



SEMINAR NASIONAL KIMIA DAN PENDIDIKAN KIMIA VI
"Pemantapan Riset Kimia dan Asesmen Dalam Pembelajaran
Berdasarkan Pendekatan Saintifik"
Program Studi Pendidikan Kimia Jurusan PMIPA FKIP UNS
Surakarta, 21 Juni 2014



**MAKALAH
PENDAMPING**

KIMIA PENDIDIKAN

ISBN : 979363174-0

INSTRUMEN PENILAIAN PRAKTIKUM KIMIA DAN ESTIMASI RELIABILITASNYA DENGAN KOEFISIEN GENERALISABILITAS

Endang Susilaningih^{1*}

[1] Jurusan Kimia FMIPA *Universitas Negeri Semarang, Kampus Sekaran Gunungpati
Semarang 50229*

e-mail endang.arkan@gmail.com

ABSTRAK

Penelitian ini bertujuan untuk menyusun instrumen penilaian praktikum Kimia yang teruji reliabilitasnya dengan menggunakan koefisien generalisabilitas berdasarkan teori Brennan. Langkah-langkah penelitian ini dimulai dari membuat *blueprint* penyusunan instrumen penilaian praktikum Kimia, menyusun kisi-kisi, menyusun rancangan lembar observasi, panduan penilaian atau rubrik, dan skoring/penskalaan, uji coba instrumen penilaian praktikum, analisis data uji coba, verifikasi instrumen, penyempurnaan instrumen penilaian praktikum Kimia, implementasi instrumen, analisis data hasil implementasi instrumen penilaian praktikum Kimia sampai diperoleh koefisien generalisabilitas sama/lebih besar dari 0,7 ($r_{xx} > 0,7$). Hasil penelitian diperoleh koefisien generalisabilitas data uji coba = 0,72 ini berarti instrumen reliabilitasnya perlu ditingkatkan, setelah direvisi dan diimplementasikan lagi diperoleh koefisien generalisabilitas = 0,95 yang berarti instrumen reliabel. Kesimpulan dari hasil penelitian ini telah diperoleh instrumen penilaian praktikum Kimia yang teruji reliabilitasnya.

Kata kunci: blueprint, instrumen penilaian, koefisien generalisabilitas, dan reliabilitas.

PENDAHULUAN

Penilaian praktikum Kimia di laboratorium merupakan salah satu penilaian proses yang hanya dapat dilakukan dengan pengamatan langsung pada saat praktikum dilaksanakan.

Berdasarkan kurikulum 2013 evaluasi pembelajaran sains adalah evaluasi otentik yang meliputi penilaian proses dan penilaian hasil belajar. Penilaian proses berupa observasi pada saat pembelajaran

untuk menilai psikomotorik siswa seperti: (aktivitas siswa pada pembelajaran, demonstrasi, mengoperasionalkan prosedur pada penggunaan media, dan presentasi oral laporan tugas) di kelas, keterampilan proses sains, keterampilan laboratorium, keterampilan generik sains, dan keterampilan menggunakan alat-alat elektronik selama praktikum Kimia di Laboratorium. Penilaian hasil belajar adalah penilaian yang dilakukan setelah selesai pembelajaran untuk menilai sejauh mana tujuan pembelajaran telah dicapai, untuk mengetahui (pemahaman konsep siswa, daya serap, dan pencapaian kompetensi siswa) setelah pembelajaran.

Praktikum Kimia merupakan pembelajaran Kimia di laboratorium, menurut Freedman (2002) dalam Chiappetta dan Koballa (2010:213-214) menjelaskan tentang aktivitas kerja laboratorium atau yang disebut dengan praktikum adalah kegiatan psikomotorik siswa yang dapat membantu para siswa memperoleh pengetahuan yang lebih baik tentang konsep-konsep, prinsip-prinsip, hukum-hukum dasar, dan teori yang merupakan hasil dari praktikum untuk memperoleh pengalaman nyata. Praktikum secara umum dapat digunakan untuk meningkatkan sikap terhadap sains, sikap ilmiah, penemuan ilmiah, pengembangan konsep, keterampilan-keterampilan teknik, dan keterampilan-keterampilan dalam kerja kelompok untuk memperoleh pengalaman nyata di laboratorium.

In general, laboratory work can be used to promote the following learning outcomes : attitudes

toward science; scientific attitudes; scientific inquiry; conceptual development; technical skills; and teamwork skills (Chiappetta dan Koballa, 2010:213-214).

Praktikum kimia adalah pembelajaran praktik kimia yang dilaksanakan di laboratorium yang aktivitasnya memprediksi, mengobservasi, mengorganisasi data, dan mendesain percobaan untuk konfirmasi konsep-konsep kimia, prinsip-prinsip, hukum-hukum dasar, dan teori-teori kimia (Freedman, 2002). Praktikum merupakan ciri khusus/karakteristik dari pembelajaran sains, sehingga tidak bisa lepas dari pembelajaran sains. Norman, R dan Iqbal, S (2007:173) menjelaskan pembelajaran di laboratorium merupakan salah satu karakteristik dari pembelajaran sains pada semua tingkatan pendidikan. Praktikum kimia di laboratorium dimaksudkan untuk mendapatkan keterampilan laboratorium, pengalaman laboratorium dan bukti-bukti nyata dari prinsip, konsep, hukum dasar dan teori. Bukti nyata yang diperoleh dari kegiatan praktikum untuk menjembatani antara teori dan dunia nyata dalam kehidupan sehari-hari. Praktikum Kimia tidak bisa lepas dari mata kuliah kimia sebagai pembelajaran di kelas dalam upaya mencapai standar kompetensi yang telah ditentukan. Hal tersebut di sebabkan oleh konsep-konsep kimia yang pada umumnya bersifat abstrak divisualisasikan dengan berbagai bentuk metode. Kegiatan praktikum merupakan salah satu contoh bentuk metode pendekatan visualisasi untuk menjembatani antara teori dengan

dunia nyata dalam kehidupan sehari-hari, memberikan stimulus, ketertarikan, dan perhatian kepada siswa agar mampu dan terampil melakukan praktikum di laboratorium untuk mendapatkan pengalaman tentang hubungan antara teori dengan dunia nyata dalam kehidupan sehari-hari (Ottander & Grelsson, 2006).

Assessment performance merupakan salah satu asesmen yang paling tepat untuk Praktikum Kimia. Asesmen ini mengharuskan siswa unjuk kerja, melakukan sesuatu, mempertunjukkan kinerja, bukan menjawab atau memilih jawaban dari sederetan jawaban yang sudah tersedia (Zainul, 2001:8). Asesmen Praktikum Kimia Analitik Dasar meliputi ujian awal dan responsi yang dilakukan dengan ujian, dan asesmen yang dilakukan tidak dengan ujian berupa diagram alur kerja, laporan praktikum, dan observasi perilaku mahasiswa selama praktikum. Hibbard (1993:1-2) menyatakan bahwa penilaian unjuk kerja adalah sistim yang digunakan untuk menentukan kualitas kinerja mahasiswa dalam menyelesaikan tugas-tugas kinerja. Asesmen dalam penelitian ini adalah asesmen Praktikum kimia yang menggunakan asesmen kinerja laboratorium untuk estimasi peningkatan kinerja/performan siswa di laboratorium, keterampilan laboratorium, dan pengalaman laboratorium setelah selesai melakukan praktikum. Asesmen kinerja merupakan proses pengambilan data secara pengamatan yang sistematik dalam menyelesaikan tugas secara individual (Berk, 1986: ix). Asesmen kinerja secara prinsip terdiri atas dua bagian, yaitu tugas-

tugas (*tasks*) dan kriteria penilaian (*rubric*). Tugas-tugas kinerja adalah tugas-tugas yang mengharuskan siswa unjuk kerja, memperlihatkan kemampuan menangani sesuatu yang kompleks melalui penerapan pengetahuan dan keterampilan dalam bentuk yang nyata (*real-world applications*). Tugas-tugas kinerja di laboratorium adalah keterampilan proses sains untuk mencapai tujuan yang telah ditentukan. Kriteria penilaian atau rubrik merupakan panduan untuk memberi skor. Rubrik harus jelas dan disepakati oleh praktisi laboratorium (siswa, asisten, guru dan Kepala Laboratorium). Diharapkan siswa secara jelas dapat berupaya menyempurnakan kinerjanya dalam mencapai tujuan yang telah ditentukan.

Asesmen kinerja merupakan penilaian subjektif yang menyangkut kualitas kinerja atau hasil kerja yang ditunjukkan oleh setiap siswa. Tentu saja dengan demikian akan terjadi penilaian subjektif, artinya secara mudah akan kehilangan reliabilitasnya dan keadilan dalam penilaian, sehingga sangat diperlukan cara-cara tertentu untuk menjamin reliabilitas, keadilan, dan kebenaran penilaian. Kriteria penilaian atau rubrik berperan penting dalam hal ini, karena rubrik dapat membantu menentukan tingkat ketercapaian kinerja siswa yang diharapkan. Sebagai kriteria penilaian dan alat penskoran, rubrik harus mencakup dimensi-dimensi kinerja, aspek-aspek atau konsep-konsep yang akan dinilai, dan gradasi kualitas kinerja mulai dari tingkat yang paling sempurna sampai dengan

tingkat yang paling buruk. Dimensi-dimensi kinerja tersebut yang akan ditentukan kualitasnya atau yang akan diberi peringkat. Setiap kategori kualitas terdapat contoh kinerja sehingga mempermudah penilai/rater untuk memberikan penilaian. Berbeda dengan tes untuk prestasi yang penyusunannya dimulai dari kisi-kisi tes yang menguraikan secara rinci tujuan/kemampuan yang ingin dicapai, pokok bahasan dan sub pokok bahasan, indikator, butir-butir tes yang dikonstruksi, sehingga dapat digunakan untuk mengukur hasil belajar siswa.

Components of a performance assessment system include: (1) Tasks that ask students to use knowledge and processes they have learned; (2) Checklists that identify the elements of the performance or product to inspect; (3) Rubrics (sets of descriptions of a process and/or a quality-rating continuum) from which to base an overall score for the work; and (4). Examples of excellence for a model of work to be done (Hibbard, 1993: 2)

Validitas dan Reliabilitas.

Validitas dan realibilitas merupakan persyaratan penting yang harus dipenuhi untuk menentukan kualitas suatu instrumen penilaian. Validitas instrumen penilaian merupakan bukti teoretik dan empirik dalam menafsirkan skor tes sesuai dengan tujuan penggunaan tes, oleh

karena itu proses validasi menyangkut pengumpulan bukti-bukti untuk menunjukkan dasar saitifk/ilmiah penafsiran skor sesuai yang direncanakan. Validitas instrumen penilaian merupakan ketepatan dalam memberikan interpretasi terhadap hasil pengukurannya. Instrumen penilaian dapat dikatakan baik apabila mampu menggambarkan ciri-ciri, sifat-sifat, atau aspek yang akan diukur sesuai dengan yang sebenarnya. Secara sederhana validitas didefinisikan sebagai ukuran seberapa cermat suatu instrumen penilaian melakukan fungsi ukurnya (Djemari, 2004:25).

Berdasarkan *Standards for Educational and Psychological Testing* (American Psychological Association, 1985) validitas suatu instrumen penilaian mengacu kepada ketercukupan (*appropriateness*), kebermaknaan (*meaningfulness*) dan kebergunaan (*usefulness*) dari simpulan/ inferensi spesifik yang dapat dibuat dari skor tes yang diperoleh. Validitas berkaitan erat dengan relevansi dan akurasi. Relevansi menunjuk pada kemampuan instrumen untuk menerapkan fungsinya yaitu untuk apa instrumen penilaian yang dimaksudkan. Akurasi menunjuk pada ketepatan instrumen mengidentifikasi aspek-aspek yang akan diukur secara tepat menggambarkan yang sebenarnya.

Reliabilitas pada prinsipnya menunjukkan sejauh mana pengukuran dapat memberikan hasil yang relatif tidak berbeda bila dilakukan pengukuran kembali

terhadap subjek yang sama. Reliabilitas diestimasi dengan koefisien yang menggambarkan derajat relatif kesalahan pengukuran dalam skor-skor tes. Kesalahan pengukuran dalam teori tes klasik (*classical test theory*) diyakini sebagai perbedaan antara skor manapun dengan suatu skor yang sebenarnya. Secara statistik reliabilitas ditunjukkan oleh rasio varians *true score* dengan varians *observe score*. Estimasi reliabilitas instrumen *performance assessment* dapat dilakukan melalui beberapa metode antara lain metode *inter raters reliability*, metode *Kappa-Cohens*, metode *Alpha-Cronbach*, metode *Coefficient Generalizability* dan *The general multifacet Model*. Estimasi reliabilitas instrumen *performance assessment* dalam penelitian ini menggunakan *Coefficient Generalizability* berdasarkan teori Brennan.

Metode **Generalizability Coefficient** adalah salah satu metode estimasi reliabilitas instrumen *performance assessment* yang dikemukakan oleh Brennan. Metode ini jarang digunakan, bahkan dapat dikatakan tidak pernah, karena formula perhitungannya melibatkan persamaan matematika yang tinggi. Kendala yang ditemukan pada para peneliti adalah karena susah menggunakan program Genova. Sebenarnya dari formula yang dikemukakan analisisnya dapat dilakukan dengan manual/kalkulator, program excel, atau program GENOVA. Desain *multifacets universes of admissible* dan *G-study* sebagai komponen observasi yang dikemukakan menggunakan konsep

matematik yang tinggi. Jenis interaksi yang ada sebagai berikut P (*person*), Q (*question*), R (*rater*) dan interaksinya menjadi berbeda-beda:

<i>Design</i>	<i>Main Effects</i>
<i>Interaction Effects</i>	
1. Px Qx R	P,Q,R
	PQ, PR, QR dan PQR.
2. (Q:R)xP	P, R, Q:R
	PR, P(Q:R)
3. P, Q, R	P, R, Q:R
	PR, P(R:Q)
4. P:Q;R	P, Q:P, Q:R:P
-	
5. (QxR):P	P, Q:P, R:P
	QR:P

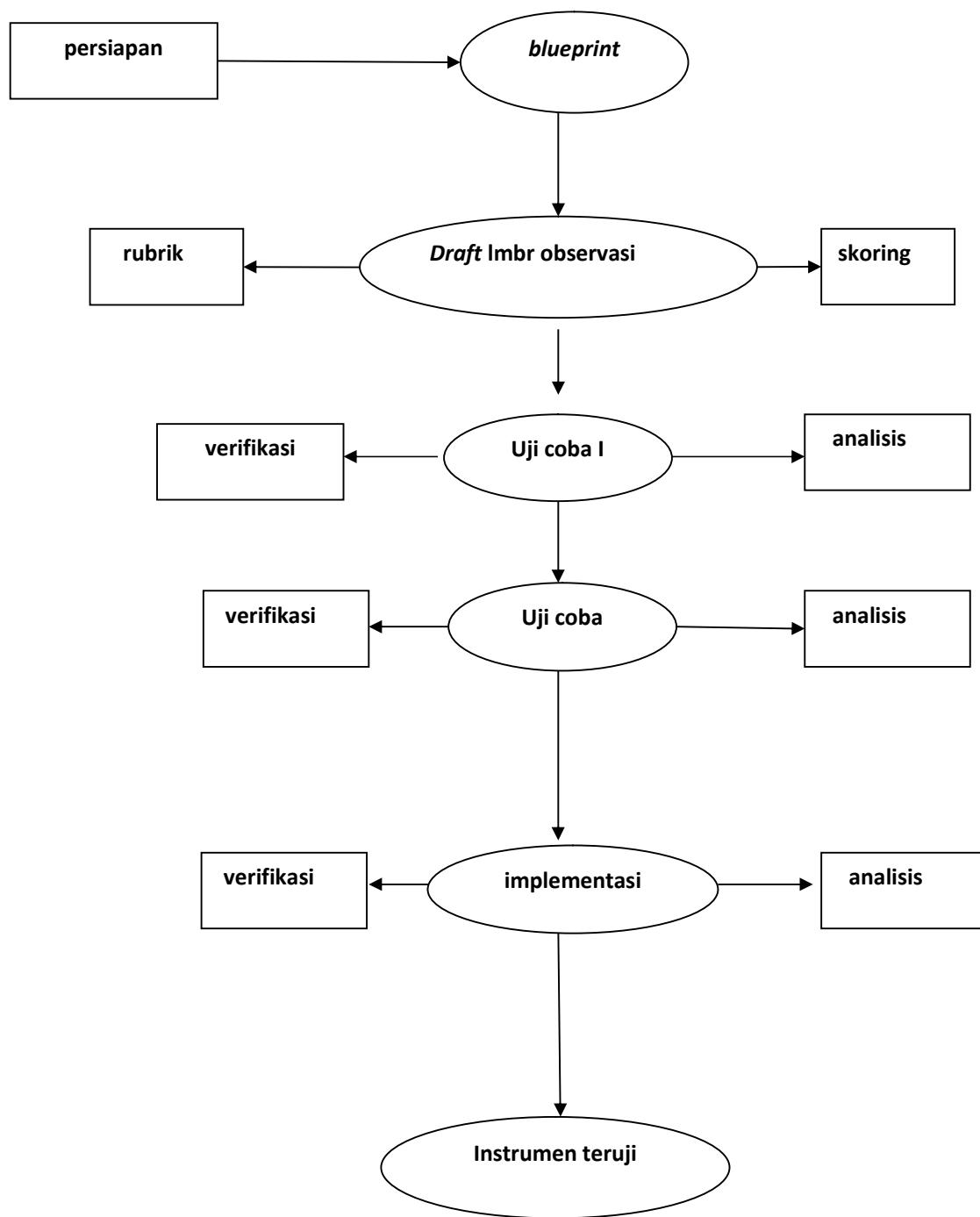
Selain desain *multifacets universes* komponen observasi, untuk estimasi reliabilitas instrumen *performance assessment* menggunakan *G-study* dan *D-study*. *G-study* digunakan untuk estimasi besarnya koefisien reliabilitas antar penilai atau antar observer pada keadaan tertentu. Hasil yang diperoleh dari *G-study* selanjutnya digunakan pada *D-study*. Peneliti mencoba menganalisis koefisien reliabilitas *performance assessment* pada penilaian praktikum Kimia menggunakan **Elements of Generalizability Theory** dengan bantuan program **GENOVA**, serta memilih *Design*, *Main Effects* dan *Interaction Effects* yang sesuai dengan penilaian keterampilan laboratorium pada praktikum kimia yang dilaksanakan pada

semester gasal dan semester genap tahun pelajaran 2013-2014.

PROSEDUR PENELITIAN

Langkah-langkah prosedur penelitian dimulai dari: (1) membuat *blueprint* berisi tujuan yang ingin dicapai pada praktikum kimia, penggalan silabi, kisi-kisi dimensi kerja yang ingin dinilai, rancangan lembar observasi, rancangan rubrik, dan gradasi tingkat ketercapaian kerja yang akan dinilai, skoring dan penskalaan; (2) menyusun instrumen *performance assessment* berdasarkan rancangan dalam *blueprint*; (3) menyusun rubrik atau panduan penilaian *performance assessment*; (4) validasi instrumen *performance assessment* oleh pakar; (5) mengujicobakan *performance assessment*. (6) mengestimasi koefisien generalisabilitas yang setara dengan

reliabilitas instrumen *performance assessment* data uji coba dengan bantuan program GENOVA; (7) menganalisis hasil estimasi reliabilitas *performance assessment*; (8) menyempurnakan instrumen *performance assessment*; (9) mengimplementasikan instrumen *performance assessment* yang telah disempurnakan; (10) mengestimasi reliabilitas instrumen *performance assessment* menggunakan data implementasi dengan cara dan teknik yang sama pada estimasi reliabilitas instrumen *performance assessment* data uji coba; (11) menganalisis hasil estimasi reliabilitas instrumen *performance assessment* data implementasi; dan terakhir (12) mengambil kesimpulan dari hasil analisis. Diagram langkah-langkah penyusunan instrumen dapat dilihat pada Gambar 1



Gambar 1

HASIL PENELITIAN

Estimasi reliabilitas instrumen *performance assessment* dari data uji coba dengan program GENOVA memilih *design* P_xQ_xR ; *Main Effects* P,Q,R; *Interaction Effects* PQ, PR, QR dan PQR dengan jumlah P (jumlah siswa)=20 orang, Q (jumlah aspek) yang dinilai= 4 aspek; dengan R (jumlah observer)=3 orang; *record correlation* =20, pilihan format menggunakan

OPTIONS RECORDS ALL

CORRELATION

EFFECT * P 20 0

EFFECT + R 3 0

EFFECT + T 4 0

FORMAT (12F6.2)

Hasil yang diperoleh koefisien generalisabilitas data uji coba= 0,72 yang masuk dalam kategori reliabilitas tinggi. Sedangkan peneliti menghendaki reliabilitas instrumen *performance assessment* praktikum kimia masuk dalam kategori sangat tinggi yaitu koefisien generalisabilitasnya mendekati 1. Penyempurnaan instrumen dilakukan pada panduan penilaian yang dilengkapi dengan gambar yang jelas dan lebih komunikatif. Diharapkan para

observer lebih mudah dan mempunyai tingkat kesepahaman yang tidak jauh berbeda dalam memberi skor. Setelah instrumen *performance assessment* disempurnakan kemudian di implementasikan ke kelas praktikum rombel berikutnya, data hasil implementasi di estimasi koefisien generalisabilitasnya dengan menggunakan desain dan teknik yang sama seperti pada estimasi koefisien generalisabilitas instrumen *performance assessment* dari data uji coba. Hasil analisis estimasi reliabilitas diperoleh $r_{xx} = 0,956$. Ini berarti instrumen *performance assessment* masuk dalam kategori sangat reliabel atau teruji reliabilitasnya, sehingga instrumen penilaian *performance assessment* ini dapat digunakan untuk menilai performa siswa pada saat praktikum melalui pengamatan langsung pada saat praktikum.

DAFTAR PUSTAKA

- [1] Berk, Ronald, A.(1986). *Performance Assessment*.London: The John Hopkins Press Ltd.
- [2] Brennan, R; L. (1983).*Elements of Generoalizability Theory* (4thed). IowaCity: ACT Publications.
- [3] Chiappetta,E.L., Koballa, Th. R.Jr. (2010). *Science Instruction in the Middle andsecondary Schools. Developing Fundamenta*

Knowledge and Skills (7th ed) New York: Allyn & Bacon .

- [4] Djemari Mardapi. (2004). *Penyusunan Tes Hasil Belajar*. Yogyakarta: Program Pasca Sarjana Universitas Negeri Yogyakarta.
- [5] Freedman, M.P.(2002). The influence of laboratory instruction on science achievement and attitude toward science across gender differences. *Journal of Women and Minorities in Science and Engineering*, 8 (2), 191-199.
- [6] Hibbard, F; K; M. (1993). *Performance Assessment in The Science Classroom*. California: McGraw-Hill Companies. Inc.
- [7] Nitco, A.J., Brookhart, S.M. (2001). *Educational assessment of student* (5th ed), Upper Saddle River NJ: Pearson Education, Inc
- [8] Norman, R., & Iqbal, S. (2007). The role of laboratory work in university chemistry. *Journal of Chemistry Education Research and Practice*, 8 (2), 172-185
- [9] Popham, W.J., (2005) *Classroom Assessment What Teachers Need to Know*, (4th ed) .Boston : Pearson Education, Inc.
- [10] Ross, J.A. 2006. The Reliability, Validity, and Utility of Self Assessment. *Journal Practical Assessment, Research & Evaluation* 11 (10), 1-13
- [11] Tari, Juan Jose, dan Heraz-Saizarbitoria, I. 2012. Self

Assessment. *International Journal for Quality Research*, 6 (4) 343-354

- [12] Zainul, A. (2000). *Alternative Assessment*. Jakarta : Pusat Antar Universitas untuk Peningkatan dan Pengembangan Aktivitas Instruksional, Direktorat Jendral Pendidikan Tinggi. Departemen Pendidikan Nasional.

PEMAKALAH : Endang

Susilaningih

NAMA PENANYA : Diah Ika R

PERTANYAAN :

1. Kelemahan dan keunggulan dari implementasi penilaian prakt. kimia?
2. Seberapa efektif jika implementasinya di sekolah karena observer hanya guru kelas itu sendiri, sedangkan di penelitian tersebut membutuhkan lebih dari satu observer !

JAWABAN :

Kelemahannya pada pemilihan observer, sedangkan keunggulan dari penelitian ini dampak teknik self assessment dapat membangun siswa bersikap jujur, disiplin membuat praktikum menjadi lebih bermakna. Efektif atau tidaknya sangat dipengaruhi oleh peran observer.

LAMPIRAN

CONTOH LEMBAR OBSERVASI

**LEMBAR OBSERVASI PENILAIAN
PRAKTIKUM TITRASI ASIDI-ALKALIMETRI**

Hari/Tanggal:

Nama

Nomor

NO	DIMENSI KERJA YANG DINILAI	TINGKAT KETERCAPIAN				BOBOT	SKOR
		1	2	3	4		
I	MENYIAPKAN PRAKTIKUM						
	1. Menyiapkan alat yang digunakan					5	
	2. Menyiapkan larutan kerja					10	
	3. Menyiapkan larutan untuk pembakuan larutan baku yang akan digunakan					5	
JUMLAH							
20							
II	KETERAMPILAN PROSES						
	1. Keterampilan merang					10	

	kai alat untuk praktikum titrasi						
2.	Keterampilan menuang larutan baku ke dalam buret					5	
3.	Keterampilan menuang larutan yang akan dianalisis ke dalam erlenmeyer.					5	
4.	Keterampilan memilih dan menetapkan indikator ke dalam larutan yang akan dianalisis					5	
5.	Keterampilan melakukan pembakuan					10	

	larutan baku yang akan digunakan					
6.	Keterampilan melakukan titrasi				10	
7.	Keterampilan melakukan pengamatan				5	
8.	Keterampilan menganalisis hasil pengamatan.				10	
60	JUMLAH					