga dicantumkan pedoman penskoran dan cara menilainya termasuk bagaimana mengubah dari data kualitatif menjadi kuantitatif. Instrumen penilaian unjuk kerja dapat berupa lembar pengamatan/observasi dengan skala rentang (*rating scale*) dan daftar cek (*check list*).

Tabel 7.6 Contoh Format Penilaian Unjuk Kerja

Contoh Format Penilaian Unjuk Kerja

Mata Pelajaran/Kelas	: Matematika/V (semester 1)
Kompetensi Dasar	: Melakukan pengukuran sudut
Indikator	: Menggambar dan mengukur besar sudut dengan satuan derajat

Contoh-1: Format penilaian melalui unjuk kerja dengan skala rentang (*rating scale*)

No	Nama Siswa	Aspek yang dinilai				Jml. skor
		Aspek 1	Aspek 2	Aspek 3	Aspek 4	
1.						
2.						
...						

Keterangan:

Aspek 1: Langkah-langkah dalam menggambar suatu sudut

Aspek 2: Langkah-langkah dalam mengukur besar suatu sudut

Aspek 3: Ketepatan dalam mengukur besar suatu sudut

Aspek 4: Ketepatan dalam menggambar suatu sudut

- Skor 4 = tanpa kesalahan
- Skor 3 = ada sedikit kesalahan
- Skor 2 = ada banyak kesalahan
- Skor 1 = tidak melakukan kerja
 - a. Skor maksimal = 16
 - b. Skor minimal = 4

Jumlah skor dapat ditransfer ke nilai dengan skala 0 s.d. 100 atau 0 s.d. 10

Contoh: Jumlah skor = 14, Nilai = $14/16 \times 100 = 87,5$

Contoh-2: Format penilaian unjuk kerja dengan daftar cek (*check list*)

No	Nama Siswa	Aspek yang dinilai				Jumlah skor	Nilai
		Aspek 1	Aspek 2	Aspek 3	Aspek 4		
1.							
2.							
3.							

Keterangan:

Aspek yang dinilai disesuaikan dengan tuntutan kompetensi dasar yang harus dikuasai siswa.

- Tanda “√” berarti ya, diberi skor 1.
- Tanda “-“ berarti tidak, diberi skor 0.

Jumlah skor maksimal = 4, minimal = 0

Jumlah skor dapat ditransfer ke nilai dengan skala 0 s.d. 10 atau 0 s.d. 100.

Contoh: Jumlah skor = 2. Nilai = $2/4 \times 10 = 5$

2. *Penilaian Produk*

Penilaian produk adalah penilaian terhadap keterampilan dalam membuat suatu produk. Dalam matematika, produk yang dapat dibuat misalnya bangun datar persegi, persegipanjang, segitiga, dan bangun datar lainnya dari bahan kertas atau kayu atau bahan lainnya; benda-benda ruang seperti kubus, balok, prisma, dan sebagainya dari bahan kertas atau kayu atau bahan lainnya. Atau membuat benda-benda ruang yang ada dan dikenal di lingkungan sekitar siswa seperti membuat *dos* untuk tempat kue (berbentuk balok atau kubus), dadu, dan sebagainya. Produk juga dapat berupa hasil kerja siswa misalnya gambar, grafik, diagram, membuat denah berskala, dan sebagainya.

Penilaian produk tidak hanya dilakukan terhadap hasil akhir produk, tetapi juga terhadap proses ketika membuat produk. Pengembangan produk meliputi tiga tahap, yaitu tahap persiapan, tahap proses pembuatan, dan tahap penilaian akhir produk. Pada setiap tahapan dalam pembuatan produk perlu dilakukan penilaian. Oleh karena itu, penilaian unjuk kerja dapat mengacu pada tahapan ini.

Tahap 1: persiapan

Penilaian pada tahap persiapan meliputi penilaian terhadap kemampuan siswa dalam merencanakan, menggali dan mengembangkan gagasan, dan mendesain produk.

Tahap 2: proses pembuatan produk

Penilaian pada tahap pembuatan produk meliputi penilaian terhadap kemampuan siswa dalam menyeleksi dan menggunakan bahan, alat, dan teknik.

Tahap 3: penilaian (appraisal) akhir produk

Penilaian pada tahap penilaian akhir produk meliputi penilaian terhadap kemampuan siswa dalam membuat produk sesuai kegunaannya dan memenuhi kriteria keindahan dengan menggunakan bahan, alat, dan teknik tertentu.

Produk dinilai secara holistik (keseluruhan) dan analitik (bagian-bagian). Penilaian dengan cara holistik didasarkan pada kesan keseluruhan dari produk. Penilaian dengan cara analitik didasarkan pada aspek-aspek produk yang biasanya dilakukan terhadap semua kriteria yang terdapat pada semua tahap proses pengembangan produk.

Tabel 7.7 Contoh Format Penilaian Melalui Produk

Format Penilaian Melalui Produk

Mata Pelajaran/Kelas : Matematika/V (semester 2)

Kompetensi Dasar : Menentukan jaring-jaring berbagai bangun ruang sederhana.

Indikator : Mampu membuat berbagai jaring-jaring kubus.

No	Nama siswa	Aspek 1					Aspek 2					Aspek 3					Jumlah skor
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	
1.																	
2.																	

Misalnya aspek yang dinilai:

Aspek 1: Keterampilan menggambar jaring-jaring kubus

Aspek 2: Terbentuknya semua jaring-jaring kubus

Aspek 3: Kemampuan membuat kubus dari semua jaring-jaring kubus

Keterangan:

1. Siswa tidak terampil menggambar, tidak terampil membuat kubus, atau tidak menemukan jaring-jaring kubus.
2. Siswa kurang terampil menggambar, kurang terampil membuat kubus, atau kurang dalam menemukan jaring-jaring kubus.
3. Siswa cukup terampil menggambar, cukup terampil membuat kubus, atau cukup banyak menemukan jaring-jaring kubus.

4. Siswa terampil menggambar, terampil membuat kubus, atau menemukan hampir semua jaring-jaring kubus.
5. Siswa sangat terampil menggambar, sangat terampil membuat kubus, atau menemukan semua jaring-jaring kubus.

Rumus penilaian:

$$\text{Nilai} = \text{Jumlah skor} \times \frac{100}{15}$$

Contoh:

Jumlah skor = 10, nilai = 66,67.

$$\begin{aligned} \text{Nilai} &= \text{Jumlah skor} \times \frac{100}{15} \\ &= 10 \times \frac{100}{15} \\ &= 66,67 \end{aligned}$$

3. *Penilaian Proyek*

Proyek adalah rencana pekerjaan dengan sasaran khusus dan saat penyelesaian yang tegas. Penilaian proyek adalah penilaian terhadap suatu tugas yang harus diselesaikan dalam periode/waktu tertentu. Tugas tersebut berupa penyelidikan (misalnya eksplorasi) terhadap sesuatu yang mencakup perencanaan, pengumpulan data, pengorganisasian, pengolahan dan penyajian data. Penilaian proyek dimaksudkan untuk mengetahui: pemahaman siswa dalam bidang tertentu, kemampuan siswa mengaplikasikan pengetahuan tertentu melalui suatu penyelidikan, kemampuan siswa memberi informasi tentang sesuatu yang menjadi hasil penelidikannya.

Dalam matematika proyek misalnya siswa diminta melakukan kegiatan mengumpulkan data, menganalisis, dan membuat laporannya; siswa diminta mengamati dan menyelidiki sifat-sifat kesejajaran rusuk-rusuk kubus dan menuliskan hasil-hasil temuannya. Kegiatan siswa yang termasuk proyek antara lain: penelitian sederhana tentang air di rumah, perkembangan harga sembako dalam suatu periode tertentu. Kegiatan siswa yang juga termasuk proyek antara lain

terkait dengan penelitian sederhana terhadap suatu masalah, misalnya mencari rata-rata banyaknya mobil yang melewati suatu jalan tertentu per jam.

Ada tiga hal yang perlu diperhatikan dalam penilaian proyek, yaitu:

- a. Kemampuan pengelolaan yang meliputi kemampuan dalam memilih topik (bila belum ditentukan secara spesifik oleh guru), mencari informasi dan mengelola waktu pengumpulan data serta penulisan laporan.
- b. Relevansi yaitu kesesuaian dengan mata pelajaran ditinjau dari segi pengetahuan, keterampilan dan pemahaman selama proses belajar.
- c. Keaslian yaitu proyek yang dilakukan siswa merupakan karya nyata siswa dengan kontribusi guru pada petunjuk dan dukungan.

Penilaian hasil karya dalam proyek dilakukan dari proses perencanaan, proses pengerjaan tugas sampai hasil akhir proyek. Oleh karena itu perlu ditetapkan hal-hal atau aspek yang perlu dinilai, seperti penyusunan desain, pengumpulan data, analisis data dan penyiapan laporan tertulis. Instrumen penilaian proyek dapat terdiri dari lembar pengamatan (observasi) dengan daftar cek (*check list*) dan skala rentang (*rating scale*).

Tabel 7.8 Contoh Format Penilaian Proyek

Format Penilaian Proyek

Mata Pelajaran/Kelas : Matematika/ VI

Kompetensi Dasar : a. Mengumpulkan dan membaca data

b. Mengolah dan menyajikan data dalam bentuk tabel

Indikator:

- a. Mengumpulkan data dengan pencatatan langsung dan dengan lembar isian
- b. Mengolah data dan menyajikannya dalam bentuk tabel

Contoh-1: Format penilaian proyek dengan skala rentang (*rating scale*)

No	Nama Siswa	Aspek yang dinilai				Nilai
		Tahap Persiapan	Tahap Pelaksanaan	Tahap Pelaporan	Jumlah Skor	
1.						
2.						
3.						

Keterangan:

- Skor 4 = tanpa kesalahan
- Skor 3 = ada sedikit kesalahan
- Skor 2 = ada banyak kesalahan
- Skor 1 = tidak melakukan
- Skor maksimal = 12, skor minimal = 4

Jumlah skor dapat ditransfer ke nilai dengan skala 0 s.d. 10 atau 0 s.d. 100

Contoh: Jumlah skor = 11. Nilai = $11:12 \times 10 = 9,2$

Contoh-2: Format penilaian proyek dengan daftar cek (*check list*)

No	Nama Siswa	Aspek yang dinilai				Nilai
		Tahap persiapan	Tahap pelaksanaan	Tahap pelaporan	Jumlah skor	
1.						
2.						
3.						

Keterangan:

- Skor 1 = ya
- Skor 0 = tidak
- Skor yang dicapai maksimal = 3
- Skor yang dicapai minimal = 0

Jumlah skor dapat ditransfer ke nilai dengan skala 0 s.d. 100 atau 0 s.d. 10

Contoh: Skor yang dicapai = 2, Nilai = $2 : 3 \times 100 = 67$

Keterangan: Aspek yang dinilai

- Tahap persiapan : persiapan format-format untuk pengumpulan data secara langsung maupun dengan lembar isian
- Tahap pelaksanaan : proses pencatatan data, pengelompokan data dan analisis data.
- Tahap pelaporan : ketepatan isi laporan dan bentuk sajian laporan.

4. Penilaian Portofolio

Portofolio adalah suatu kumpulan sistematis hasil-hasil pekerjaan seseorang. Penilaian portofolio adalah penilaian berkelanjutan yang didasarkan pada kumpulan informasi yang menunjukkan perkembangan kemampuan siswa dalam suatu periode tertentu. Informasi perkembangan siswa dapat berupa hasil karya terbaik siswa selama proses belajar, pekerjaan hasil tes, piagam penghargaan, atau bentuk informasi lain yang terkait kompetensi tertentu dalam suatu mata pelajaran. Dari informasi perkembangan itu siswa dan guru dapat menilai kemajuan belajar yang dicapai dan siswa terus berusaha memperbaiki diri.

Dalam matematika, portofolio dapat berupa dokumen hasil kerja siswa misalnya hasil pekerjaan di rumah, sertifikat hasil lomba, tugas-tugas karya ilmiah siswa, dan sebagainya. Secara teknis pengelolaan penilaian portofolio dapat ditempuh dalam berbagai variasi. Walaupun demikian pengelolaannya mengacu pada paling sedikit 7 unsur kunci, yaitu:

- a. Membuat siswa memahami makna portofolio dalam kaitan dengan pencapaian dan kemajuan hasil belajarnya;
- b. Menentukan topik pekerjaan atau karya siswa yang akan dikoleksi sebagai portofolio;
- c. Mengumpulkan dan menyimpan pekerjaan atau karya siswa yang dipilih sebagai portofolio;
- d. Memilih atau menentukan kriteria untuk menilai pekerjaan atau karya siswa yang akan dikoleksi sebagai portofolio;
- e. Membantu dan mendorong siswa agar selalu mengevaluasi dan memperbaiki hasil-hasil pekerjaan atau karya portofolio mereka;

- f. Menjadwalkan dan melaksanakan pertemuan portofolio dengan siswa;
- g. Melibatkan orang tua dan unsur lain terkait dalam program dan pelaksanaan penilaian portofolio siswa.

5. *Wawancara*

Wawancara adalah kegiatan tanya jawab antara penanya dan yang ditanya dengan tujuan untuk memperoleh informasi yang dapat digali dari responden. Penilaian matematika melalui wawancara bersifat kualitatif. Guru perlu membuat pedoman wawancara dengan memperhatikan aspek-aspek yang ingin diketahui, rangkaian pertanyaan yang akan diajukan, dan kapan wawancara akan dilakukan. Namun demikian pertanyaan dalam wawancara dapat berkembang mengikuti jawaban dari responden.

Wawancara dapat dilakukan oleh guru setelah melakukan tes tertulis dengan maksud untuk memperoleh informasi lebih lanjut tentang pemahaman siswa terhadap pemahaman konsep, pemecahan masalah, dan dalam mengkomunikasikan matematika. Misalnya guru telah melaksanakan tes tertulis dan memeriksa hasil kerja siswa, ternyata ada siswa yang hasil kerjanya "menarik" perhatian guru misalnya siswa memecahkan masalah dengan cara baru.

Guru dapat melakukan wawancara untuk mengetahui jalan pikiran siswa sehingga memperoleh cara baru tersebut, dan selanjutnya guru dapat mengetahui apakah cara baru tersebut diperoleh dengan alasan yang benar atau ada langkah-langkah pemecahan yang perlu diperbaiki. Hasil wawancara ini sekaligus dapat digunakan untuk mengetahui bagaimana cara siswa berpikir dan mengkomunikasikan pemikiran matematikanya.

6. *Penilaian Tes Lisan*

Tes lisan adalah teknik penilaian hasil belajar yang pertanyaannya dan jawabannya atau pernyataannya atau tanggapannya disampaikan dalam bentuk lisan dan secara spontan. Misalnya dalam pembelajaran matematika guru telah menjelaskan konsep persegi panjang, lalu guru mengambil beberapa contoh bangun datar berbentuk persegi dan mengatakan "ini persegi". Kemudian guru dapat mengajukan pertanyaan secara lisan: "Coba kalian ucapkan dengan bahasa kalian sendiri apa arti bangun datar persegi?". Siswa diminta menjawab soal ini secara lisan dan bersifat spontan berdasar pengamatan dan pemahaman tentang definisi persegi panjang. Guru mengajak siswa mengkonstruksi makna persegi.

Tes lisan juga dapat berupa pertanyaan yang jawabannya dapat dilakukan dengan "mencongak", misalnya dalam pembelajaran matematika di kelas III semester 2 guru mengajukan pertanyaan secara lisan: "Keliling persegi panjang dengan panjang 4 cm dan lebar 3 cm adalah ...?", guru meminta siswa menjawab secara mencongak dan menyebutkan jawabannya dengan lisan.

Tes lisan seperti ini berguna untuk mengetahui pemahaman siswa tentang konsep keliling suatu bangun datar, kecepatan berpikir, dan ketepatan siswa dalam menjawab dan menanggapi suatu pertanyaan, serta sikap kritis dalam menghadapi dan memecahkan masalah. Tes lisan ini juga dapat menjadi umpan balik bagi guru untuk mengetahui tingkat pemahaman siswa dalam menangkap konsep dan keterampilan berhitung siswa.

Tabel 7.9 Contoh Format Penilaian Tes Lisan

Contoh: Format Penilaian Tes Lisan

Mata Pelajaran/Kelas : Matematika/ I(semester 1)

Kompetensi Dasar : 1.4 Menyelesaikan masalah yang berkaitan dengan penjumlahan dan pengurangan sampai 20.

Indikator : Menyelesaikan masalah yang berkaitan dengan penjumlahan sampai 20.

No	Aspek yang dinilai	Bobot Penilaian	Skala				Skor	Rubrik
			1	2	3	4		
1.	Pemahaman terhadap apa yang diketahui	5					1. jika sama sekali tidak memahami 2. jika pemahamannya sebagian kecil 3. jika pemahamannya sebagian besar 4. jika sangat memahami
2.	Pemahaman terhadap apa yang ditanyakan	5					1. jika sama sekali tidak memahami 2. jika pemahamannya sebagian kecil 3. jika pemahamannya sebagian besar 4. jika sangat memahami
3.	Ketepatan penggunaan strategi dalam pemecahan masalah	5					1. jika sama sekali tidak tepat 2. jika sebagian kecil tepat 3. jika sebagian besar tepat 4. jika seluruhnya tepat
4.	Kecermatan dalam menyampaikan argumentasi	4					1. jika sama sekali tidak cermat 2. jika sebagian kecil cermat 3. jika sebagian besar cermat 4. jika seluruhnya cermat
5.	Kelogisan dalam urutan langkah pemecahan masalah	3					1. jika sama sekali tidak logis 2. jika sebagian kecil logis 3. jika sebagian besar logis 4. jika seluruhnya logis
6.	Kebenaran jawaban	3						1. jika sama sekali tidak benar 2. jika sebagian kecil benar 3. jika sebagian besar benar 4. jika seluruhnya benar
	Jumlah	25					

Skor maksimal = 100

Rumus penilaian: Nilai = Jumlah skor