

KOMPETENSI GURU MIN SABANG DALAM MELAKSANAKAN EVALUASI PEMBELAJARAN; Telaah atas Konstruksi Instrumen Penilaian Berbasis Kurikulum 2013

Realita

UIN Ar-Raniry Banda Aceh
realitanurdin@gmail.com

Abstract

This study aims to obtain information about the competence of MIN Sabang teachers in compiling authentic assessment instruments, and to find forms of difficulties faced by teachers in developing cognitive, affective, and psychomotor instruments. This is a qualitative research with descriptive approach method with document study questionnaire, and interview data collectio technique. Documentation data were analyzed in accordance with the rules for preparing questions, and questionnaire was analyzed quantitatively using percentage formula. The interview data were analyzed thematically. Based on the distribution of questionnaire data to 27 MIN Sabang teachers, document study and interview with 6 teachers of the school, the findings shwoed that MIN Sabang teachers had good competence in developing cognitive assessment instruments, but were less competent in developing affective and psychomotor assessment instruments. The perceived difficulties in preparing assessment instruments are in developing indicators of basic competence, determining operational verbs, compiling communicative languages, arranging aspects that are assessed, adjusting between indicators and question instruments/aspects of assessment, and determining options for answers to the question of multiple choices and matching items.

Keywords: *authentic assessment instruments*

Abstrak

Penelitian ini bertujuan untuk mendapatkan informasi mengenai kompetensi guru MIN Sabang dalam menyusun instrumen penilaian autentik, dan untuk menemukan bentuk-bentuk kesulitan yang dihadapi guru dalam menyusun instrumen kognitif, afektif, dan psikomotor. Penelitian ini merupakan penelitian kualitatif. Metode penelitian yang digunakan adalah metode deskriptif dengan teknik studi dokumen, angket, dan wawancara. Data hasil dokumentasi dianalisis sesuai dengan kaidah penyusunan soal, dan angket dianalisis secara kuantitatif dengan rumus persentase. Adapun data hasil wawancara dianalisis secara tematik. Berdasarkan sebaran data angket terhadap 27 guru MIN Sabang dan studi dokumen serta wawancara dengan 6 orang guru di sekolah tersebut, ditemukan bahwa guru-guru MIN Sabang memiliki kompetensi yang baik dalam menyusun instrumen penilaian kognitif, namun kurang berkompeten dalam menyusun instrumen penilaian afektif dan psikomotor. Adapun kesulitan yang dirasakan dalam menyusun instrumen penilaian adalah saat mengembangkan indikator dari

kompetensi dasar, menentukan kata kerja operasional, menyusun bahasa yang komunikatif, menyusun aspek-aspek yang dinilai, menyesuaikan antara indikator dan instrumen soal/aspek-aspek penilaian, dan menentukan pilihan jawaban pada soal bentuk pilihan ganda dan menjodohkan.

Kata Kunci: *instrumen penilaian autentik*

PENDAHULUAN

Evaluasi merupakan bagian integral yang tak terpisahkan dari proses pembelajaran, sehingga guru tidak hanya menilai hasil belajar siswa melalui tes, tetapi juga menilai proses pembelajaran melalui pengamatan, dan teknik-teknik penilaian lainnya. Kemampuan guru menyelenggarakan evaluasi proses dan hasil belajar, dan memanfaatkannya untuk kepentingan pembelajaran merupakan salah satu bagian dari kompetensi pedagogik¹ yang harus dimiliki oleh guru. Hal tersebut sesuai dengan Permendiknas No. 16 Tahun 2007, yang mengungkapkan bahwa standar kualifikasi akademik dan standar kompetensi guru terdiri dari kompetensi pedagogik, kepribadian, sosial dan profesional. Kompetensi pedagogik yang harus dikuasai guru mata pelajaran yang berhubungan dengan pengelolaan penilaian hasil belajar adalah melaksanakan penilaian dan evaluasi proses dan hasil belajar, dan memanfaatkan hasil penilaian dan evaluasi tersebut untuk kepentingan pembelajaran, serta melakukan tindakan reflektif untuk peningkatan kualitas pembelajaran.²

Menurut Abdul Majid, parameter untuk menilai kompetensi guru dalam penyelenggaraan kegiatan evaluasi yaitu kompetensi guru dalam menyusun alat penilaian yang sesuai dengan tujuan pembelajaran, melakukan penilaian dengan berbagai teknik dan jenis penilaian, mengumumkan hasil-hasil serta implikasinya kepada peserta didik, menganalisis hasil penilaian untuk mengidentifikasi topik/kompetensi dasar yang sulit dalam rangka remedial dan pengayaan, memanfaatkan masukan dari peserta didik dan merefleksikannya untuk pembelajaran selanjutnya, membuktikannya melalui catatan, jurnal pembelajaran, dan lain-lain, serta memanfaatkan hasil penilaian sebagai bahan penyusunan rancangan pembelajaran yang akan dilaksanakan selanjutnya.³

Berkaitan dengan penyelenggaraan evaluasi dalam kurikulum 2013, guru dituntut harus dapat melaksanakan penilaian autentik, yaitu suatu penilaian yang tidak hanya difokuskan

¹Kompetensi pedagogik meliputi pemahaman terhadap peserta didik, perencanaan dan pelaksanaan kegiatan pembelajaran, evaluasi hasil dan proses belajar, dan pengembangan peserta didik untuk mengaktualisasikan berbagai potensi peserta didik. Rusman, *Model-Model Pembelajaran; Mengembangkan Profesionalisme Guru*, Jakarta: Rajawali Pers, 2011, h. 54.

²Depdiknas, "Standar Kualifikasi Akademik dan Kompetensi Guru. Salinan Lampiran Menteri Pendidikan Nasional No. 16 Tahun 2007 tanggal 4 Mei 2007" [Online] <http://vervalsp.data.kemdikbud.go.id/prosespembelajaran/file/Permendiknas%20No%2016%20Tahun%202007.pdf> diakses tanggal 27 Juni 2019.

³Abdul Majid, *Penilaian Autentik; Proses dan Hasil Belajar*, Bandung: Remaja Rosdakarya, 2014, h. 3.

pada aspek kognitif semata, tetapi dilakukan secara menyeluruh, meliputi penilaian sikap, pengetahuan, dan psikomotor,⁴ karena dalam kurikulum 2013 ketiga kompetensi tersebut menjadi sasaran pembelajaran dalam standar kompetensi (SK) atau kompetensi inti (KI) dan kompetensi dasar (KD). Penilaian autentik dilakukan secara komprehensif ditujukan untuk menilai *input*, proses, dan *output* pembelajaran.⁵

Melalui penilaian autentik, guru dapat mengetahui kondisi siswa yang sebenarnya. Guru dapat mengetahui apakah siswa benar-benar belajar atau tidak, apakah pengalaman belajar siswa berpengaruh positif terhadap perkembangan siswa, baik intelektual maupun mental.⁶ Dengan makna lain, dengan berbagai teknik penilaian yang digunakan seperti tes, observasi, kinerja, portofolio, penilaian teman sejawat, dan lain-lain akan dapat merekam seluruh kecerdasan siswa yang meliputi ranah kognitif, afektif, psikomotor. Oleh karena itu, guru harus memiliki kompetensi yang baik dalam melaksanakan penilaian proses dan hasil, termasuk kompetensi merancang instrumen penilaian, baik dengan teknik tes maupun non tes. Hal ini sangat terkait dengan tugas guru sebagai unsur penting dalam pengembangan instrumen penilaian dan sekaligus sebagai pelaksana.

Instrumen penilaian kognitif dikembangkan berdasarkan (1) karakteristik mata pelajaran, (2) capaian dalam kurikulum, (3) konstruksi soal; yaitu kejelasan dan ketegasan rumusan pertanyaan, (4) materi; kesesuaian soal dengan standar kompetensi, kompetensi dasar, dan indikator, dan (5) bahasa soal yang tidak menimbulkan penafsiran ganda.⁷

Lebih detail, kriteria yang digunakan untuk menilai kualitas butir soal bentuk uraian adalah sebagai berikut:

- a. *Materi*; soal sesuai dengan indikator, materi sesuai dengan tujuan pengukuran, dan tingkatan mata pelajaran.
- b. *Konstruksi*; menggunakan kata untuk bertanya atau kata perintah yang menuntut jawaban uraian, ada petunjuk yang jelas tentang cara mengerjakan soal, dan ada pedoman penskorannya.
- c. *Bahasa/Budaya*; rumusan kalimat soal komunikatif, butir soal menggunakan bahasa Indonesia yang baku, butir soal tidak menggunakan kata atau ungkapan yang dapat menimbulkan penafsiran ganda atau salah pengertian, tidak menggunakan bahasa yang berlaku setempat/tabu, rumusan soal tidak mengandung kata atau ungkapan yang dapat menyinggung perasaan siswa.⁸

⁴Kunandar, *Penilaian Autentik Penilaian Hasil Belajar Peserta Didik Berdasarkan Kurikulum 2013; Suatu Pendekatan Praktis Disertai dengan Contoh*, Jakarta: RajaGrafindo Persada, 2014, h. 36.

⁵Ahmad Yani, *Mindset Kurikulum 2013*, Bandung: Alfabeta, 2014, h. 144.

⁶M. Hosnan, *Pendekatan Saintifik dan Kontektual dalam Pembelajaran Abad 21; Kunci Sukses Implementasi Kurikulum 2013*, Bogor: Ghalia Indonesia, 2014, h. 387.

⁷Abdul Majid, *Penilaian Autentik...*, h. 190.

⁸Supardi, *Penilaian Autentik Pembelajaran Afektif, Kognitif dan Psikomotor; Konsep dan Aplikasi*, Jakarta: Rajawali Pers, 2016, hal. 82-83.

Selanjutnya, kriteria atau aspek-aspek penilaian untuk soal bentuk pilihan ganda adalah sebagai berikut:

- a. *Materi*; soal sesuai dengan indikator, batasan pertanyaan dan jawaban jelas, materi yang ditanyakan sesuai dengan tujuan pengukuran, jenjang sekolah dan tingkat kelas.
- b. *Konstruksi*; pokok soal (*stem*) dirumuskan dengan bahasa singkat, jelas dan tegas, rumusan soal dan pilihan jawaban adalah pernyataan yang diperlukan saja, pokok soal bebas dari pernyataan yang bersifat negatif ganda, pilihan jawaban pada soal bersifat homogen dan logis ditinjau dari sisi materi, panjang pilihan jawaban relatif sama, opsi jawaban tidak menggunakan pernyataan “semua salah atau semua benar” dan sejenisnya, opsi jawaban yang berbentuk angka atau waktu diurutkan berdasarkan besar-kecilnya angka/waktu atau kronologisnya, soal tidak bergantung pada jawaban soal yang pada nomor sebelumnya, hanya ada satu kunci jawaban, dan bila menggunakan kata negatif maka harus digarisbawahi atau dicetak lain.
- c. *Bahasa/Budaya*, soal menggunakan bahasa Indonesia yang baik dan benar, serta komunikatif, Tidak menggunakan bahasa yang berlaku setempat/tabu, opsi jawaban tidak mengulang kata atau kelompok kata yang sama, kecuali jika merupakan satu kesatuan pengertian.⁹

Adapun kriteria atau aspek-aspek penilaian untuk soal bentuk isian singkat adalah sebagai berikut:

- a. *Materi*; soal sesuai dengan indikator, materi yang ditanyakan sesuai dengan tujuan pengukuran, jenjang sekolah dan tingkat kelas.
- b. *Konstruksi*; pernyataan disusun dengan bentuk pertanyaan langsung, dan menuntut jawaban singkat (berupa sebuah kata, angka, simbol, atau kelompok kata), tidak menggunakan kata-kata yang langsung dikutip dari buku, pertanyaan hanya ada satu jawaban benar, tempat jawaban yang dikosongkan sama panjangnya.
- c. *Bahasa/Budaya*; bahasa soal komunikatif, menggunakan bahasa Indonesia baku, dan tidak menggunakan bahasa setempat/tabu.¹⁰

Selanjutnya, kurikulum 2013 membagi kompetensi afektif menjadi dua, yaitu sikap spiritual dan sikap sosial. Sikap spiritual terkait dengan pembentukan subjek didik yang beriman dan bertakwa, dan sikap sosial terkait dengan pembentukan subjek didik yang berakhlak mulia, mandiri, demokratis, bertanggung jawab, dan lain-lain. Kompetensi sikap spiritual ada pada KI-1; yaitu menghargai dan menghayati ajaran agama yang dianutnya, sedangkan kompetensi sikap sosial ada pada KI-2: menghargai dan menghayati perilaku jujur, disiplin, tanggung jawab, peduli, santun, percaya diri, dan lain-lain.¹¹

⁹Sunarti dan Selly Rahmawati, *Penilaian dalam Kurikulum 2013; Membantu Guru dan Calon Guru Mengetahui Langkah-langkah Penilaian Pembelajaran*, Yogyakarta: Andi, 2014, hal. 36-37.

¹⁰Safari, *Teknik Analisis Butir Soal Instrumen Tes dan Non-Tes*, Jakarta: Asosiasi Pengawas Sekolah Indonesia Departemen Pendidikan Nasional, 2005, hal. 8.

¹¹Abdul Majid, *Penilaian Autentik; Proses...*, h. 164-165.

Kompetensi guru dalam merancang instrumen penilaian dapat terlihat dari naskah soal yang diujikan ke siswa pada pelaksanaan ujian tengah semester (UTS) atau ujian akhir semester (UAS). Selain itu, melalui Rencana Pelaksanaan Pembelajaran (RPP) yang disusun guru, kompetensi tersebut juga dapat diidentifikasi, karena salah satu komponen yang turut didesain dalam RPP adalah komponen penilaian. Namun, berdasarkan hasil wawancara dengan salah seorang guru Madrasah Ibtidaiyah Negeri (MIN) Sabang, terungkap bahwa soal-soal yang disusun oleh guru-guru selama ini kebanyakan diambil dari buku paket atau internet, dan RPP yang disusun oleh sebagian guru selama ini ternyata bukan karyanya sendiri, melainkan *copy paste* dari RPP orang lain. Selain itu, instrumen penilaian yang dilaksanakan di sekolah selama ini cenderung berupa tes. Walaupun penilaian afektif dan psikomotor turut serta dalam penilaian, namun teknik penilaiannya dilaksanakan tanpa disertai instrumen.¹²

Fenomena ini sesuai dengan beberapa penelitian terdahulu, seperti yang dilakukan oleh Misbahul Jannah,¹³ Rohmad Qomari¹⁴ dan Dasrieny Pratiwi dan Hening Widowati.¹⁵ Hasil penelitian Misbahul Jannah mengungkapkan rendahnya pemahaman mahasiswa PGMI UIN Ar-Raniry dalam Evaluasi Pembelajaran IPA. Penelitian Rohmad Qomari mengungkapkan fakta bahwa penyelenggaraan evaluasi belum terlaksana secara komprehensif, karena masih didominasi pada evaluasi ranah kognitif. Penelitian yang dilakukan oleh Dasrieny Pratiwi dan Hening Widowati juga menemukan bahwa mahasiswa calon guru biologi FKIP Universitas Muhammadiyah Malang memiliki kompetensi yang jauh lebih baik dalam menyusun instrumen penilaian kognitif daripada penilaian afektif dan psikomotor.

Berdasarkan landasan teoretis dan empiris di atas, penulis tertarik untuk melakukan penelitian mengenai “Kompetensi Guru MIN Sabang dalam Evaluasi Pembelajaran; Telaah Atas Konstruksi Instrumen Berbasis Kurikulum 2013” dengan tujuan untuk mengukur kompetensi guru dan kendala yang dihadapi guru dalam menyusun instrumen penilaian. Jika hasil penelitian menunjukkan rendahnya kompetensi guru, maka penelitian ini berkepentingan sebagai upaya kuratif, yaitu mengatasi kesulitan yang dialami oleh guru-guru madrasah ibtidaiyah dalam menyusun instrumen penilaian, seperti memberikan solusi dan rekomendasi kepada *stakeholder* yang terkait dengan pembinaan dan peningkatan kualifikasi guru. Namun jika hasil penelitian menunjukkan tingginya tingkat kompetensi guru madrasah ibtidaiyah dalam menyusun instrumen penilaian, maka penelitian ini diharapkan dapat menjadi *blue print* bagi guru-guru yang berada di kawasan Sabang khususnya, dan wilayah Provinsi Aceh pada umumnya.

¹²Wawancara dengan salah seorang guru MIN Sabang tahun 2018.

Jannah, Misbahul, “Kemampuan Mahasiswa PGMI UIN Ar-Raniry dalam Evaluasi Pembelajaran IPA”, *Proceeding*, Banda Aceh: ARICIS I, 2016, hal. 116.

¹⁴Rohmad Qomari, “Pengembangan Instrumen Evaluasi Afektif Domain” *Jurnal pemikiran Alternatif Pendidikan*, Vol. 3, No. 1, hal. 3. file:///C:/Users/USER/Downloads/287-Article%20Text-553-1-10-20150304.pdf

¹⁵Dasrieny Pratiwi dan Hening Widowati, “Deskripsi Kemampuan Calon Guru Biologi dalam Menyusun Perangkat Penilaian Berbasis Kurikulum 2013,” *Proceeding Seminar Nasional Pendidikan Biologi*, Malang: Universitas Muhammadiyah, 2015, hal. 190.

<http://biology.umm.ac.id/files/file/521-529%20Hening%20Widowati.pdf>. Diakses tanggal 15 Mei 2018.

Berdasarkan jenisnya, penelitian ini tergolong dalam penelitian kualitatif, yaitu penelitian yang bertujuan untuk mendeskripsikan dan menganalisis suatu fenomena atau peristiwa atau aktivitas sosial, meliputi sikap, kepercayaan, persepsi dan pemikiran yang dilakukan baik secara individu maupun kelompok.¹⁶ Metode yang digunakan dalam penelitian ini adalah metode deskriptif, yang bertujuan untuk mendeskripsikan suatu subjek penelitian berdasarkan data dari variable kelompok subjek yang diteliti dan tidak dimaksudkan untuk menguji hipotesis.¹⁷ Hal ini dikarenakan tujuan penelitian adalah untuk mendapatkan gambaran apa adanya mengenai kompetensi guru dalam menyusun penilaian autentik. Penelitian dilakukan di MIN Sabang, Kecamatan Suka Jaya Sabang Provinsi Aceh. Responden yang terlibat dalam penelitian adalah seluruh guru MIN Sabang yang peneliti jumpai saat di lokasi penelitian, yang berjumlah 27 orang guru.

Instrumen penelitian berupa studi dokumen terhadap RPP dan soal-soal Ujian Akhir Semester (UAS) buatan guru, wawancara dan angket. Studi dokumen bertujuan untuk mendapatkan data dan informasi mengenai kompetensi guru dalam membuat instrumen penilaian autentik. Instrumen wawancara untuk memperoleh data yang lebih rinci mengenai kompetensi guru dalam membuat instrumen penilaian autentik serta bentuk-bentuk kesulitan yang dihadapi dalam penyusunannya. Adapun angket untuk memperoleh informasi mengenai bentuk-bentuk kesulitan yang dihadapi guru dalam menyusun instrumen penilaian. Studi dokumen dan wawancara dilakukan terhadap 5 (lima) responden. Angket disusun berbentuk skala, dan terdiri atas 20 item. Analisis data dilakukan secara kualitatif. Data hasil dokumen akan dianalisis secara konten sesuai dengan kaidah penyusunan soal. Data hasil angket dianalisis dengan persentase. Adapun data hasil wawancara dianalisis secara tematik analisis dengan menempuh beberapa tahapan, yaitu tahap data *reduction*, tahap data *display* dan tahap data *conclusion drawing/verification*.

PEMBAHASAN

Kompetensi yang dimaksud dalam penelitian ini adalah kecakapan guru MIN Sabang dalam menyusun instrumen penilaian kognitif, afektif, dan psikomotor. Kompetensi guru MIN Sabang dalam menyusun instrumen penilaian autentik dapat diidentifikasi berdasarkan hasil analisis Rencana Pelaksanaan Pembelajaran (RPP) dan butir-butir soal Ujian Akhir Semester (UAS) buatan guru. Analisis terhadap naskah RPP dilakukan dengan tujuan untuk melihat pengembangan indikator dan kesesuaian soal dengan indikator yang telah dirumuskan dalam RPP.

1. Kompetensi Guru MIN Sabang dalam Menyusun Instrumen Penilaian Kognitif

Kompetensi guru dalam menyusun instrumen penilaian kognitif diamati melalui empat kriteria, yaitu tinjauan materi, konstruksi soal, bahasa, dan tingkatan berpikir. Untuk

¹⁶Suharsimi Arikunto, *Prosedur Penelitian: Suatu Pendekatan dan Praktek*, Jakarta: Rineka Cipta, 1998, hal. 245.

¹⁷Saifuddin Azwar, *Metode Penelitian*, Yogyakarta: Pustaka Pelajar, 2007, hal. 126.

melihat kompetensi guru MIN Sabang dalam menyusun instrumen kognitif, peneliti merangkumkan data ke dalam table berikut ini:

Tabel 2. Deskripsi Kompetensi Guru MIN Sabang dalam menyusun Soal Ujian Akhir Semester

Responden (R)	Aspek yang Dinilai	Deskripsi/Uraian
R-1 (Kelas I)	Materi	Secara umum, soal yang disusun sesuai dengan indikator, materi yang ditanyakan sudah sesuai dengan tujuan pengukuran, dan isi materi yang ditanyakan sesuai dengan tingkatan kelas I. Namun, sebagian kecil kata kerja operasional (KKO) dalam indikator perlu direvisi agar relevan dengan tuntutan kompetensi dasar.
	Konstruksi	Soal berbentuk isian singkat dan Benar-Salah. Secara umum, konstruksi soal masih memerlukan perbaikan. Sebagian soal, baik soal bentuk isian singkat maupun benar-salah tidak memiliki petunjuk soal. Soal-soal yang disusun belum sesuai dengan kaidah penyusunan soal bentuk isian singkat, seperti: memuat 2 atau 3 tempat jawaban yang kosong, jawaban yang dituntut tidak memuat satu jawaban melainkan beberapa jawaban. Demikian pula benar-salah (dalam hal ini guru menggunakan istilah Boleh/Tidak). Penempatan kata-kata “Boleh/Tidak” menyerupai soal bentuk isian singkat sehingga dapat membuat <i>testee</i> ambigu.
	Bahasa	Sebagian kecil rumusan kalimat soal tidak komunikatif, sehingga dapat menimbulkan kesalahpahaman pada <i>testee</i> . Bahkan, pada soal bentuk isian singkat, petunjuk soal bersatu dengan pertanyaan. Selain itu, beberapa soal belum menggunakan bahasa Indonesia yang baku atau sesuai dengan kaidah bahasa Indonesia, seperti spasi, penulisan tanda titik, koma dan lain-lain.
	Level berpikir	Secara garis besar, soal-soal pada mapel yang diujikan memuat tingkatan berpikir LOTS (<i>Low Order Thinking Skill</i>) yaitu pada tataran <i>cognitif</i> mengingat (C-1). Dari 20 soal UAS, hanya 1 soal yang bermuatan MOTS (<i>Middle Order Thinking Skill</i>), dan tidak ada soal-soal HOTS (<i>Higher Order Thinking Skill</i> .)
R-2 Kelas V	Materi	Soal-soal yang disusun pada umumnya sudah sesuai dengan indikator. Namun, sebagian soal tidak memiliki indikator, karena sebagian indikator yang disusun terlalu umum dan kurang sesuai dengan tuntutan Kompetensi Dasar (KD). Beberapa kata kerja operasional perlu direvisi agar sesuai dengan tuntutan KD. Isi materi yang ditanyakan sesuai dengan tingkatan kelas V.
	Konstruksi	Soal berbentuk pilihan ganda dan uraian. Secara umum, konstruksi soal sudah baik. Namun, pilihan jawaban pada beberapa soal bentuk pilihan ganda belum homogen dan logis.
	Bahasa	Sebagian kecil rumusan soal belum menggunakan bahasa Indonesia yang baku atau sesuai dengan kaidah bahasa Indonesia, seperti

		spasi, penulisan huruf kapital, tanda titik, koma, peletakan imbuhan “di” pada kata kerja, penggunaan simbol tanda tanya, dan lain-lain.
	Level berpikir	Secara garis besar, soal didominasi oleh tingkatan berpikir LOTS (L-1), namun ditemukan satu soal yang berlevel MOTS (L-2) dan satu soal berlevel HOTS (L-3)
R-3 Kelas III	Materi	Secara umum, soal yang disusun sesuai dengan indikator, materi yang ditanyakan sudah sesuai dengan tujuan pengukuran, dan isi materi yang ditanyakan sesuai dengan tingkatan kelas III.
	Konstruksi	Soal berbentuk pilihan ganda, isian singkat dan uraian. Secara umum, konstruksi soal sudah baik. Pilihan jawaban pada beberapa soal bentuk pilihan ganda sudah homogen, namun ada beberapa pilihan yang tidak logis. Pilihan jawaban yang berbentuk angka tidak disusun berdasarkan urutan besar kecilnya angka. Pada soal bentuk isian singkat, jawaban mengandung lebih dari satu jawaban karena soal terlalu umum. Sebagian soal isian singkat memiliki beberapa tempat jawaban kosong.
	Bahasa	Secara umum, bahasa soal yang digunakan sudah baik dan komunikatif, hanya sebagian kecil yang belum sesuai dengan aturan EYD.
	Level berpikir	Secara garis besar, soal didominasi oleh tingkatan berpikir LOTS (L-1), namun beberapa soal sudah berada pada level MOTS dan HOTS (L-3)
R-4 Kelas IV	Materi	Secara umum, soal disusun sesuai dengan indikator, materi yang ditanyakan sudah sesuai dengan tujuan pengukuran, dan isi materi yang ditanyakan sesuai dengan tingkatan kelas IV. Namun, indikator yang dirumuskan masih bersifat umum, sehingga diperlukan pengembangan indikator. Selain itu, sebagian kecil kata kerja operasional (KKO) dalam indikator perlu direvisi agar relevan dengan tuntutan kompetensi dasar.
	Konstruksi	Soal berbentuk pilihan ganda dan uraian. Soal-soal yang disusun, baik soal bentuk pilihan ganda maupun uraian memiliki petunjuk soal yang jelas, dan sesuai dengan kaidah penyusunan soal.
	Bahasa	Bahasa soal jelas dan komunikatif. Namun, beberapa konstruksi soal masih memerlukan perbaikan dari segi EYD (Ejaan yang Disempurnakan), seperti huruf kapital dan penulisan “di” sebagai imbuhan.
	Level berpikir	Secara garis besar, soal-soal pada mapel yang diujikan memuat tingkatan berpikir LOTS, yaitu pada tataran <i>cognitif</i> mengingat (C-1). Beberapa soal baik bentuk pilihan ganda, maupun uraian memiliki level MOTS dan juga HOTS.
R-5 Kelas VI	Materi	Semua soal yang disusun sesuai dengan indikator, materi yang ditanyakan sudah sesuai dengan tujuan pengukuran, dan isi materi yang ditanyakan sesuai dengan tingkatan kelas VI. Namun, indikator yang dirumuskan masih bersifat umum, sehingga diperlukan pengembangan indikator.

	Konstruksi	Soal berbentuk pilihan ganda. Soal-soal yang disusun, baik soal bentuk pilihan ganda maupun uraian memiliki petunjuk soal yang jelas, dan pada umumnya sesuai dengan kaidah penyusunan soal. Namun, ditemui satu soal yang belum tepat konstruksinya, yaitu opsi (pilihan) yang belum logis pada salah satu soal bentuk pilihan ganda. Selain itu, tanda titik-titik pada soal isian singkat tidak beraturan dan tidak konsisten, yang mana sebagian soal memiliki tanda titik-titik yang sangat panjang, sebagian panjang, dan sebagian tidak panjang.
	Bahasa	Bahasa soal jelas dan komunikatif. Namun, beberapa konstruksi soal masih memerlukan perbaikan dari segi EYD (Ejaan yang Disempurnakan), seperti penulisan “di” sebagai imbuhan.
	Level berpikir	Soal-soal pada mapel yang diujikan didominasi oleh soal-soal tingkatan berpikir LOTS, yaitu pada tataran <i>cognitif</i> mengingat (C-1). Namun demikian, banyak pula soal-soal yang berlevel MOTS.
R-6 Kelas II	Materi	Soal-soal yang disusun sudah sesuai dengan materi kelas 2 dan juga indikator, namun beberapa indikator masih umum, tidak spesifik, dan memerlukan pengembangan agar lebih mencerminkan karakteristik KD.
	Konstruksi	Soal berbentuk isian singkat. Sebagian konstruksi soal isian singkat belum tepat. Dalam penyusunannya, jawaban mengandung dua atau lebih jawaban. Hal ini disebabkan redaksi bahasa soal yang sangat umum sehingga siswa berpeluang menjawab pertanyaan dari berbagai sudut pandang, menurut pikirannya dan logis. Salah satu pertanyaan menuntut adanya stimulus berupa gambar, namun gambar tersebut tidak tercantum dalam soal.
	Bahasa	Secara umum, rumusan kalimat soal sudah komunikatif dan sesuai dengan kaidah bahasa Indonesia yang benar. Namun, pada beberapa kalimat terdapat pengetikan yang salah, seperti tidak adanya spasi setelah tanda baca, dan kesalahan pengetikan pada beberapa kalimat.
	Level berpikir	Secara garis besar, soal isian singkat didominasi oleh tingkatan berpikir LOTS (L-1), dan sebagian kecil memiliki tingkatan berpikir penerapan (MOTS).

Berdasarkan tabel 2 di atas, kompetensi guru dalam menyusun instrumen penilaian kognitif dapat dikatakan baik. Pada aspek materi, mulai dari R-1, sampai R-6 memiliki kompetensi yang baik dalam membuat soal yang relevan dengan indikator dan materi, serta sesuai dengan tingkatan kelas. Hanya beberapa soal yang tidak ada indikatornya seperti yang ditunjukkan oleh data dokumen Responden 2. Sebagian indikator yang dijabarkan dari KD masih sangat umum (belum spesifik) seperti yang diamati pada naskah RPP dan soal UAS buatan R-1, R-2, R4, R-5, dan R-6. Dalam penyusunan kata kerja operasional (KKO), sebagian kecil KKO yang tercantum dalam RPP guru R-1, R-2, dan R-4 masih belum sesuai dengan tuntutan KD.

Ditinjau dari aspek konstruksi soal UAS, terlihat R-1, R-3, R-5, dan R-6 memiliki keterbatasan pemahaman mengenai kaidah/aturan menyusun soal isian singkat, B-S, dan soal bentuk pilihan ganda. Pada sebagian kecil soal isian singkat buatan R-1, R-5, dan R-6, memuat 2 atau 3 tempat jawaban yang kosong, jawaban yang dituntut tidak memuat satu jawaban melainkan beberapa jawaban. Pada soal buatan guru R-5, tanda titik-titik pada soal isian singkat tidak beraturan dan tidak konsisten. Demikian juga pada soal bentuk pilihan ganda buatan guru R-3, dan R-5. Pilihan jawaban pada beberapa soal pilihan ganda sudah homogen, namun ada beberapa pilihan yang tidak logis. Pilihan jawaban yang berbentuk angka tidak disusun berdasarkan urutan besar kecilnya angka.

Pada aspek bahasa, secara umum redaksi soal sudah baik dan komunikatif, hanya sebagian kecil soal buatan guru R-1 yang belum komunikatif sehingga dapat menimbulkan ambigue dan kesalahpahaman pada peserta tes. Berdasarkan naskah soal UAS, juga dapat dianalisis bahwa para guru (R-1 sampai dengan R-6) memiliki pemahaman yang serupa mengenai penulisan “di” sebagai imbuhan atau kata depan. Pada umumnya, kata “di” dipisahkan dari kata kerja yang bersanding dengan “di”, dan kata “di” ditulis tanpa spasi ketika berfungsi sebagai kata depan. Contoh, di sekolah. Kedua kata tersebut ditulis menjadi satu kata, yaitu “disekolah”. Selain itu, beberapa soal juga belum tepat dan sesuai dengan aturan tata bahasa seperti penulisan tanda titik, koma, tanda tanya, huruf kapital dan lain-lain.

Adapun pada aspek tingkatan berpikir, soal-soal UAS buatan para guru (R-1 sampai dengan R-6) cenderung berada pada level LOTS (*Low Order Thinking Skill*). Hanya sedikit yang berlevel MOTS atau HOTS. Soal-soal yang berlevel HOTS biasanya terdapat pada soal bentuk uraian, dan itupun masih didominasi oleh soal berlevel LOTS. Dalam hal ini, dapat dikatakan bahwa kompetensi guru untuk membuat soal berlevel MOTS dan HOTS masih kurang baik.

2. Kompetensi Guru MIN Sabang dalam Menyusun Instrumen Penilaian Afektif

Kompetensi guru dalam menyusun instrumen afektif diamati melalui empat kriteria, yaitu ketermuatan aspek-aspek yang dinilai, konstruksi instrumen, bahasa, dan rubrik penskoran. Kompetensi guru MIN Sabang dalam menyusun instrumen penilaian autentik aspek afektif dapat dideskripsikan melalui tabel berikut ini:

Tabel 3. Deskripsi Kompetensi Guru MIN Sabang dalam Menyusun Instrumen Penilaian Afektif

Responden (R)	Aspek yang Dinilai	Deskripsi/Uraian
R-1 Kelas I	Ketermuatan aspek-aspek yang dinilai	Instrumen yang disusun memuat aspek sikap spiritual dan sosial. Namun, aspek-aspek sikap yang dinilai belum lengkap, dan masih bisa dikembangkan. Instrumen penilaian sikap sosial belum lengkap seperti tuntutan Kurikulum 2013 (K-13). Instrumen sikap santun, demokratis, dan bertanggung jawab serta sikap jujur belum ada.

	Konstruksi	Konstruksi instrumen sudah baik.
	Bahasa	Bahasa yang digunakan baik dan jelas, dan singkat.
	Penskoran	Teknik penskoran bervariasi, namun tidak ada teknik pengolahan skor menjadi nilai.
R-2 Kelas V	Ketermuatan aspek-aspek yang dinilai	Instrumen memuat aspek sikap spiritual dan sosial. Namun, sikap aspek-aspek yang dinilai belum lengkap, dan masih bisa dikembangkan. Instrumen penilaian sikap sosial belum lengkap seperti tuntutan Kurikulum 2013. Instrumen sikap santun, demokratis, jujur, dan bertanggung jawab belum disusun.
	Konstruksi	Konstruksi instrumen sudah baik.
	Bahasa	Bahasa yang digunakan baik dan jelas dan singkat.
	Penskoran	Teknik penskoran bervariasi. Tidak ada teknik pengolahan skor menjadi nilai.
R-3 kelas III	Ketermuatan aspek-aspek yang dinilai	Instrumen hanya memuat aspek penilaian sikap teliti, cermat, dan percaya diri. Instrumen sikap spiritual maupun sikap sosial lainnya tidak ada. Dalam instrumen tersebut, tidak ada ketermuatan nilai-nilai sikap yang diamati, cermat, maupun percaya diri.
	Konstruksi	Konstruksi instrumen sangat sederhana dan sangat umum
	Bahasa	Tidak tergambar oleh peneliti, karena tidak mencantumkan aspek-aspek yang akan dinilai dan kriteria penilaian bersifat umum.
	Penskoran	Penskoran menggunakan skala rentang 1-4. Instrumen tidak disertai dengan teknik penskoran.
R-4 kelas IV	Ketermuatan aspek-aspek yang dinilai	Instrumen hanya memuat aspek penilaian sikap teliti, bertanggung jawab dan disiplin. Instrumen sikap spiritual maupun sikap sosial lainnya tidak ada. Dalam instrumen tersebut, tidak ada ketermuatan nilai-nilai sikap teliti, bertanggung jawab, maupun sikap disiplin.
	Konstruksi	Konstruksi instrumen sangat sederhana dan sangat umum, karena tidak mencantumkan kriteria atau rubrik penilaian, sehingga instrumen tersebut masih memerlukan pengembangan.
	Bahasa	Tidak tergambar oleh peneliti, karena tidak mencantumkan aspek-aspek yang akan dinilai dan kriteria penilaian bersifat umum.
	Penskoran	Penskoran menggunakan skala rentang “Belum Terlihat, Mulai Terlihat, Mulai Berkembang dan Membudaya, namun instrumen tersebut tidak disertai dengan teknik penskoran.
R-5 Kelas VI	Ketermuatan aspek-aspek yang dinilai	Instrumen hanya memuat aspek penilaian sikap diskusi. Instrumen sikap spiritual maupun sikap sosial lainnya tidak ada.
	Konstruksi	Konstruksi instrumen sudah baik dan memuat aspek-aspek diskusi yang ingin dinilai, namun beberapa rubrik penilaian sulit dinilai karena rubrik penskoran yang disusun tidak logis.
	Bahasa	Bahasa yang digunakan singkat, dan jelas.
	Penskoran	Penskoran menggunakan skala rentang “Sangat baik, Baik,

		Cukup, dan Kurang.” Namun instrumen tersebut tidak disertai dengan teknik penskoran.
R-6 Kelas II	Ketermuatan aspek-aspek yang dinilai	Instrumen hanya memuat aspek penilaian sikap jujur, disiplin, tanggung jawab, santun, dan peduli. Instrumen sikap spiritual maupun sikap sosial lainnya tidak ada. Dalam instrumen tersebut, tidak ada ketermuatan nilai-nilai sikap jujur, disiplin, tanggung jawab, santun, dan peduli.
	Konstruksi	Konstruksi instrumen afektif yang tercantum dalam RPP sangat sederhana dan sangat umum, karena tidak mencantumkan kriteria atau rubrik penilaian. Instrumen tersebut masih memerlukan pengembangan.
	Bahasa	Bahasa yang digunakan sangat singkat. Dalam instrumen tersebut, dicantumkan aspek-aspek yang akan dinilai dan kriteria penilaian bersifat umum.
	Penskoran	Penskoran menggunakan skala rentang “Belum Terlihat, Mulai Terlihat, Mulai Berkembang dan Membudaya, namun instrumen tersebut tidak disertai dengan teknik penskoran.

Berdasarkan tabel 3. di atas, dapat digambarkan bahwa instrumen yang disusun guru sudah memuat aspek sikap spiritual dan sosial. Namun, aspek-aspek sikap yang dinilai belum lengkap, dan masih dapat dikembangkan. Instrumen penilaian sikap sosial belum lengkap seperti tuntutan Kurikulum 2013 (K-13), yaitu sikap sosial jujur, disiplin, tanggung jawab dan toleransi. Dalam instrumen afektif tersebut, tidak ada ketermuatan nilai-nilai sikap yang diamati, sehingga instrumen tersebut memerlukan pengembangan.

Pada umumnya, konstruksi instrumen sangat sederhana dan sangat umum, karena tidak mencantumkan kriteria atau rubrik penilaian. Walaupun ada kriteria dalam instrumen, namun beberapa rubrik penskoran sulit dinilai karena rubrik penskoran yang disusun tidak logis, dan instrumen tidak disertai dengan teknik penskoran. Selanjutnya, bahasa yang digunakan, baik pada kolom aspek-aspek yang dinilai, maupun aspek kriteria atau rubrik penskoran sangat singkat dan jelas.

3. Kompetensi Guru MIN Sabang dalam Menyusun Instrumen Penilaian Psikomotor

Kompetensi guru dalam menyusun instrumen psikomotor diamati melalui kriteria kesesuaian instrumen dengan indikator, ketermuatan aspek-aspek yang dinilai, bahasa, dan rubrik penskoran. Kompetensi guru MIN Sabang dalam menyusun instrumen penilaian autentik aspek psikomotor dapat dideskripsikan melalui tabel berikut ini:

Tabel 4. Deskripsi Kompetensi Guru MIN Sabang dalam Menyusun Instrumen Psikomotor

Responden (R)	Aspek yang Dinilai	Deskripsi/Uraian
R-1 Kelas I	Kesesuaian instrumen dengan indikator	Instrumen sesuai dengan materi dan indikator.
	Ketermuatan aspek-	Semua instrumen sudah memiliki ketermuatan aspek-

	aspek yang dinilai	aspek yang dinilai.
	Bahasa	Bahasa yang digunakan singkat dan jelas.
	Penskoran	Secara umum, penskoran sudah baik, hanya sebagian kecil perlu perbaikan karena sulit dibedakan saat pemberian skor tiap-tiap tingkatan, dan tidak ada teknik pengolahan skor.
R-2 Kelas V	Kesesuaian instrumen dengan indikator	Instrumen yang disusun belum sesuai dengan materi dan tuntutan indikator, dan tidak semua indikator memiliki instrumen penilaian
	Ketermuatan aspek-aspek yang dinilai	Ketermuatan aspek-aspek yang dinilai sudah baik dan relevan dengan nilai-nilai yang diinginkan.
	Bahasa	Menggunakan aturan bahasa Indonesia yang baik.
	Penskoran	Penskoran sudah baik, namun tidak diiringi dengan teknik pengolahan skor menjadi nilai.
R-3 Kelas III	Kesesuaian instrumen dengan indikator	Instrumen yang disusun sesuai dengan materi dan indikator.
	Ketermuatan aspek-aspek yang dinilai	Ketermuatan aspek-aspek yang dinilai sudah baik dan relevan dengan nilai-nilai yang diinginkan.
	Bahasa	Menggunakan aturan bahasa Indonesia yang baik.
	Penskoran	Penskoran berbentuk daftar cek.
R-4 Kelas IV	Kesesuaian instrumen dengan indikator	Instrumen sesuai dengan materi dan indikator.
	Ketermuatan aspek-aspek yang dinilai	Semua instrumen sudah memiliki ketermuatan aspek-aspek yang dinilai.
	Bahasa	Bahasa yang digunakan singkat dan jelas.
	Penskoran	Secara umum, penskoran sudah baik, hanya sebagian kecil perlu perbaikan, dan tidak ada teknik pengolahan skor.
R-5 Kelas VI	Kesesuaian instrumen dengan indikator	Instrumen sesuai dengan materi dan indikator.
	Ketermuatan aspek-aspek yang dinilai	Semua instrumen sudah memiliki ketermuatan aspek-aspek yang dinilai.
	Bahasa	Bahasa yang digunakan jelas, namun masih dapat dipersingkat.
	Penskoran	Secara umum, penskoran sudah baik, hanya sebagian kecil perlu perbaikan, dan tidak ada teknik pengolahan skor.
R-6 Kelas II	Kesesuaian instrumen dengan indikator	Instrumen yang disusun sesuai dengan materi dan indikator.
	Ketermuatan aspek-aspek yang dinilai	Semua instrumen sudah memiliki ketermuatan aspek-aspek yang dinilai.
	Bahasa	Bahasa yang digunakan jelas, dan singkat.
	Penskoran	Secara umum, penskoran sudah baik, hanya sebagian kecil instrumen masih memerlukan perbaikan, karena

		rubrik penskoran instrumen tersebut belum tepat. Selain itu, tidak ada teknik pengolahan skor dalam instrumen.
--	--	--

Data pada tabel 4 di atas menunjukkan bahwa sebagian guru sudah memiliki kompetensi yang baik dalam menyusun instrumen penilaian psikomotor. Kompetensi ini terindikasi dari adanya kesesuaian antara tema/materi instrumen dengan indikator, penggunaan bahasa yang jelas dan singkat, dan ketermuatan aspek-aspek yang dinilai. Secara umum, teknik dan rubrik penskoran sudah baik, hanya sebagian kecil perlu perbaikan dan perlu dicantumkan teknik pengolahan skor.

Berdasarkan hasil analisis terhadap data wawancara, RPP dan soal-soal UAS, kompetensi guru dalam menyusun instrumen penilaian baik kognitif, afektif, maupun psikomotor, sangat berbeda-beda antara satu sama lain. Secara umum, kompetensi guru dalam menyusun instrumen kognitif sudah baik. Hal ini terlihat dari kesesuaian soal yang disusun dengan indikator, materi yang ditanyakan sudah sesuai dengan tujuan pengukuran, dan isi materi yang ditanyakan sesuai dengan tingkatan kelas. Namun, sebagian sebagian kecil kata kerja operasional (KKO) dalam indikator memerlukan revisi sehingga relevan dengan tuntutan kompetensi dasar. Selain itu, konstruksi soal sudah baik sesuai dengan kaidah bentuk soal yang digunakan. Hanya sebagian instrumen yang belum tepat seperti pilihan jawaban pada soal bentuk pilihan ganda yang belum homogen dan logis, serta soal isian singkat yang masih bersifat umum dan menyerupai soal bentuk uraian. Pada aspek bahasa, rumusan kalimat soal komunikatif, dan sudah menggunakan bahasa Indonesia yang baku. Namun pada sebagian kecil soal, masih terdapat kesalahan pada tanda baca seperti spasi, tanda titik, koma dan lain-lain.

Bila ditinjau dari aspek tingkatan berpikir, soal-soal pada mapel yang diujikan didominasi oleh soal-soal tingkatan berpikir LOTS (*Low Order Thinking Skill*), yaitu pada tataran *cognitif* mengingat (C-1). Sebagian kecil soal berlevel MOTS (*Middle Order Thinking Skill*) dan HOTS (*Higher Order Thinking Skill*). Bahkan, ditemukan ada soal UAS yang semuanya berada pada tingkatan berpikir LOTS.

Kriteria penilaian sebagaimana penulis sebutkan di atas menunjukkan bahwa sebagian besar guru MIN Sabang sudah memiliki kompetensi yang baik dalam menyusun instrumen penilaian kognitif. Namun kompetensi guru tersebut masih perlu ditingkatkan, karena ada beberapa aspek yang masih memerlukan pengembangan, seperti dalam pengembangan indikator, pemilihan kata kerja operasional, teknik membuat rubrik penskoran, soal-soal yang memiliki tingkatan MOTS dan HOTS, dan lain-lain.

Instrumen penilaian afektif memuat beberapa aspek penilaian sikap yang diteliti, seperti bertanggung jawab, disiplin, diskusi dan lain-lain. Namun, tidak ada ketermuatan nilai-nilai sikap teliti, bertanggung jawab, maupun sikap disiplin. Walaupun sebagian guru mengaku paham mengenai teknik menyusun instrumen penilaian afektif dan cara

mengembangkan muatan sikap, namun pada bagian rubrik masih memerlukan pengembangan lebih mendalam, karena rubrik penskoran masih bersifat umum.

Berdasarkan analisis data dokumen RPP, terlihat bahwa pada domain psikomotor, sebagian besar instrumen penilaian sudah baik. Aspek-aspek yang dinilai sesuai dengan rumusan indikator dalam RPP, dan bahasa yang digunakan jelas serta singkat. Namun, data tersebut masih perlu dipertanyakan mengingat sebagian besar instrumen psikomotor yang terdapat dalam RPP diadopsi dari internet. Terlepas dari kondisi tersebut, seperti halnya instrumen penilaian afektif, instrumen penilaian psikomotor masih memerlukan pengembangan yang lebih mendalam, terutama pada aspek rubrik penskoran.

Jika ketiga domain (kognitif, afektif, dan psikomotor) dikomparasikan, terlihat bahwa kompetensi guru dalam menyusun soal ternyata lebih baik daripada menyusun instrumen penilaian afektif dan psikomotor. Sedangkan kompetensi guru dalam menyusun instrumen psikomotor lebih baik dibanding kompetensi guru dalam menyusun instrumen afektif. Hasil penelitian ini sesuai dengan hasil penelitian yang dilakukan oleh Dasrieny Pratiwi dan Hening Widowati yang mengungkapkan bahwa mahasiswa calon guru biologi FKIP Universitas Muhammadiyah Malang memiliki kompetensi yang jauh lebih baik dalam menyusun instrumen penilaian kognitif daripada penilaian afektif dan psikomotor.¹⁸

Kemampuan guru yang unggul pada domain kognitif menunjukkan adanya kesenjangan penerapan penilaian autentik antara idealitas dan realitas. Salah satu prinsip evaluasi yang tidak boleh diabaikan adalah penyelenggaraan penilaian domain kognitif, afektif, dan psikomotor secara komprehensif,¹⁹ dan hal tersebut ternyata tidak berjalan seimbang. Hal ini sesuai dengan yang dikemukakan oleh Rohmad Qomari bahwa dalam faktanya, kegiatan evaluasi yang menonjol di suatu lembaga dan satuan pendidikan adalah pelaksanaan ujian (tes) setiap kali menyelesaikan pokok bahasan tertentu.²⁰ Selain itu, guru juga sudah terbiasa dan terlatih dengan instrumen tes mulai dari tes formatif yang berkala, tes sumatif, sampai dengan tes kelulusan pada Ujian Nasional (UN).

4. Bentuk-bentuk Kesulitan yang Dihadapi Guru dalam Penyusunan Instrumen Penilaian

Kualifikasi kemampuan guru MIN Sabang sebagaimana yang peneliti uraikan di atas tentunya sangat terkait dengan kesulitan yang dihadapi guru-guru tersebut. Kesulitan dalam menyusun instrumen penilaian kognitif lebih dirasakan pada penyusunan soal berbentuk pilihan ganda. Hal ini ditunjukkan oleh hasil angket yang mengungkapkan sebanyak 51% guru mengalami kesulitan dalam menyusun soal berbentuk pilihan ganda,

¹⁸Dasrieny Pratiwi dan Hening Widowati, "Deskripsi Kemampuan Calon...", hal. 190.

¹⁹Hamzah B. Uno dan Satria Koni, *Assessment Pembelajaran*, Jakarta: Bumi Aksara, hal. 37. Lihat: Zainal Arifin, *Evaluasi Pembelajaran: Prinsip, Teknik, Prosedur*, Bandung: Remaja Rosdakarya, 2014, hal. 31. Lihat: Anas Sudijono, *Pengantar Evaluasi Pendidikan*, Jakarta: RajaGrafindo Persada, 2006, hal. 31.

²⁰Rohmad Qomari, "Pengembangan Instrumen Evaluasi... hal. 3.

sebanyak 43% berbentuk soal isian singkat, sebanyak 39% berbentuk soal benar-salah (B-S), 48% untuk soal menjodohkan, dan 49% untuk soal uraian. Kesulitan dalam menyusun soal terletak pada tahap penyusunan indikator, terutama dalam menentukan kata kerja operasional, dan ini terjadi ketika guru tidak memiliki buku paket.²¹ Berdasarkan hasil wawancara, terungkap bahwa sebagian guru belum memahami dengan baik teknik penyusunan indikator, baik untuk instrumen tes maupun instrumen penilaian afektif dan psikomotor.²²

Selain itu, bentuk kesulitan dalam menyusun instrumen penilaian adalah saat menyusun redaksi bahasa yang baik, tegas dan komunikatif, merumuskan pilihan-pilihan jawaban yang logis dan homogen pada soal bentuk pilihan ganda maupun menjodohkan. Berdasarkan hasil telaah dokumen soal, wawancara dan angket, ditemukan bahwa sebagian guru belum memahami tata cara/kaidah pembuatan soal sesuai dengan karakteristik soal. Soal bentuk uraian yang semula dianggap mudah penyusunannya, ternyata dipandang sulit oleh sebagian guru. Hal ini disebabkan pemahaman yang kurang dalam membedakan soal bentuk uraian dengan soal bentuk isian singkat. Walaupun indikator yang tertera dalam instrumen sebagian besar sudah sesuai dengan tuntutan kompetensi dasar, namun bagi guru-guru tersebut, mudah atau tidaknya menjabarkan indikator, sangat tergantung pada konten kompetensi dasar.²³

Kesulitan menyusun instrumen lebih dirasakan pada saat menyusun instrumen penilaian afektif dan psikomotor. Data hasil angket menunjukkan bahwa sebanyak 61% guru mengaku sulit membuat instrumen afektif dan sebanyak 62% merasa sulit merancang instrumen psikomotor. Berdasarkan data angket, juga diperoleh informasi bahwa pada aspek penyusunan indikator, tingkat kesulitan yang dirasakan guru MIN Sabang sebesar 64%, aspek kesesuaian antara indikator dengan komponen yang dinilai mencapai 74%, dan persentase penyusunan aspek-aspek yang dinilai mencapai 56%. Adapun dalam hal penyusunan redaksi bahasa yang baik, tegas dan komunikatif, tingkat kesulitan mencapai 43%.

Berdasarkan analisis dari berbagai data, peneliti menemukan sebagian besar guru pada MIN Sabang belum memahami teknik penilaian afektif dan psikomotor dengan baik. Penilaian afektif dilaksanakan berdasarkan format yang umum dan sederhana yang sudah ada di madrasah, tanpa ada pengembangan dan diiringi oleh pemahaman. Demikian pula pada penilaian psikomotor, sebagian guru tidak mempunyai rubrik penilaian sehingga nilai akhir yang diberikan sulit untuk dipertanggungjawabkan. Bahkan, peneliti juga menemukan adanya sebagian guru yang hanya mengira-ngira dalam memberi nilai akhir aspek psikomotor. Keterbatasan pemahaman guru mengenai penilaian berbasis kurikulum 2013 disebabkan kurangnya pelatihan-pelatihan terkait dengan kurikulum 2013, dan terbatasnya waktu untuk belajar secara serius.

Kesulitan yang dirasakan guru-guru MIN Sabang sebagaimana peneliti kemukakan

²¹Wawancara dengan R3 dan R5, tanggal 6 Agustus 2018.

²²Wawancara dengan R1, R2, R4, R5, dan R6, tanggal 6 Agustus 2018.

²³Wawancara dengan R1, R2, R3, R4, dan R6, tanggal 6 Agustus 2018.

di atas senada dengan penelitian Misbahul Jannah yang mengungkapkan bahwa kesulitan mahasiswa calon guru MI dalam menyusun evaluasi konsep IPA adalah dalam menentukan kata kerja operasional dan kurangnya pengetahuan mengenai teknik penilaian.²⁴ Berdasarkan hal tersebut, diperlukan upaya-upaya lebih intensif untuk mengatasi kesulitan guru dalam menyusun instrumen penilaian baik dari internal guru, maupun dari eksternal. Dengan demikian, guru, kepala sekolah dan *stakeholder* masing-masing memegang peran yang sangat penting untuk mengatasi problematika tersebut.

SIMPULAN

Sebagian besar guru MIN Sabang sudah memiliki kompetensi yang baik dalam menyusun instrumen penilaian aspek kognitif, namun belum begitu baik pada aspek sikap, maupun psikomotor. Penyusunan instrumen penilaian kognitif sudah sesuai dengan aturannya, meskipun dalam beberapa aspek perlu pengembangan dan perbaikan, seperti dalam pengembangan indikator, pemilihan kata kerja operasional, konstruksi soal, penyusunan soal berlevel MOTS dan HOTS, dan penggunaan bahasa Indonesia yang tegas, dan jelas sesuai dengan ejaan yang disempurnakan. Penyusunan instrumen penilaian sikap sosial belum lengkap seperti tuntutan Kurikulum 2013, yaitu sikap sosial jujur, disiplin, tanggung jawab dan toleransi. Dalam instrumen afektif tersebut, tidak ada ketermuatan nilai-nilai sikap yang diamati, sehingga instrumen tersebut memerlukan pengembangan. Selain itu, konstruksi instrumen sangat sederhana dan sangat umum, karena tidak mencantumkan kriteria atau rubrik penilaian. Walaupun ada kriteria dalam instrumen, namun beberapa rubrik penskoran sulit dinilai karena rubrik penskoran yang disusun tidak logis, dan instrumen tidak disertai dengan teknik penskoran. Namun demikian, bahasa yang digunakan baik pada kolom aspek-aspek yang dinilai, maupun aspek kriteria atau rubrik penskoran jelas dan singkat.

Kesulitan guru dalam menyusun instrumen penilaian autentik terletak pada cara mengembangkan indikator dari Kompetensi Dasar, yaitu dalam menentukan kata kerja operasional, cara menyusun bahasa yang baik sesuai dengan EYD (Ejaan yang Disempurnakan) dan redaksi bahasa yang komunikatif, menyusun aspek-aspek yang dinilai dalam instrumen afektif dan psikomotor, menyesuaikan antara indikator dan instrumen soal/aspek-aspek penilaian, dan teknik menentukan *opsi* (pilihan) jawaban yang logis dan homogen pada soal berbentuk pilihan ganda dan menjodohkan. Informasi ini memberikan gambaran tentang kondisi guru-guru MIN Sabang yang memiliki kompetensi terbatas dalam menyusun instrumen penilaian autentik, sehingga pengadaan pelatihan terkait dengan evaluasi atau asesmen sangat dibutuhkan untuk meningkatkan kompetensi pedagogik guru-guru MIN Sabang.

²⁴Misbahul Jannah, "Kemampuan Mahasiswa PGMI... hal. 116

DAFTAR PUSTAKA

- Arifin, Zainal 2014. *Evaluasi Pembelajaran, Bandung: Remaja Rosdakarya.*
- Arikunto, Suharsimi. *Dasar-dasar Evaluasi Pendidikan*, Jakarta: Bumi Aksara, 2008.
- , *Prosedur Penelitian: Suatu Pendekatan dan Praktek*, Jakarta: Rineka Cipta, 1998.
- Azwar, Saifuddin, *Metode Penelitian*, Yogyakarta: Pustaka Pelajar, 2007.
- Depdiknas, “Standar Kualifikasi Akademik dan Kompetensi Guru. Salinan Lampiran Menteri Pendidikan Nasional No. 16 Tahun 2007 tanggal 4 Mei 2007” [Online]
<http://vervalsp.data.kemdikbud.go.id/prosespembelajaran/file/Permendiknas%20No%2016%20Tahun%202007.pdf> diakses tanggal 27 Juni 2019.
- Hosnan, M. *Pendekatan Sainifik dan Kontekstual dalam Pembelajaran Abad 21; Kunci Sukses Implemetasi Kurikulum 2013*, Bogor: Ghalia Indonesia, 2014.
- Jannah, Misbahul. “Kemampuan Mahasiswa PGMI UIN Ar-Raniry dalam Evaluasi Pembelajaran IPA”. *Proceeding*, Banda Aceh: ARICIS I, 2016.
- Kunandar. *Penilaian Autentik (Penilaian Hasil Belajar Peserta Didik Berdasarkan Kurikulum 2013); Suatu Pendekatan Praktis Disertai dengan Contoh*, Jakarta: RajaGrafindo Persada, 2014.
- Majid, Abdul. *Penilaian Autentik; Proses dan Hasil Belajar*, Bandung: Remaja Rosdakarya, 2014.
- Qomari, Rohmad. “Pengembangan Instrumen Evaluasi Afektif Domain” *Jurnal pemikiran Alternatif Pendidikan*. Vol. 3, No. 1, hal. 3. file:///C:/Users/USER/Downloads/287-Article%20Text-553-1-10-20150304.pdf
- Pratiwi, Dasrieny dan Widowati, Hening. “Deskripsi Kemampuan Calon Guru Biologi dalam Menyusun Perangkat Penilaian Berbasis Kurikulum 2013.” *Proceeding Seminar Nasional Pendidikan Biologi*. Malang: Universitas Muhammdiyah, 2015.
<http://biology.umm.ac.id/files/file/521-529%20Hening%20Widowati.pdf>. Diakses tanggal 15 Mei 2018.
- Rusman. *Model-Model Pembelajaran; Mengembangkan Profesionalisme Guru*, Jakarta: Rajawali Pers, 2011.

Safari. *Teknik Analisis Butir Soal Instrumen Tes dan Non-Tes*, Jakarta: Asosiasi Pengawas Sekolah Indonesia Departemen Pendidikan Nasional, 2005.

Sudijono, Anas. *Pengantar Evaluasi Pendidikan*, Jakarta: RajaGrafindo Persada, 2006.

Sunarti dan Rahmawati, Selly. *Penilaian dalam Kurikulum 2013; Membantu Guru dan Calon Guru Mengetahui Langkah-langkah Penilaian Pembelajaran*, Yogyakarta: Andi, 2014.

Supardi. *Penilaian Autentik Pembelajaran Afektif, Kognitif dan Psikomotor; Konsep dan Aplikasi*, Jakarta: Rajawali Pers, 2016.

Uno, Hamzah B. dan Koni, Satria. *Assessment Pembelajaran*, Jakarta: Bumi Aksara, 2013.

Yani, Ahmad. *Mindset Kurikulum 2013*, Bandung: Alfabeta, 2014.