

## PENGEMBANGAN PERANGKAT PENILAIAN AUTENTIK UNTUK PEMBELAJARAN MATEMATIKA DI KELAS VII SEMESTER 1

Rolina Amriyanti Ferita

Program Studi Pendidikan Matematika Universitas Nahdlatul Ulama Kalimantan Selatan  
rolinaamriyantiferita@gmail.com

---

**Abstrak:** Penelitian ini bertujuan untuk menghasilkan perangkat penilaian autentik untuk pembelajaran matematika di kelas VII semester 1 yang autentik, reliabel dan valid meliputi instrumen dengan kompetensi dasar bilangan bulat, garis dan sudut, serta rasio dan perbandingan untuk penilaian KI 1, KI 2, KI 3, dan KI 4 berupa: indikator, instrumen penilaian, dan rubrik penilaian. Jenis penelitian ini adalah penelitian pengembangan (*research & development*). Penelitian ini dilakukan dengan menggunakan model pengembangan Plomp (2010) tiga tahap pengembangan yaitu *preliminary research*, *prototyping phase*, dan *assessment phase*. Aspek reliabilitas dari uji coba lapangan. Aspek kevalidan perangkat pembelajaran ditinjau dari penilaian para ahli. Hasil penelitian menunjukkan bahwa untuk aspek penilaian autentik berupa aspek konteks dan tugas memperoleh kategori baik, sedangkan aspek peserta didik dan indikator memperoleh kategori sangat baik. Perangkat yang dikembangkan juga telah memenuhi taraf konsistensi dan reliabilitas, dengan hasil penilaian kualitas oleh ahli memperoleh kategori valid.

*Kata Kunci: pengembangan, perangkat penilaian autentik, assessment, matematika kelas VII*

---

Pendidikan sudah menjadi kebutuhan bagi setiap manusia di era modern. Prestasi dan keterampilan menjadi kebutuhan mutlak bagi bekal kehidupan. Indonesia telah menerapkan wajib belajar 9 tahun dan akan mengembangkan menjadi wajib belajar 12 tahun seperti yang tercantum dalam Pengembangan Kurikulum 2013 oleh Kemendikbud (2012, hal. 7). Para orang tua peserta didik dan guru menyadari betapa pentingnya bekal pendidikan. Penilaian sumatif sering menjadi acuan orang tua dan guru dalam melihat kompetensi yang telah diperoleh peserta didik. Nilai ulangan ataupun nilai rapor menjadi hal yang paling umum dalam melihat prestasi peserta didik. Akan tetapi, justru kebermaknaan dari pembelajaran yang dilalui oleh peserta didik yang akan mengantarkan peserta didik belajar menuju

kesuksesan. Berdasarkan hasil TIMSS (*Trends in International Mathematics and Science Study*) 2011 dan PISA (*Program for International Student Assessment*) 2009 memperlihatkan bahwa peserta didik Indonesia yang hanya mampu sampai level 3 masih lemah dalam analisis penerapan dalam kehidupan sehari-hari. Peserta didik di lapangan pada faktanya masih kesulitan untuk jenis masalah yang berkaitan dengan kehidupan sehari-hari yang membutuhkan analisis, proses, kemudian mengevaluasi hasil yang diperoleh. Penilaian konvensional berupa *paper and pencil test* mulai ditinggalkan karena tidak autentik.

Menurut *Indiana Department of Education*, peserta didik yang mengerjakan tes dengan hasil yang buruk bisa terjadi karena tidak memahami pembelajaran dari awal,

memahami pembelajaran tetapi tidak belajar kembali saat akan tes, memahami pembelajaran dan sudah belajar tetapi tidak memahami maksud dari pertanyaan atau perintah saat tes. Tes merupakan salah satu bentuk dari penilaian. Penilaian merupakan kata yang lebih sering didengar dan merupakan kata umum dalam Bahasa Indonesia dari *assessment*. Pemahaman terhadap penilaian (*assessment*) sering tertukar antara evaluasi (*evaluation*), tes (*test*) dan pengukuran (*measurement*). Menurut Nitko & Brookhart (2011, hal. 3-7) penilaian adalah istilah yang luas didefinisikan sebagai proses untuk memperoleh informasi yang digunakan untuk membuat keputusan terhadap siswa, kurikulum, program dan sekolah dan kebijakan pendidikan.

Penilaian merupakan hal yang sangat penting berkolaborasi dengan pembelajaran sejalan dengan yang Hyde (2013, hal. 196) sampaikan yakni suatu penilaian memang diperlukan sebagai satu kesatuan utuh bersama kurikulum dan pembelajaran. Bukti sebagai pendukung hasil pencapaian masing-masing peserta didik dari waktu ke waktu juga berperan besar. Miller (2009, hal. 28) menyatakan bahwa penilaian adalah suatu proses untuk memperoleh informasi terhadap kinerja peserta didik. Proses memperoleh informasi juga sangat menentukan agar keautentikan dari penilaian terdukung.

Frey (2007, hal. 417) menyatakan bahwa penilaian autentik digunakan untuk mengukur kemampuan dalam tugas yang mewakili masalah dunia nyata. Pemahaman tentang kata autentik memang sangat kompleks bahkan kadang menimbulkan beberapa pendapat yang berbeda. Para peneliti memiliki pendapat yang berbeda-beda mengenai keautentikan (Gulikers, 2004, hal. 69). Penilaian autentik (PP no.66, 2013, hal. 2)

merupakan penilaian yang komprehensif secara menyeluruh. Autentik sering digunakan sebagai tugas cerminan dari kenyataan pengetahuan yang siswa miliki (Frey, *et al.*, 2012, hal. 12). Salah satu jenis penilaian yang digunakan dalam kurikulum 2013 adalah penilaian autentik. Beberapa penulis menggunakan istilah penilaian autentik yang juga disebut dengan penilaian alternatif (Nitko, 2011, hal. 246; Lai, 2011, hal. 1; Reynolds, 2009, hal. 248; Weber, 2001, hal. 14; Frey, 2007, hal. 406; Anderson, 2003, hal. 72). Reynolds (2009, hal. 26-27) juga mengemukakan bahwa penilaian autentik atau penilaian kinerja yang kompleks biasanya dalam bentuk penilaian kinerja. Sejalan dengan yang dikemukakan oleh Wright (2008, hal. 163) bahwa saat tugas kinerja berdasarkan pada kehidupan nyata maka dapat melampaui dasar minimal standar sekolah.

Suatu penilaian dianggap autentik ketika guru langsung menguji siswa dengan tugas intelektual yang tepercaya (Wiggins, 1990, hal. 1). Hal ini sejalan dengan yang disampaikan oleh Nitko (2011, hal. 246-247) bahwa arti kata autentik dalam penilaian autentik adalah menyajikan tugas secara langsung kepada peserta didik yang berarti bagi pendidikan mereka sehingga bermakna.

Guru di sekolah menengah dalam situasi kelas yang heterogen baik latar belakang sikap, keterampilan, maupun pengetahuan dituntut tepat dalam penilaian tetapi dengan berbagai tantangan yang mendukung semua siswa (Moon, *et al.*, 2005, hal. 121). Keadilan tidak selalu semua disamakan. Keadilan tidak ada dalam penilaian yang seragam, standar, tidak mengacu hanya pada orang tertentu, dan mutlak. Penilaian adil apabila menurut ukuran tertentu, alami, dan fleksibel sesuai dengan kemampuan tingkatan masing-masing (*Indiana Department of*

*Education*, hal. 2). Menurut Amrosetti dan Cho (2005, hal. 57), menilai kompetensi siswa selalu menjadi tugas yang tidak mudah, terutama didorong oleh kurikulum saat ini. Tantangan tersendiri bagi pendidik untuk memenuhi tujuan kompetensi yang akan dicapai, menilai kemajuan peserta didik dengan gaya belajar yang beragam, serta mempersiapkan peserta didik menjadi sumber daya manusia yang berkualitas.

Penilaian akan autentik apabila menyerupai kegiatan yang realistis dalam situasi dunia nyata (Gulikers, Bastiaens & Kirschner, 2006, hal. 8). Pedoman yang dikembangkan untuk menentukan penilaian autentik (konteks, faktor peserta didik, faktor tugas, dan indikator)

Pendekatan yang pertama berasal dari Herrington & Herrington (2006, hal. 147) tentang pedoman yang dikembangkan untuk menentukan penilaian autentik. Penilaian akan menjadi autentik apabila memenuhi kriteria sebagai berikut.

Faktor yang pertama adalah konteks yang membutuhkan keterhubungan dan transfer ke dunia di luar kelas. Membutuhkan rekaman murni dari tugas yang dikerjakan dalam kondisi biasa terjadi atau sehari-hari.

Faktor kedua adalah peserta didik yang meliputi keterampilan pemecahan masalah dan berpikir tingkat tinggi. Memproduksi pengetahuan dari pada memproduksi kembali pengetahuan. Membutuhkan kerja sama dan waktu dari peserta didik. Membutuhkan peserta didik yang efektif menggunakan pengetahuan yang diperoleh untuk membentuk capaian kinerja atau produk, serta meningkatkan kedalaman pengetahuan (Newmann & Wehlage, 1993, hal. 9).

Faktor ketiga adalah tugas yang merangsang berbagai tanggapan aktif, kompleks, tidak restruktur yang memerlukan

pengambilan keputusan, terdiri dari beberapa langkah, dan terdiri dari serangkaian tugas. membutuhkan penilaian sehingga dapat terpadu dengan aktivitas.

Faktor yang terakhir adalah indikator. Membuat beberapa indikator yang mendukung pembelajaran serta mencapai validitas dan reliabilitas dengan kriteria yang tepat untuk mencetak produk bervariasi.

*Assessment* dalam pembelajaran matematika dalam penelitian ini menyesuaikan dengan kurikulum 2013 meliputi kompetensi sikap spiritual, sikap sosial, pengetahuan, dan keterampilan. Menurut Shapiro (2011, hal. 148) menilai matematika dimulai dengan rangkaian pembelajaran untuk keterampilan berhitung di masing-masing sekolah. Ketika keterampilan berhitung berkembang, peserta didik dapat dengan tepat menerapkan pada keterampilan yang membutuhkan matematika, seperti: keuangan, waktu, geometri, pengukuran, dan penerapan konsep matematika yang lain. Peserta didik harus mampu memecahkan permasalahan dengan efektif dan menggunakan grafik. Sehingga penilaian konsep dan penerapan perlu ditambahkan dalam penilaian komputasi. Penilaian yang menyeluruh kemudian perlu analisis dan deskripsi.

Masing-masing penilaian autentik dari kompetensi peserta didik memiliki deskripsi sebagai pedoman bagi guru dan wali dari peserta didik untuk mendukung pembelajaran selanjutnya. Deskripsi sikap oleh wali kelas dari masing-masing pendidik yang mengajar melalui analisis nilai sikap setiap mata pelajaran dan proses diskusi berkesinambungan dalam periode tertentu. Deskripsi sikap antar mata pelajaran menguraikan kelebihan sikap peserta didik dan sikap yang masih perlu ditingkatkan. Contoh uraian deskripsi sikap seperti: menunjukkan

sikap baik dalam kejujuran, disiplin, toleransi, gotong royong, santun, dan percaya diri. Perlu ditingkatkan sikap tanggung jawab, melalui pembiasaan penugasan mandiri di rumah.

Deskripsi kompetensi pengetahuan juga harus sesuai dengan analisis hasil dari keseluruhan periode dengan konversi skala 4 yakni: sangat baik, baik, cukup baik, kurang. Kemdikbud Dirjen Dikdas Direktorat Pembinaan (2014, hal. 47) memberikan contoh untuk deskripsi akhir nilai mata pelajaran yaitu tingkat pemahaman rerata adalah baik. Memiliki pemahaman sangat baik pada kompetensi dasar bilangan, himpunan, dan perbandingan, tetapi masih memerlukan penguatan pada kompetensi dasar garis dan sudut. Hasil penilaian keterampilan berupa laporan oleh pendidik berupa nilai serta deskripsi pencapaian kompetensi keterampilan dan dari pendidik disampaikan kepada wali kelas, kepala sekolah, guru bimbingan dan konseling, orang tua dari peserta didik pada periode yang telah ditentukan dan dilaporkan dalam bentuk laporan setiap jangka waktu tertentu atau rapor.

Fenomena guru sudah dilatih, fasilitas sudah ditambah, buku-buku sudah disediakan, tetapi kualitas luaran (prestasi belajar siswa) tidak meningkat, menjadi suatu masalah yang harus segera dibenahi (Sudiyatno, 2010, hal. 14). Salah satu penyebab mengapa berbagai macam usaha perbaikan mutu pendidikan tidak efektif adalah karena siklus perencanaan yang terlalu panjang, implementasi yang terlalu kompleks dan tidak praktis, sehingga menyebabkan *overload* dan melelahkan. Oleh karena itu, dibutuhkan sebuah usaha perbaikan yang sederhana, tetapi terus-menerus.

Penilaian yang baik ialah penilaian yang mengukur apa yang sudah peserta didik pahami. Penilaian autentik yang akan dikembangkan juga untuk lebih mengukur

kompetensi yang telah peserta didik capai. Hasil Ujian Nasional untuk tahun 2012 yang telah dianalisis Badan Penelitian dan Pengembangan Kemendikbud RI menunjukkan bahwa daya serap peserta didik pada kemampuan memahami bangun datar, bangun ruang, sudut, serta menggunakannya dalam pemecahan masalah hanya 69,39%, bahkan daya serap peserta didik Propinsi DIY hanya 57,32%.

Salah satu upaya yang dapat dilakukan untuk memecahkan masalah tersebut adalah dengan mengembangkan perangkat penilaian autentik untuk pembelajaran matematika yang reliabel dan valid fokus pada kurikulum 2013. Hal ini sesuai dengan peraturan pemerintah nomor 66 tahun 2013 mengenai penilaian.

Oleh karena itu penting untuk melaksanakan pengembangan perangkat penilaian autentik untuk pembelajaran matematika kelas VII semester 1, dengan menggunakan kurikulum 2013 sebagai penilaian yang mendukung. Pendahuluan berisi latar belakang, rasional, dan atau urgensi penelitian.

Berdasarkan paparan ini, dapat disusun rumusan masalah sebagai berikut. (1) Bagaimana kualifikasi produk perangkat penilaian autentik bagi pembelajaran matematika kelas VII berdasarkan aspek penilaian autentik? (2) Bagaimana kualitas produk berupa perangkat pengembangan penilaian autentik untuk pembelajaran matematika kelas VII berdasarkan aspek reliabilitas? (3) Bagaimana kualitas produk berupa perangkat pengembangan penilaian autentik untuk pembelajaran matematika kelas VII berdasarkan aspek validitas?

## Metode Penelitian

Jenis penelitian ini adalah penelitian pengembangan (*research & development*). Model pengembangan yang digunakan adalah model pengembangan Plomp (2010) berupa tiga tahapan yaitu *preliminary research*, *prototyping phase*, dan *assessment phase*. Penelitian ini dilaksanakan di SMP IT Abu Bakar Yogyakarta, SMP Negeri 9 Yogyakarta, dan SMP IT Luqman Al-Hakim Yogyakarta dari bulan Mei hingga November 2014.

Subjek pada uji keterbacaan adalah 11 siswa dan 1 guru matematika SMP IT Abu Bakar Yogyakarta. Subjek pada uji coba lapangan adalah 32 siswa dan 2 guru matematika di SMP Negeri 9 Yogyakarta kemudian uji coba lapangan selanjutnya terdiri atas 83 siswa dan 3 orang guru matematika gabungan dari SMP IT Abu Bakar Yogyakarta, SMP Negeri 9 Yogyakarta, dan SMP IT LHI.

Prosedur pengembangan yang dilakukan peneliti meliputi 3 tahap, yaitu *preliminary research*, *prototyping phase*, dan *assessment phase*. Tahap pertama adalah *preliminary research*. Tahap ini merupakan tahap persiapan untuk penelitian pengembangan. Tahap ini terdiri dari studi pustaka untuk mengumpulkan literatur-literatur yang diperlukan dalam melakukan kajian teori berkenaan dengan penilaian autentik pada pembelajaran matematika yang akan dikembangkan, melakukan survei lapangan dengan tujuan untuk memperoleh data awal tentang keadaan-keadaan di lapangan misalnya perangkat penilaian autentik pada pembelajaran matematika yang digunakan oleh guru.

Tahap kedua adalah *prototyping phase*. Hasil studi kepustakaan mengenai dasar-dasar teori dan konsep mengenai penilaian autentik pada pembelajaran, survei di lapangan, dan perencanaan, maka peneliti menyusun rancangan awal perangkat penilaian autentik pada pembelajaran matematika yang meliputi,

petunjuk penggunaan, cakupan instrumen, kisi-kisi indikator, lembar penilaian, dan rubrik penilaian pada KI 1, KI 2, KI 3, dan KI 4.

Tahap ketiga adalah *assessment phase* merupakan fase penilaian berupa validasi ahli dan uji coba lapangan untuk pengembangan lebih lanjut. Uji coba yang dilakukan meliputi uji ahli, uji keterbacaan perangkat pembelajaran, dan uji coba lapangan. Uji coba lapangan untuk mengetahui reliabilitas dari perangkat yang dikembangkan dan uji ahli dilakukan untuk mendapatkan kevalidan produk berupa perangkat penilaian autentik untuk pembelajaran matematika pada kompetensi bilangan bulat, rasio dan perbandingan, serta garis dan sudut. Hasil uji coba lapangan sebagai dasar untuk mengetahui reliabilitas produk sedangkan hasil validasi sebagai dasar dalam melakukan perbaikan rancangan awal perangkat yang dikembangkan. Hasil uji coba kemudian dijadikan dasar untuk mengetahui reliabilitas.

Data hasil penelitian terbagi menjadi dua yaitu data kuantitatif dan data kualitatif. Data kuantitatif diperoleh dari hasil uji coba lapangan berupa reliabilitas, sedangkan data kualitatif diperoleh dari hasil masukan saran dari para ahli dan praktisi. Data kualitatif berupa hasil pengisian lembar validasi dan angket tersebut diklasifikasikan menjadi 5 kategori pilihan. Instrumen untuk membuktikan kevalidan perangkat pembelajaran terdiri dari lembar validasi produk perangkat penilaian autentik, lembar validasi kepraktisan untuk KI 1, KI 2, KI 3 dan KI 4. Lembar validasi digunakan untuk membuktikan kevalidan perangkat penilaian yang dikembangkan.

Analisis data dalam penelitian ini bertujuan untuk menjawab pertanyaan penelitian. Data yang berupa komentar dan

saran dianalisis secara kualitatif, yang selanjutnya digunakan sebagai masukan untuk merevisi produk yang dikembangkan. Sedangkan data yang diperoleh melalui lembar validasi perangkat. Untuk menilai kelayakan dari perangkat pembelajaran yang dikembangkan ditinjau dari reliabilitas dan validitas.

Data berupa komentar dan saran dianalisis secara kuantitatif, yang selanjutnya digunakan sebagai dasar untuk merevisi produk yang dikembangkan. Data yang berupa “ya”-“tidak” dikonversikan menjadi data kualitatif yang juga berskala 5. Kategori konversi data tersebut dilakukan berdasarkan kategori yang disajikan dalam Tabel 1 (Azwar, 2010, hal. 163).

Perangkat penilaian dikatakan valid untuk digunakan dalam uji coba jika skor kevalidan tiap perangkat pembelajaran memiliki kategori minimal valid (B). Apabila di bawah B maka perangkat penilaian tidak valid. Dengan demikian, hasil analisis data yang tidak memenuhi kategori minimal valid dalam penelitian ini akan dijadikan bahan pertimbangan untuk melakukan revisi perangkat pembelajaran sebelum diuji cobakan.

**Tabel 1. Kategori Kevalidan Produk**

Interval Skor	Kategori
$(M + 1,50s) < \bar{x}$	Sangat Valid
$(M + 0,50s) < \bar{x} < (M + 1,50s)$	Valid
$(M - 0,50s) < \bar{x} < (M + 0,50s)$	Cukup Valid
$(M - 1,50s) < \bar{x} < (M - 0,50s)$	Kurang Valid
$\bar{x} < (M - 1,50s)$	Tidak Valid

Keterangan:

$\bar{x}$  = skor rata-rata

M = Rata-rata skor ideal

$$= \frac{1}{2} (\text{skor maksimum} + \text{skor minimum})$$

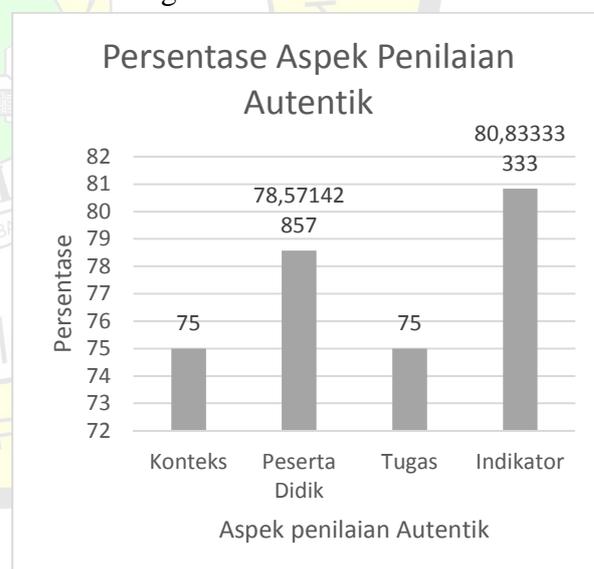
S = Simpangan baku ideal

$$= \frac{1}{6} (\text{skor maksimum} - \text{skor minimum})$$

## Hasil Penelitian dan Pembahasan

Hasil pengembangan berupa perangkat penilaian autentik untuk pembelajaran matematika meliputi petunjuk penggunaan, kisi-kisi berupa indikator, instrumen penilaian, rubrik penskoran yang autentik, reliabel, dan valid. Desain dari masing-masing produk yang dikembangkan dijelaskan pada uraian berikut ini.

Produk mengalami penyusunan dari fase *preliminary research*, *prototyping stage*, dan beberapa kali revisi pada *assessment phase*. Produk awal melalui beberapa revisi para ahli kemudian melalui tiga kali uji coba lapangan dari uji coba terbatas hingga diperluas. Perolehan analisis hasil dari produk berupa perangkat penilaian autentik dengan empat aspek penilaian autentik yang meliputi: konteks, peserta didik, tugas, dan indikator adalah sebagai berikut.

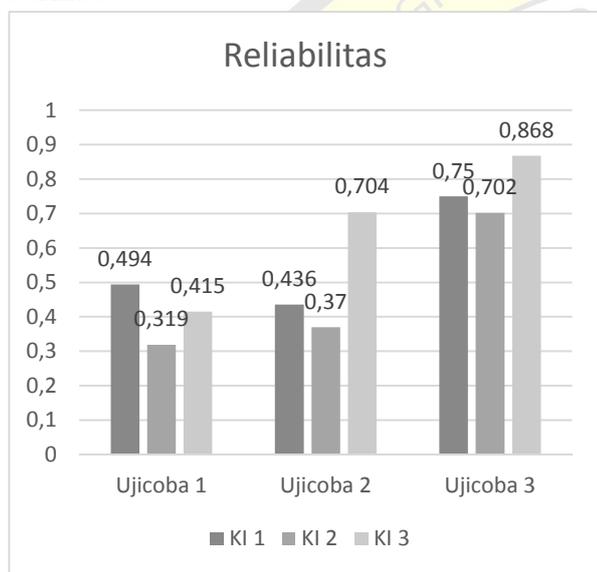


**Gambar 1. Grafik Hasil Aspek Penilaian Autentik**

Aspek penilaian autentik yang pertama yakni konteks berupa keterpaduan dengan mata pelajaran lain dan kehidupan sehari-hari memperoleh skor 12 dengan persentase 75% dalam kategori baik. Aspek peserta didik memecahkan masalah serta menggunakan

pengetahuan sebelumnya mendapat skor 44 dengan persentase 78,571% dalam kategori sangat baik. Aspek tugas yang terpadu, meliputi masalah kompleks, dan memerlukan pengambilan keputusan mencapai skor 24 dengan persentase 75% dalam kategori baik. Aspek yang terakhir adalah indikator yang meliputi validitas, kisi-kisi, instrumen, dan penskoran mencapai skor 97 dengan persentase 80,833% dalam kategori sangat baik.

Pada pelaksanaan uji coba, masing-masing instrumen dilakukan analisis reliabilitas seperti pada Gambar 2 grafik berikut.

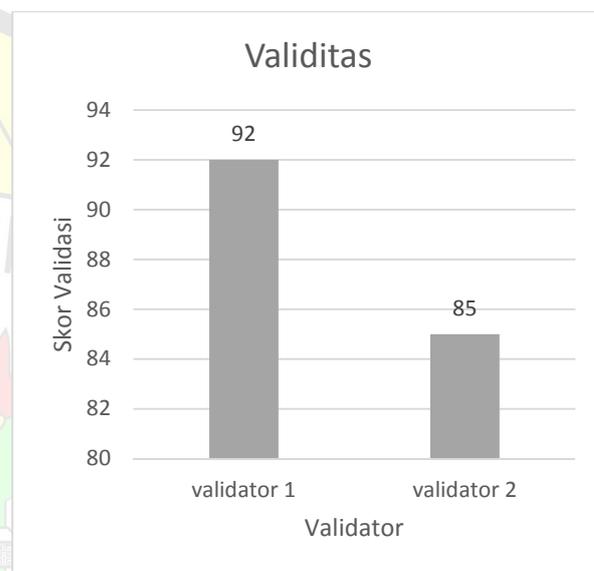


Gambar 2. Grafik Hasil Analisis Reliabilitas

Hasil analisis reliabilitas seperti pada Gambar 2, penggunaan instrumen penilaian oleh peserta didik seperti tampak pada grafik di atas telah mencapai taraf yang diharapkan pada uji coba ketiga. Pada uji coba pertama reliabilitas masih rendah bagi masing-masing instrumen KI 1, KI 2, maupun KI 3. Kemudian pada uji coba kedua dengan 32 peserta didik, instrumen penilaian KI 1 dan KI 2 reliabilitas masih rendah, hanya instrumen KI 3 saja yang mencapai kriteria  $\alpha \geq 70$ . Pada ujicoba ketiga dengan sampel 3 kelas dengan jumlah 83

peserta didik masing-masing reliabilitas dari instrumen KI 1, KI 2 dan KI 3 telah mencapai kriteria  $\alpha \geq 70$ . Reliabilitas yang sudah memenuhi kriteria dan telah memenuhi kriteria valid pada validasi instrumen.

Validitas dari instrumen sangat penting untuk benar-benar mengukur sikap, pengetahuan, dan keterampilan peserta didik. Hasil validitas instrumen adalah sebagai berikut.



Gambar 3. Grafik Hasil Validasi

Perolehan skor seperti pada Gambar 3 untuk kevalidan yakni 177 sehingga berada pada kategori valid. Validitas meliputi kesesuaian penilaian instrumen dengan kajian teori dari pengembangan perangkat penilaian autentik.

### Simpulan dan Saran

#### Simpulan

Simpulan dari penelitian ini adalah hasil pengembangan berupa perangkat penilaian autentik untuk pembelajaran matematika kelas VII semester 1 dengan kurikulum 2013. Hasil analisis produk pengembangan instrumen penilaian autentik

berupa perangkat penilaian yang meliputi: petunjuk penggunaan, cakupan instrumen, kisi-kisi indikator, lembar penilaian, dan rubrik penilaian pada KI 1, KI 2, KI 3, dan KI 4 berdasarkan aspek penilaian autentik yakni: konteks dengan skor 12 berupa persentase sebesar 75% dalam kategori baik, peserta didik dengan skor 44 berupa persentase sebesar 78,57% dalam kategori sangat baik, tugas dengan skor 24 berupa persentase sebesar 75% dalam kategori baik, dan indikator dengan skor 97 berupa persentase sebesar 80,833% dalam kategori sangat baik.

Hasil analisis dari data dari peserta didik menunjukkan bahwa konsistensi telah memenuhi pada taraf  $\alpha \geq 0,7$  yakni: 0,75 untuk KI 1, 0,702 untuk KI 2, dan 0,87 untuk KI 3 sehingga taraf reliabilitas telah terpenuhi. Hasil penilaian kualitas pengembangan instrumen penilaian oleh ahli dengan skor 177 berada pada kategori valid untuk diterapkan dalam rangka penilaian kepada peserta didik.

### Saran

Beberapa saran yang dapat dikemukakan pada penelitian ini yaitu bahwa produk pengembangan instrumen penilaian ini hanya terbatas untuk SMP kelas VII semester 1, sehingga penelitian dan pengembangan lebih lanjut mendukung instrumen penilaian yang digunakan. Pengembangan instrumen penilaian ini lebih berorientasi kepada pelaksanaan kurikulum 2013 yang secara dasar dan pelaksanaan masih mengacu pada perbaikan kurikulum dari pemerintah sehingga adanya kajian yang lebih mendalam mendukung instrumen penilaian dan sejalan dengan kurikulum yang berlaku.

### Daftar Pustaka

- Ambrosetti, D. D. & Cho, G. (2005). Synergism in learning: a critical reflection of authentic assessment. *The University of North Carolina Press*. 57-60.
- Anderson, L. W. (2003). *Classroom assessment: enhancing the quality of teacher decision making*. New Jersey, London: Lawrence Erlbaum Associates, Publishers.
- Frey, B. B. & Schmitt, V. L. (2007). Coming to term with classroom assessment. *Journal of Advanced Academic*, 18, 402-423.
- Frey, B. B., Schmitt, V. L. & Allen, J. P. (2012). Defining authentic classroom assessment. *Practical Assessment, Research & Evaluation*. Journal Volume 17 Number 2 January 2012, 1531-7714.
- Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *ETR&D*, Vol.52, No. 3, 2004, 67-86.
- Herrington, J. & Herrington, A. (2006). Authentic conditions for authentic assessment:aligning task and assessment, in critical Vision. *Proceedings of the 29th HERDSA Annual Conference*, 10-12 July 2006, 146-151.
- Hyde, D. P. (2013). What makes a good secondary assessment? On achieving the aim of assessment. *Journal Education and Practice*, Vol.4, No.12, 188-197.

- Indiana Departement of Education. (tanpa tahun). Authentic assessment. *Indiana Departement of Education*. Diambil tanggal 26 September 2013, dari [www.msdt.k12.in.us/msd/wp-content/.../10/authentic\\_assessment.pdf](http://www.msdt.k12.in.us/msd/wp-content/.../10/authentic_assessment.pdf).
- Kemdikbud. (2012). *Pengembangan kurikulum 2013*. Jakarta: Kementerian Pendidikan dan Kebudayaan RI.
- Kemdikbud. (2014). *Model penilaian pencapaian kompetensi peserta didik Sekolah Penengah Pertama*. Jakarta: Kementerian Pendidikan dan Kebudayaan RI.
- Lai, E. R. (2011). Performance-based assessment: some new thoughts on an old idea. *Bulletin Pearson Education*, May 2011 Issue 20, 1-4.
- Moon, T. R., et. al. (2005). Development of authentic assessment for the middle school classroom. *The Journal of Secondary Gifted Education*, Vol. XVI. No.2/3. Winter/Spring 2005, 119-133.
- Newmann, F.M & Wehalage, G.G. (1993). Five standard of authentic instruction. *Educational Leadership*, 50(7), 8-12.
- Nitko, A. J. & Brookhart, S. M.. (2011). *Educational assesment of students*. Upper Saddle River, NJ: Pearson Education, Inc.
- Plomp. (2010). Educational design research: an introduction. Dalam Plomp, T. & Nieveen, N. (Eds.), *An introduction to educational design research*. Proceedings of the seminar conducted at the East China Normal University, Shanghai (PR China), November 23-26, 2007. (3<sup>rd</sup> print, pp. 9-35). Netherland: Netherland Institute for Curriculum Development.
- Reynolds, C. R., Livingston, R. B., & Willson, V. (2009). *Measurement and assessment in education*. New York: Pearson.
- Shapiro, E. S. (2011). *Academic skills problems: direct assessment and intervention*. 4<sup>th</sup>ed. New York: The Guilford Press.
- Sudiyatno. (2010). *Pengembangan model penilaian komprehensif unjuk kerja siswa pada pembelajaran berbasis standar kompetensi di SMK Teknologi Industry*. Disertasi Doktor pada PPs UNY: Tidak diterbitkan.
- Weber, M. & Stewart, B. R. (2001). Authentic assessment: good or bad?. *ProQuest Agricultural Journal, The Agricultural Education Magazine*, Mar/Apr 2001, 73, 5, 14.
- Wiggins, G. P. (1993). *Assessing student performance: exploring the purpose and limits of testing*. New York: Jossey-Bass Publisher.
- Wright, R. J. (2008). *Educational assessment: test and measurements in the age of accountability*. Los Angeles: Sage Publication.