

BAB II

LANDASAN TEORI

A. Evaluasi

1. Pengertian evaluasi

Dunia pendidikan sangat erat hubungannya dengan proses evaluasi. Bloom et. al dalam Daryanto (2010: 01) menyatakan bahwa evaluasi dalam pengumpulan datanya dilakukan secara sistematis, setelah itu kemudian dapat ditetapkan perubahan pada siswa. Evaluasi dilaksanakan berdasarkan urutan-urutan atau langkah yang telah ditentukan. Perubahan pada siswa dapat dilihat dengan cara mengukur dan menilai baik dari aspek kognitif, afektif maupun psikomotor.

Sejalan dengan pendapat di atas, Arikunto (2010: 3) mengemukakan bahwa mengukur dan menilai terdapat dalam kegiatan evaluasi. Kegiatan mengukur menghasilkan data berupa angka dan bersifat kuantitatif, sedangkan menilai merupakan deskripsi dari data pengukuran dan bersifat kualitatif, seperti yang dikatakan oleh bahwa evaluasi mencakup dua kegiatan, yaitu mengukur dan menilai. Mengukur mengacu pada satu ukuran dan bersifat kuantitatif. Mengukur dilakukan menggunakan alat ukur. Sedangkan menilai mengacu pada baik atau buruk dan bersifat kualitatif. Menilai mengacu pada kualitas sesuatu yang didasarkan pada pengalaman. Pendapat yang lainnya

dikemukakan Purwanto (2010: 3-4) mengatakan ada tiga aspek yang perlu diperhatikan, yaitu:

- a. Kegiatan evaluasi merupakan proses yang sistematis

Kegiatan evaluasi dilakukan pada awal program, saat program berlangsung dan pada akhir program. Selain itu, evaluasi juga bersifat terencana dan berkesinambungan.

- b. Kegiatan evaluasi memerlukan berbagai informasi atau data terkait objek yang sedang dievaluasi.

Evaluasi memerlukan data yang sesuai dengan maksud dan tujuan yang telah ditetapkan. Data yang terkumpul harus shahih dan bersifat objektif sehingga dapat diambil keputusan yang tepat.

- c. Kegiatan evaluasi tidak dapat dilepaskan dari tujuan-tujuan evaluasi
Tujuan pembelajaran dirumuskan sebelum dilakukan kegiatan evaluasi. Setelah tujuan ditentukan maka dapat diketahui suatu kriteria yang dijadikan acuan dalam melakukan evaluasi.

Berdasarkan pendapat beberapa ahli, dapat disimpulkan bahwa evaluasi merupakan kegiatan yang mencakup pengukuran dan penilaian yang dilakukan secara sistematis sesuai dengan kriteria tertentu untuk mengetahui tingkat perkembangan siswa dan pembelajaran itu sendiri.

2. Tujuan dan Fungsi Evaluasi

Evaluasi mempunyai tujuan dan fungsi bagi pelaksanaan pembelajaran sesuai dengan kebutuhan. Arikunto (2010: 10)

mengemukakan bahwa penilaian yang diartikan sama dengan evaluasi menyebutkan ada empat tujuan dan fungsi evaluasi, yaitu:

a. Evaluasi berfungsi selektif

Evaluasi berfungsi selektif dalam pembelajaran memiliki tujuan, yaitu:

- 1) Untuk memilih siswa yang dapat diterima disekolah tertentu
- 2) Untuk memilih siswa yang dapat naik kelas atau tingkat berikutnya.
- 3) Untuk memilih siswa yang seharusnya mendapat beasiswa.
- 4) Untuk memilih siswa yang sudah berhak meninggalkan sekolah, dan sebagainya.

b. Evaluasi berfungsi diagnostik

Evaluasi mempunyai tujuan untuk mengetahui kelebihan, kelemahan dan penyebab kelemahan baik kelemahan dari segi materi, siswa, cara mengajar dan sebagainya, sehingga guru dapat mencari cara untuk mengatasinya.

c. Evaluasi berfungsi sebagai penempatan

Evaluasi ini bertujuan mengetahui kekurangan dan kelebihan siswa, untuk kemudian mengelompokkan siswa berdasarkan, kemampuan, bakat, minat dan lain sebagainya.

d. Evaluasi berfungsi sebagai pengukur keberhasilan

Evaluasi ini bertujuan untuk mengetahui tingkat keberhasilan program. Keberhasilan program tersebut dipengaruhi oleh beberapa

faktor yaitu faktor guru, metode mengajar, kurikulum, sarana dan sistem administrasi.

Berdasarkan uraian di atas, dapat diambil kesimpulan bahwa evaluasi mempunyai 4 fungsi dilihat dari kebutuhan evaluator untuk mendapatkan data sesuai program yang terkait. Empat fungsi tersebut yaitu berfungsi sebagai selektif, diagnostik, sebagai penempatan dan sebagai pengukur keberhasilan.

3. Perangkat Evaluasi

Pelaksanaan evaluasi dilakukan dengan menggunakan perangkat evaluasi. Perangkat diartikan sebagai perlengkapan, menurut Kamus Besar Bahasa Indonesia (KBBI). Dalam memperoleh data yang dibutuhkan untuk evaluasi dapat diperoleh dari perlengkapan yang mendukung kegiatan evaluasi.

4. Alat (Instrumen)

Instrumen yang digunakan dalam melakukan evaluasi harus dapat memberikan data sesuai dengan kebutuhan. Selain itu, data yang diperoleh juga harus menggambarkan keadaan sesungguhnya. Arikunto (2010: 26) mengatakan bahwa instrumen evaluasi yang baik adalah instrumen yang mampu mengevaluasi sesuatu secara objektif. Penggunaan instrumen membutuhkan cara atau teknik, yang disebut dengan teknik evaluasi. Teknik evaluasi dibagi menjadi dua, yaitu teknik tes dan teknik non tes. Penggunaan teknik tes maupun non tes disesuaikan dengan kebutuhan. Teknik tes berfungsi sebagai alat

pengumpul data yang mempunyai sifat lebih resmi dibandingkan dengan teknik non tes. Pada pembelajaran di dalam kelas, evaluasi dengan teknik tes berfungsi untuk mengukur siswa dan untuk mengukur tingkat keberhasilan program pembelajaran.

B. Matematika di Sekolah Dasar

Matematika merupakan salah satu cabang ilmu yang bermanfaat dalam kehidupan sehari-hari. Matematika juga merupakan salah satu program studi yang diajarkan di Sekolah Dasar. Russefendi dalam Suwangsih dan Tiurlina (2006:3) berpendapat bahwa matematika terbentuk karena pikiran-pikiran manusia, yang berhubungan dengan ide, proses, dan penalaran. Kegiatan dalam dunia rasio (penalaran) lebih ditekankan, bukan pada hasil eksperimen atau observasi. Mata pelajaran matematika diajarkan berdasarkan jenjangnya, mulai dari yang paling mudah hingga menuju yang paling sulit, dari yang konkret menuju yang abstrak.

Pembelajaran matematika di Sekolah Dasar memiliki batasan ruang lingkup sesuai dengan tingkat berfikir pada jenjang Sekolah Dasar. Depdiknas (2006) menyebutkan ruang lingkup materi matematika Sekolah Dasar sebagai berikut: (1) bilangan, (2) geometri, (3) pengolahan data cakupan data bilangan antara lain bilangan dan angka, perhitungan dan perkiran. Cakupan geometri antara lain bangun dua dimensi, tiga dimensi, transformasi dan simetri serta lokasi dan susunan berkaitan dengan koordinat.

Cakupan pengukuran berkaitan dengan perbandingan kuantitas suatu obyek, penggunaan suatu ukuran dan pengukuran.

Berdasarkan pemaparan di atas, maka matematika di Sekolah Dasar merupakan suatu bidang studi yang diajarkan di sekolah dasar dengan berdasarkan ruang lingkup yang ditentukan sesuai jenjang Sekolah Dasar. Matematika di sekolah dasar berkaitan dengan angka, objek, dan hitungannya yang menekankan kegiatan dalam dunia rasio. pembelajaran matematika pada jenjang Sekolah dasar, dapat membuat siswa lebih kritis, disiplin dan teliti dalam memecahkan suatu permasalahan.

C. Tes

1. Pengertian Tes

Istilah tes digunakan di berbagai bidang termasuk dalam bidang pendidikan. Guru sebagai pelaku pendidikan sebaiknya mengetahui tingkat kemampuan siswanya. Tingkat kemampuan siswa tersebut dapat diketahui dengan pelaksanaan tes. Diungkapkan oleh Arifin (2013: 118) bahwa tes merupakan suatu teknik atau cara yang digunakan dalam rangka melaksanakan kegiatan pengukuran, yang didalamnya terdapat berbagai pertanyaan, atau serangkaian tugas yang harus dikerjakan atau dijawab oleh siswa untuk mengukur aspek perilaku siswa. Rumusan yang dapat diambil dari penjelasan di atas yaitu: *Pertama*, tes disusun secara sistematis dan digunakan dalam rangka pengukuran; *Kedua*, di dalam tes terdapat berbagai pertanyaan atau pernyataan yang ditujukan kepada

siswa; *Ketiga*, tes digunakan untuk mengukur suatu aspek perilaku siswa; *Keempat* hasil tes siswa perlu diberi skor dan nilai.

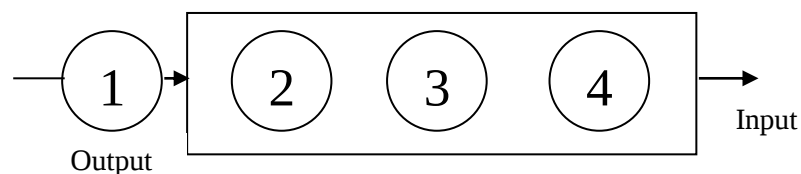
Guru dapat mengetahui tingkat kemampuan siswa melalui pelaksanaan tes. Tes terdiri dari pertanyaan atau pernyataan atau tugas yang harus dikerjakan oleh siswa. Tes merupakan prosedur yang bersifat sistematis dan objektif. Tes juga digunakan untuk mengukur perilaku dan kemampuan siswa. Tes digunakan untuk mendapatkan yang objektif.

2. Jenis Tes

Teknik tes yang digunakan dalam evaluasi pembelajaran dikelas mempunyai dua fungsi yaitu untuk mengukur siswa dan mengukur keberhasilan program pembelajaran. Ditinjau dari segi kegunaan untuk mengukur siswa, Daryanto (2010: 36-42) membagi tes menjadi tiga jenis, yaitu:

a. Tes diagnostik

Tes diagnostik digunakan untuk mengetahui kelemahan-kelemahan siswa untuk kemudian dilakukan tindakan yang tepat. Sekolah berfungsi sebagai transformasi, yaitu membentuk siswa menjadi lebih baik berdasarkan tujuan yang telah ditetapkan maka dapat digambarkan dalam diagram di bawah ini:



Daryanto (2010: 37)

Keterangan:

Tes diagnostik ke-1: dilakukan pada calon siswa sebagai input, untuk mengetahui tingkat pengetahuan dasar yang dikuasai.

Tes diagnostik ke-2: dilakukan terhadap calon siswa yang akan mulai mengikuti program. Tes diagnostik berfungsi sebagai tes penempatan (*placement test*) yaitu menentukan pengelompokan siswa berdasarkan kriteria tertentu.

Tes diagnostik ke-3: dilakukan pada siswa yang sedang belajar. Dengan melakukan diagnostik maka guru dapat mengetahui materi yang belum dikuasai siswa dan penyebab siswa tersebut belum menguasainya. Setelah itu guru dapat memberikan bantuan yang diperlukan

Tes diagnostik ke-4: dilakukan pada waktu siswa akan mengakhiri pembelajaran. Dengan tes ini guru dapat mengetahui tingkat pemahaman siswa terhadap materi yang telah diajarkan.

b. Tes formatif

Evaluasi formatif dimaksudkan untuk mengetahui bahwa siswa telah terbentuk setelah mengikuti suatu program tertentu. Evaluasi formatif digunakan di akhir program. Tes ini merupakan *post-test* atau tes akhir. Tes formatif mempunyai beberapa manfaat di dalam pendidikan. Daryanto (2010: 39-41) menyebutkan bahwa ada tiga manfaat tes formatif yaitu bagi siswa, guru, maupun program itu sendiri:

1) Manfaat bagi siswa

- a) Tes formatif digunakan untuk mengetahui materi yang telah dikuasai siswa.
- b) Tes formatif merupakan penguatan (*reinforcement*) bagi siswa, yaitu dengan melakukan tes di akhir program maka siswa akan mengingat kembali materi yang telah disampaikan sehingga akan bertambah membekas diingatan. Tes ini juga dapat memberikan motivasi kepada siswa sehingga lebih giat lagi dalam belajar.
- c) Setelah melakukan tes formatif maka siswa akan mengetahui kelemahannya, kemudian dapat memperbaiki kelemahan tersebut.
- d) Tes formatif memiliki manfaat untuk mendiagnosis, yaitu tes ini dapat menentukan letak kesalahan siswa dalam mengerjakan soal untuk kemudian diperbaiki.

2) Manfaat bagi guru

- a) Mengetahui tingkat pemahaman siswa terkait bahan ajar atau materi yang telah diberikan. Setelah itu, guru dapat memutuskan untuk selanjutnya akan menggunakan cara mengajar yang lama atau mengganti cara mengajarnya.
- b) Mengetahui bagian-bagian dari bahan pelajaran yang belum diterima siswa. Jika bahan pelajaran tersebut merupakan

prasyarat pelajaran lain maka perlu diulang atau diterangkan kembali.

- c) Tes ini dapat meramalkan sukses dan tidaknya seluruh program yang diberikan.

3) Manfaat bagi program

Manfaat bagi program dapat diketahui setelah memperoleh hasil dari dilakukannya tes, manfaat tersebut yaitu :

- a) Dapat mengetahui kesesuaian program yang telah diberikan dengan kecakapan siswa.
- b) Dapat mengetahui kebutuhan pengetahuan-pengetahuan prasyarat yang belum diperhitungkan.
- c) Dapat mengetahui alat, sarana dan prasarana yang diperlukan untuk mempertinggi hasil yang akan dicapai.
- d) Dapat mengetahui ketepatan metode, pendekatan dan alat evaluasi yang digunakan.

c. Tes sumatif

Tes sumatif hampir sama dengan tes formatif. Perbedaannya pada tes sumatif dilakukan pada program yang lebih besar. Tes formatif dapat disamakan dengan ulangan harian, sedangkan tes sumatif dapat disamakan dengan ulangan umum yang biasanya dilaksanakan pada akhir catur wulan atau akhir semester.

Mengacu pada jenis tes di atas maka tes terbagi menjadi tiga jenis berdasarkan dari segi kegunaan untuk mengukur siswa, yaitu

tes diagnosis, tes formatif dan tes sumatif. Tes diagnosis dilakukan di awal program yaitu bertujuan untuk mengetahui kemampuan awal siswa sehingga dapat diambil keputusan terkait program yang akan dilaksanakan. Tes ini dapat berupa tes seleksi masuk sekolah, tes di awal pembelajaran, dan lain sebagainya. Tes formatif dilakukan di akhir program untuk mengetahui tingkat pemahaman siswa terhadap materi yang telah disampaikan. Tes ini memiliki manfaat bagi siswa, bagi guru dan bagi program itu sendiri sehingga dapat diambil keputusan bagaimana program selanjutnya akan dilaksanakan. Contoh tes formatif yaitu ulangan harian. Tes sumatif dalam pelaksanaannya mempunyai maksud yang hampir sama dengan tes formatif, perbedaannya adalah tes ini mempunyai cakupan program yang lebih luas. Contoh tes sumatif yaitu Ulangan Tengah Semester, Ulangan Akhir Semester dan lain sebagainya.

3. Analisis tes

Analisis tes dilakukan untuk mengetahui kualitas suatu tes. Tes yang memiliki kualitas baik, maka akan membantu guru dalam melakukan evaluasi pada proses pembelajaran. Tingkat kualitas suatu tes perlu diketahui, baik secara keseluruhan maupun setiap butir soal yang menjadi bagian dari tes tersebut. Arifin (2013: 246) menyatakan bahwa cara mengetahui kualitas suatu tes dapat dilakukan melalui analisis kualitas tes. Tes diharapkan dapat menggambarkan perilaku subjek yang diteliti dan menghasilkan nilai yang objektif serta akurat, jika tes yang

digunakan guru kurang baik, maka hasil yang diperoleh juga kurang baik. Penjelasan di atas, dapat disimpulkan guru dalam menggunakan tes harus memiliki kualitas yang baik dilihat dari berbagai segi. Surapranata (2009: 1-11) menyatakan bahwa analisis soal mencakup dua cara, yaitu analisis kualitatif (*qualitatif control*) dan analisis kuantitatif (*quantitative control*), penjelasannya sebagai berikut:

a. Analisis kualitatif

Analisis merupakan salah satu tingkatan berfikir untuk memahami sesuatu yang diamati. Kuswana (2014: 115) mengatakan bahwa memahami atau analisis yaitu mendeskripsikan susunan dalam artian pesan pembelajaran, mencakup oral, tulisan dan komunikasi grafik menggunakan pengetahuan yang telah diterima. Analisis kualitatif dalam pendeskripsiannya menggunakan bahasa kualitatif yaitu berupa kata-kata yang menunjukkan kualitas, bukan berupa angka-angka yang bersifat kuantitatif.

Soal yang akan diujikan kepada siswa perlu dilakukan dengan analisis kualitatif. Analisis kualitatif dilakukan untuk mengetahui berfungsi atau tidaknya soal berdasarkan tujuan yang telah ditentukan. Surapranata (2009:1) berpendapat bahwa analisis kualitatif dapat dikategorikan dari segi materi, konstruksi dan bahasa. Segi materi atau isi mencakup analisis substansi keilmuan yang ditanyakan dalam soal serta tingkat kemampuan yang sesuai dengan soal, segi konstruksi mencakup analisis sebagai penelaahan

yang berkaitan dengan teknik penulisan soal, segi bahasa analisis berkaitan dengan penggunaan bahasa Indonesia yang baik dan benar menurut ejaan yang disempurnakan.

1) Validitas isi

Suatu tes dikatakan valid apabila memiliki validitas yang tinggi yaitu dapat mengukur sesuatu yang hendak diukur. Validitas suatu tes dapat diketahui dengan cara mengkaji isi tes tersebut. Surapranata (2009: 51) mengungkapkan bahwa suatu tes memiliki validitas isi apabila mengukur sesuai dengan domain dan tujuan khusus tertentu yang sama dengan isi pelajaran yang telah diberikan di kelas. Prosedur dalam penentuan validitas isi yaitu: (a) mendefinisikan domain yang hendak diukur; (b) menentukan domain yang diukur oleh masing-masing soal; (c) membandingkan masing-masing soal dengan domain yang telah ditetapkan.

Validitas isi ditentukan dengan melihat kesesuaian soal-soal yang digunakan dengan sampel atribut yang diukur. Terdapat dua hal yang berpengaruh pada validitas isi, yaitu tes itu sendiri dan proses yang mempengaruhi dalam merespon tes. Validitas isi dapat diperoleh dengan melihat soal-soal yang membentuk tes itu. Pemenuhan validitas isi dapat tercapai jika suatu tes dapat mengukur yang seharusnya digunakan.

2) Validitas konstruksi

Validitas konstruksi merupakan analisis validitas berdasarkan kriteria-kriteria yang telah ditentukan. Surapranata (2009: 53) menjelaskan bahwa validitas konstruksi merupakan kemampuan soal-soal untuk dapat mengukur setiap aspek berfikir yang diuraikan dalam Standar Kompetensi, Kompetensi Dasar, maupun indikator yang terdapat dalam kurikulum. Konstruksi yang dimaksud dalam validitas ini yaitu berupa rekaan psikologis yang berkaitan dengan aspek-aspek ingatan pemahaman, aplikasi, analisis, sintesis, dan evaluasi. Sejalan degan pendapat tersebut, Arikunto (2010: 68) berpendapat bahwa validitas konstruksi dapat diketahui dengan cara merinci dan memasang setiap butir soal dengan setiap aspek dalam Tujuan Instruksional Khusus.

Validitas kontsruksional dilakukan berdasarkan logika, bukan melalui pengalaman. Setiap soal yang menyusun tes tersebut harus dapat mengukur aspek berfikir dan harus sesuai dengan kriteria yang telah ditentukan.

3) Penggunaan bahasa

Penggunaan bahasa dalam penyusunan soal merupakan hal yang harus diperhatikan. Bahasa yang digunakan merupakan bahasa baku dan sesuai dengan Ejaan Yang Disempurnakan (EYD). Surapranata (2009: 2) mengartikan penggunaan bahasa dalam penyusunan soal yaitu analisis soal yang berkaitan

dengan penggunaan Bahasa Indonesia yang baik dan benar menurut ejaan yang disempurnakan.

Analisis penggunaan bahasa pada soal dilakukan dengan kriteria-kriteria tertentu. Kriteria penggunaan bahasa yang baik disebutkan oleh Badan Standar Nasional Pendidikan (2008) yaitu:

- a) Menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia.
- b) Menggunakan bahasa yang komunikatif.
- c) Menggunakan bahasa yang berlaku setempat/tabu
- a) Pilihan jawaban tidak mengulang kata atau kelompok kata yang sama, kecuali merupakan suatu kesatuan pengertian.

b. Analisis kuantitatif

Analisis kuantitatif dapat dilakukan dengan cara mengumpulkan data yang dilakukan melalui tes. Data tersebut kemudian diolah menggunakan parameter-parameter dalam analisis kuantitatif. Parameter yang digunakan dalam analisis kuantitatif yaitu validitas, reliabilitas, tingkat kesukaran soal dan daya pembeda. Menurut Surapranata (2009:10) parameter dalam analisis kuantitatif ada empat, yaitu validitas, reliabilitas, tingkat kesukaran dan daya beda.

1) Validitas

Soal tes yang baik tentunya memiliki validitas yang tinggi. Validitas dapat diketahui dengan berdasarkan kriteria yang ada. Arikunto (2012: 85) mengatakan bahwa sebuah tes dikatakan memiliki validitas jika hasilnya sesuai dengan kriterium, dalam arti memiliki kesejajaran antara hasil tes tersebut dengan kriterium. Kesejajaran tersebut dapat diketahui dengan cara menghitung data yang telah diperoleh berdasarkan langkah-langkah pengujian validitas untuk kemudian di analisis dan dilihat kesesuaiannya dengan kriterium. Validitas secara kuantitatif disebut juga validitas empirik. Sudijono (2012: 167) berpendapat bahwa validitas soal secara empirik dapat diketahui dengan menganalisis data hasil pengamatan di lapangan, jika terbukti tes hasil belajar itu dapat mengukur yang seharusnya diukur, maka tersebut memiliki validitas yang tinggi.

2) Reliabilitas

Tes harus dapat mengukur kemampuan siswa secara tepat. Alat ukur yang digunakan harus dapat memberikan ketepatan atau keajegan walaupun diujikan beberapa kali dalam waktu yang berbeda, hal ini mengacu pada pendapat Surapranata (2009: 86) bahwa reliabilitas atau keajegan skor merupakan hal yang sangat penting dan menggambarkan tingkat kualitas tes tersebut. Tes memiliki tingkat reliabilitas apabila dites berkali-

kali memberikan hasil yang tetap atau keajegan. Tingkat reliabilitas suatu tes juga digunakan untuk mengambil keputusan lulus tidaknya seorang siswa. Pendapat tersebut diperkuat oleh Sudijono (2012: 95) bahwa suatu tes memiliki reliabilitas apabila skor atau nilai yang diperoleh menunjukkan hasil yang sama atau stabil, walaupun diujikan pada waktu, tempat dan subjek yang berbeda.

Hasil pengujian suatu tes dapat memberikan hasil yang dipengaruhi beberapa faktor, hal tersebut juga mempengaruhi pengujian reliabilitas suatu tes. Sudjana (2010) menyebutkan ada dua faktor, yaitu disebabkan kelemahan tes itu sendiri dan disebabkan perubahan pada diri subjek. Suatu tes diujikan beberapa kali dan memberikan hasil yang berbeda dapat disebabkan karena soal atau pilihan jawaban yang membingungkan, namun dapat juga karena kondisi yang terjadi pada subjek, seperti kondisi fisik subjek atau siswa dalam keadaan sakit, sedang ada masalah dan sebagainya.

Instrumen atau tes dikatakan reliabel atau memiliki reliabilitas yang tinggi apabila diujikan pada kesempatan dan waktu yang berbeda tetap memberikan hasil yang sama. Reliabilitas merupakan salah satu syarat bahwa instrumen atau tes dikatakan baik. Namun terdapat faktor yang mempengaruhi

dalam pengujian reliabilitas, yaitu faktor dari tes itu dan faktor dari subjek tes.

3) Tingkat kesukaran

Tingkat kesukaran suatu soal akan mempengaruhi hasil yang diperoleh siswa. Daryanto (2010: 179) menyatakan bahwa soal yang baik adalah soal yang tidak terlalu mudah atau terlalu sukar. Soal yang terlalu mudah tidak merangsang siswa untuk mempertinggi usaha memecahkannya. Sebaliknya, soal yang terlalu sukar akan menyebabkan siswa menjadi putus asa dan tidak mempunyai semangat mencoba lagi karena diluar jangkauan. Tingkat kesukaran perlu diperhatikan oleh pembuat soal dengan mengacu pada materi dan kemampuan siswa yang akan diukur. Tingkat kesukaran merupakan nilai rata-rata dari suatu distribusi skor kelompok dari suatu soal. Pernyataan serupa juga dikatakan oleh Sudijono (2012: 370) bahwa bermutu atau tidaknya butir-butir item tes hasil belajar pertama-tama dilihat dari derajat kesukaran atau taraf kesulitan yang dimiliki oleh masing-masing item tersebut. Butir-butir soal yang mempunyai tingkat kesukaran sedang merupakan soal yang memiliki kualitas baik.

Taraf kesukaran merupakan hal yang perlu diperhatikan dalam penyusunan tes. Tes yang baik mempunyai taraf

kesukaran yang tidak terlalu sukar atau terlalu mudah sehingga akan dapat mengukur kemampuan siswa.

4) Daya pembeda

Masing-masing anak mempunyai kemampuan yang berbeda-beda, jika satu kelas ada 30 siswa, maka dikelas tersebutpun terdapat 30 siswa dengan kemampuan yang berbeda-beda. Guru penting untuk mengetahui kemampuan masing-masing siswa, baik siswa yang memiliki kemampuan tinggi ataupun rendah. Sudijono (2012: 385-386) menjelaskan bahwa daya pembeda merupakan kemampuan suatu butir item tes hasil belajar untuk dapat membedakan antara siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah. Daya pembeda juga merupakan hal yang penting, karena setiap item dalam tes harus dapat membedakan antara siswa yang kurang pandai dan siswa yang pandai.

D. Penelitian yang Relevan

Menurut penelitian yang dilakukan oleh Tutut Kurniawan, dengan judul “ Analisis Butir Soal Ulangan Akhir Semester Gasal Mata Pelajaran IPS Kelas III SD Negeri gugus Diponegoro Kecamatan Sempor Kabupaten Kebumen”. Hasil penelitian ini menunjukkan bahwa, soal pilihan ganda yang dianalisis ditinjau dari aspek materi, konstruksi, dan bahasa memiliki validitas isi berkategori sangat tinggi. Distribusi jenjang ranah kognitifnya yaitu terdapat 7 (28%) soal berkategori C1, 17 (68%) soal berkategori

C2, dan 1 (4%) soal berkategori C3. Kualitas soal pilihan ganda ditinjau dari aspek validitasnya yaitu terdapat 2 (8%) soal berkategori sangat signifikan, 8 (32%) soal berkategori signifikan, dan 15 (60%) soal berkategori tidak signifikan. Penghitungan aspek reliabilitasnya diperoleh koefisien reliabilitas sebesar 0,68 dengan kriteria rendah. Aspek tingkat kesukarannya yaitu terdapat 17 (68%) soal berkategori mudah, 7 (28%) soal berkategori sedang, dan 1 (4%) soal berkategori sukar. Aspek daya pembedanya yaitu terdapat 7 (28%) soal berkategori baik, 7 (28%) soal berkategori cukup, 10 (40%) soal berkategori jelek, dan 1 (4%) soal berkategori jelek sekali. Aspek efektivitas pengecohnya yaitu terdapat 11 (44%) soal berkategori efektif dan 14 (56%) soal berkategori jelek. Pelaksanaan tes tergolong cukup baik, karena adanya faktor suasana ruang kelas yang kurang tenang dan posisi duduk yang berdekatan.

Menurut penelitian yang dilakukan oleh Zulkifli Matondang, dengan judul “ Validitas dan Reliabilitas Suatu Instrumen Penelitian “. Mendapatkan kesimpulan hasil penelitian bahwa Instrumen merupakan suatu alat yang dipergunakan sebagai alat untuk mengukur suatu obyek ukur atau mengumpulkan data dari suatu variabel. Suatu instrumen dikatakan baik bila valid dan reliabel. Validitas terdiri atas validitas isi, konstruk, empirik. Validitas internal skor butir dikotomi dan skor butir politomi berturut-turut digunakan korelasi biserial dan korelasi product moment. Kriteria suatu butir valid atau tidak valid didasarkan pada nilai r-tabel. Reliabilitas konsistensi gabungan butir untuk skor butir dikotomi dan skor butir politomi berturut-turut digunakan KR-20 dan koefisien Alpha. Interpretasi terhadap

koefisien reliabilitas merupakan intepretasi relatif dalam artian bahwa tidak ada batasan mutlak yang menunjukkan berapa angka koefisien minimal yang harus dicapai agar suatu pengukuran dapat disebut reliabel. Namun, memberikan informasi tentang hubungan varians skor teramati dengan skor sejati kelompok individu.

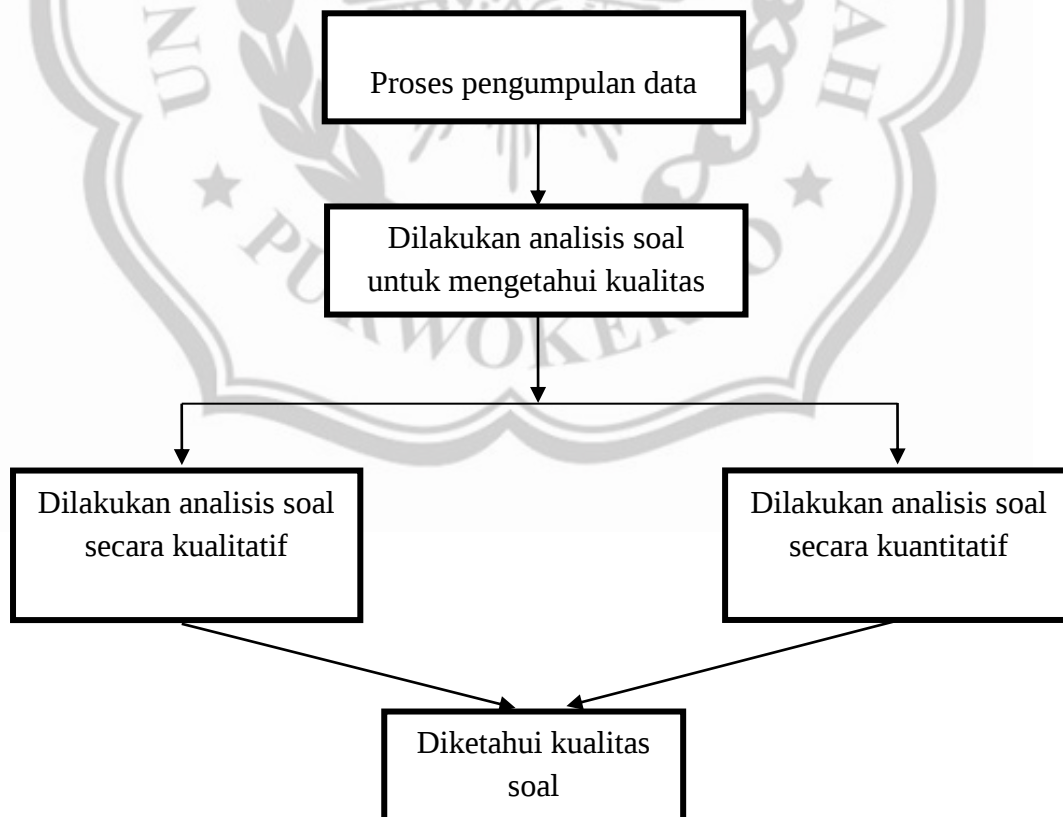
Berdasarkan penelitian Sugiyanta, dengan judul “ *An Evaluation Model Of Educational Quality Assurance At Junior Highs School* “. Hasil penelitian menunjukkan bahwa model evaluasi EQA di sekolah menengah pertama terdiri dari implementasi sistem QA dan kinerja QA. Konstruksi untuk instrumen Implementasi sistem QA terdiri dari perencanaan, pelaksanaan, monitoring dan evaluasi, dan tindakan revisi didasarkan pada analisis faktor eksploratori pada tingkat signifikansi 0.000. Konstruksi untuk instrumen kinerja EQA terdiri dari: sumber daya pengembangan; pengembangan program dan aktivitas; partisipasi, kepuasan, perubahan pengetahuan, perubahan sikap, dan perubahan perilaku komunitas sekolah; sosial, ekonomi, dan sekolah pengembangan lingkungan berdasarkan analisis faktor eksploratori pada tingkat signifikansi 0.000. Kelayakan model evaluasi berada pada kategori baik berdasarkan para ahli, dan penilaian praktisi dan bukti yang ditemukan dalam pengujian lapangan.

Dari penelitian yang relevan di atas, maka kualitas perangkat evaluasi dalam pembelajaran perlu dilakukannya analisis agar mengetahui data yang lebih akurat baik dalam segi kesesuaian soal dengan kompetensi dasar pada kurikulum, penyebaran soal, validitas soal, daya pembeda, dan tingkat kesukaran soal.

E. Kerangka Pikir

Kualitas tes yang baik dapat dilihat melalui analisis soal-soal yang dapat dilakukan oleh guru. Analisa butir soal dilakukan untuk mengkaji dan menelaah soal-soal agar sesuai dengan kebutuhan evaluasi yang akan dilaksanakan. Analisis soal sangat penting karena bias digunakan untuk menentukan dan merevisi butir soal sehingga dapat diperoleh soal yang bermutu sebelum soal digunakan dalam proses evaluasi. Setelah pengumpulan data maka diperoleh informasi bahwa soal-soal dalam tes ulangan akhir semester (UAS) yang telah digunakan dalam proses evaluasi belum dianalisis secara menyeluruh sehingga perlu dilakukan analisis baik secara kualitatif dan kuantitatif, sehingga kualitas soal-soal pada UAS dapat diketahui secara pasti.

Kerangka pikir dapat dilihat pada gambar 2.1 sebagai berikut:



Gambar 2.1. Kerangka Pikir