

PENGEMBANGAN INSTRUMEN ASESMEN OTENTIK UNJUK KERJA MATERI BANGUN RUANG DI SEKOLAH DASAR KOTA CIREBON

Rivo Panji Yudha, M.Pd

Fakultas Keguruan dan Ilmu Pendidikan
Universitas 17 Agustus 1945
Jl. Perjuangan no. 17 Cirebon
rivoyudha@yahoo.co.id

ABSTRAK

The development of an authentic assessment instrument performance is a form of performance assessment instruments in elementary school mathematics learning that aims to measure the learning outcomes from the aspects of students' skills. This assessment instrument requires the use of a valid assessment of performance, realaiabel, practical, efficient and easy to learn. Research subjects in elementary school SDN Nurrusshidiq town of Cirebon and Cirebon City Karanganyar class V on the material geometry.

This study is a model of the development of R & D using analysis of modified Borg and Gall researchers into three phases include: First, through the analysis of data needs that information obtained in the aspects of performance assessment instruments in learning mathematics are: preparation, execution and reporting of the work . Second, the lattice instrument consisting of 6 items, after tested retained as the 6th item is valid. Third, according to the input of the experts who were four people using the techniques of expert judgment, to-four experts came from three expert evaluation, and mathematical instruments and a senior Mathematics teacher. Fourth, the content validity of the test results were analyzed using the Q Score of all aspects assessed by experts covering aspects aspects Compliance with indicator, writing, language, and physical appearance ideals percentage value of 91% included in the criteria used very good and decent. Fifth, data from the instrument reliability performance on a small scale using GENOVA analysis, problem number one (0.89), Problem number two (0.90), question number three (0.82) about the number three (0.78). (4). Sixth, practicality uyang instruments were analyzed using T scores obtained an average score of 52.33 and a T score of each rater I (49), rater II (54), and a third rater score obtained from T (54) and based on the assessment teacher performance assessment instruments classified as practical .

In conclusion authentic assessment instrument performance worthy to be used by primary school mather teacher and teacher summarized the results of all development in the use of manual performance asses instrument. It is advisable for teachers to take advantage of the assessment instrument performance as a cap. assessment of students in aspects of learning math skills in elementary school.

Keywords: Instruments, Performance, Learning Mathematics, Valid, Reliable

I. PENDAHULUAN

Proses penilaian yang dilakukan selama ini pun juga semata-mata hanya menekankan

pada penguasaan konsep yang dijangar dengan tes tulis obyektif dan subyektif sebagai alat ukurnya. Hal ini didukung oleh penelitian Nuryani, dkk (2013:8) yang mengemukakan

bahwa pengujian yang dilakukan selama ini baru mengukur penguasaan materi saja dan itu pun hanya meliputi ranah kognitif tingkat rendah. Keadaan semacam ini merupakan salah satu penyebab guru enggan melakukan kegiatan pembelajaran yang memfokuskan pada pengembangan sikap dan keterampilan proses anak.

Kegiatan pembelajaran yang dilakukan umumnya hanya terpusat pada penyampaian materi dalam buku teks. Keadaan faktual ini mendorong siswa untuk menghafal pada setiap kali akan diadakan tes harian atau tes hasil belajar padahal untuk anak jenjang sekolah dasar yang harus diutamakan adalah bagaimana mengembangkan rasa ingin tahu dan daya kritis anak terhadap suatu masalah (Marjono.M: 1996), Selain itu; belum adanya penilaian unjuk kerja yang dilakukan sebagai suatu langkah mengetahui kondisi maupun tingkat kemampuan siswa, sehingga apabila pelajaran matematika disampaikan secara menarik dapat mendorong siswa untuk lebih giat dan meningkatkan pemahaman terhadap materi yang diperolehnya.

Asesmen kinerja siswa merupakan salah satu alternatif penilaian yang difokuskan pada dua aktivitas pokok, yaitu: Observasi proses saat berlangsungnya unjuk keterampilan dan evaluasi hasil cipta atau produk. Penilaian bentuk ini dilakukan dengan mengamati saat siswa melakukan aktivitas di kelas atau menciptakan suatu hasil karya sesuai dengan tujuan pembelajarannya. Kecakapan yang ditampilkan siswa adalah variabel yang dinilai. Penilaian terhadap kecakapan siswa didasarkan pada perbandingan antara kinerja siswa dengan target yang telah ditetapkan.

Proses penilaiannya dilakukan mulai persiapan, melaksanakan tugas sampai dengan hasil akhir yang dicapainya. Oleh karena itu penilaian dengan tertulis dan lisan saja tidak dapat mewakili secara keseluruhan segala penilaian yang diinginkan apalagi dengan materi pembahasan yang menuntut siswa agar dapat memecahkan masalah dan menentukan sikap, bekerja sama dengan teman sekelompoknya dan lain-lainnya.

Asmawi Zainul (2001) menekankan perlunya asesmen unjuk kerja untuk mengukur aspek lain di luar kognitif, yaitu tujuh kemampuan

dasar. Menurut Howard Gardner yang tidak mungkin dinilai hanya dengan cara-cara yang biasa. Ketujuh kemampuan dasar tersebut adalah: (1) *visual-spatial*, (2) *bodily-kinesthetic*, (3) *musical-rhythmical*, (4) *interpersonal*, (5) *Intrapersonal*, (6) *logicamathematical*, (7) *verbal linguistic*. Baru dua kemampuan yang terakhir yang banyak diukur atau dinilai orang, sementara lima kemampuan yang lainnya belum banyak diungkap, Dari keterangan di atas jelaslah bahwa proses penilaian (asesmen) terutama penilaian kinerja menjadi fokus utama penilaian.

Hasil wawancara di lapangan (terutama terhadap pembelajaran matematika di Sekolah Dasar Kota Cirebon, Sa'adah (2014) mengatakan terdapat beberapa sumber kendala yang dihadapi oleh para guru dalam menilai unjuk kerja siswa yaitu: pertama, pedoman penyekoran dalam instrumen tidak jelas sehingga sukar digunakan, komponen-komponen yang dinilai sulit untuk diamati, sehingga cenderung diabaikan; kedua, penilai (rater) umumnya hanya satu orang yaitu guru bidang studi, sedangkan komponen-komponen yang dinilai dan jumlah siswa yang dinilai cukup banyak, sehingga sulit untuk mendapat perbandingan untuk dijadikan bahan pertimbangan mengambil keputusan; ketiga, kemungkinan ada kecenderungan untuk memberi nilai tinggi atau sebaliknya, hal ini diakibatkan oleh instrumen yang digunakan belum memenuhi persyaratan validitas, reliabilitas dan kepraktisannya.

Melawati (2014) menyatakan sebagian besar guru matematika sekolah dasar disini khususnya tidak tertarik dan tidak mau menggunakan penilaian otentik atau penilaian unjuk kerja. Pada umumnya berpendapat bahwa melakukan penilaian otentik itu membuang waktu dan energi serta terlalu mahal, apalagi penilaian otentik perlu dirancang dengan baik. Pendapat tersebut tertentu tidak benar. Menilai kinerja dengan tes tertulis tentu tidak valid, karena tidak mengukur apa yang ingin dinilai, Kinerja perlu dinilai pada saat kegiatannya sedang berlangsung. Kalau penilaian kinerja dilakukan terhadap sejumlah siswa dan tidak dirancang dulu atau dilakukan asal-asalan, tentu hasilnya tidak dapat dipertanggungjawabkan karena tidak konsisten.

Dengan demikian kita mungkin berlaku tidak adil terhadap sejumlah siswa dalam menilai kinerja mereka. Menurut Wiggins (2005:2-3) merancang dan melaksanakan penilaian kinerja sangatlah efisien, karena ajeg atau konsisten (baca reliabel), tidak mahal dan tidak membuang waktu. Standar tidak dapat dibuat tanpa melakukan penilaian berbasis kinerja.

Permasalahan yang sering dihadapi guru matematika sekolah dasar selanjutnya adalah dalam melakukan asesmen unjuk kerja terletak pada validitas dan reliabilitas alat ukur yang digunakan. Penyusunan tes unjuk kerja siswa masih sangat terbatas pada pengetahuan dan pemahaman guru tentang tes berbentuk simulasi. Hasil asesmen sering dipengaruhi oleh objektivitas guru sebagai rater karena dalam melakukan asesmen dilakukan sendiri tanpa melibatkan guru yang lain sebagai kolabolator (Novi, 2014).

Berdasarkan uraian di atas dapat disimpulkan bahwa penerapan instrumen penilaian unjuk kerja yang valid reliabel, praktis dan dapat digunakan secara berulang pada tugas kinerja yang berbeda, dapat membantu guru untuk melakukan penilaian unjuk kerja siswa pada saat melakukan praktikum di laboratorium. Produk pengembangan instrumen penilaian unjuk kerja laboratorium bidang fisika, sebelum diputuskan untuk digunakan oleh guru-guru perlu diadakan penelitian. Tujuannya untuk memperoleh informasi apakah produk yang dikembangkan sudah valid, reliabel, dan praktis.

Dengan mengkaji kenyataan yang ditemukan di lapangan, nampak ada ketidaksesuaian antara pembelajaran matematika di SD dengan sistem penilaian yang digunakannya. Proses penilaian yang biasa dilakukan guru selama ini hanya mampu menggambarkan aspek penguasaan konsep peserta didik, akibatnya tujuan kurikuler mata pelajaran Matematika belum dapat dicapai dan atau tergambarkan secara menyeluruh. Selain itu, dikemukakan masih terdapat miskonsepsi tentang asesmen unjuk kerja pada responden yang diteliti, kurangnya pemahaman sebagian besar responden tentang tata cara membuat perangkat asesmen kinerja yang sesuai, dan sebagian responden yang cukup memahami

asesmen unjuk kerja dan guru juga masih kesulitan karena keterbatasan waktu dan banyaknya jumlah siswa dalam satu kelas.

Berdasarkan tujuan tersebut, Permasalahan penelitian dapat dirumuskan permasalahan sebagai berikut: (1) Bagaimanakah instrumen asesmen unjuk kerja yang selama ini ada pada pembelajaran Matematika di sekolah dasar Kota Cirebon; (2) Bagaimana Pengembangan instrumen Asesmen otentik unjuk kerja siswa pembelajaran Matematika di sekolah dasar Kota Cirebon; (3) Apakah instrumen asesmen otentik unjuk kerja pada pembelajaran matematika di sekolah dasar kelas kota Cirebon Valid dan Realiabel; (4) Apakah Kualitas implementasi asesmen otentik unjuk kerja pada pembelajaran matematika di sekolah dasar Kota Cirebon Praktis.

Manfaat dari penelitian adalah (1) Secara teoritis hasil penelitian ini diharapkan semakin memperkuat argumentasi urgensi penilaian unjuk kerja pada proses pembelajaran matematika, khusus di sekolah dasar; (2) Sebagai kajian untuk mengetahui tingkat kemampuan pemahaman siswa pada materi yang diajarkan sehingga diharapkan dapat menemukan solusi atau langkah selanjutnya; (3) Secara praktis, hasil-hasil penelitian akan dapat dijadikan sebagai pedoman yang rinci bagi guru-guru SD, khususnya pada pembelajaran Matematika dalam asesmen unjuk kerja siswa dan melakukan pelaporan hasil penilaian unjuk kerja siswa pada mata pelajaran bangun ruang secara lebih detail dan menyeluruh; (4) Manfaat lain dari hasil-hasil penelitian ini dapat digunakan sebagai bahan evaluasi dalam melakukan perbaikan kurikulum, khususnya tentang penentuan tingkat pencapaian kompetensi pada akhir tahun pembelajaran. Kedua, hasil penelitian ini dapat digunakan untuk perbaikan di bidang Matematika, khususnya materi-materi lainnya.

II. METODE PENELITIAN

Penelitian ini merupakan penelitian pengembangan dengan tujuan menghasilkan produk berupa instrumen asesmen unjuk kerja. Model pengembangan yang digunakan dalam penelitian ini adalah model pengembangan R & D analisis Borg and Gall (2007:169-170). Prosedur dalam

pengembangan ini meliputi: (1) Pendahuluan (*Define*); (2) Perencanaan (*Planning*) (3) Mengembangkan produk awal (*Develop*), (4) Uji coba awal; (5) Revisi pertama; (6) uji coba lapangan utama; (7) revisi produk kedua; (8) Uji coba produk operasional; (9) Revisi Produk final; (10) Penyajian produk akhir (*Deliver*).

Berdasarkan sepuluh langkah penelitian dan pengembangan yang dikembangkan Borg & Gall tersebut, dalam penelitian ini pada proses pelaksanaannya melakukan adaptasi yang mengacu pada model pendekatan tersebut. Adapun adaptasi penelitian pengembangan ini secara garis besar terdiri dari tiga tahap kegiatan pokok, yakni: (1) pendahuluan; (2) pengembangan produk (*develop*) dan (3) penyajian.

Tahap pendahuluan terdiri dari tiga kegiatan pokok yang meliputi kegiatan analisis kebutuhan, perancangan kisi-kisi dan pembuatan instrumen asesmen unjuk kerja. Kegiatan analisis kebutuhan bertujuan untuk mengungkap kondisi nyata guru matematika dalam melakukan asesmen unjuk kerja siswa khususnya pada pembelajaran Matematika saat ini. Analisis kebutuhan berupa hasil survey dengan menggunakan wawancara yang dilakukan terhadap para guru matematika SD. Sebagai subyek penelitian pada tahap analisis kebutuhan diperoleh informasi bahwa kendala penilaian unjuk kerja kegiatan pembelajaran matematika di sekolah dasar, salah satu kendala adalah guru masih belum memahami pedoman penyekoran dalam instrumen yang tidak jelas sehingga sukar digunakan, komponen-komponen yang dinilai sulit untuk diamati, sehingga cenderung diabaikan.

Tahap pengembangan (*develop*), pada tahap ini instrumen yang sudah dirancang dikonsultasikan kepada dosen pembimbing. Uji ahli atau Validasi, dilakukan dengan responden para ahli perancangan model atau produk. Kegiatan ini dilakukan untuk mereview produk awal memberikan masukan untuk perbaikan,

Proses validasi ini disebut dengan Expert Judgement atau Teknik Delphi. Instrumen yang sudah dihasilkan dievaluasi, apakah format yang dihasilkan sudah layak atau belum, dan bagaimana kesesuaian isi materi penilaian pembelajaran. Jika instrumen belum

layak, maka instrumen direvisi kembali sehingga instrumen menjadi layak untuk diuji cobakan. Sebelum uji coba, dilakukan validasi terhadap instrumen oleh 4 ahli evaluasi, kemudian instrumen diujicobakan untuk menilai kinerja peserta didik saat pembelajaran pada skala kecil, yaitu pada peserta didik kelas V pada skala kecil. Hal ini bertujuan untuk mengetahui apakah instrumen sudah layak digunakan atau belum untuk mengetahui bagaimana kinerja peserta didik. Hasil dari uji coba penggunaan instrumen pada kelas V dijadikan rujukan untuk pengembangan dan perbaikan instrumen selanjutnya.

Tahap penyajian (*deliver*), pada tahap ini diuji cobakan lebih luas yaitu di sekolah berbeda dengan kelas yang sama yaitu kelas V dan pada skala yang lebih besar. Uji coba produk yang sesungguhnya dilaksanakan untuk menilai kinerja peserta didik kelas V pada saat pembelajaran. Hasil dari fase ini adalah adanya kesimpulan sukses tidaknya rancangan produk yang dikembangkan bagi kepentingan pengguna dan dari tim yang terlibat.

Data Instrumen yang digunakan pada penelitian ini adalah sebagai berikut: (1) Skala Bertingkat (*Rating Scale*), skala bertingkat digunakan untuk melihat indikator keterampilan psikomotor dalam melaksanakan kegiatan praktikum. Skala bertingkat ini berisi tentang aktivitas siswa berupa keterampilan yang akan diamati. Keterampilan psikomotor yang diobservasi meliputi penyiapan alat dan bahan, menggambar jaring-jaring bangun ruang, membentuk bangun ruang, dan merapikan alat dan bahan. Skala bertingkat ini diisi oleh observer yang mengamati seluruh siswa; (2) Rubrik, Rubrik digunakan sebagai panduan penilaian yang menggambarkan kriteria yang diinginkan guru dalam menilai atau memberi tingkatan dari hasil pekerjaan siswa.

Rubrik memuat daftar karakteristik yang diinginkan yang perlu ditunjukkan dalam suatu pekerjaan siswa disertai dengan panduan untuk mengevaluasi masing-masing karakteristik tersebut. Tujuan dari penilaian rubrik yaitu siswa diharapkan secara jelas memahami dasar penilaian yang akan digunakan untuk mengukur suatu kinerja

siswa. Kedua pihak (guru dan siswa) akan mempunyai pedoman bersama yang jelas tentang tuntutan kinerja yang diharapkan; (3) Pedoman Wawancara Guru, Pedoman wawancara dalam penelitian ini digunakan untuk mengumpulkan data lebih dalam mengenai kepraktisan dan keefektifan instrument penilaian yang dikembangkan. baik berkaitan dengan instrument yang digunakan, teknik pelaksanaan penilaian, ataupun hal-hal lain yang tidak terungkap melalui skala bertingkat dan rubrik; (4) Instrumen Tes, Seperti halnya wawancara, tes juga memberikan sampel perilaku individu, hanya saja dalam tes stimulus yang direspon responden lebih terstandardisasikan daripada wawancara.

Bentuk tes yang sudah standar tersebut membantu untuk mengurangi bias yang mungkin muncul selama proses penilaian kinerja berlangsung. Respons yang diberikan biasanya dapat diubah dalam bentuk skor dan dibuat analisis kuantitatif. Hal itu membantu pengawas untuk memahami responden. Skor yang didapat kemudian diinterpretasi sesuai dengan norma yang ada; (5) Lembar Observasi Pengamatan, Pedoman pengamatan dalam asesmen unjuk kerja pada pembelajaran matematika dipergunakan untuk mengamati dan menilai setiap siswa peserta tes unjuk kerja dengan menggunakan rubrik berbentuk skala penilaian (rating scale) beserta bobot penilaiannya. Penyusunan konstruk instrumen pengamatan dalam bentuk skala penilaian berdasarkan materi yang mencerminkan keterampilan yang akan diukur. Selanjutnya ditentukan skala penilaian untuk setiap materi. Dalam penelitian ini digunakan skala empat, yaitu satu sampai empat (1-4).

III. HASIL DAN PEMBAHASAN

Pertama, pada tahap analisis kebutuhan diperoleh informasi bahwa kendala penilaian unjuk kinerja kegiatan pembelajaran matematika di sekolah dasar, yang didapat dari wawancara menggunakan lembar wawancara terhadap sepuluh guru matematika Sekolah Dasar yang ada di kota Cirebon. Salah satu kendala adalah guru masih belum memahami pedoman penyekoran dalam instrumen yang tidak jelas sehingga sukar digunakan,

komponen-komponen yang dinilai sulit untuk diamati, sehingga cenderung diabaikan. Keberhasilan pencapaian kompetensi siswa akan sangat ditentukan oleh pengalaman dan pengetahuan guru dalam mengembangkan, dan menggunakan alat ukur yang telah dikonstruksi itu dengan cara yang benar, serta kemampuan menganalisis informasi yang dihasilkan oleh alat ukur itu. Berkaitan dengan pengalaman guru dalam pengembangan perangkat instrumen yang digunakan dalam melakukan penilaian pembelajaran matematika selama ini dapat dilihat melalui jumlah respon, persentase, dan prioritas pilihan sebagaimana terangkum dalam tabel berikut.

Tabel 1 Penggunaan Perangkat/Instrumen Penilaian

Urutan	Perangkat Penilaian	Jumlah Respon
1	Kisi-Kisi dan Tugas	7
2	Lembar Pengamatan	3
3	Rubrik Penilaian	0
4	Prosedur Penilaian	0

N=10

Keberhasilan pencapaian kompetensi siswa akan sangat ditentukan oleh pengalaman dan pengetahuan guru dalam mengembangkan, dan menggunakan alat ukur yang telah dikonstruksi itu dengan cara yang benar, serta kemampuan menganalisis informasi yang dihasilkan oleh alat ukur itu. Berkaitan dengan pengalaman guru dalam pengembangan perangkat instrumen yang digunakan dalam melakukan penilaian pembelajaran matematika selama ini dapat dilihat melalui jumlah respon, persentase, dan prioritas pilihan.

Berdasarkan tabel 1 data yang diperoleh menunjukkan proposi dari 10 guru hanya 7 guru yang menyusun kisi-kisi dan suatu tes sebelum dilakukan penilaian. Sedangkan proporsi dari 10 guru hanya tiga guru yang membuat lembar pengamatan, sedangkan proporsi dari 10 guru rubrik penilaian dan prosedur penilaian tidak dibuat karena tes uraian yang digunakan oleh guru diadopsi dari tes-tes uraian yang sudah memiliki perangkat penilaian berupa prosedur penilaian. Berdasarkan tanggapan para guru dapat disimpulkan bahwa selama ini penguasaan

guru paling utama hanya menyusun kisi dan tes uraian hal ini disebabkan karena pemahaman tentang pengembangan rubrik penilaian dan prosedur penilaian belum banyak dimengerti. Selanjutnya penilai (*rater*) umumnya hanya satu orang yaitu guru bidang studi, sedangkan komponen-komponen yang dinilai dan jumlah siswa yang dinilai cukup banyak, sehingga sulit untuk mendapat pembandingan untuk dijadikan bahan pertimbangan mengambil keputusan. Terakhir kemungkinan ada kecenderungan untuk memberi nilai tinggi atau sebaliknya, hal ini diakibatkan oleh instrumen yang digunakan belum memenuhi persyaratan validitas, reliabilitas dan kepraktisannya.

Hasil wawancara dari sepuluh guru terkait dengan metode penilaian untuk mengukur kompetensi siswa pada materi bangun ruang yang digunakan oleh guru antara lain: Tes unjuk kerja, pengamatan, tes tertulis. Berdasarkan pengalamannya, pada kenyataannya setiap guru memiliki prioritas yang berbeda terhadap ketujuh metode penilaian tersebut. Tanggapan para guru terhadap metode penilaian tersebut sebagaimana terangkum dalam tabel berikut.

Tabel 2 Metode Penilaian

Uraian	Metode	Jumlah Responden
1	Tes Unjuk Kerja	0
2	Pengamatan	0
3	Tes Tulis	10

N=10

Tabel 2 menunjukkan tanggapan guru-guru bahwa, pada umumnya proporsi dari 10 guru menempatkan metode penilaian pembelajaran Matematika Materi Bangun Ruang yang mencakup metode; tes tertulis sebagai prioritas pertama yang perlu diperhatikan. Sedangkan untuk metode tes kinerja, pengamatan, tidak sama sekali digunakan oleh guru.

Tabel 3. Kisi-kisi Instrumen

Kompetensi Inti	Kompetensi Dasar	Indikator Penilaian	No Butir Soal
-----------------	------------------	---------------------	---------------

Memahami sifat-sifat bangun datar dan bangun ruang serta hubungan antar bangun	Mengidentifikasi sifat-sifat bangun ruang sederhana	Menggambar bangun ruang dari sifat-sifat bangun yang diberikan	1
			2
	Menentukan jaring-jaring berbagai bangun ruang sederhana	Membuat jaring-jaring Bangun ruang	3
			4

Kedua, Langkah berikutnya adalah penyusunan kisi-kisi instrumen, alat evaluasi yang dikembangkan dalam penelitian ini adalah instrumen unjuk kerja (*performance assesment*), untuk menilai keterampilan unjuk kerja siswa dalam melakukan pembelajaran matematika Materi bangun ruang. Penyusunan kisi-kisi instrumen penilaian unjuk kinerja kelas ini mengacu kepada Kompetensi Inti (KI), Kompetensi Dasar (KD), lebih lanjut bias dilihat di tabel 1.

Ketiga, setelah kisi-kisi dibuat langkah berikutnya adalah penyusunan instrumen unjuk kerja. Menilai instrumen unjuk kerja, guru paling tidak harus menyiapkan 2 dokumen, yaitu: 1) Soal / lembar kerja / lembar tugas / perintah kerja; 2) Instrumen pengamatan / lembar observasi berupa daftar periksa (*check list*) atau skala penilaian (*rating scale*). Lembar observasi disini adalah sebuah instrumen yang digunakan untuk mengobservasi kemunculan aspek-aspek keterampilan unjuk kerjaa yang diamati. Lembar observasi disini berupa skala penilaian (*rating scale*). Skala penilaian merupakan daftar pertanyaan / pernyataan untuk menilai kualitas pelaksanaan aspek-aspek keterampilan yang diamati dengan rentang 1 - 4.

Keempat, setelah Instrumen asesmen otentik unjuk kerja dibuat, langkah selanjutnya adalah tahap validasi ahli. Pada tahap ini instrumen yang sudah dirancang dikonsultasikan kepada dosen pembimbing. Instrumen yang sudah dihasilkan dievaluasi, apakah format yang dihasilkan sudah layak atau belum, dan bagaimana kesesuaian isi materi penilaian pembelajaran. Berikut hasil validasi dari 4 ahli evaluasi ada pada tabel 4.

Tabel 4. Data Umum Pakar Ahli Instrumen

No	Pakar	Pengalaman Kerja	Jabatan	Pendidikan
----	-------	------------------	---------	------------

(Tahun)			
1	I	27	Praktisi dan Akademisi S3
2	II	25	Praktisi dan Akademisi S3
3	III	23	Praktisi dan Akademisi S3
4	IV	30	Guru Matematika S1

Berdasarkan tabel 4 peneliti memilih 4 orang pakar dari sudut pandang berbeda dan dengan kriteria yang berbeda-beda berdasarkan keinginan peneliti tetapi homogen menurut kepentingan dan keterkaitannya dengan variabel yang ingin divalidasikan baik dari akademisi, praktisi, maupun isi, untuk menemukan variabel terpilih. Dari 4 orang pakar tersebut akan diperoleh komentar atau saran berupa kalimat variabel penelitian, penambahan dan pengurangan jumlah variabel, pengolahan data, dan sebagainya. Berikut adalah para pakar yang memenuhi syarat peneliti.

Hasil-hasil penilaian ke-empat penilai terhadap instrumen asesmen otentik unjuk kerja terangkum dalam Tabel 4.5. Aspek-aspek penilaian meliputi: kesesuaian aspek penilaian unjuk kerja dengan indikator yang ada, aspek kesesuaian aspek dengan indikator, penulisan, aspek bahasa, dan aspek penampilan fisik.

Tabel 5. Rangkuman Hasil Validasi Instrumen Asesmen Unjuk Kerja Pada Pembelajaran Matematika Sekolah Dasar Kota Cirebon

No.	Aspek Yang Dinilai	Kriteria Penilaian	Validator Ahli				Skor 20	Kualitas
			1	2	3	4		

1	Kesesuaian Aspek penilaian unjuk kerja dengan indikator yang ada	1	5	5	5	5	20	SB
		2	4	5	4	5	18	
2	Penulisan	3	5	4	4	4	17	SB
		4	5	5	3	5	18	
		5	4	5	4	5	18	
3	Bahasa	6	5	5	5	4	19	SB
		7	5	5	4	5	19	
4	Penampilan Fisik	8	5	4	3	4	16	SB
		9	5	5	3	4	17	
		10	5	5	5	5	20	SB
Jumlah Skor		4	4	4	4	4	182	SB
		8	8	0	6			

Dari hasil ujicoba yang dilakukan kepada 3 ahli instrumen dan 1 ahli Matematika maka diambil rata-rata persentase hasil uji coba yaitu untuk tugas unjuk kerja kelas V, Persentase Keidealan instrumen penilaian 91%, Uji Kesesuaian 95%, Uji Penulisan 88%, Uji Bahasa 90%, Uji Penampilan fisik 92,5%.

Selain nilai persentase keidealan hasil validasi umum dari ke-empat penilai terhadap instrumen asesmen unjuk kerja yaitu dapat dilihat dari nilai *expert judgment*, yang mendapatkan skor maksimal 20 ada pada aspek kesesuaian indikator dan penampilan fisik. Selain data yang dipaparkan pada tabel juga diperoleh data tertulis yang berasal dari kolom catatan dan data verbal yang ditranskripsikan dari hasil wawancara dengan ahli dan praktisi tentang Beberapa hal pokok yang menjadi masukan dari ke-empat penilai tersebut diantaranya adalah: (1) Tata cara penulisan bahasa masih kurang tepat misalnya penggabungan atau pemisahan kalimat; (2) Kalimat pada instrumen hendaknya langsung ke akar permasalahan, tidak bertele-tele; (3) Tampilan instrumen masih kurang menarik; (4) Instrumen Hendaknya mengukur kompetensi spesifik yang diharapkan muncul dalam pembelajaran; (5) Rubrik penilaian hendaknya dijadikan satu apabila tipe soal sama agar lebih efisien.

Kelima, Pelaksanaan Uji coba skala kecil dilakukan dengan bantuan 4 orang *rater* yang diambil dari guru matematika. Tahap-tahap pada ujicoba pertama ini sebagai berikut: (1)

sehari sebelum uji coba dilaksanakan peneliti memberikan instrument kepada para *rater* dan menjelaskan maksud yang terkandung pada butir-butir indikator; (2) Setiap *rater* mendapatkan satu eksemplar instrument dan dimohon untuk mengisi butir penilaian dengan member nilai 1,2,3,4 yang merupakan hasil penilaian. Hal ini dilakukan agar pada saat *rater* melakukan penilaian dapat terhindar dari kesalahan interpretasi terhadap butir penilaian; (3) *rater* mengadakan penilaian terhadap *ratee*, dan setiap *rater* menilai atau mengamati jumlah siswa yang ada (3) peneliti mengadakan diskusi dengan para *rater* dan mohon masukan terhadap instrumen penilaian yang digunakan. *Item* instrumen unjuk kerja yang telah dipersiapkan dan dipergunakan dalam penelitian ini, yakni masing-masing sebanyak 4 soal tes unjuk kerja yang didalamnya terdapat 4 rubrik penilaian telah diuji validitas isinya dan realibilitas menggunakan Penentuan koefisien reliabilitas instrumen asesmen unjuk kerja pada pembelajaran matematika materi bangun ruang di sekolah dasar dilakukan dengan menggunakan koefisien Genova (*Generalizability of varians*).

Program Genova mengacu pada teori *Generalizability* yang dikembangkan oleh Crick dan Brennan pada tahun 1983, yang terdiri dari *Generalizability Study* (*G study*) dan *Decision Study* (*D study*) dengan komponen variansinya adalah *person*, *rater*, dan *item*. *G study* digunakan untuk mengestimasi komponen variansi kesalahan yang diakibatkan oleh berbagai sumber variansi. *D study* digunakan untuk mengetahui koefisien reliabilitas (Tri Hartiti Retnowati, 2009: 184). Hasil koefisien GENOVA dapat dilihat pada tabel 6 dibawah ini:

Tabel 6 Estimasi Koefisien *Generalizability* Pada Penilaian Unjuk Kerja Soal Nomor Satu ($n = 26$)

Desain D Study	Sample Size			<i>Generalizability</i>	
	P	R	T	Coef.	<i>Phi</i>

001-001	26	1	1	0,51	0,016
001-002	26	2	2	0,68	0,032
001-003	26	3	3	0,76	0,048
001-004	26	4	4	0,81	0,063
001-005	26	5	5	0,84	0,078
001-006	26	6	9	0,86	0,09
001-007	26	7	7	0,88	0,11
001-008	26	8	8	0,89	0,11

Tabel 6 memberi gambaran tentang perubahan koefisien *Generalizability* untuk berbagai komposisi ukuran sampel P (*person*), R (*rater*), dan T (*item*). Pada komponen penilaian unjuk kerja siswa soal nomor 1 jika komposisinya hanya menggunakan satu indikator (*D study* design nomor 001-001 dengan P = 26, R= 1 dan T = 1) maka tingkat atau koefisien kesepahaman dan kesepakatan (reliabilitas dalam koefisien G) sebesar 0,51 Artinya penilai memiliki tingkat kesepahaman dan kesepakatan terhadap penggunaan konstruk instrumen penilaian yang dipakai sebesar 51%. Jika penilai menggunakan dua indikator (rancangan *D study* nomor 001-002, dengan P = 26, R = 8 dan T = 8) yakni indikator 1 dan 2, maka tingkat atau koefisien kesepahaman dan kesepakatan sebesar 0,68; demikian seterusnya untuk rancangan 001-008 diperoleh koefisien sebesar 0,89. Berdasarkan hasil ini maka dapat dikatakan bahwa untuk mencapai kesepahaman dan kesepakatan yang memenuhi tingkat observasi yang dapat diterima untuk faset yang lebih luas. Tingkat kesepahaman dan kesepakatan yang lebih tinggi maka jumlah indikator penilaian harus ditambah, jumlahnya tergantung pada kondisi faset yang bersangkutan dalam hal ini apabila ke 8 (delapan) indikator digunakan maka akan dicapai koefisien kesepahaman dan kesepakatan sebesar 89%.

Tabel 7. Analisis *Genova* Untuk Estimasi Komponen Variansi Tugas Unjuk Kerja Soal Nomor Dua

Desain D Study	Sample Size			<i>Generalizability</i>	
	P	R	T	Coef.	<i>Phi</i>

001-001	26	1	1	0,57	0,0116
001-002	26	2	2	0,62	0,0117
001-003	26	3	3	0,65	0,0123
001-004	26	4	4	0,70	0,0131
001-005	26	5	5	0,79	0,0139
001-006	26	6	9	0,85	0,0148
001-007	26	7	7	0,88	0,025
001-008	26	8	8	0,89	0,038

Tabel 7 memberi gambaran tentang perubahan koefisien *Generalizability* untuk berbagai komposisi ukuran sampel P (person), R (rater), dan T (item). Pada komponen penilaian unjuk kerja siswa soal nomor 1 jika komposisinya hanya menggunakan satu indikator (D study design nomor 001-001 dengan P = 26, R= 1 dan T = 1) maka tingkat atau koefisien kesepahaman dan kesepakatan (reliabilitas dalam koefisien G) sebesar 0,57 Artinya penilai memiliki tingkat kesepahaman dan kesepakatan terhadap penggunaan konstruk instrumen penilaian yang dipakai sebesar 51%. Jika penilai menggunakan dua indikator (rancangan D study nomor 001-002, dengan P = 26, R = 8 dan T = 8) yakni indikator 1 dan 2, maka tingkat atau koefisien kesepahaman dan kesepakatan sebesar 0,62; demikian seterusnya untuk rancangan 001-008 diperoleh koefisien sebesar 0,90. Berdasarkan hasil ini maka dapat dikatakan bahwa untuk mencapai kesepahaman dan kesepakatan yang memenuhi tingkat observasi yang dapat diterima untuk faset yang lebih luas. Tingkat kesepahaman dan kesepakatan yang lebih tinggi maka jumlah indikator penilaian harus ditambah, jumlahnya tergantung pada kondisi faset yang bersangkutan dalam hal ini apabila ke 8 (delapan) indikator digunakan maka akan dicapai koefisien kesepahaman dan kesepakatan sebesar 90%

Tabel 8. Analisis *Genova* Untuk Estimasi Komponen Variansi Tugas Unjuk Kerja Soal Nomor tiga

Desain D Study	Sample Size			Generalizability	
	P	R	T	Coef.	Phi
001-001	26	1	1	0,38	0,38
001-002	26	2	2	0,65	0,65
001-003	26	3	3	0,77	0,77

001-004	26	4	4	0,82	0,81
001-005	26	5	5	0,38	0,38
001-006	26	6	9	0,65	0,65
001-007	26	7	7	0,77	0,77
001-008	26	8	8	0,82	0,81
001-009	26	9	9	0,51	0,007
001-010	26	10	10	0,38	0,38
001-011	26	11	11	0,65	0,65
001-012	26	12	12	0,82	0,81

Tabel 8 memberi gambaran tentang perubahan koefisien *Generalizability* untuk berbagai komposisi ukuran sampel P (person), R (rater), dan T (item). Pada komponen penilaian unjuk kerja siswa soal nomor 1 jika komposisinya hanya menggunakan satu indikator (D study desain nomor 001-001 dengan P = 26, R= 1 dan T = 1) maka tingkat atau koefisien kesepahaman dan kesepakatan (reliabilitas dalam koefisien G) sebesar 0,38 Artinya penilai memiliki tingkat kesepahaman dan kesepakatan terhadap penggunaan konstruk instrumen penilaian yang dipakai sebesar 38%. Jika penilai menggunakan dua indikator (rancangan D study nomor 001-002, dengan P = 26, R = 12 dan T = 12) yakni indikator 1 dan 2, maka tingkat atau koefisien kesepahaman dan kesepakatan sebesar 0,65; demikian seterusnya untuk rancangan 001-012 diperoleh koefisien sebesar 0,82. Berdasarkan hasil ini maka dapat dikatakan bahwa untuk mencapai kesepahaman dan kesepakatan yang memenuhi tingkat observasi yang dapat diterima untuk faset yang lebih luas. Tingkat kesepahaman dan kesepakatan yang lebih tinggi maka jumlah indikator penilaian harus ditambah, jumlahnya tergantung pada kondisi faset yang bersangkutan dalam hal ini apabila ke 12 (dua belas) indikator digunakan maka akan dicapai koefisien kesepahaman dan kesepakatan sebesar 82%

Tabel 9. Analisis *Genova* Untuk Estimasi Komponen Variansi Tugas Unjuk Kerja Soal Nomor empat

Desain D Study	Sample Size			Generalizability	
	P	R	T	Coef.	Phi
001-001	26	1	1	0,57	0,013

001-002	26	2	2	0,62	0,026
001-003	26	3	3	0,78	0,040

Tabel 9 memberi gambaran tentang perubahan koefisien *Generalizability* untuk berbagai komposisi ukuran sampel P (*person*), R (*rater*), dan T (*item*). Pada komponen penilaian unjuk kerja siswa soal nomor 1 jika komposisinya hanya menggunakan satu indikator (D *study* desain nomor 001-001 dengan P = 26, R= 1 dan T = 1) maka tingkat atau koefisien kesepahaman dan kesepakatan (reliabilitas dalam koefisien G) sebesar 0,56 Artinya penilai memiliki tingkat kesepahaman dan kesepakatan terhadap penggunaan konstruk instrumen penilaian yang dipakai sebesar 56%. Jika penilai menggunakan dua indikator (rancangan D *study* nomor 001-002, dengan P = 26, R = 2 dan T = 2) yakni indikator 1 dan 2, maka tingkat atau koefisien kesepahaman dan kesepakatan sebesar 0,522; demikian seterusnya untuk rancangan 001-003 diperoleh koefisien sebesar 0,78. Berdasarkan hasil ini maka dapat dikatakan bahwa untuk mencapai kesepahaman dan kesepakatan yang memenuhi tingkat observasi yang dapat diterima untuk faset yang lebih luas. Tingkat kesepahaman dan kesepakatan yang lebih tinggi maka jumlah indikator penilaian harus ditambah, jumlahnya tergantung pada kondisi faset yang bersangkutan dalam hal ini apabila ke 3 (tiga) indikator digunakan maka akan dicapai koefisien kesepahaman dan kesepakatan sebesar 78%.

Keenam, kepraktisan suatu tes merupakan indikator kualitas suatu alat ukur tergolong baik atau tidak. Hasil analisa kepraktisan instrumen penilaian unjuk kerja berdasarkan data yang diperoleh dengan memberikan angket koesioner kepada keempat penilai (guru) yang mengujicobakan penggunaan instrumen. Secara empirik ketujuh penilai diminta mengisi koesioner dengan lima jenis pertanyaan dalam bentuk rubrik dengan rentang nilai 1 sampai 5, yang berhubungan dengan kepraktisan penggunaan. Data hasil jawaban responden selanjutnya dianalisis secara statistik dengan menggunakan formula T Skor. Berdasarkan analisis data dapat diinterpretasikan bahwa melalui uji skor baku (standar) dengan skor T terhadap kepraktisan penggunaan instrumen penilaian unjuk pada pembelajaran matematika dilihat pada tabel 7.

Tabel 7. Data Penilaian Rater Terhadap Kepraktisan Instrumen Asesmen Otentik Unjuk Kinerja

Perso n	x	x - \bar{x}	(x - \bar{x}) ²	Z=(x - \bar{x})/S D	Sko r T	Kriter ia
Rater I	49	- 1,33	1,7 7	- 1,15	49	Prakti s
Rater II	51	0,67	0,4 5	0,58	54	Prakti s
Rater III	51	0,67	0,4 5	0,58	54	Prakti s
Juml ah	15 1				157	
Mean		50,3 3	2,6 5		52,3 3	Prakt is
SD			1,1 5			

Terlihat dari Tabel 7 secara umum guru-guru menilai instrumen Unjuk Kerja memiliki subyektivitas, kesistematian, konstruksi, kebahasaan dan kepraktisan yang baik. Hal ini tergambar dari skor T dari masing *rater* I sebesar 49, *rater* II diperoleh skor T sebesar 54, dan dari *rater* III diperoleh skor T sebesar 54. Dengan demikian menurut kriteria kepraktisan *formula* Glicman instrumen Unjuk Kerja ini dapat dikatakan secara umum dinilai praktis oleh para guru (*rater*) dalam menilai tingkat kualitas kepraktisan instrumen unjuk kerja siswa Sekolah Dasar pada pembelajaran Matematika materi bangun ruang.

IV. Kesimpulan dan Saran

1. Simpulan

Berdasarkan hasil penelitian pengembangan yang telah dilakukan, dapat ditarik kesimpulan sebagai berikut: (1) asesmen unjuk kerja yang selama ini ada adalah guru sekolah dasar dikota Cirebon masih memakai tes tulis uraian dan lembar

tugas; (2) Validasi instrumen penilaian dilakukan melalui uji ahli dan validitas empiris. Hasil penilaian yang diperoleh dari validasi ahli menyatakan bahwa penilaian unjuk kerja ini layak digunakan sebagai bentuk penilaian. Secara keseluruhan hasil dari pengujian indeks korelasi skor butir dengan skor total dan hasil uji reliabilitas dinyatakan dalam rincian sebagai berikut: (2.1) Pengujian di skala kecil soal no satu sampai dengan empat masing-masing dikategorikan valid serta memiliki nilai koefisien GENOVA sebesar 0.89; 0.90; 0.82; dan 0.78 yang dikategorikan Valid dan Realiabel; (2.2) Pengujian di Skala besar pada tahap 2 menggunakan *Intraclass Correlation Coefficients* didapat nilai ICC hasil analisis menunjukkan rata-rata kesepakatan antar *rater* sebesar 0,935 sedangkan untuk orang *rater* konsistensinya adalah 0.940 yang artinya instrumen mempunyai kualitas stabilitas yang Cukup tinggi; (3) Kualitas penilaian unjuk kerja yang telah dikembangkan menurut para ahli masing-masing memiliki kategori sangat baik (SB) dengan persentase keidealan sebesar 91% dan menurut guru Matematika SD masing-masing memiliki kualitas sangat baik (SB) dengan persentase keidealan sebesar 90,51 %. Dan bisa diartikan bahwa instrumen unjuk kerja praktis. Berarti instrumen asesmen otentik unjuk kerja pada pembelajaran Matematika yang dikembangkan telah memenuhi syarat validitas, reliabilitas dan kepraktisan, sebagai alat evaluasi yang dapat digunakan lebih lanjut oleh para guru Matematika di sekolah dasar Kota Cirebon.

Beberapa saran yang dikemukakan terkait dengan hasil penelitian ini adalah sebagai berikut.(1) Untuk meningkatkan kualitas proses dan hasil belajar matematika, dapat digunakan instrumen penilaian unjuk kerja pada materi bangun ruang, karena aspek yang dinilai sudah mencakup hampir keseluruhan tugas-tugas kinerja siswa dalam melakukan proses menggambar bangun ruang dan jaring-jaring bangun ruang di kelas. Disamping itu instrumen ini telah memenuhi syarat validitas dan reliabilitas bahkan bernilai praktis dalam penggunaannya. Hal ini akan memberikan manfaat positif terhadap proses penilaian yang objektif, terukur, dan komprehensif atas kemampuan akhir hasil belajar siswa. (2) Untuk mengatasi kesulitan melakukan penilaian unjuk kerja siswa dalam

melakukan praktikum di laboratorium karena jumlah siswa yang relatif banyak pada setiap kelas, dapat dilakukan teknik penilaian pendampingan penilaian dilakukan dalam *team teaching*. Cara lain untuk mengatasi problema ini adalah penilaian diri oleh siswa (*self assesment*) atau teman sejawat (*peer assesment*). (3) Penulis menyadari bahwa penelitian ini masih bersifat sederhana, dimana pengembangan instrumen dilakukan hanya untuk bidang Matematika. Oleh karena itu secara lebih luas penelitian model ini perlu dikembangkan pada bidang keahlian yang lain dengan jumlah sampel yang lebih banyak serta dengan kriteria rubrik dan pembobotan penyekoran yang lebih otentik relevan, sehingga instrumen penilaian unjuk kerja yang dikembangkan dapat memenuhi syarat validitas dan reliabilitas yang lebih signifikan, dan tingkat kepraktisan menggunakan instrumen yang lebih baik.

DAFTAR PUSTAKA

- Arifin, Z. 2009. *Evaluasi Pembelajaran Prinsip Teknik Prosedur*. Bandung: Rosda Karya.
- Arifin, Z. 2012. *Evaluasi Pembelajaran*. Bandung: Remaja Rosda Karya.
- Azwar, S. 2012. *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Pelajar
- Berk, R.A. 1986. *Performance assessment: Methods and applications*. London: The John Hopkins University Press.
- Borg and Gall. 2007. *Education Research, an introduction*. New York & London Longman Inc. Dari NetLibrary, (Online), (<http://www.netlibrary.com>), diakses 8 Juli 2013.
- Carin, A.A. & Sund, R.B. 1989. *Teaching Science Through Discovery*. Columbus: Merrill Publishing Company.
- Cavendish, S. et al. 1990. *Observing Activities: Assessing Science in the Primary Classroom*. London: Paul Chapman Publishing Ltd.
- Chotimah Umi. 2013. *Pengembangan Instrumen Penilaian Domain Afektif Pada Matapelajaran PKn di Sekolah Menengah Pertama*.

- [http://eprints.unsri.ac.id/1076/1/4_Laporan_Penelitian_\(Pengembangan_instrumen_dst\)_UC.pdf](http://eprints.unsri.ac.id/1076/1/4_Laporan_Penelitian_(Pengembangan_instrumen_dst)_UC.pdf) (14 mei 2013)
- Djaali & Muljono, Pudji. 2008. *Pengukuran dalam Bidang Pendidikan*. Jakarta: Grasindo.
- Depdiknas., 2003, *Standar Kompetensi Mata Pelajaran Sains SMP dan MTs*, Jakarta.
- Fuadi, 2012, *Pengembangan Instrumen Penilaian Unjuk Kerja Pembelajaran IPA Unjuk Standar Kompetensi Materi Cahaya Kelas V*, Tesis, Semarang, Program Pascasarjana Universitas Semarang.
- Gabel, D.L. 1993. *Handbook of Research on Science Teaching and Learning*. NewYork: Maccmillan Company.
- Galton, M. & Harlen, W. 1990. *Assessing Science in the Primary School: Written Task*. Lon-don: Paul Chapman Publishing Ltd.
- Ghozali. 2009. *Aplikasi Analisis Multivariate dengan Program SPSS*. Semarang : BP UNDIP
- Gibson, Ivancevich dan Donnelly. 1997. *Organisasi: Perilaku, Struktur, Proses*. Jakarta: Binarupa Aksara.
- Guildford, J.P. 1954. *Psychometric Methods*. New York: McGraw Hill Book Company.
- Hadi, S. 2000. *Metodologi Research*. Yogyakarta : Andi Yogyakarta.
- Hadi, S. 1991. *Analisis Butir untuk Instrumen Angket, Tes dan Skala Nilai dengan Basica*. Yogyakarta: Andi Offset.
- Harlen, W. & Galton, M. (Eds.). 1990. *Observing Activities - Assessing Science in The Pri-mary Classroom*. London: Paul Chapman Publishing Ltd
- Hamid H, 2008. *Evaluasi Kurikulum Sekolah*. PascaSarjana Universitas Indonesia dengan PT. Remaja Rosdakarya
- Haryati, M. 2008. *Model & Teknik Penilaian pada Tingkat Satuan Pendidikan* Jakarta: Gaung Persada Pers.
- Herman, Aschbacher., & Winters. 1992. *Select or design assessments that elicit established outcomes*. Dia mbil pada tanggal 17 Juli 2006, dari <http://www.norel.org/sdrs/areas/issues/methods/assement/as7refs.htm>.
- Heruman. 2008. *Model Pembelajaran Matematika Di Sekolah Dasar*. Bandung: Remaja Rosdakarya.
- Iskandar, T. 2000. *Penerapan Penilaian Kinerja dalam Kegiatan Laboratorium pada Konsep Reproduksi Tumbuhan Biji di Madrasah Aliyah*. Tesis Magister pada PPS UPI: tidak diterbitkan.
- Popham, W.J. 1995. *Classroom assessment: What teachers need to know*, Boston-USA: Ally and Bacon.
- Richard et. al., 2012, “*Mesuring Science Interest: Rasch Validation of The Science Interest Survey*”, *International Journal of Sains and Mathematics Education*, Vol. 10: 643 – 668
- Robbins, Stephen P. 2002. *Prinsip-Prinsip Perilaku Organisasi*. Jakarta: Erlangga.
- Rust, C. 2002. *Purposes and Principles of Assessment*. Oxford. Oxford Centre for Staff and Learning Development OCSLD
- Sa’dijah, C. 2009. *Asesmen Kinerja Dalam Pembelajaran Matematika*. <http://jurnaljpi.files.wordpress.com/2009/09/vol-4-no-2-cholis-sadijah.pdf>. akses: 22 Februari 2012
- Slavin. 2009. *Cooperativ learning*. Nusa Media. Bandung
- Streiner DL, Norman GR, Blum HM (1989). *PDQ Epidemiology*. Toronto: BC Decker, Inc.
- Streiner, DL. & Norman, GR. (2000) *Health measurement scales: A practical guide to their development and use*. Oxford: Oxford University Press.
- Stiggins. 1994. *Student Centered Classroom Assesment*. Macmillan College Publishing Company: New York.k
- Suastra,. 2007. *Pengembangan Penilaian Otentik(Authentic Assesment) Dalam Pembelajaran Fisika*. Makalah. Disampaikan pada Seminar dengan Tema ”Pengembangan Model Pembelajaran Inovatif dan Asesmen sebagai Antisipasi Pelaksanaan KTSP

- di SMP/SMA”, Tanggal 24 sampai dengan 25 September 2007 di Singaraja.
- Sudrajat, Ahmad. 31 Januari 2008. *Teori Pendidikan dan Kurikulum*, (<http://akhmadsudrajat.wordpress.com>, diunduh 15 Desember 2013).
- Suharsimi, A. 2010. *Dasar-Dasar Evaluasi Pendidikan Edisi Revisi*. Jakarta: PT Bumi Aksara.
- Suharsimi, A. 2012, *Dasar – Dasar Evaluasi Pendidikan Edisi 2*, Jakarta: Bumi Aksara
- Sukardi. 2003. *Metodologi penelitian pendidikan: Kompetensi dan prakteknya*. Jakarta: PT Bumi Aksara.
- Susila, I.K. 2012. *Pengembangan Instrumen Penilaian Unjuk Kerja (Performance Assessment) Laboratrium pada Mata Pelajaran Fisika Sesuai Kurikulum Tingkat Satuan Pendidikan SMA Kelas X di Kabupaten Gianyar*, (Online), (http://pasca.undiksha.ac.id/e-journal/index.php/jurnal_ep/article/download/375/167), diakses 19 Mei 2013.
- Suwandi, S. 2010. *Model Asesmen dalam Pembelajaran*. Surakarta: Yuma Pustaka.