

BAB 2 KAJIAN TEORI

2.1 *Trends in International Mathematics and Science Study (TIMSS)*

Trends in International Mathematics and Science Study (TIMSS) bertujuan untuk mengukur prestasi matematika dan ilmu pengetahuan alam siswa kelas IV serta kelas VIII di negara-negara peserta (Sari, 2015: 303). Studi TIMSS menyediakan informasi yang berharga bagi suatu negara dalam hal memonitor dan mengevaluasi pembelajaran matematika dan sains dari waktu ke waktu pada setiap tingkatan.

Mullis (2009: 14) menyebutkan manfaat bagi suatu negara yang berpartisipasi dalam TIMSS sebagai berikut:

1. Memiliki data tentang konsep, proses dan sikap siswa terhadap matematika dan sains secara komprehensif dalam lingkup internasional.
2. Mampu menilai perkembangan kemampuan matematika dan sains siswa dalam rentang waktu tertentu secara internasional.
3. Mampu mengidentifikasi aspek yang mempengaruhi perkembangan pengetahuan dan keterampilan matematika dan sains siswa.
4. Memonitor efektivitas pembelajaran pada siswa kelas IV SD dan VIII SMP.
5. Memahami dalam konteks apa saja siswa belajar secara optimal. Studi TIMSS memungkinkan untuk melakukan komparasi sejumlah variabel kebijakan, kurikulum, metode mengajar dan sarana belajar yang berkaitan dengan tingginya prestasi siswa.
6. Menjadikan hasil TIMSS untuk merumuskan isu kebijakan internal sebagai upaya peningkatan kualitas pendidikan.

Bagi Indonesia, manfaat yang dapat diperoleh antara lain untuk mengetahui posisi prestasi siswa Indonesia bila dibandingkan dengan prestasi siswa di negara lain dan faktor-faktor yang mempengaruhinya. Oleh karena itu, hasil studi ini diharapkan dapat digunakan sebagai masukan dalam perumusan kebijakan untuk peningkatan mutu pendidikan di Indonesia.

Dalam TIMSS kerangka penilaian kemampuan dibidang matematika yang diuji digunakan istilah domain konten dan domain kognitif, dimensi konten merupakan kumpulan materi matematika yang dinilai dalam TIMSS. Dimensi konten pada siswa kelas IV SD berbeda dengan kelas VIII SMP berdasarkan keluasan materi yang telah diajarkan di kelas. Dalam Mullis (2013: 19) pada siswa

kelas VIII SMP, dimensi konten matematika meliputi bilangan, aljabar, geometri, data dan peluang.

1. Bilangan

Bilangan merupakan konsep penting dalam kehidupan sehari-hari dan memerlukan pemahaman tentang kuantitas untuk dapat mengoperasikannya. Suatu bilangan dapat direpresentasikan dalam banyak simbol tertulis yang berbeda-beda dan siswa perlu mengenali perbedaan masing-masing interpretasi, membangun keterhubungan diantarnya, dan bernalar menggunakannya, ada tiga topik utama dalam domain bilangan beserta kompetensi yang dinilai oleh TIMSS.

Tabel 3. Kompetensi Domain Bilangan Studi TIMSS

Topik	Kompetensi Yang Dinilai
Bilangan cacah	<ol style="list-style-type: none"> Mendemonstrasikan pemahaman terhadap bilangan cacah dan operasinya (operasi aritmatika, letak nilai, sifat komutatif, asosiatif dan distributif). Menghitung menggunakan bilangan cacah pada situasi masalah tertentu. Menemukan dan menggunakan faktor suatu bilangan, mengidentifikasi bilangan prima, mengevaluasi perangkat dan akar pangkat suatu bilangan hingga 144.
Pecahan, dimensi dan bilangan bulat	<ol style="list-style-type: none"> Mengidentifikasi, membandingkan, atau mengurutkan bilangan rasional (pecahan, desimal dan bilangan bulat), menggunakan berbagai model dan representasi (seperti garis bilangan), dan mengetahui bahwa ada bilangan yang tidak rasional. Menghitung dengan bilangan rasional (pecahan, desimal dan bilangan bulat) termasuk kumpulan permasalahan pada suatu situasi.
Rasio, proporsi dan persen	<ol style="list-style-type: none"> Mengidentifikasi dan menemukan rasio senilai dan membuat model dari suatu situasi menggunakan rasio dan membagi suatu kuantitas menggunakan rasio yang diketahui. Mengonversi di antara persen, proporsi dan pecahan. Menyelesaikan permasalahan yang melibatkan bentuk persen dan proporsi.

(Mullis, 2013: 20)

2. Aljabar

Aljabar merupakan sesuatu yang lekat dalam kehidupan kita, permasalahan dalam kehidupan nyata yang dihadapi oleh siswa dapat diselesaikan dengan model aljabar dan dijelaskan keterhubungannya dengan konsep aljabar. Kompetensi yang dinilai oleh TIMSS pada konten aljabar terbagi menjadi tiga topik sebagai berikut.

Tabel 4. Kompetensi Domain Aljabar Studi TIMSS

Topik	Kompetensi Yang Dinilai
Bentuk dan operasi aljabar	<ol style="list-style-type: none"> 1. Menemukan nilai dari suatu bentuk aljabar ketika diketahui nilai variabelnya. 2. Menyederhanakan bentuk aljabar meliputi penjumlahan, perkalian, bentuk pangkat, dan membandingkan untuk menentukan kesetaraannya. 3. Menggunakan bentuk aljabar untuk merepresentasikan suatu masalah.
Persamaan dan pertidaksamaan	<ol style="list-style-type: none"> 1. Menuliskan bentuk persamaan dan pertidaksamaan dari suatu masalah. 2. Menyelesaikan persamaan, pertidaksamaan dan sistem persamaan linier dua variabel.
Relasi dan fungsi	<ol style="list-style-type: none"> 1. Mengeneralisir pola dari suatu barisan, di antara dua suku yang berdekatan, atau di antara dua barisan bilangan menggunakan bentuk bilangan, kata dan aljabar. 2. Menginterpretasikan, menghubungkan dan membangun representasi fungsi dari suatu tabel, grafik, atau kata-kata. 3. Mengidentifikasi fungsi linier atau non linier, membedakan perbedaan sifat suatu fungsi berdasarkan tabel, grafik atau persamaan dan menginterpretasikan makna dan gradien dan perpotongan di sumber-y pada fungsi linier

(Mullis, 2013: 21)

3. Geometri

Geometri merupakan bagian penting dari matematika banyak aplikasi geometri yang relevan dengan pekerjaan dan kehidupan sehari-hari. Kompetensi yang dinilai oleh TIMSS pada konten geometri terbagi menjadi tiga topik sebagai berikut.

Tabel 5. Kompetensi Domain Geometri Studi TIMSS

Topik	Kompetensi Yang Dinilai
Bentuk geometri	<ol style="list-style-type: none"> 1. Mengidentifikasi berbagai jenis sudut dan menggunakan hubungan antara sudut pada garis dan gambar geometri. 2. Mengidentifikasi sifat geometri dimensi dua dan tiga termasuk simetri lipat dan simetri putar. 3. Mengidentifikasi segitiga dan segiempat yang kongruen serta ukurannya yang bersesuaian dan mengidentifikasi segitiga dan segiempat yang sebangun dan menggunakan

	<p>sifat-sifatnya.</p> <p>4. Menghubungkan bentuk dimensi tiga dengan representasi dimensi duanya.</p> <p>5. Menggunakan sifat-sifat geometri, termasuk teorema <i>phytagoras</i> untuk menyelesaikan masalah.</p>
Pengukuran geometri	<p>1. Melukis dan mengestimasi ukuran dari suatu sudut, ruas garis dan keliling, dan mengestimasi luas permukaan volume.</p> <p>2. Memilih dan menggunakan rumus pengukuran yang tepat untuk menentukan keliling, luas, luas permukaan dan volume dan menentukan ukuran dari bidang gabungan.</p>
Lokasi dan perpindahan	<p>1. Menentukan letak titik koordinat dan menyelesaikan masalah yang menyertakan titik pada bidang kartesius.</p> <p>2. Mengenali dan menggunakan transformasi geometri (translasi, refleksi dan rotasi) pada bentuk dua dimensi.</p>

(Mullis, 2013: 22)

4. Data dan peluang

Kemampuan matematika siswa kelas VIII SMP pada materi data dan peluang diharapkan telah memiliki pemahaman awal tentang konsep data dan peluang. Dalam kehidupan sehari-hari peluang merupakan materi yang sangat erat kaitannya dalam mengambil sebuah keputusan. Kompetensi yang dinilai oleh TIMSS pada konten data dan peluang terbagi menjadi tiga topik sebagai berikut.

Tabel 6. Kompetensi Domain Data dan Peluang Studi TIMSS

Topik	Kompetensi Yang Dinilai
Karakteristik data	<p>1. Mengidentifikasi dan membandingkan karakteristik beberapa dan ditinjau dari rata-rata, modus, range, dan bentuk distribusi.</p> <p>2. Mengkalkulasi, menggunakan atau menginterpretasikan rata-rata, median, modus dan range untuk memecahkan masalah.</p>
Interprestasi data	<p>1. Membaca data dari suatu jenis tampilan data visual</p> <p>2. Menggunakan dan menginterpretasikan data untuk menyelesaikan masalah</p> <p>3. Mengidentifikasi dan mendeskripsikan pendekatan dalam mengatur dan menampilkan data yang dapat berakibat kesalahan interpretasi.</p>
Peluang	<p>1. Menentukan peluang suatu kejadian dalam istilah umum.</p>

	2. Menggunakan data, termasuk data eksperimen untuk memperkirakan peluang kejadian dimasa depan.
	3. Jika proses dilakukan secara acak, dapat menentukan peluang suatu kejadian.

(Mullis, 2013: 23)

Pada penelitian ini peneliti melakukan pengembangan soal pada domain kognitif aljabar dengan 3 topik yaitu bentuk dan operasi aljabar, persamaan dan pertidaksamaan, relasi dan fungsi. Prestasi matematika siswa berdasarkan studi TIMSS tidak hanya dinilai berdasarkan kemampuan siswa mengenali semua domain konten dengan baik, tetapi siswa juga harus menunjukkan kemampuan pada domain kognitif. Seperti yang diungkapkan Budiman dan Jailani (dalam Hamidy, 2016: 68) soal-soal matematika yang dikembangkan oleh studi TIMSS menuntut siswa untuk berfikir tingkat rendah sampai tingkat tinggi. Tabel 7 menunjukkan proporsi kemampuan yang diuji pada dimensi kognitif dalam studi TIMSS.

Tabel 7. Proporsi Kemampuan yang Diuji pada Domain Kognitif dalam Studi TIMSS

Domain	Proporsi	Topik
Knowing	35 %	<i>recall</i> yaitu memahami definisi, sifat-sifat, terminologi, serta notasi-notasi dalam matematika (contoh : $a \times b = ab$, $a + a + a = 3a$)
		<i>Recognize</i> mengenal bilangan, ekspresi, jumlah, dan bentuk serta mengenal entitas matematika
		<i>Classify/order</i> mengklasifikasikan objek, bangun, bilangan, berdasarkan sifat-sifat tertentu.
		<i>Compute</i> menghitung prosedur-prosedur algoritmik, +, -, x, :, pada bilangan bulat, pecahan, dan desimal serta melaksanakan prosedur aljabar sederhana.
		<i>Retrieve</i> mengambil informasi dari grafik, tabel, atau sumber lain yang sederhana.
		<i>Measure</i> , yaitu menggunakan instrumen-instrumen pengukuran dan memilih unit pengukuran yang sesuai.
Applying	40 %	<i>Determine</i> memilih operasi, metode serta strategi yang tepat dalam memecahkan masalah dimana prosedur, metode atau algoritma untuk menyelesaikan masalah tersebut sudah diketahui
		<i>Represent/model</i> menyajikan informasi matematika atau data dalam bentuk tabel atau grafik, membuat persamaan, pertidaksamaan, menggunakan model matematika untuk memecahkan masalah rutin, menghasilkan representasi setara untuk entitas matematika yang diberikan atau yang saling berhubungan.
		<i>Implement</i> menerapkan strategi dan operasi untuk memecahkan masalah yang melibatkan konsep dan prosedur matematika.
Reasoning	25 %	<i>Analyze</i> mendeskripsikan atau menggunakan hubungan antar bilangan, ekspresi aljabar, jumlah dan bentuk.

		<i>Integrate/synthesize</i> membuat hubungan dari elemen-elemen pengetahuan, representasi terkait dan prosedur untuk memecahkan masalah.
		<i>Evaluate</i> mengevaluasi alternatif strategi pemecahan masalah dan solusi pemecahannya.
		<i>Draw conclusions</i> membuat kesimpulan yang valid berdasarkan informasi dan bukti.
		<i>Generalize</i> membuat pernyataan yang mewakili hubungan lebih umum dan istilah lebih luas yang berlaku.
		<i>Justify</i> memberikan argumen matematis untuk mendukung strategi atau solusi.

(Mullis, 2009: 42)

Untuk penelitian ini domain kognitif yang dipilih peneliti dalam mengembangkan soal matematika berbasis studi TIMSS dapat dilihat pada Tabel 8 berikut:

Tabel 8. Kemampuan pada Domain Kognitif dalam Studi TIMSS yang Akan di Kembangkan

Domain	Proporsi	Topik
Knowing	35 %	<i>Compute</i> , menghitung prosedur-prosedur algoritmik, $+$, $-$, \times , \div pada bilangan bulat, pecahan, dan desimal serta melaksanakan prosedur aljabar sederhana.
		<i>Classify/order</i> , mengklasifikasikan objek, bangun, bilangan, berdasarkan sifat-sifat tertentu.
Applying	40%	<i>Determine</i> , memilih operasi, metode serta strategi yang tepat dalam memecahkan masalah dimana prosedur, metode atau algoritma untuk menyelesaikan masalah tersebut sudah diketahui.
		<i>Implement</i> , menerapkan strategi dan operasi untuk memecahkan masalah yang melibatkan konsep dan prosedur matematika.
		<i>Represent/ model</i> , menyajikan informasi matematika atau data dalam bentuk tabel atau grafik, membuat persamaan, pertidaksamaan, menggunakan model matematika untuk memecahkan masalah rutin, menghasilkan representasi setara untuk entitas matematika yang diberikan atau yang saling berhubungan.
		<i>Determine</i> , memilih operasi, metode serta strategi yang tepat dalam memecahkan masalah dimana prosedur, metode atau algoritma untuk menyelesaikan masalah tersebut sudah diketahui.
Reasoning	25%	<i>Analyze</i> , mendeskripsikan atau menggunakan hubungan antar bilangan, ekspresi aljabar, jumlah dan bentuk.

		<i>Integrate/ synthesize</i> , membuat hubungan dari elemen-elemen pengetahuan, representasi terkait dan prosedur untuk memecahkan masalah.
--	--	---

Bentuk soal-soal dalam TIMSS adalah pilihan ganda dengan 4 pilihan jawaban, isian singkat dan uraian. Isian singkat dan uraian sering disebut “*constructed response*”. Untuk soal uraian akan diberi skor 2 untuk jawaban yang lengkap dan benar, skor 1 untuk jawaban yang benar namun kurang lengkap dan skor 0 untuk jawaban yang salah atau tidak menjawab (Sari, 2015: 305).

TIMSS tidak hanya mengukur kemampuan literasi sains dan matematika siswa saja, namun juga dikaitkan dengan sejumlah pertanyaan dalam bentuk angket kepada siswa dan orang seputar latar belakang kehidupan sosial ekonomi, sekolah, pengalaman belajar dan lingkungan belajar. Kepala sekolah dan guru yang mengajar siswa yang bersangkutan juga memberikan informasi seputar bagaimana pembelajaran di kelas, sistem kurikulum, serta fasilitas yang diberikan sekolah kepada siswa didiknya (Kemikbud, 2015: 1).

2.2 Kualitas Soal Mengacu pada Standar TIMSS

Kerangka penilaian TIMSS disusun berdasarkan domain kognitif dan domain isi. Untuk soal yang diberikan siswa harus memberikan tanggapan terhadap soal-soal tersebut.

Soal-soal TIMSS dikembangkan melalui pertemuan yang melibatkan masukan dan ahli matematika dan ahli pengukuran internasional. Para penyelenggara dan ahli dari 10 negara, memastikan bahwa soal-soal yang diujikan mencerminkan pemikiran saat ini dan merupakan prioritas dalam bidang matematika. Soal-soal tersebut dikembangkan dan ditinjau berulang-ulang dengan satu uji coba yang melibatkan 43 negara. Berbagai upaya telah dilakukan untuk memastikan bahwa soal-soal yang diberikan dapat mewakili kurikulum negara-negara peserta.

Menurut Martin (2000: 74) soal TIMSS memiliki kualitas/kepraktisan tinggi harus memiliki beberapa kriteria, antara lain:

1. Kata-kata yang digunakan dalam soal tidak ambigu.
2. Tuntutan membaca yang tepat.
3. Grafik yang jelas.
4. Kunci soal dan panduan penilaian yang tidak berubah-ubah (tetap).

Dalam soal uraian siswa dituntut untuk memikirkan jawaban tanpa ada opsi pilihan jawaban oleh karena itu soal uraian lebih valid dari pada soal pilihan ganda. Kualitas jawaban sangat bergantung pada kemampuan penilai dalam menilai soal.

Martin (2000: 18) menjelaskan cara penulisan soal TIMSS bentuk uraian adalah:

1. Bahasa dalam soal disesuaikan dengan tingkat kelas yang diuji.
2. Soal harus jelas, hindari menggunakan kata-kata yang samar seperti membahas, komentar dan lain-lain. Hal tersebut menyebabkan variasi jawaban dan tanggapan yang luas.
3. Siswa harus mampu menyelesaikan soal dalam waktu 1-3 menit.
4. Jika menggunakan soal yang berhubungan dengan kehidupan sehari-hari, maka harus disesuaikan dengan tingkat kelas yang diuji.
5. Hindari pertanyaan yang menimbulkan jawaban yang luas dan tidak akurat.
6. Beri indikasi, tepat, sejauh atau tingkat detail jawaban yang diharapkan.
7. Tulis jawaban pertanyaan yang ada dengan bahasa, pengetahuan, dan keterampilan siswa secara tepat.
8. Buatlah paduan penilaian untuk soal-soal yang telah dibuat.

Soal-soal yang telah dibuat akan ditinjau kembali oleh *TIMSS & PIRLS International Study Center Staff*, koordinator dan konsultan peninjauan soal matematika dan sains, *The Science and Mathematics Item Review Committee (SMIRC)*, dan *The National Research Coordinator*. Dalam peninjauan soal uraian, para pengulas harus:

1. Memeriksa penilaian mereka tentang klasifikasi TIMSS sesuai dengan yang ditunjukkan oleh tim penulis soal.
2. Memeriksa setiap entri *checklist* dalam *Constructed-response Item and Scoring Guide Review Checklist*.
3. Menuliskan garis besar jawaban soal untuk siswa. Meninjau panduan penilaian untuk soal dan membandingkannya. Memastikan diri setuju dengan jumlah skor poin yang dialokasikan. Melihat tipe penyelesaian yang paling mungkin digunakan siswa.
4. Mengidentifikasi dan mencatat setiap masalah dalam soal dan mencetak paduan penilaian.

Untuk penilaian validitas tes sesuai dengan standar TIMSS menurut Martin (2000: 28), antara lain:

1. *Is the mathematics/ science correct*

2. *Task clear to students*
3. *Free of cultural, gender, or geographical bias*
4. *Seems to be OK for transiation*
5. *No unfamiliar factors contributing to difficulty*
6. *Clear expectations for full-credit response*
7. *Task can be completed in a reasonable time*
8. *Scoring guide has appropriate number of score points*
9. *Scoring guide descriptors clear*
10. *Content classification correct*
11. *Cognitive classification correct*

Pada penelitian ini soal yang dibuat sudah sesuai dengan validitas tes sesuai dengan standar TIMSS karena soal telah memenuhi seluruh kriteria validitas tes sesuai TIMSS.

2.3 Validitas Tes

Menurut Anatasi dan Urbina (dalam Purwanto, 2014: 114) validitas berhubungan dengan apakah tes mengukur apa yang mesti diukurnya dan seberapa baik dia melakukannya. Langkah yang harus dilakukan agar instrumen memiliki validitas yang tinggi adalah dengan cara uji coba instrumen. Suatu instrumen penelitian dikatakan baik apabila memenuhi syarat valid dan reliabel. Instrumen yang valid ialah instrumen yang mampu mengukur apa yang diinginkan oleh peneliti dan dapat mengungkapkan data dari variabel yang diteliti secara tepat.

Untuk tes evaluasi dengan bentuk uraian menurut Basuki dan Hariyanto (2014: 133) komponen yang ditelaah adalah:

- a. Materi:
 1. Soal sesuai dengan indikator (menurut tes tulis untuk bentuk uraian).
 2. Materi yang ditanyakan sesuai dengan komposisi (urgensi, relevansi, kontinuitas, keterpakaian sehari-hari tinggi).
 3. Batasan pertanyaan dan jawaban yang diharapkan sudah sesuai.
 4. Isi materi yang ditanyakan sesuai dengan jenjang, jenis sekolah atau tingkat kelas.
- b. Konstruksi
 1. Menggunakan kata tanya atau perintah yang menuntut jawaban uraian.
 2. Ada petunjuk yang jelas tentang cara mengerjakan soal.
 3. Ada pedoman penskorannya.
 4. Tabel, gambar, grafik, peta, atau yang sejenisnya disajikan dengan jelas dan terbaca.
- c. Bahasa/budaya.

1. Menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia.
2. Rumusan kalimat soal komunikatif.
3. Tidak menggunakan bahasa yang berlaku setempat/tabu.
4. Tidak menggunakan kata/ungkapan yang menimbulkan penafsiran ganda atau salah pengertian.
5. Rumusan soal tidak mengandung kata/ungkapan yang dapat menyinggung perasaan siswa.

Menurut Kunandar (2014: 237) komponen yang telaah untuk soal uraian adalah:

- a. Materi
 1. Sesuai dengan indikator.
 2. Batasan dan pertanyaan dan jawaban yang diharapkan jelas.
 3. Isi materi sesuai dengan tujuan pengukuran.
 4. Isi materi yang ditanyakan sudah sesuai dengan jenjang, jenis sekolah dan tingkat kelas.
- b. Konstruksi
 5. Rumusan kalimat soal atau pertanyaan harus menggunakan kata tanya atau menuntut jawaban terurai.
 6. Ada petunjuk yang jelas tentang cara mengerjakan jawaban soal.
 7. Ada pedoman penskoran.
 8. Tabel, gambar, grafik, peta atau sejenisnya disajikan dengan jelas dan terbaca.
- c. Bahasa
 9. Rumusan soal komunikatif.
 10. Butir soal menggunakan bahasa Indonesia yang baik dan benar.
 11. Rumusan soal tidak menggunakan kata-kata/kalimat yang menimbulkan penafsiran ganda atau salah pengertian.
 12. Tidak menggunakan bahasa yang berlaku setempat.
 13. Rumusan soal tidak mengandung kata-kata yang dapat menyinggung siswa.

Jadi, perangkat evaluasi atau instrumen evaluasi yang dikembangkan dikatakan valid jika perangkat tersebut memiliki keterkaitan antara komponen-komponennya secara tepat, pada penelitian ini peneliti menggunakan komponen penelaahan soal uraian menurut Basuki dan Hariyanto.

2.4 Reliabilitas Tes

Menurut Yusuf (2015:74) “Reliabilitas suatu instrumen merujuk kepada kadar stabilitas dan konsistensi instrumen tersebut. Suatu alat ukur dikatakan reliabel, apabila alat ukur itu diujikan kepada subjek atau objek yang sama secara berulang-ulang, hasilnya akan tetap sama, konsisten, stabil atau relatif sama (tida

berbeda jauh secara statistik)”. Artinya, kapan pun alat penilaian tersebut digunakan akan memberikan hasil yang rekatif sama.

2.5 Praktikalitas Tes

Menurut menurut Arifin (2013: 264) “ Kepraktisan mengandung arti kemudahan suatu tes, baik dalam mempersiapkan, menggunakan, mengolah dan menafsirkan maupun mengadministrasikannya”. Kepraktisan merupakan syarat suatu tes standar, namun dalam kenyataannya banyak tes yang dibuat tidak menunjukkan kepraktisan. Akibatnya, ketika tes digunakan orang lain, maka orang tersebut akan merasa kesulitan. Sedangkan Purwanto (2009: 141) mengemukakan bahwa suatu tes dikatakan mempunyai kepraktisan yang baik jika kemungkinan untuk menggunakan tes itu besar.

Dimiyati dan Mudjiono (2013: 198) mengemukakan faktor- faktor yang mempengaruhi kepraktisan instrumen evaluasi meliputi:

1. Kemudahan mengadministrasi.
2. Waktu yang disediakan untuk melancarkan evaluasi.
3. Kemudahan menskor.
4. Kemudahan interpretasi dan aplikasi.
5. Tersedianya bentuk instrumen evaluasi yang ekuivalen atau sebanding.

Menurut Purwanto (2009: 141) kriteria untuk mengukur praktis-tidaknya suatu tes dapat dilihat dari:

1. Biaya yang diperlukan untuk menyelenggarakan tes itu,
2. Waktu yang diperlukan untuk menyusun tes itu,
3. Sukar-mudahnya menyusun tes itu,
4. Sukar-mudahnya menilai (*scoring*) hasil tes itu,
5. Sulit-tidaknya menginterpretasikan (mengolah) hasil tes itu,
6. Lamanya waktu yang diperlukan untuk melaksanakan tes itu.

Pada penelitian kali ini peneliti menggunakan modifikasi kriteria kepraktisan Purwanto antara lain:

1. Soal yang dibuat dapat diujikan dengan pengawasan pihak atau guru lain atau tanpa pengawasan guru yang bersangkutan.
2. Biaya untuk melakukan tes ini relatif ekonomis.
3. Ada pedoman petunjuk pengerjaan soal.
4. Waktu pengerjaan soal oleh siswa sesuai dengan tingkat kesulitan soal dan kemampuan siswa.
5. Alternatif jawaban dan pedoman penilaian yang disediakan rinci dan jelas.
6. Pemeriksaan soal tidak membutuhkan waktu yang lama.

2.6 Analisis Butir Soal Uraian

Menurut Sudjana (2009: 135) “Analisis butir soal adalah pengkajian pertanyaan-pertanyaan tes agar diperoleh perangkat pertanyaan yang memiliki kualitas yang memadai”. Suatu tes yang baik berarti tes tersebut memiliki butir-butir soal yang baik pula, oleh karena itu pengujian tingkat kebaikan suatu soal tes tidak terlepas dari pengujian atau analisis kebaikan butir-butir soalnya. Adapun analisis butir soal uraian sebagai berikut:

2.6.1 Tingkat Kesukaan

Bermutu atau tidaknya butir soal suatu tes hasil belajar dapat diketahui dari tingkat kesukaran atau taraf kesulitan yang dimiliki oleh masing-masing butir soal tersebut. Butir soal dapat dinyatakan sebagai butir soal yang baik, apabila butir soal tersebut tidak terlalu sukar dan tidak terlalu mudah dengan kata lain tingkat kesukaran butir soal itu adalah sedang atau cukup (Sudijono, 2011: 370).

Menurut Sudjana (2009: 135) “Tingkat Kesukaran soal dipandang dari kesanggupan atau kemampuan siswa dalam menjawabnya, bukan dilihat dari sudut guru sebagai pembuat soal. Persoalan yang penting dalam melakukan analisis tingkat kesukaran soal salah satunya menentukan proporsi dan kriteria soal yang termasuk mudah, sedang dan sukar”.

Soal yang terlalu mudah tidak merangsang siswa untuk lebih berusaha memecahkannya. Sedangkan soal yang terlalu sukar akan menyebabkan siswa menjadi putus asa dan membuat siswa menjadi malas untuk mencoba karena mereka berpikir soal tersebut di luar jangkauan mereka.

Menurut Arikunto (2015: 223) “Bilangan yang menunjukkan sukar atau mudahnya sesuatu soal disebut indeks kesukaran (*difficulty index*). Besarnya indeks kesukaran antara 0,00 sampai dengan 1,0. Indeks kesukaran ini menunjukkan taraf kesukaran soal. Soal dengan indeks kesukaran 0,0 menunjukkan bahwa soal itu terlalu sukar, sebaliknya indeks 1,0 menunjukkan bahwa soalnya terlalu mudah”.

2.6.2 Daya Pembeda

Sudijono (2011: 386) mengatakan bahwa mengetahui daya pembeda butir soal penting sekali, sebab salah satu dasar untuk membuat butir soal tes hasil belajar yang baik adalah adanya anggapan bahwa kemampuan antara siswa yang satu dengan siswa yang lain itu berbeda-beda, dan butir soal soal tes hasil belajar

itu haruslah mencerminkan adanya perbedaan kemampuan yang terdapat dikalangan siswa tersebut.

Sejalan dengan pernyataan di atas maka kegiatan analisis terhadap daya pembeda butir soal itu ditujukan untuk menjawab pertanyaan : “Apakah jawaban siswa yang dianggap pandai pada umumnya betul?” dan “ Apakah jawaban siswa yang dianggap kurang pandai pada umumnya salah?”. Jika jawaban atas pertanyaan itu “iya”, maka butir soal yang bersangkutan dapat dianggap butir soal yang baik, dalam arti bahwa butir soal tersebut telah menunjukkan kemampuannya dalam membedakan siswa kelompok atas dengan siswa kelompok bawah. Sebaliknya, jika jawaban atas pertanyaan itu “tidak” maka butir soal yang bersangkutan dinyatakan sebagai butir soal yang jelek, sebab hasil yang dicapai dalam tes itu justru bertentangan dengan tujuan tes itu sendiri (Sudijono, 2011: 386).

Menurut Arikunto (2015: 232) butir-butir soal yang baik adalah butir-butir soal yang mempunyai indeks diskriminasi 0,4 sampai dengan 0,7. Sedangkan indek diskriminasi yang bernilai negatif semuanya tidak baik dan sebaiknya dibuang saja.

2.6.3 Validitas Butir Soal (Item)

Menurut Arikunto (2015: 90) jika seorang peneliti atau guru mengetahui bahwa validitas soal tes misalnya terlalu rendah atau rendah saja, maka selanjutnya ingin mengetahui butir-butir tes manakah yang menyebabkan soal secara keseluruhan tersebut jelek karena memiliki validitas rendah, untuk keperluan inilah dicari validitas butir soal. Sehingga perlu dilakukan telaah dan tindakan yang tepat pada soal-soal yang menyebabkan tes tersebut menjadi jelek.

Jadi, perangkat evaluasi atau instrumen evaluasi yang dikembangkan dikatakan valid jika perangkat tersebut memiliki keterkaitan antara komponen-komponennya secara tepat.