

Authentic assessment: Evaluation and its application in science learning

Maria Hanifah *, Alfred Irambona

Universitas Negeri Yogyakarta. Jalan Colombo No 1, Karangmalang, Yogyakarta, 55281, Indonesia
Burundi University. Boulevard Mwezi Gisabo, Bujumbura, Burundi

* Corresponding Author. E-mail: maria.hanifah05@gmail.com

Received: 6th September 2018; Revised: 25th February 2019; Accepted: 22th August 2019

Abstract

The research aims to reveal: (1) the suitability of the implementation of science authentic assessment in Kulonprogo Regency, and (2) the forms of science authentic assessment applied. The study was used the CIPP Stufflebeam evaluation model. The quantitative data were analyzed with T-score, while the qualitative data were analyzed by using Miles & Huberman. The results of this research are as follows. (1) The implementation of science authentic assessment is wholly fairly effective be seen from T-score analysis 52.44 from the score extent 20-80. (2) The forms of science authentic assessment are practice, portfolio, teacher transcript journal, and daily test, whereas project activity, self assessment, and peer assessment are not done for the heat matter and its movement with T-score 45.14 wholly fairly effective from the score extent 20-80.

Keywords: penerapan, penilaian otentik IPA, K-13

How to Cite: Hanifah, M., & Irambona, A. (2019). Authentic assessment: Evaluation and its application in science learning. *Psychology, Evaluation, and Technology in Educational Research*, 1(2). doi:<http://dx.doi.org/10.33292/petier.v1i2.4>



INTRODUCTION

Kurikulum pada dasarnya dituntut untuk selalu berubah sesuai dengan perkembangan jaman. Kurikulum baru membutuhkan tantangan untuk mewujudkan penerapannya, dengan menargetkan tiga elemen sukses. Tiga elemen sukses tersebut adalah peserta didik, guru, dan materi (Sparapani, Perez, Gould, Hilman, & Clark, 2014, p.12). Elemen sukses tersebut menjadi subjek evaluasi. Kurikulum berubah dengan demikian, sudah dilakukan evaluasi sebelumnya dan diperoleh kebijakan. Penerapan kebijakan harus diimbangi dengan kemampuan dan kesiapan pelaksanaannya yaitu guru, peserta didik, dan kepala sekolah yang berlaku sebagai pelaksana langsung kurikulum. Perbedaan penerapan Kurikulum yang digunakan di Indonesia pada tahun 2013 sampai sekarang (K-13) secara khusus berimbas juga pada penilaian IPA SMP di Kabupaten Kulonprogo, baik perencanaan, pelaksanaan, dan evaluasinya

K-13 menekankan guru harus melaksanakan penilaian otentik dalam pembelajarannya. Penilaian otentik di K-13 juga mengacu pada peraturan yang sudah ada tentang Standar Penilaian Pendidikan di Indonesia. Penilaian otentik menurut Kunandar (2014, pp.35-36) adalah suatu kegiatan pengumpulan data yang memberikan gambaran perkembangan belajar peserta didik baik proses maupun hasilnya dengan instrumen penilaian.

Penilaian otentik IPA di SMP Kabupaten Kulonprogo dapat diketahui secara aktual dalam penerapan K-13 dengan melakukan penelitian evaluasi. Sekolah percontohan yang telah melaksanakan K-13 selama 3 semester diwajibkan menjalankan K-13 yang kemudian akan dievaluasi penilaian otentiknya. Materi yang digunakan adalah kalor dan perpindahannya. Evaluasinya menggunakan model CIPP.

Penilaian otentik dilakukan karena memiliki manfaat yang sangat luas. Penilaian otentik memiliki karakteristik yang harus dilakukan, yaitu memiliki relevansi dunia nyata, membuat peserta didik merelevansi tugas dalam dunia nyata, memberikan tugas yang kompleks secara

berkesinambungan, peserta didik mengamati tugas-tugas perspektif yang berbeda, terdapat refleksi diri, ada kerja sama dalam team, menghasilkan hasil karya yang bernilai, menghasilkan pemikiran pemecahan masalah yang kompetitif (Basuki & Hariyanto, 2015, p.170). Sehingga penilaian otentik akan menciptakan peserta didik yang memiliki sikap aktif, inovatif, sosialis, dan *scientific*.

Penilaian otentik merupakan penilaian yang dilakukan secara komprehensif untuk menilai mulai dari masukan (*input*), proses, dan keluaran (*output*) pembelajaran. Penilaian menjadi suatu perkembangan yang dasar ketika guru dan peserta didik menggunakan informasi dari penilaian tersebut sebagai informasi pengajaran serta pemberharuan dari pembelajaran. Penilaian pada umumnya dapat menggambarkan tes di akhir bab, rapot semester, menghasilkan ulangan keterampilan dasar, dan menyusun sebuah rapot laboratorium akhir (Atkin, J. M., Black, P., & Coffey, J., 2001, pp. 12-24). Penilaian akan mencakup seluruh proses yang dialami peserta didik dan diakhiri dengan pemberitahuan hasil penilaian yang dilakukan.

Aspek komponen yang harus diperhatikan dalam melakukan penilaian otentik diantaranya adalah otentik dari instrumen yang digunakan artinya menggunakan instrumen otentik yang bervariasi (tes tertulis, tes lisan, tes proyek, dan tes kinerja); otentik dari aspek yang diukur artinya mengukur aspek hasil belajar dari kompetensi sikap, pengetahuan, dan keterampilan; otentik dari aspek kondisi peserta didik melihat *input*, proses, dan *output* (Kunandar, 2014, p.42).

Penilaian otentik haruslah berjalan berkesinambungan yang menyangkup keseluruhan pelaksanaan penilaian. Penilaian otentik dibuat untuk melengkapi penilaian tes standar yang bisa digunakan untuk melengkapi tes baku yang ada (Pantiwati, 2013, p.5). Penilaian otentik dapat menyangkup tes, tugas dan proyek selama pembelajaran. Penilaian otentik memastikan bahwa semua peserta didik diberi kesempatan untuk menunjukkan apa yang bisa mereka lakukan, sehingga guru memiliki informasi yang dibutuhkan untuk membangun sebuah penilaian yang seimbang dan benar dari masing-masing peserta didik. Penilaian otentik dapat memfasilitasi keseluruhan kebutuhan peserta didik dalam perkembangannya dan dirangkum dalam penilaian portofolio. Portofolio memberikan respon peserta didik dapat memandang tujuan pembelajaran, dan meluaskan pandangan pembelajaran peserta didik (Nezakatogoo, 2011, p.232).

Penilaian otentik juga mencakup 3 aspek kemampuan. Pemerintah Indonesia menjabarkan dalam aspek penilaian sikap dapat dilakukan dengan observasi, penilaian diri, penilaian teman sejawat, dan jurnal. Penilaian keterampilan dengan tes praktik, proyek, dan portofolio. Penilaian pengetahuan dapat dilihat dari tes tulis, lisan, dan penugasan.

IPA merupakan proses ilmiah yang bersifat empiris, sistematis, dan logis serta sikap ilmiah seperti sikap ingin tahu, menghargai pembuktian, sabar, kritis, tidak putus asa, kreatif, dan berdaya cipta. IPA adalah ilmu yang dipelajari secara *scientific* dengan berbagai metode, model dapat berupa fakta, konsep, prinsip, hukum, dan teori (Lukum, 2015, p.64).

Evaluasi adalah kegiatan untuk mengetahui pencapaian suatu program. Evaluasi memberi informasi untuk kebijakan dalam dua cara (Lukum, 2015, p.64). Pertama evaluasi memberi informasi bagi pembuat kebijakan tentang keadaan pendidikan atau pencapaian belajar suatu grup tertentu. Kedua, informasi evaluasi digunakan sebagai piranti administratif untuk menerapkan kebijakan. Evaluasi pembelajaran IPA bertujuan untuk mendeskripsikan program pembelajaran IPA SMP di Kabupaten Kulonprogo. Evaluasi program penilaian adalah kegiatan mengevaluasi hal yang dilakukan dalam proses penilaian di pembelajaran IPA sesuai standar penilaian meliputi perencanaan, pelaksanaan, dan proses penilaian serta dampaknya terhadap peserta didik.

Evaluasi pada khususnya adalah proses penting yang mengkover seluruh penilaian otentik. Evaluasi yang dilakukan dalam penelitian IPA tidak ada ukuran pasti untuk menentukan suatu keputusan akhir sebagai *judgements*. Hasil dari evaluasi ini dapat diadopsi sebagai bahan rekomendasi di akhir (Levin, Mirin, & Rozalis, 2000, p.424). Evaluasi membutuhkan intelektual yang didukung dengan logika dalam melangkah sehingga memberikan keputusan akhir.

Hakikat IPA harusnya memiliki dimensi untuk menciptakan jalan pikir peserta didik, jalan untuk melakukan investigasi, jalan menumbuhkan pengetahuan, sehingga terbentuk interaksi antara teknologi dan sosial (Chiapetta & Koballa, 2010, p.105). IPA disini mengembangkan proses ilmiah untuk membentuk pola pikir peserta didik dan membutuhkan teknik penilaian khusus agar terukur hakikat IPA tersebut.

Penilaian yang sedang berlangsung harusnya menjadi kegiatan yang terintegrasi dengan proses pengajaran dan pembelajaran. Guru menilai setiap proses yang dialami oleh peserta didik (Chiapetta

& Koballa, 2010, p.6). Penilaian pembelajaran IPA tidak hanya berlangsung di akhir pembelajaran, namun dilakukan selama proses pembelajaran. Penilaian dalam pembelajaran IPA di K-13 haruslah menggunakan penilaian otentik. Penilaian otentik ini menerapkan konsep dan teori yang dimiliki peserta didik dengan kemampuan dan keterampilan yang dimiliki dapat diamati langsung (Mardapi, 2012, pp.166-167). Penilaian IPA tersebut menekankan kemampuan peserta didik yang aktif dalam setiap pembelajaran.

Hakikat IPA pada dasarnya adalah mengembangkan kompetensi keterampilan-keterampilan proses ilmiah, memelihara sikap-sikap ilmiah, dan produk ilmiah (mengembangkan konsep-konsep yang berkaitan dengan pengalaman sehari-hari) menurut Ali, Suastra, & Sudiatmika (2013, pp.12-14) IPA berkaitan dengan cara mencari tahu tentang alam secara sistematis sehingga IPA bukan hanya penguasaan kumpulan pengetahuan berupa fakta, konsep, prinsip, hukum, dan teori, tetapi juga merupakan suatu proses penemuan. Materi IPA menyediakan sekumpulan integrasi dari berbagai materi yaitu fisika, kimia, biologi, ilmu bumi, dan astronomi (Hewitt, Lyons, Suchocki, & Yeh, 2007, p.xvi). IPA dapat disimpulkan menekankan pada pembelajaran yang dilakukan melalui kegiatan proses ilmiah yaitu observasi atau eksperimen berdasarkan materi-materi IPA.

Tujuan utama dari pelajaran IPA adalah melibatkan semua pikiran peserta didik dan memberi pemahaman akhir kepada peserta didik untuk belajar lebih dari topik bahasan yang disampaikan. Seorang guru sains haruslah memahami rencana dari pembelajaran yang akan dilakukan dan apa saja yang akan dipelajari oleh peserta didik (Chiappetta & Koballa, 2010: 5) Pembelajaran IPA yang baik membentuk *knowledge outcomes*, *reasoning outcomes*, *skill outcomes*, *product outcomes*, dan *affective outcomes* (Stiggins, 1994: 3). Pembelajaran pada dasarnya akan membentuk pengetahuan, memberikan materi pemecahan masalah, menghasilkan, menciptakan, dan mengalikasikan kemampuan dalam menguasai materi. Rencana dari pembelajaran dilengkapi dengan penilaian selama proses pembelajaran. Penilaian yang sedang berlangsung harus menjadi kegiatan yang terintegral dengan proses pengajaran dan pembelajaran. Guru menilai setiap proses yang dialami oleh peserta didik (Chiappetta & Koballa, 2010: 6). Penilaian pembelajaran IPA tidak hanya berlangsung di akhir pembelajaran, namun dilakukan selama proses pembelajaran.

Penilaian yang dilakukan sepanjang hayat adalah penilaian otentik. Penilaian otentik merupakan penilaian yang melibatkan keaktifan peserta didik. Penilaian yang dilakukan mengalami evolusi dengan keterlibatan peserta didik secara langsung, maksudnya penilaian yang tidak dilakukan satu kali selama periode waktu. Penilaian otentik harus multifaset, terus menerus, interaktif, dan kolaboratif. Penilaian otentik juga akan memberikan wawasan bagi para guru dan peserta didik (Murphy, 1994, pp.23-24). Penilaian otentik ini harus berjalan terus menerus sehingga keaktifan peserta didik dapat terasah dengan baik dan tertampung perkembangannya oleh guru.

Penilaian otentik haruslah berkesinambungan maksudnya penilaian yang mencakup keseluruhan pelaksanaan penilaian. Penilaian otentik dibuat untuk melengkapi penilaian tes standar, jadi penilaian otentik oleh guru digunakan untuk melengkapi tes baku yang ada (Pantiwati, 2013, p.5). Penilaian otentik dapat mencakup tes baku, tugas dan proyek selama pembelajaran. Penilaian otentik memastikan bahwa semua peserta didik diberi kesempatan untuk menunjukkan apa yang bisa mereka lakukan, dan guru memiliki informasi yang dibutuhkan untuk membangun sebuah penilaian yang seimbang dan benar dari masing-masing peserta didik. Penilaian otentik dapat memfasilitasi keseluruhan kebutuhan peserta didik dalam perkembangannya dan dirangkum dalam penilaian portofolio.

Model evaluasi yang digunakan penelitian ini adalah *CIPP model*. Model *CIPP* menurut Stufflebeam & Shinkfield (2002, p.159) memberikan arti "*evaluation is the process delineating, obtaining, and providing descriptive and judgment information ... in order to guide decision making serve needs for accountability and promote understanding of the involved phenomena*". Stufflebeam mengartikan model evaluasi CIPP dapat memberikan gambaran, hasil, dan menyediakan informasi pertimbangan yang bermanfaat untuk mengambil keputusan yang bertanggung jawab.

Evaluasi merupakan proses penting yang mengkover seluruh penilaian otentik. Evaluasi yang dilakukan dalam penelitian IPA tidak ada ukuran pasti untuk menentukan suatu keputusan akhir sebagai *judgements*. Hasil dari evaluasi ini dapat diadopsi sebagai bahan rekomendasi di akhir (Levin, 2000: 424). Evaluasi membutuhkan intelektual yang didukung dengan logika dalam melangkah sehingga memberikan keputusan akhir.

Penelitian ini dapat dimanfaatkan oleh guru sebagai pertimbangan untuk meningkatkan kualitas pembelajaran IPA dalam menanggapi kebijakan perubahan kurikulum. Bagi sekolah, diharapkan hasil

penelitian dapat digunakan sebagai pertimbangan evaluasi pembelajaran dan memberi masukan dalam rangka pelaksanaan K-13 pada tahap berikutnya. Bagi pengambil kebijakan, hasil penelitian ini dijadikan pertimbangan untuk mengatasi polemik permasalahan perubahan kurikulum yang beredar di masyarakat, sehingga menghasilkan persiapan yang lebih baik. Bagi para peneliti, diharapkan penelitian ini sebagai kajian konseptual yang menjawab permasalahan kurikulum pembelajaran IPA dalam K-13, sehingga dapat digunakan sebagai acuan pembuatan penelitian R & D.

METODE

Penelitian ini merupakan jenis penelitian evaluasi program. Program yang diteliti adalah program penilaian IPA di K-13. Penelitian ini menggunakan teknik deskriptif kualitatif dan kuantitatif. Evaluasi ini dilakukan di SMP piloting se-Kabupaten Kulonprogo yang telah melaksanakan K-13. SMP piloting yang melaksanakan K-13 akan diteliti pada pembelajaran IPA di kelas VII. Evaluasi ini dilakukan selama kurang lebih 3 bulan pada akhir semester genap.

Populasi yang akan berperan sebagai responden ialah guru, dan peserta didik di Kabupaten Kulonprogo yang menerapkan K-13. Guru IPA yang dievaluasi adalah guru IPA yang mengajar di kelas VII. Peserta didik yang diberi perlakuan evaluasi adalah peserta didik kelas VII. Sampel kelas VII diambil 2 kelas dengan teknik *purposive sampling* kelas yang diampu guru terpilih sebagai pemilihan kelas yang akan diamati proses pembelajarannya. Guru berdasarkan pertimbangannya akan memilihkan dua kelas yang akan digunakan sebagai penelitian. Teknik sampel yang digunakan dengan rumus Issac dan Michael. Rumus tersebut terdapat pada Persamaan 1.

$$S = \frac{\lambda^2 \cdot N \cdot P \cdot Q}{d^2 (N-1) + \lambda^2 \cdot P \cdot Q} \quad (1)$$

Keterangan:

S = Jumlah sampel

N = Jumlah populasi = 682

λ^2 = Chi Kuadrat tergantung dengan derajat kebebasan dan tingkat kesalahan, dk = 1, taraf 5%, chi kuadrat nya 3,841

P = Q = 0,5

d = Perbedaan antara rata-rata sampel dengan rata-rata populasi yaitu 0,05

Sampel yang digunakan dengan jumlah seluruh populasi peserta didik kelas VII yang menggunakan K-13 yaitu 682, dengan sebaran SMP PL 1 Kalibawang jumlah peserta didik kelas VII yaitu 55 peserta didik. SMP N 1 Wates jumlah peserta didik kelas VII yaitu 192. SMP Negeri 1 Galur berjumlah 150 peserta didik untuk kelas VII. SMP Negeri 1 Saminggaluh berjumlah 97 peserta didik untuk kelas VII. SMP Negeri 2 Lendah berjumlah 188 peserta didik kelas VII.

Jumlah populasi kelas VII yang sebanyak 682 dapat diperoleh jumlah sampel yang diambil sebanyak 246 menurut perhitungan S sampel (Sugiyono, 2012, p.129). Jumlah sampel ini kemudian dibagi secara merata untuk 5 sekolah *pilot proyek* hingga diperoleh sampel masing-masing sekolah sebesar 49 peserta didik dan khusus SMP Negeri 1 Wates digunakan 50 peserta didik sebagai sampel karena memiliki jumlah peserta didik terbanyak. Rincian sampel dapat dilihat pada Tabel 1.

Tabel 1. Rincian Sampel Penelitian

No	Nama Sekolah	Subjek Penelitian	Jumlah Responden
1.	SMP Negeri 1 Wates	Guru IPA	1
		Peserta didik	50
2.	SMP Negeri 1 Galur	Guru IPA	1
		Peserta didik	49
3.	SMP Negeri 1 Saminggaluh	Guru IPA	1
		Peserta didik	49
4.	SMP Negeri 2 Lendah	Guru IPA	1
		Peserta didik	49
5.	SMP Panggudi Luhur 1 Kalibawang	Guru IPA	1
		Peserta didik	49
Jumlah			246

Prosedur

Data yang dikumpulkan dalam penelitian ini adalah data informasi lisan berupa fakta, keterangan tertulis, dan angka-angka yang mendukung permasalahan perenapan penilaian otentik IPA di Kabupaten Kulonprogo. Teknik pengumpulan data yang digunakan oleh karena itu adalah studi dokumentasi, teknik angket, teknik observasi dan wawancara. Teknik pengumpulan data ini adalah triangulasi data. Triangulasi data digunakan agar tidak ada ketimpangan data satu dengan yang lain, sehingga data dapat dipertahankan keakuratannya dengan menganalisis ketiga data yang digunakan. Kisi-kisi instrumen angket guru dapat dilihat pada Tabel 2.

Tabel 2. Contoh Kisi-Kisi Instrumen Angket Guru

Aspek	Subaspek	Indikator	Butir		Jumlah
			Positif	Negatif	
Perencanaan	Analisis kemampuan guru	1. Pemahaman standar penilaian	1,2	-	2
		2. Pelatihan penyusunan penilaian otentik	3,4,5	-	3
	Analisis sarana prasarana	3. Penyediaan kertas dalam penilaian.	6	-	1
		4. Penyediaan alat dan bahan praktikum.	7	-	1
	Perencanaan instrumen	5. Penyusunan rencana penilaian otentik dalam perangkat pembelajaran RPP.	8	-	1
		6. Penyusunan rencana penilaian otentik dalam perangkat pembelajaran media.	9	-	1
		7. Penyusunan rencana penilaian otentik dalam perangkat pembelajaran LKS.	10	-	1
		8. Penyusunan rubrik penilaian	11	-	1
		9. Penyusunan kisi-kisi penilaian.	12	-	1
		10. Perumusan instrumen secara jelas.	-	13	1
		11. Penyusunan indikator penilaian diri dapat memahami peserta didik dalam memetakan kemampuannya.	14	-	1
		12. Penyusunan rencana penilaian otentik dalam kegiatan praktikum.	15	-	1
		13. Penyusunan penilaian otentik dalam kegiatan proyek.	16	-	1
		14. Penyusunan rencana penilaian otentik dengan portofolio yang dikumpulkan.	-	17	1
Pelaksanaan	Persiapan pelaksanaan penilaian	15. Menyampaikan aspek yang akan dinilai.	18	-	1
		16. Menyampaikan waktu pelaksanaan penilaian	19	-	1
		17. Menyampaikan bentuk/jenis penilaian yang digunakan.	20	-	1
	Penilaian sikap	18. Menjelaskan petunjuk instrumen penilaian diri.	21	-	1
		19. Penilaian antarteman terhadap 3 (tiga) teman sekelas atau sebaliknya.	22	-	1
		20. Guru mencatat seluruh hasil temuan selama pembelajaran.	23	-	1
		21. Terdapat ulangan harian di akhir bab dalam bentuk ulangan atau penugasan.	24	-	1
	Penilaian pengetahuan keterampilan	22. Melakukan kegiatan percobaan	25	-	1
		23. Peserta didik mencatat temuannya	26	-	1
		24. Peserta didik menyimpulkan temuannya	27	-	1
		25. Peserta didik mempresentasikan hasil percobaan di kelas	-	28	1
		26. Guru menjelaskan kegiatan dari percobaan.	29	-	1
		27. Guru menjelaskan hal-hal yang berbahaya dari percobaan (<i>sefty-lab</i>)	30	-	1
		28. Bertindak sebagai fasilitator dan menawarkan bimbingan kepada peserta didik.	31	-	1
		29. Menyusun kegiatan proyek secara mandiri.	32	-	1
		30. Mengumpulkan dan menyimpan hasil kerja	33	-	1

Aspek	Subaspek	Indikator	Butir		Jumlah
			Positif	Negatif	
Pelaporan	Prinsip penilaian	peserta didik dalam 1 map atau folder dengan mengingatkan peserta didik untuk memberi tanggal pembuatan.			
		31. Memberikan waktu perbaikan bagi karya yang kurang dengan memberikan jangka waktu.	-	34	1
		32. Guru memeriksa dan mengembalikan pekerjaan peserta didik dan selanjutnya memberikan umpan balik dan komentar yang sifatnya mendidik.	35	-	1
		33. Transparasi hasil penilaian.	37	36	2
		34. Penggunaan instrumen penilaian.	38	-	1
		35. Melaksanakan program pengayaan.	39, 40	-	1
		36. Melaksanakan program remedial.	41	-	1
		37. Melaksanakan program perbaikan.	42	-	1
		38. Analisis hasil pelaporan.	43,44	-	2
		39. Penyusunan pelaporan hasil belajar peserta didik.	45,46	-	2
	Standar pelaporan	Jumlah	41	5	46

Kisi-kisi instrumen angket peserta didik digunakan sebagai pedoman membuat lembar angket peserta didik. Lembar angket peserta didik digunakan sebagai data proses pelaksanaan penerapan penilaian otentik. Lembar angket ini berjumlah 23 butir. Kisi-kisi instrumen angket peserta didik dapat dilihat pada Tabel 3.

Tabel 3. Kisi-Kisi Instrumen Angket Peserta Didik

Aspek	Subaspek	Indikator	Butir		Jumlah
			Positif	Negatif	
Pelaksanaan	Persiapan penilaian	1. Disampaikan aspek yang akan dinilai.	1	-	1
		2. Disampaikan waktu pelaksanaan penilaian.	2	-	1
		3. Disampaikan bentuk/jenis penilaian yang digunakan.	3	-	1
	Penilaian sikap	4. Dijelaskan cara kerja instrumen penilaian diri	-	4	1
		5. Melaksanakan penilai antarteman terhadap 3 (tiga) teman sekelas atau sebaliknya	5	-	1
	Penilaian pengetahuan	6. Terdapat ulangan harian di akhir bab dalam bentuk ulangan atau penugasan.	6	-	1
		7. Melakukan kegiatan percobaan dan mencatat temuannya	-	7	1
		8. Peserta didik menyimpulkan temuannya	8	-	1
	Penilaian keterampilan	9. Peserta didik mempresentasikan hasil percobaan di kelas	-	9	1
		10. Guru menjelaskan kegiatan dari percobaan.	10	-	1
		11. Guru menjelaskan hal-hal yang berbahaya dari percobaan (sefty lab)	11	-	1
		12. Bertindak sebagai fasilitator dan menawarkan bimbingan kepada peserta didik.	12	-	1
		13. Menyusun kegiatan proyek secara mandiri.	13,14	-	2
		14. Mengumpulkan dan menyimpan hasil kerja peserta didik dalam 1 map atau folder dengan memberi tanggal pembuatan.	15,16	-	2
		15. Diberikan waktu perbaikan bagi karya yang kurang dengan memberikan jangka waktu.	-	17	1
		16. Diperiksa dan dikembalikan tugasnya dan selanjutnya diberikan umpan balik dan komentar yang sifatnya mendidik.	18	-	1

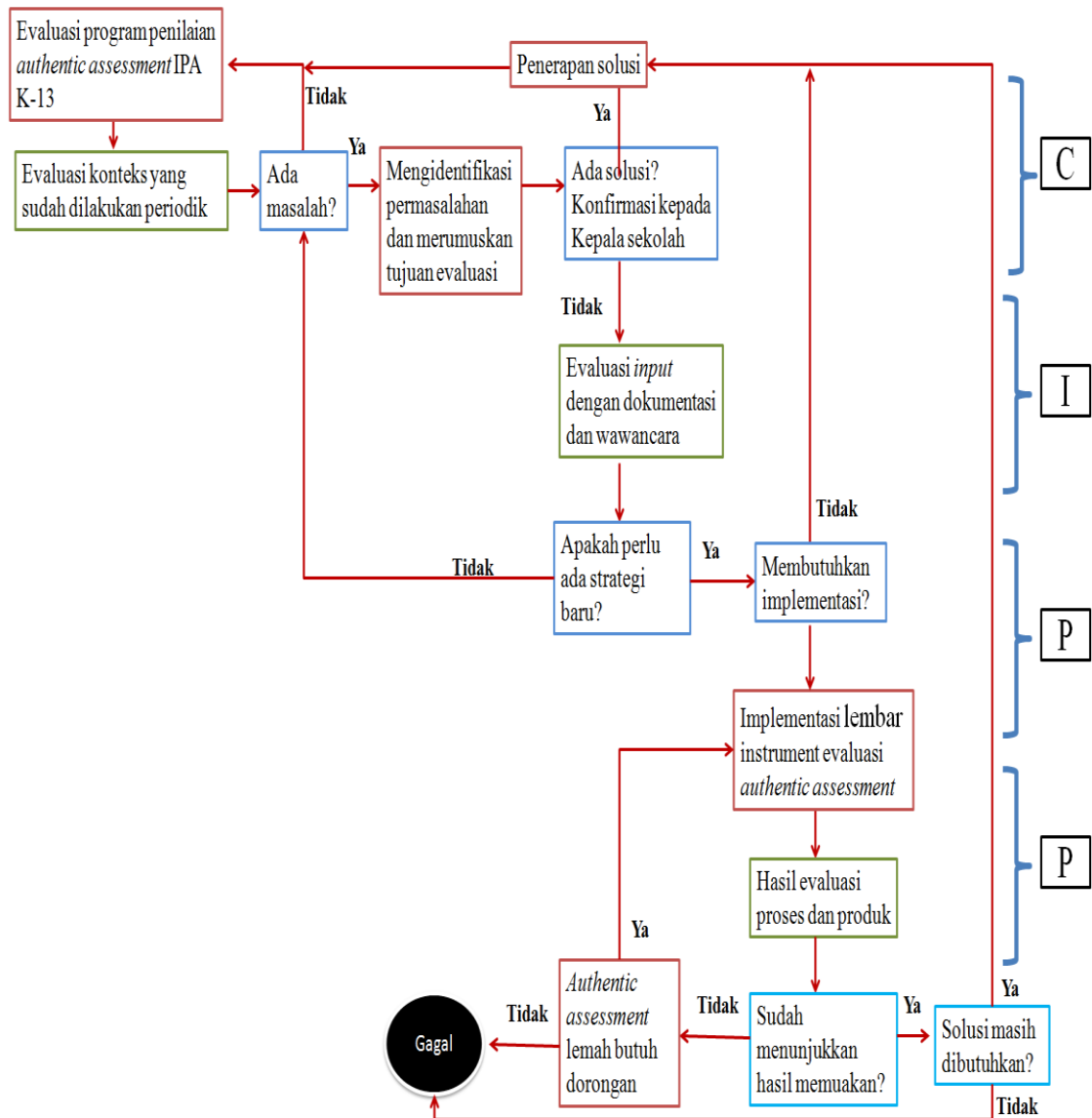
		17. Transparansi hasil penilaian.	-	19	1
	Prinsip penilaian	18. Melaksanakan program remedial dan pengayaan.	20	-	1
		19. Pendidik melaksanakan ujian ulangan bagi peserta didik yang mengikuti pembelajaran remedial.	21	-	1
Pelaporan	Standar pelaporan	20. Penyerahan hasil belajar peserta didik.	22,23	-	2
Jumlah			18	5	23

Data, Instrumen, dan Teknik Pengumpulan Data

Instrumen pengumpulan data dibuat berdasarkan indikator. Indikator yang dirancang sesuai dengan kebutuhan penelitian untuk masing-masing subjek penelitian. Skala yang digunakan adalah skala likert dengan skala 4 dengan bentuk *check list* (Widoyoko, E.,P., 2012, p.105). Skala linkert ini digunakan untuk lembar angket.

Rincian dari pengelompokkan penggunaan instrumen penelitian sebagai berikut. (1) Data *context* dilakukan pada guru dengan instrumen lembar angket dan pedoman wawancara untuk guru. (2) Data *input* atau data perencanaan diperoleh dari penelitian tahap *input* dilakukan pada guru dengan instrumen lembar angket, dokumentasi, dan pedoman wawancara. Instrumen berupa kuosioner yang akan diisi guru dilanjutkan wawancara dan studi dokumentasi. Analisis perangkat juga dilakukan sebagai data tambahan dengan dokumentasi. (3) Data *process* diperoleh dengan lembar observasi, dan lembar angket. Evaluasi proses dilakukan pada peneliti dengan cara observasi serta dokumentasi kegiatan penilaian untuk memperkaya informasi proses penilaian. Guru dan peserta didik diberikan observasi saat pelaksanaan pembelajaran. Untuk mengecek ketepatan hasil observasi secara menyeluruh diberikan lembar angket pada guru dan siswa. (4) Data *product* diperoleh dari hasil penyimpulan semua kegiatan *context*, *input*, dan *process* dengan instrumen lembar angket pada guru dan peserta didik, serta instrumen wawancara pada peserta didik dengan dokumentasi hasil akhir pembelajaran. Model alur penelitian dapat dilihat pada Gambar 1.

Model Alur Penelitian Evaluasi CIPP



Gambar 1. Model alur penelitian

Seluruh instrumen yang digunakan dilakukan validitas dan uji reabilitas. Validitas akan menunjukkan sah tidaknya instrumen yang peneliti disusun. Reliabilitas adalah mengukur tingkat keajekan instrumen digunakan berkali-kali. Reliabilitas digunakan untuk menentukan baik tidaknya instrumen. Reliabilitas instrumen untuk lembar observasi dan lembar angket dapat direliabilitas dengan menentukan nilai estimasi *alpha Cronbach*. Lembar angket peserta didik dianalisis untuk menentukan *reliability estimasi*.

Validitas isi dari instrumen penelitian dapat menggunakan analisis rasional. Analisis rasional menggambarkan butir instrumen indikator dari variabel yang diukur dalam penelitian. Langkah dalam menyusunnya yaitu butir instrumen divalidasi berdasarkan indikator-indikatornya dengan diserahkan pada validator (*expert judgment*) kemudian revisi hasil validasi ahli. Data yang diperoleh dari validasi ahli berupa penskoran 1-5. Data tersebut dihitung kevalidannya dengan menghitung koefisien validitas isi dengan *Aiken's V*. Data tersebut akan dikelompokkan berdasarkan kriteria keberhasilan. Rumus yang ditunjukkan oleh Aiken dalam Azwar (2012: 113) dijelaskan pada Persamaan 1 dan 2.

$$V = \frac{s}{[n(c-1)]} \quad (2)$$

$$S = r - l_0 \quad (3)$$

Keterangan:

lo = angka penilaian validitas yang terendah

c = angka penilaian validitas tertinggi

r = angka yang diberikan oleh penilai

n = jumlah penilai

Agar instrumen yang dibuat peneliti dapat dikatakan memiliki validasi isi yang tinggi maka instrumen divalidasi oleh 7 ahli sama dengan raters (n) 7. Validasi terhadap tujuh orang tersebut dengan skala penilaian 1-5 atau *number of rating categories* 5. Maka dari itu, rentang yang diperbolehkan (valid) dari hasil analisis Aiken terendah adalah 0,82 (Aiken, 1985, p.137). Berikut ini adalah hasil analisis Aiken tiap instrumen.

Teknik Analisis Data

Analisis data yang digunakan adalah deskriptif kuantitatif dan deskriptif kualitatif.

Analisis Deskriptif Kuantitatif

Data kuantitatif diperoleh dari data angket, observasi, dan dokumentasi, kemudian ditentukan tingkat penerapan penilaian otentik dengan mengkatogorikan dalam skala lima dapat dilihat pada Tabel 1. Kategori penilaian tersebut digunakan untuk menghitung data ketercapaian penerapan penilaian otentik IPA yang sudah berjalan. Data mentah yang dapat diolah menjadi data baku menggunakan z -score dan diubah agar tidak menghasilkan nilai negatif menjadi T -score.

Model evaluasi yang digunakan adalah CIPP. Model CIPP ini akan diukur keberhasilannya dalam mengukur evaluasi dengan kriteria keberhasilan. Kriteria keberhasilan yang digunakan dapat menggunakan skala 5. Kriteria penilaian (Rosana, 2015: 361) dapat dirumuskan seperti Tabel 4.

Tabel 4. Rumus Interval Nilai dan Kriterion Penilaian

No	Interval Nilai	Kriteria Penilaian
1	$X > \bar{X} + 1,5 SBx$	Sangat efektif
2	$\bar{X} + 0,5 SBx < X \leq \bar{X} + 1,5 SBx$	Efektif
3	$\bar{X} - 0,5 SBx < X \leq \bar{X} + 0,5 SBx$	Cukup efektif
4	$\bar{X} - 1,5 SBx < X \leq \bar{X} - 0,5 SBx$	Kurang efektif
5	$X \leq \bar{X} - 1,5 SBx$	Sangat Kurang efektif

Keterangan:

$\bar{X}$ = Nilai rata-rata skor ideal

SBx = Simbangan baku skor ideal

X = Skor responden

SBx diperoleh dari 1/6 kali pengurangan skor ideal dengan skor terendah ideal. $\bar{X}$ diperoleh dari 1/2 kali penjumlahan skor tertinggi ideal dengan skor terendah ideal. Skor maksimum ideal pada setiap instrumen apabila responden memilih dengan skor tertinggi setiap butir, sedangkan skor minimum ideal jika responden tidak memilih semua kriteria dari butir pernyataan yang ada.

Evaluasi dengan model CIPP ini akan diperoleh data kuantitatif yang dianalisis menggunakan rumus z -score Arikunto (2012, p.303). Z -score adalah angka perbandingan yang membedakan skor seseorang dengan mean dan standar deviasi. Z -score digunakan untuk mengetahui penyetaraan beberapa instrumen yang menggunakan skala yang berbeda-beda. Menghitung z -score harus mengetahui rata-rata dari kelompok kelas (rata-rata masing-masing data instrumen) dan standar deviasi kelas tersebut (instrumen tersebut). Rumus interval nilai dapat dilihat pada Tabel 5.

Tabel 5. Rumus Interval Nilai dan Kriterion Penilaian

No.	Interval Nilai	Kriteria Penilaian
1	$X > 65$	Sangat efektif
2	$55 < X \leq 65$	Efektif
3	$45 < X \leq 55$	Cukup efektif
4	$35 < X \leq 45$	Kurang efektif
5	$X \leq 35$	Sangat kurang efektif

Perhitungan T -score dan Z -score

$$Z = \frac{x - \bar{x}}{SD} \quad (4)$$

Keterangan:

Z = nilai baku atau standar skor

$\bar{X}$ = Rata-rata kelas data instrumen

X = Data indikator instrumen

Dengan dihasilkan angka-angkat z -score yang bernilai dari -3 sampai 3, kemudian angka-angka tersebut dibulatkan menjadi angka kisaran 50 dengan T -score. Skor pada masing-masing indikator dihitung T -score. T -skor adalah angka yang menggunakan $M = 50$ dan $SD = 10$. Skala T -score dihitung dari z -score dikalikan 10 (Arikunto, 2012: 306). Rumus T -score disajikan sebagai berikut.

$$T = 10Z + 50 \quad (5)$$

Analisis Deskriptif Kualitatif

Data kualitatif diperoleh dari data wawancara dan interpretasi dari hasil perhitungan T -score. Hasil interpretasi data akan dijelaskan secara kualitatif dengan analisis teori Miles, Huberman, dan Saldana (2014, p.33).

HASIL DAN PEMBAHASAN

Penerapan penilaian otentik IPA di Kabupaten Kulonprogo secara keseluruhannya berjalan cukup efektif. Nilai T -score yang diperoleh dalam sebaran CIPP yaitu 52,44 dalam rentang 20-80. Penilaian otentik dilakukan secara komprehensif untuk menilai mulai dari masukan (*input*), proses, dan keluaran (*output*) pembelajaran.

Perencanaan penilaian otentik berdasarkan angket guru, wawancara, dan dokumen menunjukkan bahwa guru IPA se-Kabupaten Kulonprogo telah melaksanakan perencanaan pembelajaran dengan baik. Data wawancara terhadap 5 guru, bahwa guru telah menyusun perangkat penilaian terdapat pada RPP yang disiapkan. Perencanaan penilaian tersebut tidaklah disusun guru secara mandiri, guru menyiapkan secara bersama-sama dalam MGMP atau kegiatan lainnya. Hasil analisis seluruh instrumen perencanaan yang digunakan, guru memiliki T -score 50,32 berkategori cukup efektif pada rentang 20-80.

Pelaksanaan penilaian otentik IPA guru hanya melaksanakan beberapa penilaian keterampilan dan penilaian pengetahuan. Penilaian sikap dilaksanakan guru hanya 1 kali dalam 1 semester atau dalam 1 tahun sehingga tidak teramati dalam lembar observasi. Pelaksanaan penilaian otentik IPA memiliki T -score 49,45 bernilai cukup efektif pada rentang 20-80. Tabel 6 menunjukkan data penerapan penilaian otentik IPA.

Tabel 6. Data T -Score Penerapan Penilaian Otentik IPA

Model	T -score rentang (20-80)				Σ	Kesimpulan
	Angket Guru	Angket Peserta Didik	Observasi	Dokumen		
Bobot	4	5	5	4		
C	48,51	-	-	57,77	53,14	Cukup efektif
I	52,94	-	-	48,33	50,64	Cukup efektif
P	48,49	49,67	50	-	49,45	Cukup efektif
P	55,62	56,22	-	57,78	56,51	Efektif
Hasil					52,44	Cukup efektif

Hasil perencanaan didukung dengan data wawancara perencanaan penilaian otentik terhadap 5 guru, bahwa guru telah menyusun perangkat penilaian terdapat pada RPP yang disiapkan. Perencanaan penilaian tersebut tidaklah disusun guru secara mandiri, guru menyiapkan secara bersama-sama dalam MGMP atau kegiatan lainnya

Pelaksanaan penilaian otentik IPA juga didukung dengan data wawancara terhadap 5 guru, bahwa guru memiliki komitmen dalam melaksanakan penilaian otentik, namun tidak secara keseluruhan guru selalu melaksanakan ulangan setiap babnya. Satu guru tidak melaksanakan ulangan setiap babnya. Penilaian otentik lainnya berdasarkan hasil observasi hanya kegiatan portofolio, praktik, dan jurnal yang dilaksanakan guru setiap babnya. Berdasarkan wawancara tersebut dapat disimpulkan untuk indikator 5 yaitu guru cukup efektif dalam melaksanakan penilaian otentik IPA. Hasil pelaporan juga didukung dengan data wawancara terhadap 5 guru, bahwa guru melaksanakan pelaporan penilaian otentik, namun tidak secara keseluruhan guru selalu melaksanakan analisis hasil ulangan

setiap babnya. Berdasarkan wawancara tersebut dapat disimpulkan untuk indikator 5 yaitu guru efektif dalam pelaporan penilaian otentik IPA.

Bentuk penilaian otentik IPA yang diterapkan dalam K-13 di Kabupaten Kulonprogo diperoleh dengan instrumen angket guru, angket peserta didik, dan observasi. Keseluruhan penilaian otentik yang dilakukan cukup efektif dengan nilai *T-score* yaitu 45,14. Penilaian sikap yang terdiri tiga subaspek. Subaspek penilaian sikap memiliki *T-score* bernilai cukup efektif pada penilaian diri dan dua lainnya kurang efektif. Hal ini terjadi karena tidak ada guru yang melaksanakan penilaian diri dan antar teman selama pembelajaran kalor dan perpindahannya. Penilaian pengetahuan berjalan dengan cukup efektif, dengan nilai *T-score* 54,79. Kemudian lembar portofolio dan lembar kinerja juga berjalan cukup efektif dengan nilai *T-score* ≥ 50 .

Jenis-jenis penilaian otentik IPA yang dilaksanakan di Kabupaten Kulonprogo dalam bab kalor dan perpindahannya adalah penilaian kinerja, portofolio, jurnal catatan guru, dan ulangan harian. Penilaian otentik IPA secara keseluruhan berjalan cukup efektif dengan nilai *T-score* 45,14 pada rentang 20-80. Penilaian dalam pembelajaran IPA sesuai standar penilaian mengukur aspek pengetahuan menggunakan teknik tes tertulis dan tes lisan, sedangkan aspek keterampilan dapat diukur dengan observasi, unjuk kinerja, penugasan, portofolio, dan jurnal. Aspek sikap diukur dengan observasi, lembar penilaian antarteman, dan penilaian sendiri. Tabel 7 menunjukkan data aspek pelaksanaan penilaian otentik IPA.

Tabel 7. Data *T-Score* Aspek Pelaksanaan Penilaian Otentik IPA

Jenis Penilaian	Sub-Penilaian	Angket Guru	Angket Peserta Didik	Observasi	Σ	Keterangan
Penilaian Sikap	Penilaian Diri	50,6	56,89	35,53	47,46	Cukup efektif
	Penilaian Antar Teman	33,88	36,77	35,53	35,5	Kurang efektif
	Catatan Guru	37,22	-	46,62	44,5	Kurang efektif
Penilaian Pengetahuan	Ulangan Harian	47,26	57,89	57,72	54,79	Cukup efektif
Penilaian Keterampilan	Penilaian Kinerja	47,26	54,57	58,64	53,93	Cukup efektif
	Penilaian Proyek	35,55	26,71	35,53	32,39	Sangat kurang efektif
Remidi dan Pengayaan	Penilaian Portofolio	49,49	49,69	50,78	50,02	Cukup efektif
	Pengayaan	45,58	55,55	41,08	47,53	Cukup efektif
	Remidi Soal Berbeda	37,22	-	41,08	39,36	Kurang efektif
	Remidi Soal Sama	40,57	49,61	52,17	47,94	Cukup efektif
Total					45,14	Cukup efektif

Kegiatan penilaian lainnya seperti proyek, penilaian diri, dan penilaian antar teman yang harusnya dilaksanakan, namun tidak dilaksanakan dalam materi kalor dan perpindahannya. Penilaian proyek tidak dilaksanakan berdasarkan wawancara guru dikarenakan kerja proyek dalam materi kalor dan perpindahannya yaitu membuat kompor surya sulit bagi siswa kelas VII. Hal ini dikarenakan kondisi bahan dari kaca yang disusun membentuk mangkuk itu membutuhkan biaya besar dan keahlian khusus. Penilaian proyek berdasarkan analisis skor *T* memiliki nilai 32,39 bermakna sangat kurang efektif pada rentang 20-80. Nilai kerja proyek ini paling kecil dari penilaian lainnya. Hal ini menandakan bahwa kegiatan proyek paling kurang efektif dalam penilaian bab materi kalor dan perpindahannya. Kegiatan proyek secara keseluruhan maka dari itu tidaklah dijalankan di 5 sekolah yang menjadi penelitian.

Penilaian diri dan penilaian antar teman juga berjalan cukup efektif dan kurang efektif, dikarenakan kelima guru melaksanakan penilaian sikap tersebut hanya 1 kali dalam 1 semester atau 1 kali dalam 1 tahun. Maka dari itu nilai skor *T* yang diperoleh dari penilaian diri adalah 47,46 berkategori cukup efektif dan penilaian antar teman sebesar 35,5 berkategori kurang efektif pada rentang 20-80. Berdasarkan nilai *T* yang diperoleh dua penilaian sikap tersebut dapat disimpulkan penilaian antar teman lebih cukup efektif dalam penerapan pelaksanaannya.

Pendidik sesuai standar penilaian menjamin pelaksanaan ulangan harian atau ujian bebas dari kemungkinan terjadinya tindak kecurangan. Pendidik sesuai standar penilaian harus melaksanakan ulangan harian di akhir bab dalam bentuk ulangan atau penugasan. Penilaian otentik dijalankan sesuai

dengan prinsip penilaian yang digunakan sebagai acuan dalam merencanakan suatu penilaian. Prinsip penilaian yang baik menurut standar penilaian pendidikan memiliki cakupan, yaitu sahih, objektif, adil, terpadu, terbuka, berkesinambungan, sistematis, akuntabel, dan edukatif. Prinsip tersebut digunakan standar pembuatan penilaian.

SIMPULAN

Penerapan penilaian otentik IPA pada K-13 di Kabupaten Kulonprogo secara keseluruhan berjalan cukup efektif. Hal ini dilihat dari hasil analisis *T-score* bernilai yaitu 52,44 pada rentang 20-80. Hasil penerapan penilaian otentik IPA dapat memberikan informasi dari bentuk-bentuk penilaian otentik IPA yang diterapkan. Bentuk-bentuk penilaian otentik IPA di K-13 yang sudah diterapkan di Kabupaten Kulonprogo yaitu lembar kinerja, portofolio, jurnal catatan guru, dan ulangan harian. Bentuk penilaian lainnya seperti lembar proyek, penilain diri, dan penilaian antar teman tidak dilaksanakan pada materi kalor dan perpindahannya, hanya dilaksanakan 1 kali dalam semester atau 1 tahun. Penilaian otentik IPA di Kabupaten Kulonprogo secara keseluruhan berjalan cukup efektif dengan nilai *T-score* 45,14 pada rentang 20-80.

Saran yang dapat diberikan dengan masalah sebagian besar guru belum menguasai instrumen penilaian. Pemerintah perlu melaksanakan pelatihan mengenai penyusunan instrumen penilaian otentik IPA, sehingga guru dapat memahami hakikat dari pelaksanaan penilaian otentik IPA. Pembelajaran kerja proyek pembuatan kompor surya tidak dilaksanakan oleh guru IPA di SMP Kabupaten Kulonprogo. Pemerintah perlu memotivasi guru dan memberikan pelatihan dalam menyusun kegiatan proyek kompor surya supaya kegiatan proyek kompor surya dapat terlaksana.

Saran berikutnya terkait kreativitasnya dalam mengajar baik dalam penugasan dan beragamnya kegiatan pembelajaran perlu guru tingkatkan, sehingga penilaian otentik terasa menjadi kegiatan yang efektif. Penilaian tersebut memiliki aspek terlalu banyak untuk diamati. Perlu dibuatkan penilaian otentik yang mudah dan bisa dilakukan oleh satu orang guru baik aspek sikap, keterampilan, dan pengetahuan. Kepraktisan tersebut dapat dibuat seperti yang dilakukan salah satu guru responden penelitian yang menilai kegiatan praktikum peserta didik.

DAFTAR PUSTAKA

- Ali, Suastra & Sudiatmika. (2013). Pengelolaan pembelajaran IPA ditinjau dari hakikat sains pada SMP di Kabupaten Lombok Timur. *e-Jurnal Program Pascasarjana Universitas Pendidikan Ganesha*, Volume 3.
- Arikunto, S. (2012). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Arikunto, S., & Jabar, C.S.A. (2010). *Evaluasi program pendidikan*. Jakarta: Bumi Aksara.
- Atkin, J. M., Black, P., & Coffey, J. (2001). *Committee on classroom assessment and the national science education standards, center for education, national research council*. Washington, D.C.: National Academy Press.
- Basuki, I., & Hariyanto. (2014). *Asesmen pembelajaran*. Bandung: PT Remaja Rosdakarya.
- Blerkom, M. (2009). *Measurement and statistics for teachers*. New York: Routledge.
- Chiappetta, E.L., & Koballa, T.R. (2010). *Science instruction in the middle and secondary schools developing fundamental knowledge and skills seventh edition*. United States of America: Pearson Education.
- Fitzpatrick, J., & Sanders, J. (2011). *Program evaluation alternative approaches and practical guidelines*. United States: Pearson.
- Kemdikbud. (2014). *Peraturan menteri pendidikan dan kebudayaan republik indonesia nomor 104 Tahun 2014 tentang penilaian hasil belajar oleh guru pada pendidikan dasar dan pendidikan menengah*.
- Kunandar. (2014). *Penelitian otentik (penilaian hasil belajar peserta didik berdasarkan kurikulum 2013) suatu pendekatan praktis disertai dengan contoh*. Jagakarsa: PT Rajagrafindo Persada.
- Levin, Mirin, & Rozalis. (2000). *Abduction: a logical criterion for programme and project evaluation. Jurnal Evaluation SAGE Publications Vol 6(4): 415-432*.

- Lukum, A. (2015). Evaluasi program pembelajaran IPA SMP menggunakan model countenance stake. *Jurnal Penelitian dan Evaluasi Pendidikan Volume 19, No 1, Juni 2015 (25-37)* .
- Mardapi, D. (1999). *Pengukuran penilaian & evaluasi pendidikan*. Buku Beta.
- Mardapi, D. (2012). *Performance assessment dalam perspektif*.
- Miles, Huberman, & Saldana. (2014). *Qualitative data analysis a methods sourcebook edition 3*. USA: Sage.
- Murphy, A. A. (1994). A model for authentic assessment utilizing portfolios. *Jurnal UMI Northern Illinois University*.
- Nezakatgoo, B. (2011). The effects of portfolio assessment on writing of EFL students. *Journal English Language Teaching* Vol. 4, No. 2.
- Pantiwati, Y. (2013). Authentic assessment for improving cognitive skill, critical-creative thinking and meta cognitive awareness. *Journal of Education and Press* Vol. 4, No. 14.
- Rosana, D. (2014). Pendekatan saintifik dalam pembelajaran IPA secara terpadu. *Seminar Nasional Unes*, (pp. 4-21). Semarang.
- Sparapani, E.F., Perez, D.C., Gould, J., Hilman, S., & Clark, L. (2014). A global curriculum? Understanding teaching and learning in the United States, Taiwan, India, and Mexico. *SAGE open*, 1-15 .
- Stiggins, R. (1994). *Student centered classroom assessment*. New York: Macmillan College Publishing Company
- Stufflebeam, D., & Shinkfield, A. (2004). *Systematic evaluation a self-instruction guide to theory and practice*. United states of America: Nijhoff Publishing.
- Subali, B. (2012). *Prinsip assessment dan evaluasi pembelajaran*. Yogyakarta: UNY Pres