



JURNAL BASICEDU

Volume 6 Nomor 2 Tahun 2022 Halaman 3076 - 3085

Research & Learning in Elementary Education

<https://jbasic.org/index.php/basicedu>



Analisis Instrumen Penilaian Kelas Domain Kognitif Guru Mata Pelajaran Agama Islam di Madrasah Ibtidaiyah

Ahmad Sholeh

Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia

E-mail: sholeh_76@yahoo.com

Abstrak

Penelitian ini bertujuan ingin mengungkap profesionalisme guru agama Islam dalam melakukan penilaian pembelajaran domain kognitif di kelas dengan memfokuskan pada bagaimana model instrument dan teknik para guru mata pelajaran agama Islam dalam melakukan penilaian kognitif pembelajaran kelas serta bagaimana kualitas instrument kognitif penilaian kelas yang dilakukannya tersebut. Jenis penelitian ini adalah penelitian survey yang bersifat deskriptif eksploratif. penelitian ini dilakukan di Madrasah Ibtidaiyah Sunan Kalijaga Malang. Karena terbatasnya penelitian ini maka yang menjadi sampel adalah para guru agama Islam pada kelas 3 dan kelas 5. Pengumpulan data dilakukan dengan menggunakan dokumentasi/lembar penilaian dan wawancara. Sedangkan untuk memecahkan permasalahan digunakan teknik analisis deskriptif. Hasilnya adalah; 1. Model dan teknik pembuatan instrumen penilaian kelas yang digunakan para guru agama Islam MI Sunan Kalijaga Malang sudah sesuai dengan aturan teoritis yang ada. 2. Masih terdapatnya soal yang dibuat oleh guru agama Islam itu perlu direvisi dan ada yang perlu dibuang. Namun kalau dilihat dari reliabilitas soal memiliki reliabilitas yang tinggi. Oleh karena itu ke depannya masih perlu digiatkan lagi pelatihan-pelatihan kepada guru agama Islam guna pembuatan instrument penilaian kelas yang baik.

Kata Kunci: Instrumen, penilaian kelas, kognitif, guru agama Islam.

Abstract

This study aims to reveal the professionalism of Islamic religious teachers in assessing cognitive domain learning in the classroom by focusing on how the instrument models and techniques for Islamic religious teachers in carrying out cognitive assessments of classroom learning and how the quality of the cognitive instruments of classroom assessments they do. This type of research is survey research that is descriptive exploratory. This research was conducted at Madrasah Ibtidaiyah Sunan Kalijaga Malang. Due to the limitations of this study, the samples were Islamic religious teachers in grade 3 and grade 5. Data was collected using documentation/assessment sheets and interviews. Meanwhile, to solve the problem using descriptive analysis techniques. The results of this study showed that; 1. Models and techniques for making class assessment instruments used by Islamic religion teachers at MI Sunan Kalijaga Malang are following existing theoretical rules. 2. There are still questions made by the Islamic religion teacher that need to be revised and some need to be discarded. However, if viewed from the reliability of the questions, they have high reliability. Therefore, in the future, training for Islamic religious teachers still needs to be intensified to make good class assessment instruments.

Keywords: instrument, classroom assessments, cognitive, Islamic religious teachers.

Copyright (c) 2022 Ahmad Sholeh

✉ Corresponding author :

Email : sholeh_76@yahoo.com

DOI : <https://doi.org/10.31004/basicedu.v6i2.2129>

ISSN 2580-3735 (Media Cetak)

ISSN 2580-1147 (Media Online)

PENDAHULUAN

Dalam Undang-Undang Guru dan Dosen disebutkan bahwasanya “guru adalah pendidik profesional yang mendidik, mengajar, membimbing, mengarahkan, melatih, menilai, dan mengevaluasi peserta didik pada pendidikan anak usia dini pada jalur pendidikan formal, pendidikan dasar, dan pendidikan menengah” (Siti Mariyah, 2019). Berdasarkan pernyataan di atas jelaslah bahwa seorang guru agama Islam dalam menjalankan tugasnya harus mampu mengevaluasi atau menilai para siswanya dengan baik agar proses pembelajaran tersebut diarahkan sesuai dengan tujuan yang diharapkan. Penilaian menurut Scriven itu ada dua macam yaitu; penilaian proses belajar dan hasil belajar. Penilaian proses belajar lebih disebut dengan penilaian formatif yang difokuskan untuk mengidentifikasi kemajuan atau perkembangan belajar para siswa yang berguna untuk umpan balik sebagai perbaikan pembelajaran. Sedangkan penilaian hasil belajar itu lebih dikenal dengan penilaian sumatif yang difokuskan pada hasil belajar para siswa setelah akhir program pembelajaran (Ina Magdalena, Fatimah Mulyani, Dwi Nur Faridah, Nuri Fitriyani, 2020).

Berdasarkan Peraturan Menteri Pendidikan dan Kebudayaan Nomor 24 tahun 2017 tentang Standar Penilaian Pendidikan terkait dengan penilaian dilaksanakan untuk memantau proses belajar mengajar. Penilaian dalam proses belajar ini diharapkan digunakan oleh seorang guru untuk dapat menyiapkan fasilitas pembelajaran yang variatif dan inovatif dalam rangka meningkatkan kualitas kemampuan para siswanya dalam menguasai ilmu pengetahuan dan perbaikan pendidikan yang dilaksanakan di sekolah maupun secara nasional (Hartini et al., 2021).

Penilaian dalam pembelajaran ini haruslah dilaksanakan oleh setiap guru agar dapat mengetahui kekurangan maupun kelebihan pada saat melakukan pembelajaran. Penilaian dalam pembelajaran ini juga dapat digunakan untuk mengambil suatu keputusan akademik terhadap siswa, efektifitas dan efisiensi kurikulum dan kebijakan akan pendidikan yang ada. Dalam hal ini Arikunto & Jabar menjelaskan bahwa, penilaian adalah suatu proses kegiatan yang dilakukan dengan mengumpulkan suatu informasi yang berkaitan dengan proses kerja seseorang, yang pada akhirnya informasi tersebut digunakan untuk menentukan sebuah alternatif keputusan yang tepat (Tambakmulyo et al., 2021).

Realitas di lapangan menunjukkan bahwa para guru dalam melakukan proses penilaian kelas masih banyak menekankan pada penguasaan konsep dengan lebih banyak menggunakan instrumen penilaian yang mengukur ranah kognitif rendah. Padahal menurut Anderson dan Krathwohl “Ranah kognitif itu berisi perilaku yang menekankan aspek intelektual, seperti pengetahuan, dan keterampilan berpikir yang mencakup kemampuan berpikir tingkat rendah atau *Lower Order Thinking Skills* (LOTS) mengingat (C1), memahami (C2), dan mengaplikasikan (C3), lalu ada tiga aspek dari kemampuan berpikir tingkat tinggi *Higher Order Thinking Skills* (HOTS) yaitu kemampuan menganalisis (C4), mengevaluasi (C5) dan mencipta (C6) berdasarkan taksonomi Bloom yang telah direvisi” (Nabilah et al., 2020).

Menurut (Dinni, 2018) revisi taksonomi Bloom ini pada hakekatnya lebih memfokuskan pada bagaimana aspek kognitif bisa lebih aplikatif dan hidup bagi seorang guru dan praktek pembelajaran dengan harapan agar dapat membantu seorang guru dalam merumuskan suatu tujuan dari pembelajaran dan strategi penilaian yang lebih efisien. Disamping itu, dengan harapan juga untuk meningkatkan kualitas para siswa. Oleh karena itu, seorang guru harus dapat merancang suatu pembelajaran yang dapat meningkatkan kemampuan kognitif (Nabilah et al., 2020).

Seiring dengan berubahnya kurikulum menjadi kurikulum 2013 membawa konsekuensi logis terhadap perubahan paradigma penilaian pembelajaran yang dianut para guru sebelumnya. Berubahnya sistem pembelajaran dari yang berpusat kepada guru ke arah pembelajaran yang lebih mengaktifkan siswa dan bermakna, mengharuskan guru untuk melakukan inovasi terhadap pola pembelajaran yang dilakukannya. Perubahan pada persiapan mengajar, membawa konsekuensi tambahan tugas guru untuk mengubah

perencanaan mengajar, pelaksanaan mengajar, dan tindak lanjut terhadap hasil kegiatan pembelajaran yang disesuaikan dengan kurikulum yang berlaku termasuk pada perubahan system evaluasi yang dilakukannya.

Mengingat perbedaan kompetensi dasar guru (termasuk guru agama Islam) pada berbagai jenjang pendidikan diwujudkan dalam kadar kompetensi yang berbeda, terutama kompetensi dalam melakukan penilaian. Adanya kompetensi yang berbeda ini menimbulkan suatu permasalahan dalam penilaian hasil belajar siswa. Permasalahan ini diduga terutama menyangkut kegiatan perencanaan, pelaksanaan, kemampuan menganalisis hasil penilaian dan tindak lanjut hasil penilaian yang dilakukan.

Akibat dari perkembangan paradigma penilaian dari yang bersifat konvensional yang lebih menekankan pada penilaian kognitif semata ke system penilaian yang lebih komprehensif dan bermakna, maka kompetensi guru dalam melakukan penilaian paling tidak dipengaruhi oleh dua factor, yaitu factor yang ada dalam diri pribadi guru agama Islam dan factor yang berada di luar diri guru (*factor eksternal*). Faktor-faktor yang diduga mempengaruhi kompetensi guru agama nantinya dalam melakukan penilaian kognitif adalah : pengalaman mengajar, pengalaman penataran, beban kerja dan motivasi berprestasi.

Mengingat penting dan urgennya penilaian pembelajaran di kelas pada domain kognitif, maka profesionalisme guru agama Islam sangat dituntut dalam penerapannya di sekolah. Namun yang menjadi permasalahannya adalah bagaimana profesionalisme guru agama Islam dalam melakukan penilaian pembelajaran pada domain kognitif di kelas?.

Untuk menjawab permasalahan tersebut diperlukan suatu studi tersendiri melalui penelitian. Sebenarnya berdasarkan kajian literatur yang ada, sudah ada beberapa penelitian yang terkait dengan permasalahan ini, di antaranya “Analisis Teknik Evaluasi Guru pada Pembelajaran Agama Islam di Madrasah Aliyah” (Hasnia et al., 2021), “Analisis Instrumen Tes Soal Pilihan Ganda Ujian Sekolah (US) Mata Pelajaran Al Islam Kelas XII SMA Muhammadiyah se-kota Palembang” (Amrullah, 2018), “Evaluasi Guru Pendidikan Agama Islam dalam Pembelajaran Daring” (Wahyuningsih & Sriyanto, 2021), “Evaluasi Penerapan Kurikulum 2013 Pada Mata Pelajaran Aqidah Akhlak Di Madrasah Tsanawiyah Negeri 1 Kolaka Kabupaten Kolaka” (Fitria, 2013) dan “Pelaksanaan Evaluasi Formatif dalam Mengukur Hasil Belajar Mata Pelajaran IPS di Kelas V SDK TSM Posona” (Sawania et al., 2016). Namun penelitian ini sangat berbeda karena mencoba mengungkap profesionalisme guru agama Islam dalam melakukan penilaian pembelajaran domain kognitif di kelas dengan memfokuskan pada bagaimana model instrument dan teknik para guru mata pelajaran agama Islam dalam melakukan penilaian kognitif pembelajaran kelas serta bagaimana kualitas *instrument* kognitif penilaian kelas yang dilakukan oleh para guru mata pelajaran agama Islam tersebut dalam penerapan kurikulum 2013.

METODE

Jenis penelitian ini adalah penelitian survey yang bersifat deskriptif eksploratif. Jenis penelitian ini adalah penelitian survey yang bersifat *Ex Post Facto*. Karena akan menganalisis bagaimana guru instrument yang digunakan guru PAI dan Dalam hal ini peneliti hanya menganalisis data yang sudah diperoleh dari lapangan dengan setting yang natural. Dengan demikian penelitian ini dilakukan penelaahan terhadap keadaan yang sedang ada pada diri subjek penelitian.

Populasi penelitian ini adalah para guru agama Islam yang ada di Madrasah Ibtidaiyah Sunan Kalijaga Malang yang mengajar mata pelajaran agama Islam seperti al-qur`an, fiqh, akidah ahlak, dan sejarah kebudayaan Islam. Karena terbatasnya penelitian ini maka yang menjadi sampel adalah para guru agama Islam pada kelas 3 dan kelas 5 (lima). Kelas 3 dianggap mewakili kelas rendah dan kelas 5 dianggap peneliti mewakili kelas tinggi.

Pengumpulan data dilakukan dengan menggunakan; Pertama, dokumentasi/lembar penilaian. Teknik ini peneliti gunakan terutama untuk mendapatkan data dari jawaban-jawaban siswa terhadap jawaban instrumen

penilaian guru yang ada guna untuk dijadikan data analisis. Kedua, wawancara. Teknik ini peneliti gunakan untuk mengungkap dan menemukan data yang belum bisa dijawab oleh teknik dokumentasi seperti bagaimana proses guru membuat instrumen penilaian kelas atau bagaimana teknik yang telah ditempuh oleh guru dalam pembuatan instrumen penilaian kelas. Yang peneliti wawancarai adalah guru Feqh, Sejarah Kebudayaan Islam, Aqidah Akhlak dan al-Qur`an Hadits.

Sedangkan untuk menganalisis data digunakan teknik analisis deskriptif. Adapun untuk melakukan analisis data terhadap kualitas alat penilaian pembelajaran kelas digunakan analisis butir soal (Teori Tes Klasik) dengan program MICROCAT ITEMAN versi 3.0. Untuk lebih jelasnya, dalam menilai validasi instrumen penilaian kelas domain kognitif para guru agama Islam dapatlah ditabelkan sebagai berikut :

Tabel I
Tabel untuk Iteman
(Menilai Validasi Soal Per-Butir)

Tingkat Kesukaran Butir (Prop. Correct)	
0.00 - 0.24	Sukar
0.25 - 0.75	Sedang
0.76 - 1.00	Mudah
Daya Pembeda (deskriminan)	
0.40 - 1.00	Baik
0.20 - 0.39	Cukup Baik
-1.00 - 0.19	Tidak Baik
Distraktor (Biserial)	
Lebih - 0.30	Baik
Kurang - 0.30	Belum Baik

HASIL DAN PEMBAHASAN

Untuk mencapai efektifitas pembelajaran mata pelajaran agama Islam, di antaranya mata pelajaran fiqh, sejarah kebudayaan Islam, qur`an Hadits, aqidah akhlak di kelas, seorang guru agama Islam haruslah mengetahui berbagai kesulitan yang dihadapi oleh siswa, karena setiap siswa mempunyai pemerolehan dan pemahaman yang berbeda-beda tentang mata pelajaran-mata pelajaran di atas yang dipelajarinya. Di sinilah fungsi penilaian sebagai bahan acuan guna meenentukan tolak ukur akan keberhasilan suatu proses pembelajaran yang dilakukannya. Di samping itu, dapat juga dijadikannya sebagai bahan perbaikan terhadap pembelajaran tersebut.

Menurut (Supriyadi, 2020) pada hakekatnya penilaian kelas itu bisa dilakukan secara formal maupun informal dan harus dilakukan secara berkelanjutan. Penilaian kelas ini sebgai bahan masukan guru agama Islam dalam menyempurnakan suatu proses pembelajaran. Di samping itu juga dapat digunakan sebagai penguasaan siswa terhadap subtansi materi pelajaran dan kendala para siswa dalam mempelajari suatu materi pelajaran tersebut. Oleh karena itu penilaian kelas harus terus dikembangkan guna memperbaiki mutu pendidikan.

Dengan diberlakukannya Kurikulum 2013, dalam satu tahun program pengajaran itu dibagi menjadi dua semester. Evaluasi pembelajaran dalam rangka mengukur tingkat keberhasilan pengajaran dilakukan setiap satu semester. Berdasarkan hasil penelitian para guru agama Islam di MI Sunan Kalijaga Malang tidak mengadakan ulangan semester bersama dengan guru di lain lembaga. Sehingga evaluasi dilakukan oleh guru

yang bersangkutan dan jadwalnya tergantung kalender pendidikan yang ada. Di setiap semester para guru agama Islam di MI Sunan Kalijaga Malang ini melakukan penilaian pembelajaran kelas dengan tiga (tiga) model, yaitu: Pertama, model ulangan sisipan. Model penilaian ini dilakukan beberapa kali di setiap semesternya berdasarkan kebutuhan dalam melihat ketercapaian kompetensi dasar yang ada. Kedua, model ulangan tengah semester. Proses ujian ini dilakukan sesuai dengan waktu yang ditetapkan bersama. Para guru dituntut untuk membuat soal sendiri dan dikoreksi sendiri. Ketiga, model ulangan akhir semester. Model ini sama dengan model ulangan tengah semester.

Dalam kenyataannya, sesuai dengan hasil penelitian para guru agama Islam di MI Sunan Kalijaga Malang ketika dalam pelaksanaan evaluasi atau penilaian yang ada tidak pernah mengalami kesulitan-kesulitan yang berarti. Ada beberapa faktor yang mengakibatkan guru agama Islam tidak mengalami kendala dalam penilaiannya di antaranya : *Pertama*, adanya kerjasama antara guru-guru agama Islam satu rumpun yang saling memberi masukan antara satu dengan lainnya, *Kedua*, sebelumnya alat penilaian/evaluasi yang digunakan sudah dirancang jauh sebelum di mulai proses pembelajaran di awal semester. Hal ini sesuai dengan pernyataan guru SKI sebagai berikut : “saya selalu mempersiapkan soal-soal penilaian jauh-jauh sebelumnya agar supaya ketika membuatnya lebih berkonsentrasi dan tidak diganggu oleh kesibukkan saya yang ada. Di samping itu kita bisa berdiskusi dengan guru-guru agama Islam yang serumpun dengan waktu yang tidak terbatas”.

Adapun model instrument yang digunakan dalam penilaian kognitif kelas oleh para guru agama Islam MI Sunan Kalijaga Malang adalah; *Pertama*, Tes Obyektif. Model soal yang digunakan untuk tes ini adalah soal *multiple choice* di mana para siswa hanya memilih jawaban yang benar dari jawaban yang disediakan dengan melingkari, tanda silang atau yang lainnya. Menurut Azwar keunggulan tes pilihan ganda, diantaranya; *Pertama*, bersifat kompherhensif, *Kedua*, lebih cepat dan mudah dalam memeriksa jawabanya dan memberi skornya, *Ketiga*, bersifat lebih efisien dan hemat bahan dalam menyediakan lembar jawabannya, *Keempat*, secara empirik kualitas butir soalnya dapat dianalisis, *Kelima*, tinggi objektifitasnya, dan keenam, secara umum reabilitas yang memuaskan (Kadir, 2015). Walaupun demikian menurut (Suharsimi Arikunto, 2013) tes *mutiple choice* juga mempunyai beberapa kelemahan, di antaranya; 1) Menyusun soalnya itu lebih sulit, dikarenakan butuh ketelitian untuk menghindari bentuk lainnya yang bermakna sama dengan item-item lainnya. 2) Lebih cenderung hanya mengukur ingatan dan daya pengenalan kembali dari pada mengukur proses mental yang tinggi. 3) Peluang main untung-untungan lebih banyak dan 4) Para siswa mempunyai kesempatan kerjasama lebih banyak pada saat mengerjakannya. *Kedua*, tes essay, meliputi : *Pertama*. Bentuk soal uraian terbatas dan *Kedua*. Bentuk soal uraian bebas. *Ketiga*, tes lisan, yakni tes yang pelaksanaannya dilakukan oleh guru agama Islam dengan mengadakan tanya jawab secara dengan para siswanya. Kelemahannya adalah bersifat subjektivitas dan butuh waktu yang banyak. *Keempat*, Praktek. Untuk menilai tes praktek ini para guru agama Islam MI Sunan Kalijaga Karang Besuki pada umumnya menggunakan format pengamatan.

Selanjutnya teknik yang dilakukan oleh guru agama Islam MI Sunan Kalijaga Malang dalam penilaian pembelajaran adalah sebagai berikut : *Pertama*, perencanaan tes. Pada tahap ini para guru agama Islam memilih dan mempersiapkan teknis tes yang disesuaikan dengan apa yang mau diukurnya. *Kedua*, penyusunan tes. Pada tahap ini para guru agama menyusun soal disesuaikan dengan kisi-kisi yang dibuatnya sebelumnya dengan memperhatikan tingkat kesukaran dan daya distraktor soal dengan mengacu pada pedoman pembuatan soal yang baik. Hal ini sesuai yang dikatakan guru al-Qur`an Hadits sebagai berikut; “Dalam membuat instrument penilaian, tahap awal dalam evaluasi yang saya lakukan adalah membuat kisi-kisi soal, membuat soal-soal, baru membuat atau menentukan pedoman penilaian (soal remidi) dan ditelaah guru serumpun”. *Ketiga*, pengumpulan data yang meliputi: melaksanakan tes, dan mengoreksi serta memberi skor yang berbeda terhadap tes objektif dan esai, dan *Keempat*, membuat laporan hasil penilaian yang diberikan kepada para siswa dan pihak sekolah. Hasil penelitian ini menguatkan temuan penelitiannya (Dea

Mustika, Ambiyar, 2021) yang menyatakan bahwa “Proses penilaian hasil belajar siswa yang guru lakukan meliputi pada lima tahapan yaitu mengkaji silabus, pembuatan kisi-kisi penilaian, melaksanakan kegiatan penilaian, analisis hasil penilaian dan menyusun laporan penilaian”.

Sedangkan di dalam analisis kuantitatif ini peneliti hanya mengambil sampel pada soal mata pelajaran Aqidah akhlak dan Sejarah Kebudayaan Islam di kelas III dan mata pelajaran Feqh dan al-Qur`an Hadits di kelas V. Berdasarkan hasil penelitian yang dianalisis dengan program *MicroCAT ITEMEN* version 3.00 terhadap instrument penilaian kelas Aqidah Akhlak, SKI, Qur`an Hadits dan Fiqh dapatlah dinyatakan bahwa ;

- 1) Untuk instrumen mata pelajaran aqidah akhlak didapatkan sejumlah 4 soal (20%) termasuk soal sukar, 16 soal (80%) soal sedang, 0 soal (0%) soal mudah. Daya beda soal dari sejumlah 20 soal termasuk soal baik ada 14 soal (70%), soal cukup baik 2 soal (10%), soal tidak baik 4 soal (20%). Sedangkan Berdasarkan data hasil efektifitas distraktor terdapat 14 soal (70%) distraktor efektif atau baik dan 6 soal (30 %) distraktor belum baik atau tidak efektif.
- 2) Untuk instrumen mata pelajaran SKI didapatkan sejumlah 3 soal (20%) termasuk soal sukar, 11 soal (73%) soal sedang, 1 soal (7%) soal mudah. Daya beda soal dari sejumlah 15 soal termasuk soal baik ada 11 soal (73%), soal cukup baik 1 soal (7%), soal tidak baik 3 soal (20%). Sedangkan Berdasarkan data hasil efektifitas distraktor terdapat 9 soal (60%) distraktor efektif atau baik dan 6 soal (30 %) distraktor belum baik atau tidak efektif.
- 3) Untuk instrumen mata pelajaran Feqh didapatkan sejumlah 20 soal terdapat 12 soal (60%) termasuk soal sukar, 3 soal (15%) soal sedang, 5 soal (25%) soal mudah. Daya beda soal dari sejumlah 20 soal termasuk soal baik ada 13 soal (65%), soal cukup baik 2 soal (10%), soal tidak baik 5 soal (25%). Sedangkan Berdasarkan data hasil efektifitas distraktor terdapat 13 soal (65%) distraktor efektif atau baik dan 7 soal (35 %) distraktor belum baik atau tidak efektif.
- 4) Untuk instrumen mata pelajaran al-Qur`an Hadits didapatkan sejumlah 10 soal terdapat 3 soal (30%) termasuk soal sukar, 5 soal (50%) soal sedang, 2 soal (20%) soal mudah. Daya beda soal dari sejumlah 10 soal termasuk soal baik ada 8 soal (80%), soal cukup baik 1 soal (10%), soal tidak baik 1 soal (10%). Sedangkan Berdasarkan data hasil efektifitas distraktor terdapat 7 soal (70%) distraktor efektif atau baik dan 3 soal (30 %) distraktor belum baik atau tidak efektif.

Dari hasil analisis di atas dapat dinyatakan bahwa ternyata tingkat kesukaran soal-soal Aqidah Akhlak, SKI, al-Qur`an Hadits dan Feqh yang dibuat guru agama Islam MI Sunan Kalijaga Malang sebagian besar mempunyai indeks tingkat kesukaran tidak seimbang antara yang kaegori sulit-sedang-mudah. Menurut Nana Sudjana, sebaiknya sebuah paket soal yang diberikan kepada siswa memiliki keseimbangan antara sukar : sedang : mudah dengan perbandingan 3: 4: 3 atau 2 : 5 : 3 (Nuswowati et al., 2011). Oleh sebab itu perlu ada revisi.

Dengan berdasarkan analisa kualitatif ada soal-soal yang tergolong dapat direvisi, karena ini ada yang bersifat tumpang tindih antara satu soal dengan soal yang lain. Soal itu adalah :

Kutipan soal Aqidah Akhlak no 4 dan 8 :

4. Bacaan untuk mengusir dari gangguan jin dan syetan di sebut....

- a. ta`awudz
- b. basmalah
- c. tasbih
- d. tachmid

8. Agar kita terhindar dari godaan syetah maka kita baca kalimat....

- a. ta`awudz
- b. hamdalah

- c. basmalah
- d. tasbih

Disamping itu ada pula instrument soal yang belum ada kunci jawaban yang jelas. Hal ini terdapat di soal SKI kelas III, yaitu :

Kutipan soal no. 2 :

1. Sebelum Abrahah menghancurkan ka'bah dia mengutus salah satu utusannya untuk merampas harta milik kakek nabi Muhammad SAW, yang bernama....
- a. Hunata
 - b. Aswad
 - c. Abu Sufyan
 - d. Abu Tholib

Berdasarkan hasil analisis data kuantitatif dapat diketahui bahwa masih ada pengecoh (distraktor) tidaklah berfungsi dengan baik, karena semua pilihan jawaban (option) belum dipilih dari 5% seluruh pengikut tes. Seperti pada soal :

Kutipan soal no. 14 mata pelajaran Feqh :

14. Pak Danu bernadzar jika lulus ujian PNS akan berkorban satu ekor kambing pada hari raya idhul adha yang akan datang, maka hukum berkorban pak Danu adalah.....

- a. Wajib
- b. Sunnah
- c. Mubah
- d. Makruh

Kunci jawaban : A

Hasil print out analisis soal no 14;

14	0-14	0.441	0.112	0.089	A	0.441	0.112	0.089	*
					B	0.059	0.537	0.268	?
					C	0.206	0.022	0.016	
					D	0.294	-0.329	-0.249	
					Other	0.000	-9.000	-9.000	

Sebaiknya soal no 14 direvisi menjadi :

14. Pak Danu bernadzar jika lulus ujian PNS akan berkorban satu ekor kambing pada hari raya idhul adha yang akan datang, maka hukum berkorban pak Danu adalah.....

- a. Sunnah
- b. Wajib
- c. Mubah
- d. Makruh

Kunci jawabannya menjadi : B

Berdasarkan hasil analisis daya beda soal diketahui terdapat adanya soal yang tidak perlu digunakan lagi untuk disimpan di bank soal. Di antara beberapa soal yang mempunyai daya beda jelek, karena nilai daya bedanya berharga (-) sehingga soal tersebut harus dibuang. Hal ini dikarenakan soal tersebut akan tidak dapat membedakan antara kelompok yang pandai dan tidak pandai. Apalagi ditambah dengan daya distraktornya juga tidak berfungsi dengan baik. Hal ini terdapat pada soal;

Kutipan soal Feqh no. 17 :

17. Waktu paling akhir penyembelihan hewan qurban adalah....

- tanggal 12 dzulhijjah sebelum matahari terbenam
- tanggal 13 dzulhijjah sebelum matahari terbenam
- tanggal 13 dzulhijjah sebelum shoal dhuhur
- akhir bulan dzulhijjah

Kunci jawaban ; B

17	0-17	0.500	-0.285	-0.227	A	0.235	0.271	0.196	?
					B	0.500	-0.285	-0.227	*
					C	0.118	-0.052	-0.032	
					D	0.147	0.177	0.115	
					Other	0.000	-9.000	-9.000	

CHECK THE KEY

B was specified, A works better

Kutipan soal Qur`an Hadits no. 7 :

7. Lafal Surat al-Alaq ayat 18 adalah

Dengan kunci jawaban : **B**

7	0-7	0.118	-0.077	-0.047	A	0.265	0.745	0.553	?
					B	0.118	-0.077	-0.047	*
					C	0.500	-0.373	-0.298	
					D	0.118	-0.405	-0.248	
					Other	0.000	-9.000	-9.000	

CHECK THE KEY

B was specified, A works better

Kalau kita perhatikan dari hasil print out soal no 17 dan 7 terdapat kata-kata “**CHECK THE KEY**” ini menandakan bahwa kunci jawaban dari kedua soal tersebut penempatannya tidak tepat. Seharusnya kalau kunci jawabannya soal no. 17 itu di jawaban: A (**B was specified, A works better**) dan soal no. 7 itu di jawaban : A (**B was specified, A works better**).

Walaupun instrumen yang dibuat oleh guru agama Islam MI Sunan Kalijaga Malang berdasarkan analisis ada yang masih perlu direvisi dan ada yang perlu dibuang, namun kalau dilihat dari reliabilitas soal mempunyai Koefisien Alpha : 0,831 itu mata pelajaran SKI, 0,828 untuk mata pelajaran Aqidah Akhlak, 0,849 untuk mata pelajaran Feqh dan 0,710 untuk mata pelajaran al-Qur`an Hadits, maka soal-soal yang dibuat oleh guru agama Islam MI Sunan Kalijaga Malang bisa digolongkan memiliki reliabilitas yang tinggi, artinya soal-soal tersebut memiliki kehandalan yang tinggi. Kehandalan yang dimaksud dalam hal ini meliputi ketepatan/kecermatan hasil pengukuran dan kejegan/kestabilan dari hasil pengukuran. Ini sesuai yang dikatan oleh Nunnally (1978) berpendapat bahwa soal uraian yang memiliki koefisien reliabilitas 0,6 – 0,7 dan untuk soal pilihan ganda yang memiliki koefisien reliabilitas 0,70 – 0,90 sudah dapat dikatakan baik. Feldt dan Brehmman (1989) mengatakan bahwa suatu instrumen yang memiliki koefisien reliabilitas $r_{xy} \geq 0,7$ sudah dikatakan reliable (Badrun Kartawagiran, 2008).

Menurut peneliti masih terdapatnya soal tes yang tidak baik dari segi tingkat kesukaran soal, daya diskriminan soal dan daya distraktor soal di atas di sebabkan juga adanya keterbatasan atau kelemahan analisis dari alat analisis yang digunakan dengan menggunakan program *MicroCAT ITEMAN* version 3.00 ini. Kelemahan ini bersumber dari teori yang digunakan yaitu teori tes klasik. Ini terjadi disebabkan oleh keterbatasan model pengukuran secara klasik bila dibandingkan dengan teori jawaban butir soal adalah seperti yang dijelaskan Hambleton, Swaminathan & Rogers (1991), sebagai berikut: “adanya sifat group dependent dan item dependent, juga indeks daya pembeda, koefisien validitas, koefisien reliabilitas skor tes yang keseluruhannya tergantung kepada peserta tes yang mengerjakan tes tersebut. Group dependent artinya hasilnya pengukuran tergantung pada kemampuan peserta yang mengerjakan tes. Jika tes diujikan kepada

kelompok peserta dengan kemampuan tinggi, tingkat kesulitan butir item akan rendah. Sebaliknya jika tes diujikan kepada kelompok peserta dengan kemampuan rendah, tingkat kesulitan butir soal akan tinggi. Item dependent artinya hasil pengukuran tergantung pada tes mana yang diujikan. Jika tes yang diujikan mempunyai tingkat kesulitan tinggi, estimasi kemampuan peserta tes akan rendah. Sebaliknya, jika tes yang diujikan mempunyai tingkat kesulitan rendah, estimasi kemampuan peserta tes akan tinggi” (Nursakinah Oktaviana Sasmita, 2018).

KESIMPULAN

Dengan merujuk pada hasil penelitian dan pembahasan, maka dapat ditarik kesimpulan sebagai berikut: 1. Model dan teknik pembuatan Instrumen penilaian kelas yang digunakan para guru agama Islam MI Sunan Kalijaga Malang sudah sesuai dengan aturan teoritis yang ada. Hal ini dikarenakan para guru agama Islam MI Sunan Kalijaga Malang selalu menganalisis kualitas soalnya baik secara kualitatif maupun kuantitatif terhadap soal yang mereka buat. 2. Masih terdapatnya soal yang dibuat oleh guru agama Islam MI Sunan Kalijaga Malang itu perlu direvisi dan ada yang perlu dibuang. Namun kalau dilihat dari reliabilitas soal mempunyai Koefisien Alpha: 0,831 itu mata pelajaran SKI, 0,828 untuk mata pelajaran Aqidah Akhlak, 0,849 untuk mata pelajaran Feqh dan 0,710 untuk mata pelajaran al-Qur`an Hadits, maka soal-soal yang dibuat oleh guru agama Islam bisa digolongkan memiliki reliabilitas yang tinggi. Oleh karena itu ke depannya masih perlu digiatkan lagi pelatihan-pelatihan kepada guru agama Islam guna pembuatan instrument penilaian kelas yang baik.

DAFTAR PUSTAKA

- Amrullah. (2018). Analisis Instrumen Tes Soal Pilihan Ganda Ujian Sekolah (Us) Mata Pelajaran Al Islam Kelas Xii Sma Muhammadiyah Se-Kota Palembang. *Ad-Man-Pend (Jurnal Administrasi Manajemen Pendidikan)*, 1, 11–20. <https://doi.org/https://doi.org/10.32502/Amp.V1i1.956>
- Badrun Kartawagiran. (2008). *Item Dan Tes Analisis*. Diktat Short Course Pgsd/Mi Uny.
- Dea Mustika, Ambiyar, I. A. (2021). Proses Penilaian Hasil Belajar Kurikulum 2013 Di Sekolah Dasar. *Jurnal Basicedu*, 5(6), 6158–6167. <https://doi.org/https://doi.org/10.31004/basicedu.v5i6.1819>
- Dinni, H. N. (2018). Hots (High Order Thinking Skills) Dan Kaitannya Dengan Kemampuan Literasi Matematika. *Prisma*, 1, 170–176.
- Fitria. (2013). Evaluasi Penerapan Kurikulum 2013 Pada Mata Pelajaran Aqidah Akhlak Di Madrasah Tsanawiyah Negeri 1 Kolaka Kabupaten Kolaka. *Journal Of Chemical Information And Modeling*, 53(9), 1689–1699. <https://doi.org/10.5281/zenodo.4398943>
- Hartini, P., Setiadi, H., & Ernawati. (2021). Instrumen Penilaian Berbasis Lots Dan Hots Buatan Guru Kelas Vi. *Jurnal Penelitian Dan Penilaian Pendidikan (Jppp)*, 3(1), 14–24.
- Hasnia, H., Said, M., & Nursaeni, N. (2021). Analisis Teknik Evaluasi Guru Pada Pembelajaran Agama Islam Di Madrasah Aliyah. *Iqro: Journal Of Islamic ...*, 4(1), 31–40. <http://ejournal.iainpalopo.ac.id/index.php/iqro/article/view/1851>
- Ina Magdalena, Fatikah Mulyani, Dwi Nur Faridah, Nuri Fitriyani, A. H. D. (2020). Analisis Sistem Penilaian Kurikulum 2013 Di Sdn Bencongan 01. *Jurnal Edukasi Dan Sains*, 2(3), 333–341.
- Kadir, A. (2015). Menyusun Dan Menganalisis Tes Hasil Belajar. *Jurnal Al-Ta'dib*, 8(2), 70–81.
- Nabilah, M., Sitompul, S. S., & Hamdani, H. (2020). Analisis Kemampuan Kognitif Peserta Didik Dalam Menyelesaikan Soal Momentum Dan Impuls. *Jurnal Inovasi Penelitian Dan Pembelajaran Fisika*, 1(1), 1. <https://doi.org/10.26418/jippf.v1i1.41876>
- Nursakinah Oktaviana Sasmita. (2018). Dampak Dari Multidimensionalitas Butir Soal Terhadap Estimasi

3085 *Analisis Instrumen Penilaian Kelas Domain Kognitif Guru Mata Pelajaran Agama Islam di Madrasah Ibtidaiyah – Ahmad Sholeh*
DOI : <https://doi.org/10.31004/basicedu.v6i2.2129>

True Score Dengan Pendekatan Model Bifaktor. *Jurnal Pengukuran Psikologi Dan Pendidikan Indonesia (Jp3i)*, 4(4). <https://doi.org/10.15408/Jp3i.V4i4.9300>

Nuswowati, M., Binadja, A., Efti, K., & Ifada, N. (2011). Pengaruh Validitas Dan Reliabilitas Butir Soal Ulangan Akhir Semester Bidang Studi Kimia Terhadap Pencapaian Kompetensi. *Jurnal Inovasi Pendidikan Kimia*, 4(1), 566–573.

Sawania, Arsy, R. F., & Amiruddin. (2016). Pelaksanaan Evaluasi Formatif Dalam Mengukur Hasil Belajar Mata Pelajaran Ips Di Kelas V Sdk Tsm Posona. *Jurnal Kreatif Tadulako Online*, 5(4), 109–119.

Siti Mariyah. (2019). Peningkatan Kompetensi Guru Dalam Menyusun Instrumen Penilaian Hasil Belajar Melalui Supervisi Akademik Teknik Kelompok. *Ideguru: Jurnal Karya Ilmiah Guru*, 4(1), 1–9. <https://doi.org/10.51169/ideguru.V4i1.62>

Suharsimi Arikunto. (2013). *Dasar-Dasar Evaluasi Pendidikan*. Pt. Bumi Aksara.

Supriyadi, E. (2020). Penyelenggaraan Penilaian Hasil Belajar Smk Rujukan Program Keahlian Teknik Ketenagalistrikan Di Daerah Istimewa Yogyakarta. *Jurnal Edukasi Elektro*, 3(2), 70–80. <https://doi.org/10.21831/Jee.V3i2.29119>

Tambakmulyo, N., Kebumen, K., Ambarwati, D., Wiarsih, C., & Irawan, D. (2021). Strategi Guru Dalam Penyelenggaraan Pembelajaran Pada Masa Pandemi Covid-19 Di Sd. *Prosiding Seminar Nasional*, 9(1), 215–224. <https://jurnal.uhnp.ac.id/psn-uhnp/article/view/126>

Wahyuningsih, P. G., & Sriyanto, S. (2021). Evaluasi Guru Pendidikan Agama Islam Dalam Pembelajaran Daring. ... : *Jurnal Studi Islam*, 2(2), 159–169. <http://jurnalnasional.ump.ac.id/index.php/alhamra/article/view/10940>