
Pengukuran, *Assessment* dan Evaluasi dalam Pembelajaran Matematika

Nur Fitriani Zainal¹⁾

¹⁾Institut Agama Islam Negeri (IAIN) Manado

Email : nur.zainal@iain-manado.ac.id

ABSTRACT

Evaluation is one of the important components and stages that must be taken by the teacher to determine the effectiveness of learning. The results obtained from the evaluation can be used as a consideration for teachers in improving and perfecting learning programs and activities. The terms used in the evaluation system, namely measurement, assessment, and evaluation. This research is a library research that examines references about measurement, assessment, and evaluation in mathematics learning. In learning activities, evaluation is preceded by assessment, while assessment is preceded by measurement. In measurement, the activity is carried out comparing the results of observations with the criteria. Then, the measurement results are interpreted and described in the assessment process. Next, evaluate, which is to determine the value or implications of behavior. The types of instruments used in assessment are tests and non-tests. Steps in analyzing non-test results, namely: a) transferring qualitative questionnaire results into quantitative data and b) comparing total scores (average scores) of each student with total neutral scores (mean neutral scores). On the other hand, steps are used to analyze test results, namely: a) scoring and b) interpretation of test scores.

Keywords : Measurement, Assessment, and Evaluation

ABSTRAK

Evaluasi merupakan salah satu komponen penting dan tahap yang harus ditempuh oleh guru untuk mengetahui keefektifan pembelajaran. Hasil yang diperoleh dari evaluasi dapat dijadikan pertimbangan bagi guru dalam memperbaiki dan menyempurnakan program dan kegiatan pembelajaran. Istilah-istilah yang digunakan dalam sistem evaluasi, yaitu pengukuran, penilaian, dan evaluasi. Penelitian ini merupakan *library research* yang mengkaji referensi tentang pengukuran, *assessment*, dan evaluasi dalam pembelajaran matematika. Dalam kegiatan pembelajaran, evaluasi didahului oleh penilaian (*assessment*), sedangkan penilaian didahului oleh pengukuran. Dalam pengukuran, dilakukan kegiatan membandingkan hasil pengamatan dengan kriteria. Kemudian, hasil pengukuran ditafsirkan dan

dideskripsikan dalam proses penilaian. Selanjutnya, evaluasi, yaitu menetapkan nilai atau implikasi perilaku. Jenis instrumen yang digunakan pada penilaian adalah tes dan non-tes. Langkah dalam menganalisis hasil non-tes, yaitu: a) mentransfer data kualitatif hasil angket menjadi data kuantitatif dan b) membandingkan total skor (rerata skor) tiap siswa dengan total skor netral (rerata skor netral). Di sisi lain, langkah yang digunakan untuk menganalisis hasil tes, yaitu: a) menskor dan b) penafsiran skor tes.

Kata Kunci : Pengukuran, *Assessment*, dan Evaluasi

PENDAHULUAN

Dalam UU No. 20 Tahun 2003 tentang Sistem Pendidikan Nasional, khususnya pada pasal 1 ayat ke 16, evaluasi pendidikan adalah kegiatan pengendalian, penjaminan, dan penetapan mutu pendidikan terhadap berbagai komponen pendidikan pada setiap jalur, jenjang, dan jenis pendidikan sebagai bentuk pertanggung jawaban penyelenggaraan pendidikan. Di samping itu, pada pasal 58, evaluasi hasil belajar peserta didik dilakukan oleh pendidik untuk memantau proses, kemajuan, dan perbaikan hasil belajar peserta didik secara berkesinambungan.

Oleh karena itu, dalam sistem pembelajaran, evaluasi merupakan salah satu komponen penting dan tahap yang harus ditempuh oleh guru untuk mengetahui keefektifan pembelajaran. Hasil yang diperoleh dari evaluasi dapat dijadikan pertimbangan bagi guru dalam memperbaiki dan menyempurnakan program dan kegiatan pembelajaran. Istilah yang digunakan dalam sistem evaluasi, yaitu pengukuran, penilaian, dan evaluasi.

Berdasarkan Peraturan Menteri Pendidikan dan Kebudayaan No. 66 tahun 2013 tentang penilaian hasil belajar oleh satuan pendidikan dan penilaian hasil belajar oleh pemerintah, penilaian pendidikan diartikan sebagai proses pengumpulan dan pengolahan informasi untuk mengukur pencapaian hasil belajar peserta didik mencakup penilaian otentik, penilaian diri, penilaian berbasis protfolio, ulangan, ulangan harian, ujian tengah semester, ulangan akhir semester, ujian tingkat kompetensi, ujian mutu tingkat kompetensi, ujian nasional, dan ujian sekolah/madrasah. Hal tersebut, merupakan landasan guru di sekolah dalam memberikan ulangan harian, ujian akhir semester, tes tertulis dan tes lainnya sebagai wujud penjaminan mutu pendidikan.

METODE PENELITIAN

Penelitian ini merupakan *library research*. Dengan mengkaji referensi yang berkaitan dengan pengukuran, *assessment*, dan evaluasi dalam pembelajaran matematika.

HASIL DAN PEMBAHASAN

Pengertian Pengukuran, Assesment, dan Evaluasi

Ahmann dan Glock (Arifin, 2012 : 3) menjelaskan istilah pengukuran, sebagai berikut: “ *in the last analysis measurement is only a part, although a very substantial part of evaluation. It provides information upon which an evaluation can be based ... Educational measurement is the process that attempts to obtain a quantified representation of the degree to which a trait is possessed by a pupil.*

 Pengukuran merupakan bagian penting dari evaluasi. Pengukuran memberikan informasi yang dapat menjadi dasar evaluasi. Pengukuran dalam pendidikan merupakan proses untuk memperoleh gambaran terukur mengenai sejauh mana siswa menguasai suatu kompetensi.

Widoyoko (2014 : 2) menyebutkan pengukuran merupakan kuantifikasi atau penetapan angka tentang karakteristik atau keadaan individu menurut aturan-aturan tertentu. Asmawi Zainul dan Noehi Nasution (Hamzah, 2014 : 19) juga memberikan definisi yang serupa mengenai pengukuran, yaitu pemberian angka kepada suatu atribut atau karakteristik tertentu menurut aturan atau formulasi yang jelas. Dengan demikian, secara sederhana pengukuran dapat dikatakan sebagai kegiatan pemberian atau penetapan angka pada objek yang diukur yang disesuaikan dengan kriteria-kriteria tertentu sesuai dengan objek tersebut.

Pengukuran dapat dibedakan menjadi 3 macam (Sudijono, 2013 : 4), yaitu: 1) pengukuran yang dilakukan bukan untuk menguji sesuatu, seperti pengukuran yang dilakukan oleh penjahit pakaian; 2) pengukuran yang dilakukan untuk menguji sesuatu, misalnya pengukuran untuk menguji daya tahan lampu pijar; dan 3) pengukuran untuk menilai, yang dilakukan dengan menguji, misalnya mengukur kemajuan belajar peserta didik dalam rangka mengisi nilai rapor yang dilakukan dengan menguji mereka dalam bentuk tes hasil belajar.

Adapun istilah penilaian merupakan alih bahasa dari istilah *assessment*, bukan dari istilah *evaluation*. Sudijono (2014: 4-5) menyebutkan bahwa penilaian berarti menilai sesuatu. Adapun menilai mengandung arti mengambil keputusan terhadap sesuatu dengan mendasarkan diri atau berpegangan pada ukuran baik atau buruk, sehat atau sakit, pandai atau bodoh, dan sebagainya. Jadi, penilaian itu bersifat kualitatif. Widoyoko (2014:3) juga memberikan definisi mengenai penilaian, yaitu kegiatan menafsirkan data hasil pengukuran berdasarkan kriteria maupun aturan-aturan tertentu.

Kusaeri dan Suprananto (2012:8) mendefinisikan penilaian sebagai suatu prosedur sistematis dan mencakup kegiatan mengumpulkan, menganalisis, serta menginterpretasikan informasi yang dapat digunakan untuk membuat kesimpulan tentang karakteristik seseorang atau objek. Adapun menurut Wiggins (1994), *assessment* merupakan sarana yang secara kronologis membantu guru dalam memonitor siswa. Definisi tersebut sejalan dengan pengertian penilaian pendidikan dalam Peraturan Menteri Pendidikan dan Kebudayaan Nomor 66 Tahun 2013 yaitu proses pengumpulan dan pengolahan informasi untuk menentukan pencapaian hasil belajar siswa (Permendikbud No. 66 Tahun 2013). Berdasarkan berbagai uraian di atas, dapat disimpulkan bahwa asesmen atau penilaian adalah suatu kegiatan membandingkan atau menerapkan hasil pengukuran untuk memberikan nilai terhadap objek penilaian.

Tujuan penilaian diarahkan pada empat hal, yaitu: 1) penelusuran (*keeping track*), yaitu untuk menelusuri agar proses pembelajaran tetap sesuai dengan rencana; 2) pengecekan (*checking-up*), yaitu untuk mengecek adakah kelemahan-kelemahan yang dialami siswa selama proses pembelajaran; 3) pencarian (*finding-out*), yaitu untuk mencari dan menemukan hal-hal yang menyebabkan terjadinya kelemahan dan kesalahan dalam proses pembelajaran; dan 4) penyimpulan (*summing-up*), yaitu untuk menyimpulkan apakah siswa telah menguasai seluruh komponen yang ditetapkan dalam kurikulum atau belum (Kusaeri dan Suprananto, 2012:9).

Adapun evaluasi memiliki makna yang berbeda dengan penilaian dan pengukuran. Stufflebeam & Shinkfield (Widoyoko, 2014:3) menyebutkan bahwa evaluasi merupakan suatu proses menyediakan informasi yang dapat dijadikan sebagai pertimbangan untuk menentukan harga dan jasa dari tujuan yang dicapai, desain,

implementasi, dan dampak untuk membantu keputusan, membantu pertanggungjawaban dan meningkatkan pemahaman terhadap fenomena. Dalam rangka kegiatan belajar mengajar, Norman E. Gronlund menyatakan evaluasi sebagai suatu proses sistematis dalam menentukan tingkat pencapaian tujuan instruksional oleh siswa (Suherman dan Sukjaya, 1990:1).

Selanjutnya, Suherman dan Sukjaya (1990:3) menyatakan bahwa evaluasi sifatnya lebih luas dari pengukuran. Evaluasi meliputi aspek kuantitatif dan kualitatif. Evaluasi selain menyangkut pengukuran, berlanjut dengan pemberian nilai (*valuing*) berupa keputusan-keputusan maupun nilai tingkah laku yang diukur. Sejalan dengan Suherman dan Sukjaya, Sudijono (2013:5) menyebutkan bahwa evaluasi adalah kegiatan atau proses untuk menilai sesuatu yang mencakup dua kegiatan, yaitu kegiatan pengukuran dan penilaian. Dengan demikian, dapat disimpulkan bahwa evaluasi merupakan suatu proses sistematis yang mencakup kegiatan pengukuran dan penilaian untuk dijadikan pertimbangan dalam mengambil keputusan.

Keterkaitan antara Pengukuran, Penilaian, dan Evaluasi

Suherman dan Sukjaya (1990:3) mengatakan bahwa istilah evaluasi, penilaian, dan pengukuran dapat dibedakan. Pengukuran (*measurement*) menunjuk pada segi kuantitas (*how much*), penilaian menunjuk pada segi kualitas (*what value*), dan evaluasi berkenaan dengan keduanya, yaitu pengukuran dan penilaian. Antara evaluasi, penilaian, dan pengukuran terdapat hubungan yang erat yang tidak dapat dipisahkan. Norman E. Gronlund (Suherman dan Sukjaya, 1990:4) melukiskan hubungan ketiganya, yaitu: 1) Evaluasi adalah deskripsi kuantitatif siswa yang ditetapkan dengan penentuan nilai, dan 2) Evaluasi adalah deskripsi kualitatif siswa yang ditetapkan dengan penentuan nilai.

Widoyoko (2014:4) menyatakan bahwa pengukuran, penilaian, dan evaluasi bersifat hierarki. Dengan demikian, dapat disimpulkan bahwa evaluasi didahului oleh penilaian (*assessment*), sedangkan penilaian didahului oleh pengukuran. Dalam pengukuran, dilakukan kegiatan membandingkan hasil pengamatan dengan kriteria. Kemudian, hasil pengukuran ditafsirkan dan dideskripsikan dalam proses penilaian. Selanjutnya, evaluasi, yaitu menetapkan nilai atau implikasi perilaku.

Unsur-Unsur yang Dinilai dalam Pembelajaran Matematika pada Tingkat Satuan Pendidikan SD

Dalam Kurikulum Tingkat Satuan Pendidikan (KTSP) atau Kurikulum 2013 (K. 13), yang termasuk unsur-unsur penilaian hasil belajar matematika di Sekolah Dasar adalah ulangan harian, ulangan tengah semester, tugas, dan ulangan akhir semester. Adapun objek evaluasi hasil belajar matematika pada jenjang Sekolah Dasar mengacu kepada dua dari tiga ranah dalam taksonomi tujuan pendidikan milik Benjamin S. Bloom, yakni ranah proses berpikir dan ranah nilai atau sikap.

1. Ranah Proses Berpikir (*Cognitive Domain*)

Ranah proses berpikir adalah ranah yang mencakup kegiatan mental (otak). Menurut Bloom (Sudijono, 2013:49), segala upaya yang menyangkut aktivitas otak adalah termasuk dalam ranah kognitif. Sudijono (2013:50-52) juga memaparkan keenam jenjang proses berpikir dalam ranah kognitif, yaitu:

a. Pengetahuan (*Knowledge*)

Kemampuan seseorang untuk mengingat-ingat kembali (*recall*) atau mengenali kembali tentang nama, istilah, ide, gejala, rumus-rumus dan sebagainya, tanpa mengharapkan kemampuan untuk menggunakananya.

b. Pemahaman (*Comprehension*)

Kemampuan seseorang untuk mengerti atau memahami sesuatu setelah sesuatu itu diketahui dan diingat. Dengan kata lain, memahami adalah mengetahui tentang sesuatu dan dapat melihatnya dari berbagai segi.

c. Penerapan (*Application*)

Kesanggupan seseorang untuk menerapkan atau menggunakan ide-ide umum, tata cara ataupun metode-metode, prinsip-prinsip, rumus-rumus, teori-teori dan sebagainya, dalam situasi yang baru dan konkret.

d. Analisis (*Analysis*)

Kemampuan seseorang untuk merinci atau menguraikan suatu bahan atau keadaan menurut bagian-bagian yang lebih kecil dan mampu memahami hubungan diantara bagian-bagian atau faktor-faktor yang satu dengan faktor-faktor lainnya.

e. Sintesis (*Synthesis*)

Kemampuan berpikir yang merupakan kebalikan dari proses berpikir analisis. Sintesis merupakan suatu proses yang memadukan bagian-bagian atau unsur-unsur

secara logis, sehingga menjelma menjadi suatu pola yang berstruktur atau berbentuk pola baru.

f. Evaluasi (*Evaluation*)

Kemampuan seseorang untuk membuat pertimbangan terhadap suatu situasi, nilai atau ide, misalnya jika seseorang dihadapkan pada beberapa pilihan, maka ia akan mampu memilih satu pilihan yang terbaik, sesuai dengan patokan-patokan atau kriteria yang ada.

2. Ranah Nilai atau Sikap (*Affective Domain*)

Ranah afektif adalah ranah yang berkaitan dengan sikap dan nilai. Ranah afektif ini oleh David R. Krathwohl dan kawan-kawan dikelompokkan lebih rinci lagi ke dalam lima jenjang (Sudijono, 2013:54-56), yaitu:

a. Menerima atau Memperhatikan (*receiving/attending*)

Kepekaan seseorang dalam menerima rangsangan (stimulus) dari luar yang datang kepada dirinya dalam bentuk masalah, situasi, gejala dan lain-lain. Termasuk dalam jenjang ini misalnya adalah: kesadaran dan keinginan untuk menerima stimulus, mengontrol dan menyeleksi gejala-gejala atau rangsangan yang datang dari luar.

b. Menanggapi (*responding*)

Kemampuan yang dimiliki oleh seseorang untuk mengikut-sertakan dirinya secara aktif dalam fenomena tertentu dan membuat reaksi terhadapnya dengan salah satu cara.

c. Menghargai (*valuing*)

Memberikan nilai atau memberikan penghargaan terhadap suatu kegiatan atau obyek, sehingga apabila kegiatan itu tidak dikerjakan, dirasakan akan membawa kerugian atau penyesalan.

d. Mengorganisasikan (*organization*)

Mengatur atau mengorganisasikan merupakan pengembangan dari nilai ke dalam satu sistem organisasi, termasuk di dalamnya hubungan satu nilai dengan nilai lain, pemantapan dan prioritas nilai yang telah dimilikinya.

e. Karakterisasi dengan suatu nilai atau komplek nilai (*characterization by a value or value complex*)

Keterpaduan semua sistem nilai yang telah dimiliki seseorang, yang mempengaruhi pola kepribadian dan tingkah lakunya.

Jenis-Jenis Instrumen yang Digunakan

Ada dua jenis instrumen yang digunakan dalam mengukur hasil belajar matematika pada tingkat satuan Sekolah Dasar, yaitu:

1. Instrumen Tes

Tes merupakan suatu teknik atau cara yang digunakan dalam rangka melaksanakan kegiatan pengukuran, yang di dalamnya terdapat berbagai pertanyaan, pernyataan, atau serangkaian tugas yang harus dikerjakan atau dijawab oleh peserta didik untuk mengukur aspek perilaku peserta didik (Arifin, 2012:118). Adapun menurut Djemari (Widoyoko, 2014:45), tes merupakan salah satu cara untuk menaksir besarnya kemampuan seseorang secara tidak langsung, yaitu melalui respons seseorang terhadap stimulus atau pertanyaan. Dengan demikian, tes dapat diartikan sebagai alat yang digunakan untuk mengukur atau menaksir tingkat kemampuan seseorang.

Ada dua bentuk tes yang digunakan dalam lembaga pendidikan jika dilihat dari segi sistem penskorannya, yaitu tes objektif dan tes subjektif (Widoyoko, 2014:46).

a. Tes objektif

Tes objektif adalah bentuk tes yang mengandung kemungkinan jawaban atau respon yang harus dipilih oleh peserta tes. Kemungkinan jawaban atau respon telah disediakan oleh penyusun butir soal. Peserta hanya memilih alternatif jawaban yang telah disediakan. Dengan demikian, pemeriksaan atau penskoran jawaban peserta tes sepenuhnya dapat dilakukan secara objektif oleh pemeriksa.

Secara umum ada tiga tipe tes objektif (Widoyoko, 2014:50), yaitu benar salah (*true false*), menjodohkan (*matching*), dan pilihan ganda (*multiple choice*). Tes hasil belajar matematika siswa di SD menggunakan tes objektif tipe pilihan ganda (*multiple choice*). Tes pilihan ganda (*multiple choice*) merupakan tes dimana setiap butir soalnya memiliki jumlah alternatif jawaban lebih dari satu (Widoyoko, 2014:59). Pada umumnya alternatif jawaban berkisar antara 2 (dua) atau 5 (lima). Setiap tes

pilihan ganda terdiri dari dua bagian, yaitu pernyataan atau disebut juga *stem* dan alternatif jawaban atau disebut juga *option*.

b. Tes subjektif

Pada umumnya, tes subjektif berbentuk uraian (esai). Asmawi Zaenul dan Noehi Nasution (Widoyoko, 2014:78-79) menyebutkan bahwa tes bentuk uraian adalah butir soal yang mengandung pertanyaan atau tugas yang jawaban penggerjaan soal tersebut harus dilakukan dengan cara mengekspresikan pikiran peserta tes. Ciri khas tes uraian adalah jawaban terhadap soal tersebut tidak disediakan oleh penyusun soal, tetapi harus disusun oleh peserta tes. Butir soal tipe uraian hanya terdiri dari pertanyaan atau tugas dan jawaban sepenuhnya harus dipikirkan oleh peserta tes. Jumlah butir soal dalam tes uraian sekitar 5-10 butir soal.

Secara umum tes uraian dapat dibagi menjadi dua bentuk, yaitu tes uraian bebas (*Extended Response Test*) dan tes uraian terbatas (*restricted response*). Tes hasil belajar matematika siswa pada umumnya menggunakan tes uraian bebas (*Extended Response Test*). Tes uraian bebas (*Extended Response Test*) merupakan bentuk tes uraian yang memberi kebebasan kepada peserta tes untuk mengorganisasikan dan mengekspresikan pikiran dan gagasannya dalam menjawab soal tes (Widoyoko, 2014:79).

Dalam mengembangkan tes hasil atau prestasi perlu ditempuh delapan langkah di bawah ini, yaitu: a) menyusun spesifikasi tes, b) menulis soal tes, c) menelaah soal tes, d) melakukan uji coba tes e) menganalisis butir soal f) memperbaiki tes, g) merakit tes, h) melaksanakan tes, dan i) menafsirkan hasil tes (Mardapi, 2008).

2. Instrumen Non-Tes

Terkait dengan aspek afektif dalam pembelajaran matematika, ada beberapa komponen afektif yang penting untuk diukur. Untuk mengukur komponen-komponen afektif tersebut, perlu dibuat instrumen yang berbeda dengan instrumen tes. Untuk mengukur hasil belajar yang berkenaan dengan *soft skills* dan *vocational skills*, terutama yang berhubungan dengan apa yang dapat dibuat atau dikerjakan oleh peserta didik daripada apa yang diketahui atau dipahaminya, digunakan alat ukur non-tes.

Instrumen non-tes merupakan salah satu instrumen evaluasi yang pada tingkat satuan pendidikan SD disebut teknik penilaian untuk memperoleh gambaran mengenai karakteristik, sikap, atau kepribadian (Hamzah, 2014:158). Selanjutnya, Hamzah

(2014:158) menyebutkan beberapa instrumen evaluasi nontes, seperti angket, pedoman wawancara, lembar observasi/pengamatan, portofolio, dan jurnal rubrik. Dalam mengukur hasil belajar yang berkaitan dengan ranah afektif pada tingkat satuan pendidikan SD digunakan instrumen non-tes berupa angket.

Noehi Nasoetion (Hamzah, 2014:159) mendefinisikan angket sebagai alat untuk mengumpulkan data yang berupa pernyataan yang disampaikan kepada responden yang dijawab secara tertulis. Berdasarkan bentuk butir angket yang disusun, angket dibedakan menjadi angket tertutup dan angket terbuka (Hamzah, 2014:160). Pada tingkat satuan pendidikan SD digunakan angket tertutup, yaitu angket dengan butir pernyataan disertai dengan pilihan jawaban yang sudah tersedia. Responden hanya memilih jawaban yang dianggap paling sesuai dengan dirinya.

Salah satu komponen afektif yang akan diukur menggunakan intrumen non-tes berupa angket adalah konsep diri. Konsep diri adalah pandangan seseorang tentang dirinya sendiri dalam hal kelayakan dan ketidak-layakan, nilai-nilainya serta keinginannya (Campbell, 1967: 510). Angket konsep diri ini menggunakan metode penyekalaan likert. Skala Likert merupakan metode penyekalaan pernyataan sikap yang menggunakan distribusi jawaban sebagai dasar penentuan nilai skalanya (Kusaeri dan Suprananto, 2012:221).

Teknik Pengumpulan Data

Pengumpulan data dalam penilaian hasil belajar matematika pada tingkat satuan SD dilakukan dengan dua cara yaitu teknik tes dan teknik non-tes. Untuk teknik tes, pengumpulan data dilakukan dengan cara pemberian instrumen tes kepada siswa. Teknik tes yang digunakan adalah tes tertulis (*paper and pencil tes*), yaitu jenis tes dimana *tester* dalam mengajukan butir-butir pertanyaan atau soal dilakukan secara tertulis dan *tester* memberikan jawabannya secara tertulis pula (Sudijono, 2013:75). Tes tertulis pada umumnya dilaksanakan secara berkelompok di suatu ruangan tertentu. Siswa akan menjawab soal yang telah disusun oleh guru berdasarkan standar isi, kompetensi dasar dan indikator.

Adapun pada teknik non-tes, pengumpulan data dilakukan dengan cara pemberian angket (*questioner*) konsep diri yang berisi sejumlah pernyataan yang dibuat berdasarkan indikator konsep diri. Siswa diminta menjawab setuju atau tidak

terhadap isi pernyataan dalam lima macam kategori jawaban, yaitu “sangat tidak setuju” (STS), “tidak setuju” (TS), “ragu-ragu” (R), “setuju” (S), “sangat setuju” (SS). Pada angket konsep diri ini, digunakan metode penyekalaan Likert. Skala untuk tiap butir pernyataan ditentukan oleh distribusi respon setuju atau tidak setuju dari sekolompok responden yang bertindak sebagai kelompok uji coba. Gable (Kusaeri dan Suprananto, 2012:221) menyarankan banyaknya responden uji coba sekitar 6 sampai 10 kali lipat banyaknya dari pernyataan yang akan dianalisis.

Proses Penilaian

Penilaian hasil belajar merupakan komponen penting dalam kegiatan pembelajaran. Upaya untuk meningkatkan kualitas pembelajaran dapat ditempuh melalui peningkatan kualitas sistem penilaiannya. Sistem penilaian yang baik akan mendorong pendidik untuk menentukan strategi mengajar yang baik dan memotivasi peserta didik untuk belajar yang lebih baik.

Penilaian hasil pembelajaran dilakukan minimal setelah satu kompetensi dasar dipelajari. Bila muatan pada satu kompetensi dasar cukup padat, penilaian hasil belajar dapat dilakukan lebih dari satu kali. Fokus penilaian tidak harus pada semua indikator pencapaian kompetensi yang telah ditetapkan, namun dapat dipilih yang berkenaan dengan indikator-indikator esensial dan mencerminkan hasil akhir pencapaian kompetensi dasarnya.

Hamzah (2014:294) menyebutkan bahwa sistem penilaian terbagi dua, yaitu *Criterion Referenced Test* (CRT) atau yang dikenal dengan Penilaian Acuan Patokan (PAP) dan *Norm Referenced Test* (NRT) atau Penilaian Acuan Norma (PAN). Terkait dengan kedua sistem penilaian tersebut, Woodworth (Suherman dan Sukjaya, 1990:248) menjelaskan bahwa ada dua jenis pedoman yang bisa digunakan untuk menentukan nilai (mengubah skor menjadi nilai) sebagai hasil evaluasi, yaitu:

1. dengan cara membandingkan skor yang diperoleh seorang individu (siswa) dengan suatu standar yang sifatnya mutlak (PAP); dan
2. dengan cara membandingkan skor yang diperoleh seorang individu (siswa) dengan skor yang diperoleh siswa lainnya dalam kelompok tes tersebut (PAN).

Proses atau langkah-langkah penilaian hasil belajar dimuat dalam lima langkah. Langkah pokok yang harus ditempuh sebagai prosedur evaluasi yang disebutkan Muchtar Buchari (Suherman dan Sukjaya, 1990:16-18), yaitu:

1. Perencanaan (*Planning*)

Tahap perencanaan meliputi kegiatan merumuskan tujuan evaluasi yang akan dilaksanakan, merencanakan metode evaluasi yang akan dipakai, menyusun alat evaluasi yang akan digunakan, menentukan kriteria penilaian yang akan digunakan, dan menetapkan frekuensi evaluasi.

2. Pengumpulan Data (*Collecting*)

Tahap pengumpulan data, terdiri dari melaksanakan pengukuran, pemeriksaan hasil dan pemberian skor.

3. Verifikasi Data (*Verification*)

Pada tahap verifikasi data, data yang telah dihimpun dikelompokkan dan disaring sebelum diolah lebih lanjut. Verifikasi data dimaksudkan untuk dapat memisahkan data yang dapat memperjelas gambaran yang diperoleh mengenai diri individu atau sekelompok individu yang sedang dievaluasi dengan data yang akan mengaburkan gambaran yang akan diperoleh jika data itu ikut diolah (Sudijono, 2013:61).

4. Analisis Data (*Analysis*)

Pada tahap analisis data, data yang telah diverifikasi diolah dan dianalisis dengan maksud untuk memberikan makna terhadap data yang telah berhasil dihimpun (Sudijono, 2013:61).

5. Penafsiran (*Interpretation*)

Penafsiran atau interpretasi dimaksudkan sebagai pernyataan atau keputusan tentang hasil evaluasi. Atas dasar interpretasi terhadap data hasil evaluasi ini dapat dikemukakan kesimpulan-kesimpulan tertentu yang mengacu kepada tujuan diadakannya evaluasi.

Proses Analisis Hasil Non-Tes

Derajat penilaian siswa terhadap suatu pernyataan terbagi ke dalam lima kategori yang tersusun secara bertingkat, mulai dari “Sangat Tidak Setuju” (STS), “Tidak Setuju” (TS), “Ragu-Ragu” (R), “Setuju” (S), dan “Sangat Setuju” (SS). Dalam

menganalisis hasil angket, skala kualitatif tersebut ditransfer ke dalam skala kuantitatif. Nilai untuk setiap skala tersebut diperoleh dari hasil uji coba instrumen pada sekelompok responden dengan karakteristik yang mirip dengan siswa yang akan dinilai.

Setelah data jawaban angket setiap siswa telah diubah kedalam bentuk kuantitatif, seorang siswa sudah dapat digolongkan kedalam kelompok siswa yang memiliki konsep diri positif atau kelompok siswa dengan konsep diri negatif. Ada dua cara penggolongan yang dapat dilakukan (Suherman dan Sukjaya, 1990:237), yaitu:

1. Membandingkan total skor tiap siswa dengan total skor alternatif jawaban netral (kategori “Ragu-Ragu”) dari semua butir pernyataan. Jika skor total siswa lebih besar dari total skor netral, maka siswa digolongkan memiliki konsep diri positif. Sebaliknya, jika skor total siswa lebih kecil dari total skor netral, maka siswa digolongkan memiliki konsep diri negatif.
2. Membandingkan rerata skor yang diperoleh tiap siswa dengan rerata skor alternatif jawaban netral dari semua butir pernyataan. Jika rerata skor siswa lebih besar dari rerata skor netral, maka siswa digolongkan memiliki konsep diri positif. Sebaliknya, jika rerata skor siswa lebih kecil dari rerata skor netral, maka siswa digolongkan memiliki konsep diri negatif.

Proses Analisis Hasil Tes

Langkah-langkah pokok yang harus dilakukan untuk menganalisis data hasil tes, yaitu:

1. Menskor, yaitu memberi skor pada hasil tes yang dapat dicapai oleh peserta didik.
 - a. Penyekoran soal bentuk pilihan ganda.

Rumus untuk mencari skor dalam tes tipe pilihan ganda ada 2 macam, yaitu:

- 1) Sistem Denda

Rumus skor dengan sistem denda adalah:

$$SK = B - \frac{S}{P-1},$$

dengan ketentuan:

SK = skor yang diperoleh peserta tes

B = jumlah jawaban yang benar

S = jumlah jawaban yang salah

P = banyaknya pilihan

1 = bilangan tetap

2) Sistem Tanpa Denda

Rumus skor dengan sistem tanpa denda adalah:

$SK = B,$

dengan ketentuan:

SK = skor yang diperoleh peserta tes

B = jumlah jawaban yang benar

(Widoyoko, 2014:66-67)

b. Penyekoran untuk soal bentuk uraian

Ketika melakukan penyekoran butir tes uraian, kita harus membuat batasan yang jelas tentang suatu jawaban yang bagus. Batasan-batasan jawaban tersebut dikenal dengan rubrik. Rubrik penyekoran biasanya diklasifikasikan dalam analitik dan holistik. Rubrik penyekoran analitik mengidentifikasi jawaban dari berbagai aspek atau dimensi yang berbeda dan skor guru pada masing-masing dimensi diletakkan secara terpisah. Sedangkan rubrik penyekoran holistik, guru memberi skor tunggal berdasarkan pada keseluruhan jawaban siswa (Kusaeri dan Suprananto, 2012: 139).

2. Penafsiran Skor Tes

Untuk menafsirkan skor tes diperlukan suatu acuan. Di bidang penilaian pendidikan dikenal dua acuan yang biasa digunakan, yaitu acuan norma (*norm reference*) dan acuan kriteria (*criterion reference*). Menafsirkan skor dengan menggunakan acuan norma, berarti membandingkan kemampuan seorang siswa dengan siswa yang lainnya di dalam kelompoknya. Sedangkan pada acuan kriteria, kemampuan seorang siswa tidak dibandingkan dengan siswa yang lainnya dalam suatu kelompok melainkan dibandingkan dengan tingkat kemampuan tertentu yang dijadikan sebagai kriteria.

Pada penilaian hasil belajar untuk tingkat satuan pendidikan SD, digunakan penilaian acuan patokan atau kriteria. Jika nilai yang diperoleh siswa lebih besar atau sama dengan Kriteria Ketuntasan Minimal maka siswa dianggap tuntas. Sebaliknya,

jika nilai yang diperoleh siswa lebih kecil dari Kriteria Ketuntasan Minimal, maka siswa dianggap tidak tuntas.

Sebelum instrumen tes digunakan, terlebih dahulu dilakukan analisis keseluruhan dan analisis butir soal tes.

1. Analisis Keseluruhan Tes

Analisis keseluruhan tes terdiri dari 2, yaitu validitas dan reliabilitas. Validitas adalah suatu ukuran yang menunjukkan tingkat-tingkat kevalidan suatu instrumen. Sebuah tes dikatakan valid apabila tes tersebut mengukur apa yang hendak diukur. Validitas secara garis besar dapat dibedakan menjadi dua, yaitu validitas internal dan validitas eksternal. Validitas internal atau validitas logis merujuk pada kondisi sebuah instrumen yang memenuhi syarat berdasarkan hasil penalaran atau rasional. Validasi logis dilakukan oleh ahli atau pakar. Sedangkan validitas eksternal atau validitas empiris (uji coba) didasarkan pada fakta atau pengalaman. Cara untuk menghitung validitas untuk tes tipe pilihan ganda menggunakan koefisien korelasi biserial (y_{pbi}) dan tes tipe uraian menggunakan *Product Moment*.

Sedangkan reliabilitas adalah tingkat konsistensi suatu tes, yakni sejauh mana tes dapat dipercaya untuk menghasilkan skor yang relatif tidak berubah walaupun dites pada situasi yang berbeda-beda. Menurut Kusaeri dan Suprananto (2012: 177), faktor yang mempengaruhi reliabilitas skor tes diantaranya: (1) semakin banyak jumlah butir soal, semakin ajeg suatu tes, (2) semakin lama waktu tes, semakin ajeg, (3) semakin sempit range kesukaran butir soal, semakin besar keajegan, (4) soal-soal yang saling berhubungan akan mengurangi keajegan, dan (5) semakin objektif dalam pemberian skor, semakin besar keajegan. Cara untuk menghitung reliabilitas tes tipe pilihan ganda menggunakan Kuder-Richardson (K-R.20 dan K-R.21) dan tes tipe uraian menggunakan Alfa Cronbach.

2. Analisis Butir Soal

Analisis butir soal merupakan suatu kegiatan yang dilakukan untuk meyakinkan bahwa butir-butir soal bermutu dan memenuhi kriteria yang ditentukan. Analisis butir soal mencakup tingkat kesukaran, daya beda, dan penyebaran pilihan jawaban atau frekuensi jawaban pada pada setiap pilihan jawaban.

a. Tingkat Kesukaran (TK)

Tingkat kesukaran soal adalah peluang menjawab benar suatu soal pada tingkat kemampuan tertentu yang biasanya dalam bentuk indeks. Menurut Aiken (Kusaeri dan Suprananto, 2012:174), indeks tingkat kesukaran umumnya dinyatakan dalam bentuk proporsi yang besarnya berkisar dari 0 sampai 1. Semakin besar indeks tingkat kesukaran yang diperoleh dari hasil hitungan, berarti semakin mudah soal itu. Rumus yang digunakan untuk mengitung TK, yaitu (Nitko dalam Kusaeri dan Suprananto, 2012:174):

- 1) Tes tipe pilihan ganda:

$$TK = \frac{\text{jumlah siswa yang menjawab benar butir soal}}{\text{jumlah siswa yang mengikuti tes}}$$

- 2) Tes tipe uraian:

$$Mean = \frac{\text{jumlah skor siswa peserta tes pada suatu soal}}{\text{jumlah siswa yang mengikuti tes}}$$

$$TK = \frac{Mean}{skor maksimum yang ditetapkan}$$

Klasifikasi tingkat kesukaran soal dapat menggunakan keriteria berikut ini:

No	Range Tingkat Kesukaran	Kategori	Keputusan
1	0,00 – 0,30	Sukar	Ditolak/direvisi
2	0,31 – 0,70	Sedang	Diterima
3	0,71 – 1,00	Mudah	Ditolak/direvisi

b. Daya Beda (DB)

Daya pembeda soal adalah kemampuan suatu butir soal dapat membedakan antara siswa yang telah menguasai materi yang ditanyakan dan siswa yang belum menguasai materi yang diujikan. Daya pembeda soal memiliki manfaat, yaitu 1) untuk meningkatkan mutu setiap butir soal melalui data empiriknya; 2) untuk mengetahui seberapa jauh masing-masing butir soal dapat mendekripsi atau menbedakan kemampuan siswa, yaitu siswa yang telah memahami atau siswa yang belum memahami materi yang diajarkan.. Rumus untuk menghitung daya pembeda soal (DP), yaitu (Kusaeri dan Suprananto, 2012:176):

- 1) Tes tipe pilihan ganda:

$$DP = \frac{(BA - BB)}{\frac{1}{2}N} \text{ atau } DP = \frac{2((BA - BB))}{N}$$

dengan ketentuan:

- DP = daya pembeda soal,
 BA = jumlah jawaban benar pada kelompok atas,
 BB = jumlah jawabn benar pada kelompok bawah,
 N = jumlah siswa yang mengerjakan tes.

- 2) Tes tipe uraian

$$DP = \frac{\text{mean skor kelompok atas} - \text{mean skor kelompok bawah}}{\text{skor maksimum soal}}$$

Adapun kriteria yang dapat digunakan adalah sebagai berikut (Cracker dan Algina dalam Kusaeri dan Suprananto, 2012:177) :

No	Range Daya Pembeda	Kategori	Keputusan
1	0,40 – 1,00	Sangat memuaskan	Diterima
2	0,30 – 0,39	Memuaskan	Diterima
3	0,20 – 0,29	Tidak memuaskan	Ditolak/direvisi
4	0,00 – 0,19	Sangat tidak memuaskan	Direvisi total

c. Penyebaran (distribusi) Jawaban

Penyebaran pilihan jawaban dijadikan dasar dalam penelaah soal. Hal ini dimaksudkan untuk menghitung berfungsi tidaknya pilihan jawaban yang tersedia. Suatu pilihan jawaban (pengecoh) dapat dikatakan berfungsi apabila: (1) pengecoh paling tidak dipilih oleh 5% peserta tes atau siswa, dan (2) pengecoh lebih banyak dipilih oleh kelompok siswa yang belum memahami materi yang diujikan.

KESIMPULAN DAN SARAN

Berdasarkan pembahasan, diperoleh kesimpulan sebagai berikut:

1. Pengukuran (*Measurement*) merupakan suatu kegiatan pemberian atau penetapan angka pada objek yang diukur yang disesuaikan dengan kriteria-kriteria tertentu sesuai dengan objek tersebut. Penilaian (*Asessment*) adalah suatu

kegiatan membandingkan atau menerapkan hasil pengukuran untuk memberikan nilai terhadap objek penilaian. Evaluasi merupakan kegiatan atau proses untuk menilai sesuatu yang mencakup dua kegiatan, yaitu kegiatan pengukuran dan penilaian.

2. Unsur-unsur yang dinilai dalam pembelajaran matematika adalah ulangan harian, ulangan tengah semester, tugas, dan ulangan akhir semester.
3. Jenis instrumen yang digunakan adalah tes dan non-tes. Instrumen tes terdiri dari tes objektif tipe pilihan ganda dan tes subjektif tipe uraian. Sementara intrumen non-tes menggunakan angket konsep diri dengan skala Likert.
4. Teknik pengumpulan data pada instrument tes menggunakan tes tertulis dan non-tes menggunakan angket (*questioner*).
5. Ada beberapa langkah dalam menganalisis hasil non tes, yaitu: a) mentransfer data kualitatif hasil angket menjadi data kuantitatif; dan b) membandingkan total skor (rerata skor) tiap siswa dengan total skor netral (rerata skor netral).
6. Ada dua langkah yang digunakan untuk menganalisis hasil tes, yaitu: a) menskor, b) penafsiran skor tes.

REFERENSI

- Arifin, Z. 2012. *Evaluasi Pembelajaran*. Bandung: PT. Remaja Rosdakarya.
- Campbell, P.B.1967. *School and Self-Concept*. (http://www.ascd.org/ASCD/pdf/journals/ed_lead/el_196703_campbell.pdf), diakses tanggal 20 November 2019.
- Hamzah, A. 2014. *Evaluasi Pembelajaran Matematika*. Jakarta: Rajawali Pers.
- Kusaeri dan Suprananto. 2012. *Pengukuran dan Penilaian Pendidikan*. Yogyakarta: Graha Ilmu.
- Mardapi, Djemari. 2008. *Teknik Penyusunan Instrumen Tes Dan Non Tes*. Yogyakarta: Mitra Cendekia
- Sudijono, A. 2013. *Pengantar Evaluasi Pendidikan*. Jakarta: Rajawali Pers.
- Suherman, E dan Sukjaya, Y. 1990. *Petunjuk Praktis untuk Melaksanakan Evaluasi Pendidikan Matematika*. Bandung: Wijayakusumah 157.

Wiggins, G.1984. *A True Test: Toward More Authentic and Equitable Assessment*. Phi Delta Kappan 70, (9) 703 – 713. (<https://pdfs.semanticscholar.org/8614/15814ce733d118899dd5163ce289302b2760.pdf>), diakses tanggal 20 November 2019.

UU No. 20 Tahun 2003 tentang Sistem Pendidikan Nasional.

Permendikbud No. 66 tahun 2013 tentang Penilaian Hasil Belajar oleh Satuan Pendidikan dan Penilaian Hasil Belajar oleh Pemerintah.