

## **PENILAIAN HARIAN KURIKULUM 2013 PADA PEMBELAJARAN PENDIDIKAN AGAMA ISLAM**

Oleh:

**ILMA KHARISMATUNISA**

*UIN Kiai Haji Achmad Siddiq Jember*

Email: [ilmakharisma1@gmail.com](mailto:ilmakharisma1@gmail.com)

**MOH. SAHLAN**

*UIN Kiai Haji Achmad Siddiq Jember*

Email: [mohsahlan@uinkhas.ac.id](mailto:mohsahlan@uinkhas.ac.id)

### **Abstrak**

Perubahan pendidikan di Indonesia berupa salah satu bentuk dari dan juga merupakan kurikulum 2013 yang berfungsi sebagai penyempurna kurikulum sebelumnya. Kajian ini bertujuan untuk mengetahui penilaian harian kurikulum 2013 pada Pendidikan Agama Islam (PAI). Melalui *library research*, evaluasi diketahui dengan menggunakan pengukuran informasi dan informasi hasil penilaian. Hasilnya diukur dengan memberikan skor (angka). Kemudian, skor tersebut dinilai dan ditafsirkan oleh aturan tertentu untuk menentukan tingkat kemampuan pribadi. Selain itu, hasil dari proses penilaian ini selanjutnya dievaluasi untuk menentukan tingkat pencapaian pribadi atau terprogram. Secara umum, ada dua teknik penilaian pendidikan, yaitu tes dan non-tes. Berdasarkan hasil penelitian kualitatif dengan memanfaatkan sumber berupa data atau dokumen, penelitian ini menggambarkan bagaimana persiapan instrumen berdasarkan kognitif, afektif, dan psikomotor evaluasi domain. Suatu set tes dan non-tes yang baik sebagai pengukur prestasi harus memiliki kriteria; validitas, kepraktisan, kehandalan, dan ekonomi. Selanjutnya, analisis tes yang sesuai pada evaluasi pembelajaran Pendidikan Agama Islam (PAI) terdiri dari tingkat kesukaran soal atau indeks kesulitan, daya pembeda, analisis pengecoh, analisis homogenitas item soal, dan efektivitas fungsi opsi.

### **Kata Kunci:**

Penilaian Harian, Kurikulum 2013, Pendidikan Agama Islam

## **PENDAHULUAN**

Belajar dari sejarah tentang bagaimana hasil pendidikan nasional sejak satu abad kebangkitan nasional pertama yang lalu, ternyata persoalan pendidikan bukan hanya menyangkut peserta didik sebagai subjek namun juga menyangkut berbagai aspek. Dalam upaya mengatasi persoalan pendidikan tersebut penyelenggaraan pendidikan nasional sejak kemerdekaan telah mengalami pergantian kurikulum sebanyak 10 kali, yaitu; kurikulum 1947,1952, 1964,1968,1975,1984, 1994 dan suplemen kurikulum 1999, kurikulum 2004 (KBK), kurikulum 2006 (KTSP), dan Kurikulum 2013.

Perubahan kurikulum tersebut didasari bahwa disatu sisi, merupakan perubahan dan perkembangan disegala aspek kehidupan perlu direspon oleh kinerja pendidikan yang profesional dan bermutu tinggi. Mutu pendidikan yang demikian itu sangat diperlukan untuk mendukung terciptanya manusia yang cerdas dan berkehidupan yang damai, terbuka, dan berdemokrasi, serta mampu bersaing secara terbuka di era global, sehingga dapat meningkatkan kesejahteraan seluruh warga negara Indonesia.

Oleh karena itu, kinerja pendidikan menuntut adanya pembenahan dan penyempurnaan terhadap aspek substantif yang mendukungnya, yakni kurikulum. Pada sisi lain perubahan kurikulum juga didasari pada kesadaran bahwa perkembangan dan perubahan yang terjadi dalam kehidupan bermasyarakat, berbangsa dan bernegara di Indonesia tidak terlepas dari pengaruh perubahan global, perkembangan ilmu pengetahuan dan teknologi, serta seni dan budaya. Kurikulum sebagai suatu produk pemikiran sudah barang tentu tidak mungkin dapat diberlakukan untuk sepanjang zaman. Kurikulum selalu mempunyai keterbatasan-keterbatasan menurut ukuran ruang dan waktu ketika kurikulum tersebut dimunculkan atau diberlakukan (Hasibuan, 2010: 15). Perubahan global secara terus menerus yang terjadi saat ini menuntut perlunya perbaikan sistem pendidikan nasional termasuk penyempurnaan kurikulum selaku pedoman, acuan atau rujukan penyelenggaraan pendidikan dengan melalui pembelajaran, sehingga pendidikan nantinya diharapkan dapat mewujudkan masyarakat yang mampu bersaing dan menyesuaikan diri dengan perubahan zaman.

Untuk melihat tingkat pencapaian mutu dan tujuan pendidikan, terutama kualitas proses dan kualitas out put pembelajaran, diperlukan suatu bentuk evaluasi. Tujuan dilakukannya evaluasi adalah sesuai dengan apa yang telah dimuat dalam Undang-Undang Republik Indonesia Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional Pasal 57 ayat (1), bahwa evaluasi dilakukan dalam rangka pengendalian mutu pendidikan secara nasional sebagai bentuk akuntabilitas penyelenggara pendidikan kepada pihak-pihak yang

berkepentingan, diantaranya terhadap peserta didik, lembaga, dan program pendidikan (Sukardi, 2009: 1).

Berbagai upaya telah dilakukan dalam usaha mencapai peningkatan kualitas proses dan kualitas out put pembelajaran hendaknya dilaksanakan, khusus peningkatan kualitas proses dan kualitas out put pembelajaran Pendidikan Agama Islam sangat ditentukan oleh tiga unsur, yaitu: guru, peserta didik, dan kurikulum. Ketiga unsur tersebut dapat diasumsikan bahwa (1) Guru, sesuai dengan fungsinya bertugas mengoptimalkan kemampuan peserta didik dalam belajar atau sering diistilahkan dengan mengajar. Dalam mengoptimalkan kemampuan peserta didik profesionalisme, kemampuan dan pemahaman guru yang memadai dalam mengajar sangat menentukan peningkatan kualitas pembelajaran sekaligus keberhasilan program pembelajaran mencapai tujuan sebagaimana telah direncanakan di dalam silabus yang telah disusun sebelumnya, (2) Peserta didik, dengan segala karakteristiknya dalam proses pembelajaran diharapkan secara maksimal dapat mencapai tujuan belajar. Karakteristik peserta didik yang dimaksud khususnya dalam penelitian ini diantaranya menyangkut derajat afeksi atau sikap peserta didik, motivasi, personal skill, dan social skill yang dapat dioptimalkan untuk meningkatkan kualitas maupun out put pembelajaran, dan (3) Kurikulum, adalah merupakan pedoman atau media serta sekaligus merupakan salah satu fasilitas penunjuang untuk mencapai tujuan yang diinginkan. Oleh karena itu, maka perlu dilakukan evaluasi pada unsur-unsur tersebut untuk mengidentifikasi kualitas proses dan kualitas out put pembelajaran khususnya pada pelaksanaan program pembelajaran Pendidikan Agama Islam.

Pada umumnya, guru melakukan penilaian di kelas terikat dengan aktivitas belajar mengajar dalam upaya menghimpun data, fakta, dan dokumen belajar siswa dengan tujuan untuk melakukan perbaikan program pembelajaran. Guru yang profesional memanfaatkan penilaian prosedur dan prestasi belajar untuk memperbaiki perencanaan pembelajaran dan pelaksanaan pembelajaran (Ridwan Abdullah Sani, 2016: 15-16).

Kegiatan penilaian prosedur dan prestasi belajar tersebut membutuhkan informasi yang bervariasi dari kelompok peserta didik, guru juga dapat menerapkan metode dan teknik yang bervariasi dalam melaksanakan penilaian dengan mengumpulkan catatan pertemuan, pengamatan, portofolio, catatan harian, ujian, data hasil interview, survey dan sebagainya. Penilaian yang tepat akan dapat memberikan cerminan dan refleksi proses pembelajaran yang dialami oleh peserta didik. Penilaian seperti itu dapat menunjukkan perilaku belajar peserta didik secara lengkap, serta menunjukkan perilaku peserta didik dalam

kehidupan nyata. Gambaran lengkap tentang peserta didik juga dicerminkan dalam perilaku peserta didik pada saat istirahat, berkomunikasi dengan guru, bergaul dengan teman, berinteraksi dengan orang lain, mengikuti pelajaran, membuat tugas, menghasilkan produk.

## **METODE**

Penelitian ini menggunakan jenis penelitian literature review atau kajian literature dengan pendekatan kualitatif. Penelitian literature review merupakan penelitian dengan mencari dan mengumpulkan temuan-temuan penelitian terdahulu dengan fokus penelitian yang sama, kemudian menyusun kerangka pemikiran yang jelas tentang pemecahan dari pertanyaan yang disampaikan di pendahuluan.

Bahan-bahan yang digunakan dalam makalah ini bersandar dari berbagai referensi atau literatur yang signifikan dengan tema pembahasan yang dibahas. Validitas dan relevansi referensi yang digunakan dapat dipercaya dan dibuktikan. Jenis data yang diterima berupa data sekunder. Metode penelitian ini dilakukan dengan menggunakan studi pustaka atau kepustakaan dengan menelusuri berbagai rujukan yang terkait dengan topik utama permasalahan. Literatur yang digunakan merupakan literature yang telah diriset validitasnya dan mendukung dalam penguraian masalah. Strategi analisis yang digunakan adalah dengan memfokuskan pada interpretasi dan kontekstualisasi data yang penilaian harian Kurikulum 2013 pada pembelajaran Pendidikan Agama Islam.

## **HASIL DAN PEMBAHASAN**

### **A. Konsep Penilaian Harian Kurikulum 2013**

Kelompok pelajaran Pendidikan Agama Islam yang mengandung isi materinya penuh dengan isi norma dan nilai-nilai di dalamnya, tentunya memerlukan penilaian yang dilakukan bukan hanya terfokus pada satu aspek saja (kognitif), akan tetapi harus menyeluruh baik dari aspek pengetahuan, sikap, dan keterampilan. Dari ketiga aspek tersebut dalam penilaiannya, harus berdasarkan atas konsep keterpaduan materi, dan proses pengelolaan pendidikan yang meliputi keselarasan antara lingkungan pendidikan, yaitu: madrasah, family, dan masyarakat (Majid dan Andayani, 2004:189).

Menurut E. Mulyasa menyatakan bahwa penilaian adalah kelengkapan aktivitas pengukuran (penghimpunan data dan informasi), pengolahan, penafsiran, dan pertimbangan untuk membuat ketetapan tentang jenjang hasil belajar yang diperoleh peserta didik sesudah melakukan aktivitas belajar dalam usaha mencapai tujuan pembelajaran yang telah ditetapkan (Mulyasa, 2012:201).

Heri Gunawan mengutip Hamalik, menerangkan bahwa penilaian dalam pendidikan yakni seperangkat aktivitas, maupun proses untuk menentukan nilai sesuatu yang berhubungan dengan dunia pendidikan. Penilaian adalah prosedur pengumpulan, dan pengolahan laporan untuk menilai perolehan dampak belajar peserta didik (Salinan Lampiran Permendikbud RI No. 23 Tahun 2016 Tentang Standar Penilaian Pendidikan: 2). Penilaian terhadap prosedur dan hasil pembelajaran merupakan bagian yang tidak terpisah dari perencanaan maupun pelaksanaan proses pembelajaran guru. Penilaian pembelajaran pada Kurikulum 2013 diarahkan pada penilaian autentik.

Secara lazim penilaian autentik kerap disitir dengan *authentic assessment*. *Authentic assessment* adalah suatu penilaian hasil belajar yang mengharuskan peserta didik untuk memperlihatkan prestasi, dan hasil belajar, berupa kemampuan dalam kehidupan nyata dalam bentuk kinerja, ataupun hasil kerja (Supardi, 2015: 24). Menurut Suyadi, *authentic assessment* yakni proses yang dilakukan pendidik untuk menghimpun informasi tentang kelanjutan belajar yang dilakukan peserta didik. Penilaian ini dibutuhkan untuk mendeteksi apakah peserta didik sungguh-sungguh belajar atau tidak, memahami atau tidak, menguasai atau tidak, apakah pengalaman belajar peserta didik memiliki pengaruh yang positif terhadap kelanjutan baik intelegensi maupun mental peserta didik. Penilaian yang autentik dilakukan secara tergabung dengan prosedur pembelajaran. Penilaian ini dilakukan sebagai kontinu selama proses pembelajaran berjalan. Oleh karena itu, penilaian dipusatkan pada proses belajar, bukan pada hasil belajar. Secara lebih umum tentang penilaian autentik didefinisikan sebagai penilaian yang dibuat secara menyeluruh untuk mengevaluasi sejak dari input, proses, maupun output pembelajaran.

Berdasarkan beberapa pengertian para ilmuan tentang *authentic assessment* maka pemakalah dapat menyimpulkan bahwa penilaian autentik mengawasi keseimbangan antara penilaian kompetensi sikap, kompetensi pengetahuan, dan kompetensi keterampilan yang disesuaikan dengan progres karakteristik peserta didik sesuai dengan tingkatannya.

Menurut Kunandar dalam bukunya yang bertema Penilaian Autentik, beliau menjelaskan bahwa karakteristik penilaian autentik meliputi: (1) "Bisa digunakan untuk formatif maupun sumatif". Artinya, penilaian autentik bisa dilakukan untuk menilai keberhasilan kompetensi terhadap satu kemampuan dasar (formatif) maupun keberhasilan terhadap standar kompetensi, atau kemampuan dasar dalam satu semester (sumatif); (2) Mengukur keterampilan dan informasi, maksudnya, penilaian autentik itu ditujukan pada pencapaian kemampuan yang memfokuskan pada aspek kemampuan (*skill*) dan kemampuan

(*performance*), tidak hanya mengukur kemampuan yang sifatnya mengingat fenomena (hafalan dan ingatan); (3) Berkesinambungan dan terintegrasi. maksudnya, dalam membuat penilaian autentik perlu secara berkelanjutan (terus-menerus), dan merupakan satu kesatuan secara sempurna sebagai sarana untuk mengumpulkan informasi terhadap pencapaian kemampuan peserta didik; (4) Dapat digunakan sebagai feed back. maksudnya, penilaian autentik yang dilakukan oleh pendidik dapat digunakan sebagai umpan balik atas pencapaian kompetensi peserta didik secara komprehensif.

Berdasarkan karakteristik di atas penting untuk menjadi perhatian ketika melaksanakan penilaian autentik dalam kegiatan pembelajaran, *pertama*, instrumen penilaian yang digunakan bervariasi sesuai dengan karakteristik kemampuan yang akan diperoleh. *Kedua*, aspek kemampuan belajar dinilai secara menyeluruh melingkupi berbagai aspek penilaian baik dalam ranah pengetahuan, ranah sikap, dan ranah keterampilan. *Ketiga*, penilaian dilaksanakan pada tahap pertama, proses maupun terakhir, baik afektif, kognitif, maupun skill sebagai input, proses, maupun output belajar siswa (Supardi, 2015: 27-28).

Berasaskan Permendikbud No. 104 Tahun 2014 tentang Penilaian prestasi belajar oleh guru pada Pendidikan Dasar Dan Pendidikan Menengah bagian pedoman penilaian prestasi belajar oleh Pendidik melampirkan prinsip-prinsip Penilaian Autentik sebagai berikut: (a) Materi penilaian dikembangkan dari kurikulum; (b) Bersifat lintas muatan atau mata pelajaran; (c) Berkaitan dengan kemampuan peserta didik; (d) Berbasis kinerja peserta didik; (e) Memotivasi belajar peserta didik; (f) Menekankan pada kegiatan dan pengalaman belajar peserta didik; (g) Memberi kebebasan peserta didik untuk mengkonstruksi responnya; (h) Menekankan keterpaduan sikap, pengetahuan, dan keterampilan; (i) Mengembangkan kemampuan berpikir divergen; (j) Menjadi bagian yang tidak terpisahkan dari pembelajaran; (k) Menghendaki balikan yang segera dan terus menerus; (l) Menekankan konteks yang mencerminkan dunia nyata; (m) Terkait dengan dunia kerja; (n) Menggunakan data yang diperoleh langsung dari dunia nyata; (o) Menggunakan berbagai cara dan instrumen (Permendikbud RI No. 104 Tahun 2014 Tentang Penilaian Hasil Belajar oleh Pendidik Pada Pendidikan Dasar dan Pendidikan Menengah bagian Pedoman Penilaian Hasil Belajar oleh Pendidik).

Berkaitan dengan prinsip-prinsip penilaian autentik di atas, maka proses penilaian yakni bagian yang tidak dapat terpisah dari prosedur pembelajaran dan mencerminkan masalah dunia nyata/sehari-hari. Sehingga dalam merancang penilaian autentik, perlu memperhatikan prinsip-prinsip, sebagai berikut: penilaian wajib menggunakan beraneka dimensi, metode dan kriteria yang sesuai dengan keunikan dan esensi pengalaman belajar; penilaian wajib bersifat holistik

mencakup semua segi dari harapan pembelajaran (sikap, keterampilan, dan pengetahuan).

### **B. Komponen Penilaian Harian Kurikulum 2013**

Dalam komponen penilaian ini terdapat dua hal yaitu:

#### 1. Penyiapan Tugas Autentik

Tugas autentik yaitu perintah yang secara jelas dimuatkan kepada pembelajar untuk menilai pencapaian kompetensi yang dibelajarkan ketika kegiatan pembelajaran masih berlangsung ataupun ketika sudah berakhir. Tugas autentik kerap disamakan dengan penilaian autentik walau sebenarnya cakupan makna yang kedua lebih luas. Pemilihan tugas autentik awal-awal haruslah melihat pada kompetensi mana yang akan dinilai pencapaiannya. kedua, dan inilah yang khusus penilaian autentik, pemilihan tugas tugas itu harus merefleksikan keadaan ataupun kepentingan yang sebenarnya di dunia nyata. Maka, dalam sebuah penilaian autentik tentu termuat dua hal sekaligus: searah dengan standar (kompetensi) dan relevan (bermakna) dengan kehidupan nyata. Dua hal tersebut mestinya menjadi rujukan kita ketika membuat tugas-tugas autentik untuk menilai pencapaian kompetensi pembelajaran kepada siswa (Nurgiyantoro, 2011:31).

#### 2. Pengembangan Rubrik Penilaian

Setelah menetapkan kriteria, guru perlu mengembangkan rubrik sebagai pedoman penskoran. Pedoman penskoran ini perlu memiliki deskriptor yang menunjukkan tingkat kinerja dari masing-masing tingkatan unjuk kerja.

### **C. Teknik Penilaian Harian Kurikulum 2013**

Menurut Supardi tentang teknik yang digunakan bagi penilaian autentik, berdasarkan lampiran Permendikbud No. 66 Tahun 2013 berhubungan dengan standar penilaian pendidikan dilakukan untuk mengukur kompetensi sikap, kompetensi pengetahuan dan kompetensi keterampilan (Yubali Ani, 2013:24).

*Pertama*, penilaian pada kompetensi sikap, Pendidik melakukan penilaian dengan cara melalui teknik observasi, penilaian diri, dan penilaian teman sejawat (*peer evaluation*) oleh siswa, dan jurnal. Instrumen yang diterapkan untuk teknik pengamatan, penilaian pribadi, penilaian antar siswa yakni menggunakan daftar cek, atau skala penilaian (*rating scale*) yang disertai rubrik, untuk jurnal berbentuk notulen pendidik. Jurnal merupakan kumpulan dokumen catatan guru atau tenaga kependidikan di lingkungan sekolah tentang sikap dan perilaku baik maupun buruk, semasa dan di luar prosedur pembelajaran mata pelajaran (Permendikbud No. 104 Tahun 2014: 15).

Contohnya pada teknik penilaian diri yakni laporan diri peserta didik tentang aktivitas yang dilakukan dalam kehidupan sehari-hari, misalnya menyangkut praktik pengalaman ibadah shalat fardu, shalat jum'at, shalat sunnah, puasa ramadhan, puasa sunnah, membaca Al-Qur'an, mengikuti kegiatan pengajian di TPA, pengajian remaja masjid, dan sebagainya (Sukiman, 2012:135-137).

*Kedua*, penilaian kompetensi pengetahuan, guru menilai kompetensi pengetahuan, menggunakan teknik tes tulis (menyeleksi jawaban pilihan ganda, dua pilihan benar atau salah, ya atau tidak, menjodohkan, sebab akibat, menyediakan jawaban seperti isi atau melengkapi), tes lisan (pertanyaan yang menuntut siswa untuk menjawab secara lisan), dan penugasan (tugas yang dilakukan secara individu maupun divisi).

Sebagaimana telah dilampirkan di dalam Permendikbud No. 66 tahun 2013 menjelaskan bahwa guru mengevaluasi kompetensi pengetahuan melalui teknik ujian tulis, ujian lisan, dan penugasan: 1) Instrumen yang digunakan untuk teknik ujian tulis berbentuk soal pilihan ganda, isian jawaban singkat, betul atau salah, mempertemukan, dan uraian. Untuk instrumen uraian dipenuhi pedoman penskoran. 2) Instrumen yang digunakan untuk teknik ujian lisan berbentuk daftar pertanyaan. 3) Instrumen yang digunakan untuk teknik penugasan berbentuk pekerjaan rumah dan proyek yang dilaksanakan secara individu atau kelompok searah dengan kriteria tugas (Permendikbud No. 66 tahun 2013 tentang Standar Penilaian Pendidikan: 4).

*Ketiga*, penilaian pada kompetensi keterampilan, guru menilai kompetensi keterampilan menggunakan penilaian kinerja yaitu penilaian yang menuntut siswa mendemonstrasikan suatu kompetensi spesifik dengan menggunakan teknik ujian praktik, proyek, dan penilaian portofolio. Instrumen yang digunakan berbentuk daftar cek ataupun skala penilaian (*rating scale*) yang disempurnakan dengan rubrik (Supardi, 2015:24).

#### **D. Validitas dan Reliabilitas Penilaian Harian Kurikulum 2013**

Seperangkat tes yang baik sebagai alat pengukur menurut Arikunto (2008: 59) harus memenuhi persyaratan tes, yaitu memiliki; validitas, praktibilitas, reliabilitas, dan ekonomis. Hal senada juga, alat ukur yang baik adalah alat ukur yang memiliki bukti kesahihan dan keandalan (Widayati, 2009: 185). Menurut Azwar (2000: 5-6), validitas berasal dari kata *validity* yang mempunyai arti sejauh mana ketepatan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya. Senada dengan Subali (2010: 41) dan Arifin (2012: 314), suatu alat ukur (dalam hal ini adalah tes) dinyatakan sah (*valid*), jika alat ukur tersebut benar-benar mampu memberikan informasi empirik sesuai dengan apa yang diukur.



Bukti kesahihan atau validitas alat ukur dilihat pada kesesuaian antara definisi operasional dari konsep yang akan diukur dengan materi pertanyaan pada alat ukur. Bukti kesahihan alat ukur meliputi kesahihan isi, konstruk, dan kriteria. Gronlund (1985: 7981) mengemukakan ada tiga faktor yang mempengaruhi validitas hasil tes, yaitu: 1) Faktor instrumen evaluasi, terkait prosedur penyusunan instrumen, seperti silabus, kisi-kisi soal, petunjuk mengerjakan soal dan pengisian lembar jawaban, kunci jawaban, penggunaan kalimat efektif, bentuk alternatif jawaban, tingkat kesukaran, daya pembeda, dan distraktor, 2) Faktor administrasi evaluasi dan penskoran, seperti alokasi waktu untuk pengerjaan soal yang tidak proporsional, memberikan bantuan kepada peserta didik dengan berbagai cara, peserta didik saling menyontek ketika ujian, kesalahan penskoran, termasuk kondisi fisik dan psikis peserta didik yang kurang menguntungkan, dan 3) faktor jawaban dari peserta didik, seperti kecenderungan peserta didik untuk menjawab secara cepat tetapi tidak tepat, keinginan melakukan coba-coba, dan penggunaan gaya bahasa tertentu dalam menjawab soal bentuk uraian.

Berikut ini dipaparkan beberapa ragam validitas yang sering digunakan dalam pengujian instrumen, dalam hal ini tes, yaitu:

1. Validitas isi, validitas kurikuler, validitas perumusan, validitas rasional, atau validitas logis, berkenaan dengan keterkaitan tes dengan kurikulum atau isi materi pembelajaran yang sedang berlaku. Validitas kurikuler ini dapat dilakukan dengan beberapa cara, antara lain mencocokkan materi tes dengan silabus dan kisi-kisi, melakukan diskusi dengan pakar/ahli, praktisi dan/atau sesama guru, atau mencermati kembali substansi dari konsep yang akan diukur. Menurut Purwanto (2011), pengujian validitas isi dapat dilakukan dengan meminta pertimbangan para ahli (*experts judgement*), di mana butir-butir yang telah direspon oleh para ahli (dua orang) dalam bentuk skor kemudian dikorelasikan dengan teknik statistika dan jika signifikan berarti disepakati para ahli, sehingga butir-butir dikatakan valid.
2. Validitas empiris atau validitas statistik, yaitu mencari hubungan antara skor tes dengan suatu kriteria tertentu yang merupakan suatu tolok ukur di luar tes yang bersangkutan. Ada dua macam validitas empiris, yaitu: 1) validitas prediktif, yaitu kemampuan suatu tes dapat memprakirakan perilaku peserta didik pada masa yang akan datang, dan 2) Validitas konkuren ialah jika kriteria eksternal sudah ada saat pengujian, misalnya, skor tes dari perangkat tes buatan guru dengan buatan hasil KKG/MGMP. Pengujian validitas empiris menurut Purwanto (2011: 121-122) dapat dilakukan dengan mencari koefisien korelasi antar hasil pengukuran dari perangkatperangkat tes.

3. Validitas konstruk. Konstruk adalah konsep yang dapat diobservasi (*observable*) dan dapat diukur (*measurable*). Validitas konstruk adalah kemampuan tes dalam mengobservasi dan mengukur fungsi psikologis yang merupakan deskripsi perilaku peserta didik. Pengujian validitas konstruk menurut Purwanto (2011: 126-127) dapat dilakukan dengan cara: 1) telaah butir, 2) expert judgment, 3) konvergensi dan diskriminabilitas, 4) multitrait-multimethod, dan 5) analisis faktor.

Reliabilitas berkait dengan konsistensi, keandalan, keajegan, ataupun stabilitas. Suatu alat ukur dikatakan reliabel bila memberikan hasil yang sama pada berkali-kali pengulangan pengukuran (Subali, 2010: 43-44) dan reliabilitas berlaku pada tingkat suatu perangkat tes, bukan untuk masing-masing butir tes. Kerlinger (1986: 443) mengemukakan reliabilitas dapat diukur dari tiga kriteria, yaitu *stability*, *dependability*, dan *predictability*. *Stability* menunjukkan keajegan suatu tes dalam mengukur gejala yang sama pada waktu yang berbeda. *Dependability* menunjukkan kemantapan/keandalan suatu tes. *Predictability* menunjukkan kemampuan tes untuk meramalkan hasil pada pengukuran gejala selanjutnya. Bukti keandalan atau reliabilitas suatu alat ukur dapat dilihat pada besarnya indeks keandalan.

Selanjutnya, Gronlund (1985: 100) mengemukakan ada empat faktor yang dapat mempengaruhi reliabilitas, yaitu: 1) Panjang tes (*length of test*), semakin panjang suatu tes (banyak butir) akan lebih tinggi tingkat reliabilitas suatu tes, 2). Sebaran skor (*spread of scores*), semakin besar sebaran skor akan membuat tingkat reliabilitas menjadi lebih tinggi, 3) tingkat kesukaran (*difficulty index*), di mana tingkat kesukaran soal yang ideal untuk meningkatkan koefisien reliabilitas adalah soal yang menghasilkan sebaran skor berbentuk genta atau kurva normal, dan 4) objektifitas (*objectivity*), di mana objektifitas prosedur tes yang tinggi akan memperoleh reliabilitas hasil tes yang tidak dipengaruhi oleh prosedur penskoran.

Setelah memperhatikan faktor-faktor tersebut di atas, seperangkat tes dapat diuji dengan menggunakan beberapa klasifikasi metode pengujian seperti berikut: 1) koefisien stabilitas eksternal, seperti metode tes ulang (*test-retest*) dan metode paralel, 2) koefisien konsistensi internal, meliputi metode belah dua (*split half*), metode Flanagan, metode Rulon, metode Kuder-Richardson, anova Hoyt, dan Cronbach alpha (Purwanto, 2011: 156-176), dan 3) reliabilitas interrater, seperti pendekatan Hoyt, pengukuran observasi, dan teori generalibilitas (Mardapi, 2012: 86-92).

Guna meningkatkan derajat validitas dan reliabilitas, dapat dilakukan analisis butir soal. Berikut ini setidaknya analisis butir yang dilakukan pada penilaian harian pada pembelajaran PAI:

1. Tingkat kesukaran soal (*difficulty index*), yaitu pengukuran seberapa besar derajat kesukaran suatu soal. Jika suatu soal memiliki tingkat kesukaran seimbang (*proporsional* dalam sebaran tingkat kesukaran), maka dapat dikatakan bahwa soal tersebut baik. Menurut Allen dan Yen (1979: 121), Gregory (2007: 153), dan Mardapi (2002: 116) tingkat kesulitan butir soal sebaiknya terletak pada interval 0,3 sampai 0,8 karena pada interval ini informasi tentang kemampuan peserta didik akan diperoleh secara maksimal. Untuk menghitung tingkat kesukaran soal bentuk objektif dapat dilakukan dengan menggunakan rumus tingkat kesukaran, kemudian menggunakan tabel batas tingkat kesukaran. Dalam analisis soal secara klasikal, tingkat kesukaran dapat diperoleh dengan beberapa cara, antara lain: skala kesukaran linier, skala bivariat, indeks davis, dan proporsi menjawab benar. Cara menghitung tingkat kesukaran untuk soal bentuk uraian adalah menghitung berapa persen peserta didik yang gagal menjawab benar atau ada di bawah batas lulus (*passing grade*) untuk setiap butir soal.
2. Daya Pembeda (*discriminating power*), yaitu kemampuan butir soal dalam membedakan peserta didik yang sudah menguasai kompetensi dengan peserta didik yang belum/kurang menguasai kompetensi berdasarkan kriteria tertentu. Semakin tinggi koefisien daya pembeda suatu butir soal, semakin mampu butir soal tersebut membedakan antara peserta didik yang menguasai kompetensi dengan peserta didik yang kurang menguasai kompetensi. Teknik yang digunakan untuk menghitung daya pembeda soal bentuk uraian adalah menghitung perbedaan dua rata-rata (*mean*), yaitu antara rata-rata dari kelompok atas dengan rata-rata dari kelompok bawah untuk tiap-tiap soal. Menurut teori klasik, ada hubungan antara tingkat kesukaran dan daya pembeda butir (Kartowagiran, 2008: 189).
3. Analisis Pengecoh. Butir soal pilihan ganda dikatakan baik jika pengecohnya akan dipilih secara merata oleh peserta didik yang menjawab salah. Sebaliknya, butir soal yang kurang, pengecohnya tidak dipilih atau dipilih secara tidak merata. Menurut pendapat Fernandes (1984) menjelaskan bahwa distraktor dikatakan baik apabila paling tidak dipilih oleh 2% dari seluruh peserta. Sementara itu, Nitko (1996) mengatakan distraktor dikatakan berfungsi manakala paling tidak dipilih oleh seorang peserta tes dari kelompok rendah. Pemilih dari kelompok rendah harus lebih banyak daripada kelompok atas (Kartowagiran, 2008: 190). Distraktor juga dapat dikatakan berfungsi manakala peserta tes dari kelompok atas dapat membedakan antara distraktor dan kunci jawaban sehingga yang memilih kunci jawaban lebih banyak daripada yang memilih distraktor.

4. Analisis Homogenitas Soal, dilakukan dengan menghitung koefisien korelasi antara skor tiap butir soal dengan skor total, di mana skor setiap butir soal adalah 1 atau 0, sedang skor total tiap peserta didik akan bervariasi. Teknik korelasi yang dapat digunakan adalah korelasi Pearson/ product moment atau korelasi point biserial. Butir soal dikatakan homogen, apabila koefisien korelasinya sama atau lebih besar dari harga kritik korelasi pada tabel, atau sebaliknya. Butir soal yang tidak homogen kemungkinan besar mengukur aspek lain di luar materi/bahan yang diajarkan, karena tidak sesuai dengan kompetensi yang telah ditetapkan, sehingga butir soal yang demikian sebaiknya direvisi atau dibuang.
5. Efektifitas Fungsi Opsi. Analisis butir perlu juga dicari apakah suatu opsi (alternatif jawaban) dari setiap soal berfungsi secara efektif atau tidak, dengan langkah-langkah sebagai berikut : a). menentukan jumlah peserta didik (N), b). menentukan jumlah sampel (n), baik untuk kelompok atas maupun kelompok bawah, yaitu  $27\% \times N$ , c). membuat tabel pengujian efektifitas opsi, d). menghitung jumlah alternatif jawaban yang dipilih peserta didik, baik untuk kelompok atas maupun kelompok bawah, dan e). Menentukan efektifitas fungsi opsi dengan kriteria tertentu.

#### **E. Problematika Implementasi Penilaian Harian Kurikulum 2013**

Dalam penelitian yang telah dilakukan oleh Resti Utami Hidayati, menyebutkan problem yang dihadapi seorang guru dalam implementasi penilaian autentik ialah, bahwa guru belum menerapkan instrumen penilaian autentik dengan baik, dan guru masih kesulitan dalam memilah, dan menilai antara ketiga aspek baik itu aspek sikap, pengetahuan dan keterampilan.

Problematika muncul disebabkan karena, karakteristik siswa yang tidak mendukung diantaranya kebanyakan siswa yang nilainya tengah di bawah KKM, kemampuan siswa yang berbeda-beda dalam menghafal, siswa yang tidak disiplin dalam mengerjakan ataupun mengumpulkan tugas, dan membolos saat jam pelajaran karena alasan organisasi, serta kemampuan berfikir siswa yang beragam.

Sedangkan problematika yang muncul itu dari guru sendiri, seperti kurang dikembangkannya kreativitas guru dalam mengimplementasikan RPP ketika situasi kelas tidak kondusif, guru kesulitan dalam mengidentifikasi siswa yang tidak aktif, guru tidak menggunakan instrumen penilaian sikap dan hanya menggunakan ingatannya, guru kesulitan menyemangati siswa untuk belajar, serta guru menilai sikap peserta didik hanya berdasarkan nilai pengetahuannya. Pelaksanaan RPP dan penilaian autentik juga terhambat karena alokasi waktu yang terbatas (Hidayati, 2018:131-132).

Berbeda dengan penelitian di atas dalam penelitian yang dilakukan oleh Efi Tria Astuti (2017:37-38) menemukan bahwa penyusunan soal yang terlalu banyak, format penilaian autentik yang terlalu banyak, format penilaian yang terlalu rumit membuat seorang guru kerepotan dalam melakukan penilaian kepada setiap peserta didik. Serta kendala waktu untuk menyusun dan melaksanakan penilaian autentik dirasa kurang cukup oleh guru. Berikut ini adalah paparan tentang implementasi penilaian autentik dalam pembelajaran PAI beserta kendala dalam menerapkan penilaian:

1. Penilaian sikap. Bentuk penilaian yang diimplementasikan dalam ranah sikap mencakup observasi guru, penilaian antar diri, penilaian antar sejawat, maupun jurnal. Idealnya semua bentuk penilaian itu dapat dilaksanakan secara keseluruhan dalam pembelajaran PAI, akan tetapi pada realitasnya guru masih belum mampu melaksanakan secara maksimal sesuai dengan bentuk dan tahap penilaian yang telah ditetapkan. Mayoritas guru belum maksimal melaksanakan semua bentuk penilaian ranah sikap, kenyataannya adalah instrumen observasinya tidak tersedia sehingga aspek-aspek yang dinilai oleh guru ialah kurang terarah dengan baik. Dan juga karena keterbatasan kemampuan guru untuk menyediakan semua instrumen penilaian yang sesuai dengan bentuk penilaian pada aspek sikap. Karena idealnya guru harus membuat sendiri instrumen penilaiannya berdasarkan kompetensi yang ingin di capai dalam mata pelajaran yang diampunya.
2. Penilaian pengetahuan. Pada ranah ini mayoritas kemampuan guru dalam melaksanakan sebuah penilaian pada aspek pengetahuan sudah di lakukan dengan baik yaitu melalui penilaian harian, PTS, dan PAS yang di sampaikan melalui rapor yang akan di terima siswa pada setiap satu semester.
3. Bentuk penilaian keterampilan sendiri menggunakan tes proyek dan portofolio serta menciptakan sebuah produk. Kendalanya ialah banyaknya format penilaian sehingga membuat guru terbebani dalam melakukan penilaian yang mana guru harus menilai secara detail dalam proses pembelajaran yang dikaitkan dengan aktivitas sesungguhnya mereka di luar madrasah.

## **KESIMPULAN**

Berdasarkan Permendikbud No. 104 Tahun 2014, penulis mencoba membandingkan baik dari persamaan dan perbedaannya. Dari persamaannya, pada pelaksanaan penilaian hasil belajar siswa oleh pendidik dinilai dalam model penilaian autentik dan non-autentik tetapi penilaian autentik merupakan pendekatan pokok dalam penilaian prestasi belajar bagi guru. Selanjutnya dari

segi perbedaannya dalam bentuk penilaiannya, penilaian autentik mencakup penilaian berdasarkan pengamatan, tugas ke lapangan, portofolio, proyek, produk, jurnal, kegiatan laboratorium, dan unjuk kegiatan, serta penilaian pribadi. Teknik penilaian autentik memperhatikan 3 hal yaitu kemampuan sikap, pengetahuan, dan keterampilan. Komponen penilaian autentik sendiri yang perlu diperhatikan ialah penyiapan tugas autentik dan rubrik penilaian autentik. Berdasarkan analisis dan pembahasan di atas, peneliti dapat menyimpulkan bahwa evaluasi hasil belajar Pendidikan Agama Islam (PAI) meliputi beberapa hal, yaitu: (1) tingkat kesukaran soal (*difficulty index*); (2) daya pembeda (*discriminating power*); (3) analisis pengecoh; (4) analisis homogenitas soal; dan (5) efektifitas fungsi opsi.

## **DAFTAR PUSTAKA**

- Ani, Yubali, "Penilaian Autentik Kurikulum 2013", Seminar Nasional Implementasi 2013, 24
- Arifin, Zainal, (2012), *Evaluasi Pembelajaran*. Jakarta: Subdit Kelembagaan Direktorat Pendidikan Tinggi Islam Direktorat Jenderal Pendidikan Islam, Kementerian Agama RI
- Arikunto, Suharsimi. (2008). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara
- Astuti, Efi Tria, *Problematika Implementasi Penilaian Autentik Kurikulum 2013 Dalam Pembelajaran Pendidikan Agama Islam di SD Negeri Ploso 1 Pacitan*, Jurnal Al-Idaroh, Vol.1, No 2, September 2017
- Azwar, Saifuddin. 2000. *Reliabilitas dan Validitas*. Yogyakarta: Sigma Alpha
- Gronlund, N.E. (1985). *Measurement and Evaluation in Teaching*. New York : Mc Millan Publishing Co., Inc.
- Hidayati, Resti Utami, (2018), *Problematika Guru Dalam Pelaksanaan Penilaian Autentik Pada Mata Pelajaran Rumpun Pendidikan Agama Islam Di MAN 1 Banyumas*, IAIN Purwokerto
- Lias, Hasibuan. (2010). *Kurikulum & Pemikiran Pendidikan*. Jakarta: GP Press.
- Majid, Abdul dan Dian Andayani, 2004, *Pendidikan Agama Islam Berbasis Kompetensi*, Bandung: PT Remaja Rosdakary

- Mulyasa, E., *Manajemen Pendidikan Karakter*, Jakarta: Bumi Aksara, 2012
- Nurgiyantoro, Burhan, (2011), *Penilaian Otentik*, Yogyakarta: Gadjah Mada University Press
- Permendikbud RI No. 23 Tahun 2016 Tentang Standar Penilaian Pendidikan
- Sani, Ridwan Abdullah, 2016, *Penilaian Autentik*, Jakarta: bumi aksara
- Subali, Bambang. 2010. *Penilaian, Evaluasi, dan Remediasi Pembelajaran Biologi*. Yogyakarta: Jurusan Biologi Fakultas MIPA
- Sukardi, H.M. (2009). *Evaluasi Pendidikan: Prinsip & Operasionalnya*. Jakarta: Bumi Aksara
- Sukiman, (2012), *Pengembangan Sistem Evaluasi*, Yogyakarta: Insan Madani
- Supardi, 2015, *Penilaian Autentik Pembelajaran Afektif, Kognitif, dan Psikomotorik (Konsep dan Aplikasi)*, Jakarta: Rajawali Pers
- Widayati Catharina Sri Wahyu. 2009. *Komparasi Beberapa Metode Estimasi Kesalahan Pengukuran*. Jurnal Penelitian dan Evaluasi Pendidikan Tahun 13, Nomor 2

*Ilma Kharismatunisa dan Moh. Sahlan*