

MODEL ASESMEN MUSABAQAH TILAWAH AL-QURAN (MTQ) CABANG TILAWAH

Bahrudin, Kumaidi

UIN Sunan Gunung Djati Bandung, Universitas Muhammadiyah Surakarta
bahrudintegral@yahoo.co.id, kuma_426@yahoo.com

Abstrak

Penelitian ini bertujuan untuk mengembangkan model penilaian MTQ Cabang Tilawah, karena model penilaian yang selama ini digunakan dalam MTQ tidak sesuai dengan teori asesmen. Jenis penelitian ini adalah *research and development*, yang dimulai dengan *research* dan diteruskan dengan pengembangan model. *Research* dilakukan untuk memperoleh informasi mengenai kelemahan-kelemahan model penilaian MTQ Cabang Tilawah yang digunakan selama ini dan aspek-aspek yang mendukung perlunya dilakukan pengembangan model penilaian tersebut. Pengembangan model mengacu pada prosedur Borg & Gall (1983), yang dilakukan melalui tiga tahap, yaitu: pra-pengembangan, pengembangan, dan diseminasi. Hasil penelitian menunjukkan bahwa model penilaian yang dikembangkan dalam penelitian ini layak digunakan dalam penilaian MTQ Cabang Tilawah. Penilaian tentang kelayakan tersebut didasarkan pada validitas dan reliabilitas instrumen yang digunakan dalam uji coba. Dari analisis rasional mengenai isi instrumen yang dilakukan oleh sejumlah pakar diperoleh penilaian bahwa instrumen-instrumen penilaian yang dikembangkan dalam penelitian ini adalah valid. Kemudian dari estimasi reliabilitas skor gabungan hasil uji coba diperoleh koefisien reliabilitas sebesar 0,96. sehingga hasil penilaian yang dikembangkan dalam penelitian ini adalah reliabel.

Kata kunci: *MTQ, asesmen, tajwid, fasih, suara, irama*

AN ASSESSMENT MODEL OF MUSABAQAH TILAWAH AL-QURAN (MTQ) OF THE TILAWAH SECTION

Bahrudin, Kumaidi

UIN Sunan Gunung Djati Bandung, Universitas Muhammadiyah Surakarta
bahrudintegral@yahoo.co.id, kuma_426@yahoo.com

Abstract

This study aims to develop a model of an assessment of MTQ of the tilawah section because the assessment of MTQ that has been applied so far is not relevant to assessment theories. This was a research and development study, started with research and continued with developing a model. The research was conducted to obtain information about weaknesses of the model of the assessment of MTQ of the tilawah section having been applied so far and aspects supporting the needs for developing an assessment model. The model development referred to the procedure by Borg & Gall (1983), consisting of three steps, i.e.: pre-development, development, and dissemination. The results of the study show that the assessment model developed in the study is appropriate to be applied in the assessment of MTQ of the tilawah section. The evaluation of the appropriateness is based on the validity and reliability of the instruments in the tryout. From the analysis of the rationale of the instrument contents done by several experts, the instruments developed in the study are considered valid. Based on the estimation of the reliability of the combined scores from the tryout, the reliability coefficient is 0.96, the assessment instruments developed in the study are reliable.

Keywords: *MTQ, assessment, tajwid, fashahah, voice, rhythm*

Pendahuluan

Musabaqah Tilawah Al-Quran (MTQ) adalah nama sebuah kompetisi yang memperlombakan semua cabang lomba al-Quran. Salah satu cabang yang diperlombakan adalah tilawah al-Quran. Oleh karena itu, dalam MTQ terdapat salah satu lomba yang bernama MTQ Cabang Tilawah. Tilawah al-Quran adalah suatu jenis lomba membaca al-Quran dengan bacaan *mujanwad* dan *murattal*, yakni bacaan al-Quran yang mengandung nilai ilmu membaca, seni membaca dan adab membaca menurut pedoman yang telah ditentukan (Tim LPTQ, 2011, p.41).

Sebagai sebuah lomba, MTQ Cabang Tilawah tidak bisa dilepaskan dari kegiatan penilaian dan pengukuran. Penilaian dalam kegiatan ini dilakukan untuk menentukan kualitas tilawah (bacaan al-Quran) setiap peserta. Penilaian kualitas tilawah tersebut didasarkan pada skor yang didapatkan oleh para peserta. Besar dan kecilnya skor para peserta menggambarkan kualitas tilawah mereka. Peserta yang mendapatkan skor paling tinggi dalam MTQ Cabang Tilawah dapat dinilai sebagai yang paling bagus kualitas tilawahnya, sehingga ia layak untuk ditetapkan sebagai juara.

Penilaian dan pengukuran meskipun merupakan substansi yang berbeda, namun keduanya memiliki keterkaitan yang bersifat hirarkis (Griffin & Nix, 1991, p.3). Maksudnya, kegiatan yang satu baru akan terjadi setelah kegiatan yang lain. Dalam konteks ini, penilaian selalu dilakukan setelah pengukuran.

Pengukuran adalah proses penetapan skor terhadap individu atau karakteristiknya menurut aturan tertentu (Ebel & Frisbie, 1986, p.14), atau kegiatan menetapkan angka (skor) dengan cara yang sistematis untuk menyatakan keadaan individu (Allen & Yen, 1979, p.2). Dalam MTQ Cabang Tilawah, keadaan individu dimaksud berupa kemampuan setiap peserta untuk melantunkan ayat-ayat al-Quran sesuai kaidah-kaidah tajwid yang dibawakan dengan suara yang indah.

Adapun penilaian dapat diartikan sebagai kegiatan menafsirkan data hasil pengukuran. Data hasil pengukuran, yang berupa skor-skor bisa berbeda antara individu yang satu dengan individu yang lain, sehingga hasil penafsirannya pun akan bersifat individual. Karena itulah penilaian berfokus pada individu. Dalam dunia pendidikan misalnya, penilaian berfokus pada prestasi belajar yang dicapai oleh individu setiap peserta didik (Mardapi, 2008, p.3). Sedangkan dalam MTQ Cabang Tilawah, penilaian berfokus pada kualitas bacaan al-Quran setiap individu peserta MTQ.

Penilaian dalam MTQ Cabang Tilawah dilakukan terhadap empat komponen, yaitu: (1) *tajwid*; (2) *fashahah*; (3) suara; dan (4) lagu (Tim LPTQ, 2011, pp.53-54). Penilaian terhadap keempat komponen tersebut didasarkan pada sejumlah norma penilaian yang telah ditetapkan oleh LPTQ. Norma-norma penilaian tersebut diatur dalam buku "Pedoman Musabaqah al-Quran 2010" (buku pedoman terbaru yang terbit pada tahun 2011) pada halaman 60-61.

Norma-norma penilaian di atas jika dicermati mengandung sejumlah kelemahan, sehingga model penilaian MTQ Cabang Tilawah dalam beberapa aspek tidak sesuai dengan teori pengukuran/asesmen. Aspek-aspek penilaian MTQ Cabang Tilawah yang tidak sesuai dengan teori pengukuran tersebut terlihat dalam uraian singkat di bawah.

Pertama, selama ini penilaian dalam MTQ Cabang Tilawah menggunakan *maqra'* (sejumlah ayat lomba) yang berbeda-beda untuk setiap pesertanya. Padahal penilaian dalam MTQ Cabang Tilawah dilakukan terhadap seluruh elemen bacaan yang termuat dalam setiap *maqra'* tersebut. Dengan demikian, penggunaan *maqra'* yang berbeda-beda sama dengan menilai atau mengukur kualitas tilawah setiap peserta dengan menggunakan alat ukur yang tidak sama. Penggunaan *maqra'* yang berbeda-beda bisa menyebabkan hasil penilaian menjadi kurang adil, karena *maqra'* yang berbeda-beda di samping memuat jenis dan jumlah elemen bacaan yang berbeda, juga kemungkinan

besar memiliki tingkat kesulitan yang juga berbeda-beda.

Kedua, penilaian dalam MTQ Cabang Tilawah dilakukan secara tidak proporsional. Secara sederhana, penilaian dikatakan proporsional jika skor akhir dari sebuah tes yang seluruh jawabannya salah adalah nol. Ketidakproporsionalan penilaian dalam MTQ ini terlihat pada dua aspek, yaitu: (1) penetapan nilai maksimal untuk setiap bidangnya yang jauh tidak sebanding dengan jumlah elemen *maqra'* yang dinilai, dan (2) pemotongan skor untuk setiap kesalahannya terlalu besar. Pemotongan skor yang terlalu besar untuk setiap kesalahannya dapat menyebabkan skor maksimal setiap bidang penilaian MTQ habis karena kesalahan yang relatif kecil (sedikit).

Ketiga, model penilaian MTQ Cabang Tilawah yang digunakan selama ini dapat mendorong para hakim penilai tidak konsisten dalam melakukan penilaian. Inkonsistensi penilaian ini dapat terjadi terutama pada bidang *tajwid* dan *fashahah*. Jumlah item penilaian di kedua bidang ini sangat banyak, yakni meliputi semua elemen bacaan yang termuat dalam sebuah *maqra'* yang jumlahnya ratusan bahkan mencapai ribuan. Elemen bacaan yang jumlahnya sangat banyak itu diberi nilai maksimal masing-masing 30 point (Tim LPTQ, 2011, pp.59-60). Nilai maksimal ini akan dipotong sebesar 3 point setiap kali peserta melakukan kesalahan *jali*, dan 1 point untuk setiap kesalahan *khafi* (Tim LPTQ, 2011, p.60). Ini artinya, nilai maksimal di kedua bidang itu akan habis terpotong hanya karena peserta salah dalam membaca 10 huruf, atau 10 harakat. Bagaimana jika kesalahan yang dilakukannya lebih dari itu? Jika ini terjadi, maka akan ada sejumlah kesalahan yang tidak bisa dilakukan pemotongan skor, karena nilai maksimal sudah habis terpotong oleh beberapa kesalahan sebelumnya. Dengan demikian, model penilaian MTQ Cabang Tilawah yang berlaku selama ini dapat mendorong para hakim/penilai untuk tidak konsisten dalam melakukan penilaian.

Keempat, model penilaian MTQ Cabang Tilawah membatasi kemandirian para

hakim dalam melakukan penilaian. Penilaian dalam MTQ Cabang Tilawah dilakukan oleh beberapa orang hakim untuk setiap bidangnya. Oleh karena itu, wajar jika sering terjadi perbedaan skor di antara para hakim tersebut. Potensi terjadinya perbedaan skor ini disadari oleh LPTQ sebagai penyelenggara MTQ. Untuk menghindari terjadinya perbedaan skor yang terlalu tajam, LPTQ menetapkan aturan bahwa selisih antara skor tertinggi dan skor terendah yang diberikan oleh para hakim MTQ Cabang Tilawah tidak lebih dari 3 point. Ketika terjadi perbedaan skor lebih dari 3 point, maka Majelis Hakim akan bermusyawarah untuk melakukan perubahan terhadap sebagian skor yang diberikan oleh salah seorang hakim (Tim LPTQ, 2011, p.34).

Aturan-aturan penilaian tersebut dapat membuat hasil penilaian MTQ Cabang Tilawah tidak objektif, karena sebagian skor peserta terjadi akibat hasil kompromi Majelis Hakim, bukan lagi hasil penilaian hakim terhadap performa bacaan peserta. Selain itu, aturan tersebut juga akan mengurangi kemandirian para hakim dalam melakukan penilaian. Padahal kemandirian dalam penilaian sangat penting.

Dari uraian-uraian tersebut dapat disimpulkan bahwa model penilaian MTQ Cabang Tilawah yang digunakan selama ini dalam beberapa aspeknya tidak sesuai dengan prinsip-prinsip penilaian yang telah lama diterapkan dalam dunia pendidikan. Ketidakesesuaian tersebut antara lain terlihat pada pemakaian instrumen ganda melalui penggunaan *maqra'* yang berbeda-beda untuk setiap pesertanya; dan penilaian yang tidak proporsional; teknik penilaian yang dapat mendorong para hakim berlaku tidak konsisten dalam melakukan penskoran, dan intervensi Majelis Hakim melalui wewenangnya yang dapat mengubah sebagian skor yang diberikan hakim.

Sejumlah penyimpangan yang terdapat dalam model penilaian MTQ di atas mendorong penulis untuk melakukan penelitian dan pengembangan. Penelitian dilakukan untuk mengkaji secara lebih mendalam mengenai kelemahan-kelemahan model pe-

nilaian MTQ Cabang Tilawah yang digunakan selama ini dan aspek-aspek yang mendukung mengenai perlunya diadakan pengembangan model penilaian tersebut.

Selain oleh peneliti sendiri, penelitian tentang model penilaian MTQ Cabang Tilawah, terutama di Indonesia, bisa dikatakan belum dilakukan, setidaknya hingga artikel ini dibuat. Dari survei yang diadakan peneliti di sejumlah lembaga yang terkait dengan penyelenggaraan MTQ, seperti LPTQ dan Kemenag, dan di sejumlah perpustakaan Islam tidak diperoleh satu pun hasil penelitian tentang model penilaian MTQ. Di Indonesia, kajian tentang MTQ pada umumnya ditulis dalam bentuk artikel, misalnya artikel yang ditulis oleh M. Nasir yang berjudul "MTQ dan Kepedulian Universal." Namun, artikel yang ditulis pada 9 Mei 2008 ini sama sekali tidak menyinggung apa pun yang terkait dengan penilaian MTQ Cabang Tilawah. Sedangkan Shihab (2002) hanya mengkritisi sebagian aspek dari penilaian MTQ Cabang Tafsir. Substansi penilaian MTQ Cabang Tafsir tentu saja sangat berbeda dengan penilaian MTQ Cabang Tilawah.

Selain Indonesia, negara Islam yang giat menyelenggarakan MTQ adalah Mesir. Di negara yang merupakan salah satu kiblat MTQ Internasional ini terdapat hasil penelitian tentang tilawah yang dilakukan oleh Michael Frishkopf. Frishkopf melakukan kajian tentang tilawah al-Qur'an dengan judul "*Mediated Qur'anic Recitation and the Contestation of Islam Contemporary Egypt*." Dalam penelitian ini Frishkopf menjelaskan tilawah al-Quran sebagai salah satu media penyiaran Islam di Mesir dewasa ini. Ia membahas secara panjang lebar tentang tilawah al-Quran dan menjelaskan secara rinci mengenai perbedaan antara tilawah *mujawwad* dan tilawah *murattal* dari berbagai segi, misalnya dari segi timbre (warna suara), ekspresi, modulasi, melodi, aksen, pause (jeda), pengulangan bacaan dan sebagainya (Frishkopf, 2007, p.10). Namun, perlu di tegaskan bahwa Frishkopf dalam penelitiannya sama sekali tidak menyinggung tentang model penilaian atau apa pun namanya yang berkait-

an dengan penilaian tilawah, baik tilawah yang bercorak *mujawwad* maupun yang bercorak *murattal*.

Belum banyaknya hasil penelitian tentang MTQ merupakan tantangan dan sekaligus peluang yang sangat besar bagi peneliti untuk mengadakan penelitian tentang berbagai aspek yang berkaitan dengan MTQ. Namun demikian, bertolak dari sejumlah kelemahan yang terdapat dalam model penilaian MTQ Cabang Tilawah sebagaimana tersebut, maka permasalahan yang mendesak untuk diteliti adalah pengembangan model penilaian MTQ tersebut.

Pengembangan model penilaian MTQ Cabang Tilawah memerlukan penelitian pendahuluan untuk mengkaji berbagai model penilaian yang ada dalam literatur dan memilih model yang paling cocok untuk dikembangkan dalam penilaian MTQ Cabang Tilawah. Peneliti memandang bahwa model penilaian MTQ Cabang Tilawah dapat dikembangkan berdasarkan teori pengukuran dan asesmen, sebagaimana yang dilakukan oleh para pakar pendidikan terhadap model penilaian dari berbagai disiplin ilmu. Oleh karena itu, permasalahan yang diangkat dalam penelitian dan pengembangan ini adalah "Bagaimana mengembangkan model penilaian MTQ Cabang Tilawah yang sesuai dengan teori pengukuran" tersebut.

Bertolak dari sini, tujuan penelitian ini adalah untuk mendapatkan model asesmen MTQ Cabang Tilawah yang sesuai dengan teori pengukuran, dan untuk mengetahui aspek-aspek yang perlu dikaji dalam pengembangan model asesmen MTQ Cabang Tilawah berdasarkan teori pengukuran tersebut. Ada beberapa manfaat yang diharapkan dapat dipetik dari penelitian ini. *Pertama*, hasil penelitian ini diharapkan dapat mengungkapkan adanya sejumlah kelemahan yang terdapat dalam model penilaian MTQ yang digunakan selama ini. *Kedua*, hasil penelitian ini diharapkan menjadi salah satu literatur yang dapat dijadikan rujukan oleh dewan hakim MTQ dalam melakukan penilaian, dan oleh Lembaga (LPTQ) dalam menyusun pedoman penilaian yang baru dalam penilaian MTQ. *Ketiga*, hasil penelitian

ini diharapkan menjadi salah satu alternatif model penilaian MTQ Cabang Tilawah di masa-masa mendatang.

Metode Penelitian

Penelitian ini menggunakan prosedur penelitian pengembangan (*research & development*). *Research & development* secara general dapat diartikan sebagai kegiatan yang diawali dengan penelitian dan diteruskan dengan pengembangan. *Research* dilakukan selain untuk mendapatkan sejumlah informasi mengenai kelemahan-kelemahan yang terdapat dalam model penilaian MTQ Cabang Tilawah yang digunakan selama ini, juga untuk memperoleh informasi yang diperlukan untuk melakukan analisis dan penilaian mengenai perlu atau tidaknya dilakukan pengembangan model penilaian MTQ Cabang Tilawah yang lebih sesuai dengan teori pengukuran. Temuan-temuan dalam *research* kemudian dijadikan sebagai bahan untuk melakukan pengembangan model MTQ Cabang Tilawah tersebut.

Pengembangan model penilaian MTQ Cabang Tilawah ini mengikuti prosedur yang dikemukakan oleh Borg & Gall (1983). Secara singkat pengembangan model penilaian ini dilakukan melalui tiga tahap, yaitu: (1) tahap pra-pengembangan model; (2) tahap pengembangan model; dan (3) tahap desiminasi atau implementasi model.

Tahap pra-pengembangan ditempuh dengan melakukan sejumlah kegiatan, yaitu: (1) mengkaji buku "Pedoman Musabaqah al-Quran 2010" (buku pedoman MTQ terbaru), (2) menentukan spesifikasi penilaian MTQ Cabang Tilawah, dan (3) menentukan strategi pengembangan indikator penilaian kualitas tilawah, strategi pembuatan instrumen, dan strategi penyusunan pedoman penilaian MTQ.

Tahap pengembangan dimulai dengan penyusunan draf instrumen penilaian MTQ Cabang Tilawah, yang meliputi instrumen penilaian tajwid, *fashahah*, suara dan lagu. Draft-draf instrumen ini kemudian ditelaah oleh para pakar untuk diungkap berbagai kekurangan dan kelemahannya. Hasil telaah

para pakar ini dijadikan sebagai bahan untuk merevisi keempat draf tersebut. Draft-draf instrumen yang telah direvisi selanjutnya diujicobakan pada sejumlah subjek coba (sampel). Uji coba dilakukan sebanyak tiga kali dan setiap selesai uji coba dilakukan penelaahan oleh sejumlah pakar untuk melihat kekurangan atau kelemahan yang mungkin masih terdapat di dalam draft-draf tersebut. Sedangkan perevisian draft-draf instrumen didasarkan pada hasil penilaian dan masukan pakar ketika menelaah draft-draf dimaksud.

Subjek uji coba dalam penelitian ini ditentukan secara random. Subjek uji coba I berjumlah 10 orang peserta STQ tingkat Jawa Barat tahun 2009 (menggunakan rekaman), sedangkan subjek uji coba II dan III masing-masing berjumlah 11 dan 30 orang mahasiswa UIN Sunan Gunung Djati Bandung. Setiap uji coba melibatkan 12 orang hakim (*rater*) yang menilai empat komponen penilaian MTQ Cabang Tilawah dengan setiap komponennya dinilai oleh tiga orang hakim.

Instrumen yang digunakan untuk mengumpulkan data dalam penelitian pengembangan ini terdiri dari angket, *maqra'* dan lembar penskoran. Angket digunakan untuk mengumpulkan data yang berkenaan dengan analisis kebutuhan pengembangan model asesmen MTQ Cabang Tilawah. Sementara *maqra'* dan lembar penskoran digunakan untuk mengumpulkan data tentang tingkat kualitas tilawah para subjek coba dan terutama untuk uji model yang di dalamnya mencakup estimasi validitas dan reliabilitas hasil penilaian.

Penentuan validitas instrumen dilakukan melalui *judgment* pakar, sedangkan estimasi reliabilitasnya menggunakan pendekatan analisis varians yang prosedurnya didasarkan pada *generalizability theory* melalui konsep *G-Study*". Adapun desain yang digunakan adalah desain *multifacet*, karena penilaian dalam MTQ Cabang Tilawah setidaknya melibatkan tiga kondisi pengukuran, yaitu: peserta, *rater*, dan butir. Harga koefisien reliabilitas diperoleh dari perbandingan antara varian skor murni dan varian skor

amatannya yang dikomputasi dengan menggunakan teknik yang dikembangkan oleh Thorndike, sebagaimana dijelaskan dalam bukunya yang berjudul “*Applied Psycometrics*” terbitan tahun 1982. Kemudian nilai koefisien yang dihasilkan dibandingkan dengan kriteria reliabilitas minimal yang diperkenankan, yaitu 0,70 (Linn, 1990, p.143).

Penelitian ini menggunakan pendekatan kuantitatif, karena data utamanya berupa skor-skor penilaian MTQ Cabang Tilawah di bidang tajwid, *fashahah*, suara dan lagu. Sedangkan teknik analisis datanya menggunakan teknik analisis deskriptif. Teknik analisis ini digunakan untuk menjelaskan secara faktual karakteristik model penilaian MTQ Cabang Tilawah yang digunakan selama ini dan karakteristik model penilaian MTQ Cabang Tilawah hasil pengembangan. Analisis deskriptif juga digunakan untuk menjelaskan strategi pengembangan instrumen-instrumen penilaian MTQ Cabang Tilawah yang mencakup strategi penentuan indikator-indikator penilaian di setiap bidangnya beserta item-item pengukuran yang merupakan penjabaran dari indikator-indikator tersebut, serta untuk aspek validitas dan reliabilitas instrumen-instrumen penilaian dikembangkan tersebut.

Hasil Penelitian dan Pembahasan

Penelitian ini dilakukan untuk mengembangkan model asesmen (penilaian) Musabaqah Tilawah al-Quran (MTQ) Cabang Tilawah. Secara faktual, produk penelitian ini berupa seperangkat instrumen penilaian unjuk kerja (*performance assessment*) yang dapat digunakan untuk menilai performa bacaan al-Quran para peserta. Penggunaan instrumen unjuk kerja ini didasarkan pada kenyataan bahwa tilawah atau pembacaan al-Quran adalah kegiatan melafalkan huruf-huruf yang terhimpun dalam ayat-ayat al-Quran menurut kaidah tajwid. Seperti halnya pelafalan huruf-huruf dari teks yang lain, pelafalan huruf-huruf yang terhimpun dalam ayat-ayat al-Quran juga dilakukan dengan cara mendemonstrasikan pengucapan huruf-huruf yang dilafalkan/dibaca. Mendemonstrasikan pengucapan huruf-huruf

dari teks apa pun tidak lain merupakan sebuah performansi atau sebuah unjuk kerja.

Sebuah performansi hanya cocok diukur dengan tes performansi, dan tilawah al-Quran adalah sebuah performansi, sehingga penilaian mengenai kualitas tilawah para peserta MTQ juga sebaiknya menggunakan *performance assessment*. Sebagaimana dikatakan Stiggins (1994, p.171) bahwa *performance assessment* diperlukan terutama untuk menilai keterampilan (*skill*) dan karya cipta siswa. Karakteristik penilaian tersebut tampaknya cocok digunakan dalam penilaian MTQ Cabang Tilawah. Penilaian dalam MTQ cabang ini di samping dilakukan untuk menilai kemampuan para peserta dalam membaca ayat-ayat al-Quran, juga untuk menilai keterampilan olah vokal atau suara, dan sebagainya.

Selanjutnya, autentisitas yang menjadi ciri dalam penilaian performansi juga terlihat dalam penilaian MTQ Cabang Tilawah. Sebagaimana dikatakan oleh Stiggins (1991: 85) bahwa dalam penilaian unjuk kerja (performansi), guru menghendaki respons siswa yang “*authentic*” atau yang asli berupa aktivitas yang dapat diamati. Dalam MTQ Cabang Tilawah, pengukuran dilakukan secara langsung terhadap performa para peserta dalam melakukan seni baca al-Quran, sehingga hasil pengukuran benar-benar menggambarkan kemampuan asli para peserta tersebut.

Objek dari penilaian performansi adalah segala yang berkaitan dengan ‘*observable performance*’ (performansi yang dapat diamati). *Observable performance* dalam penilaian MTQ Cabang Tilawah tentu berbeda dengan dalam penilaian pembelajaran. Dalam pembelajaran, performa yang memungkinkan untuk diobservasi boleh jadi berkenaan dengan proses kognitif yang kompleks, misalnya melakukan analisis, memecahkan masalah, melakukan percobaan, membuat keputusan, mengukur, bekerja sama dengan yang lain, pernyataan oral, atau mengemukakan suatu produk. Di samping itu, penilaian performansi juga dapat digunakan untuk mengakses cara berpikir (*habit of mind*), cara bekerja, dan perilaku nilai (*behaviors of value*) dari siswa dalam kehidupan nyata. Jenis

penilaian ini sangat sesuai dengan efektivitas pembelajaran (Baker, 1998, p.248).

Dalam penilaian MTQ Cabang Tilawah, *observable performance* meliputi semua aspek yang dinilai dalam lomba yang ada dalam diri para peserta. Hampir seluruh objek penilaian dalam MTQ Cabang Tilawah dapat bersifat *observable*, yakni observasi dalam arti luas, meliputi pengamatan melalui penglihatan, pendengaran dan perasaan. Pengamatan melalui penglihatan dapat dilakukan ketika menilai perilaku peserta MTQ dari awal hingga akhir lomba. Pengamatan melalui penerangan pendengaran dilakukan dalam penilaian aspek tajwid, *fashahah*, suara, dan lagu. Sedangkan pengamatan melalui penerangan perasaan terutama dilakukan dalam menilai aspek seni baca al-Quran. Dalam praktiknya, ketiga jenis pengamatan tersebut dilakukan oleh dewan hakim secara simultan.

Alat penunjang dalam melaksanakan asesmen kinerja pada penilaian MTQ Cabang Tilawah tidak jauh berbeda dengan yang ada dalam asesmen pembelajaran. Jika dalam asesmen pembelajaran, asesmen kinerja menggunakan lembar observasi atau sebuah format pengamatan kinerja atau penampilan siswa, maka dalam asesmen MTQ Cabang Tilawah pun sama. Dalam lembar pengamatan tertera aspek-aspek yang diamati sesuai dengan target yang diinginkan. Berdasarkan deskriptor-deskriptor yang nampak selama proses pengamatan, ditentukanlah skor kinerja peserta dengan berpedoman pada kriteria penilaian yang telah ditetapkan sebelumnya.

Secara garis besar, kriteria penilaian dalam MTQ Cabang Tilawah dapat dibedakan menjadi dua, yaitu: ketepatan bacaan dan keindahan bacaan. Instrumen yang digunakan untuk menilai kedua kriteria tersebut tentu berbeda, karena ketepatan bacaan menyangkut benar atau salahnya bacaan al-Quran yang dilakukan oleh para peserta, sedangkan keindahan bacaan lebih berkaitan dengan unsur seni dalam membaca al-Quran.

Ada empat instrumen yang dikembangkan dalam penelitian ini untuk menilai

kedua kriteria di atas, yaitu: instrumen penilaian tajwid, instrumen penilaian *fashahah*, instrumen penilaian suara, dan instrumen penilaian lagu. Instrumen penilaian tajwid dan *fashahah* digunakan untuk menilai ketepatan bacaan, sedangkan instrumen penilaian suara dan lagu digunakan untuk menilai keindahan bacaan.

Untuk memperoleh gambaran tentang keempat instrumen penilaian unjuk kerja (*performance assessment*) di atas, berikut ini akan diuraikan mengenai tiga hal, yaitu: (1) pengembangan indikator instrumen penilaian tajwid, *fashahah*, suara, dan lagu, (2) pengembangan item-item penskoran dari setiap indikator penilaian tajwid, *fashahah*, suara, dan lagu, dan (3) teknik penskoran dari keempat bidang penilaian MTQ Cabang Tilawah tersebut.

Pengembangan Indikator Instrumen Penilaian Tajwid, Fashahah, Suara, dan Lagu

Penilaian dalam MTQ Cabang Tilawah dilakukan terhadap empat bidang, yaitu: (1) bidang *tajwid*; (2) bidang *fashahah*; (3) bidang suara; dan (4) bidang lagu (Tim LPTQ, 2011, p.54). Keempat bidang inilah yang juga menjadi komponen penilaian MTQ Cabang Tilawah dalam penelitian pengembangan ini. Penilaian terhadap keempat bidang ini dilakukan melalui sejumlah indikator yang menjadi cakupan masing-masing. Dalam penelitian ini, indikator-indikator penilaian tersebut dikembangkan melalui dua cara, yaitu: (1) mendeduksi isyarat-isyarat al-Quran dan Hadis Nabi, dan (2) mengadaptasi indikator-indikator penilaian MTQ Cabang Tilawah yang termuat dalam buku "Pedoman Musabaqah al-Quran".

Mendeduksi isyarat-isyarat al-Quran dan hadis di sini maksudnya menjabarkan pernyataan maupun petunjuk yang diberikan oleh kedua sumber ajaran Islam tersebut mengenai aspek-aspek kualitas tilawah al-Qur'an menjadi rincian indikator penilaian tilawah. Dalam konteks ini, ada sejumlah ayat al-Quran dan hadis yang dapat dijadikan sebagai rujukan untuk merumuskan indikator-indikator penilaian MTQ Cabang Tilawah. Ayat-ayat dimaksud antara lain

surat al-Furqan (25) ayat 32 dan surat al-Muzammil (73) ayat 4 yang memerintahkan umat Islam agar membaca al-Quran secara *tartil*.

Istilah *tartil* secara terminologi berarti membaca al-Quran secara perlahan-lahan dan jelas, mengeluarkan setiap huruf dari *makbraj* atau tempat keluarnya dan menerapkan sifat-sifatnya, serta men-*tadabburi* maknanya (Annuri, 2010, p.30). *Tartil* juga berarti membaca al-Quran dengan menajwidkan huruf-hurufnya dan mengetahui tempat-tempat *waqf* (berhenti) yang benar” (Utsman, 1994, p.49; Nashr, 1349 H, p.7), atau membaca al-Quran dengan memperjelas huruf-hurufnya, berhenti (*waqf*) dan memulai (*ibtida*), sehingga pembaca dan pendengarnya dapat memahami dan menghayati kandungan pesan-pesannya (Shihab, 2002, p.516).

Perlahan-lahan dalam membaca al-Quran mengisyaratkan pentingnya pengaturan tempo bacaan. Kejelasan bacaan, penerapan *waqf* dan *ibtida* secara benar, serta pemahaman dan perenungan makna menunjukkan pentingnya aspek *fashahah*. Sedangkan bacaan yang sesuai dengan ilmu tajwid menunjukkan pentingnya penerapan semua kaidah tajwid dalam membaca al-Quran. Keseluruhan aspek *tartil* ini dapat dijadikan sebagai indikator-indikator penilaian kualitas tilawah, terutama yang menyangkut penilaian bidang tajwid dan *fashahah*, serta sebagian indikator penilaian bidang suara.

Selain *tartil*, aspek keindahan dalam membaca al-Quran juga menjadi komponen yang penting. Hal itu diisyaratkan oleh beberapa hadis Nabi yang memerintahkan agar umat Islam membaca al-Quran dengan suara yang indah, atau dilagukan (Nawawi, 2010, p.343; Al-Baqi (1), 1996, p.236). Keindahan sebuah bacaan ditentukan oleh banyak faktor, antara lain kualitas suara dan lagu/nada. Keindahan bacaan al-Quran juga tidak luput dari kedua faktor tersebut. Bahkan dapat dikatakan bahwa keduanya adalah faktor keniscayaan. Artinya, untuk menghasilkan bacaan al-Quran yang terdengar indah maka pembaca harus memiliki kualiti

tas suara yang bagus dan mampu membawakannya dengan lagu/nada yang serasi. Tanpa kedua aspek tersebut, maka bacaan al-Quran akan terdengar kurang, atau bahkan tidak indah. Dengan demikian, kualitas suara dan lagu akan menentukan keindahan bacaan.

Patut diakui bahwa meskipun al-Quran dan hadis mengisyaratkan komponen suara dan lagu merupakan aspek penting dalam tilawah al-Quran, namun keduanya bukanlah kitab ilmiah yang menjelaskan segala sesuatu secara teoritis dan detail. Oleh karena itu, untuk menentukan indikator-indikator penilaian suara dan lagu diperlukan ijtihad dan kajian terhadap buku-buku yang relevan. Kajian terhadap buku “Pedoman Musabaqah al-Quran 2010” menjadi pilihan yang paling tepat, karena buku ini secara lengkap telah menjelaskan indikator-indikator penilaian dari keempat bidang penilaian MTQ Cabang Tilawah.

Salah satu bagian yang penting dalam buku ini adalah penjelasannya mengenai bidang-bidang penilaian MTQ Cabang Tilawah dan aspek-aspek penilaian yang terdapat di masing-masing bidang tersebut. Dalam buku tersebut dijelaskan tentang aspek-aspek (indikator-indikator) penilaian bidang tajwid, *fashahah*, suara dan lagu. Aspek-aspek penilaian bidang *tajwid* meliputi: (1) *makbarij al-huruf*, (2) *sifat al-huruf*, (3) *abkam al-huruf*; dan (4) *al-madd wa al-qasbr*. Aspek-aspek penilaian bidang *fashahah* meliputi: (1) *al-waqf wa al-ibtida*, (2) *mura`at al-huruf wa al-harakat*, dan (3) *mura`at al-kalimat wa al-ayat* (Tim LPTQ, 2011, p.53). Aspek-aspek penilaian bidang suara meliputi: (1) vokal dan keutuhan suara, (2) kejernihan suara, (3) kehalusan/kelembutan, (4) keserasian nada, dan (5) pengaturan nafas. Sedangkan aspek-aspek penilaian dalam bidang lagu meliputi: (1) lagu pertama dan penutup; (2) jumlah lagu; (3) peralihan; keutuhan dan tempo lagu; (4) irama dan gaya, dan (5) variasi (Tim LPTQ, 2011, pp.53-54).

Bidang-bidang penilaian MTQ Cabang Tilawah dan aspek-aspek penilaian yang ada di dalamnya merupakan hasil ijtihad (curahan pemikiran) para ulama yang ahli di bi-

dangnya selama bertahun-tahun. Meskipun demikian, dalam penelitian ini ada sejumlah aspek (indikator) penilaian di beberapa bidangnya yang perlu diadaptasi agar hasil penilaian MTQ Cabang Tilawah menjadi lebih baik, yakni lebih fair, objektif dan cermat. Aspek-aspek penilaian yang perlu diadaptasi itu terdapat dalam bidang penilaian *fashahah*, suara, dan lagu. Sedangkan aspek-aspek yang ada dalam penilaian bidang tajwid tidak diadaptasi, karena keseluruhan dinilai sudah dapat diterima.

Adaptasi indikator penilaian bidang *fashahah* dilakukan dengan meniadakan salah satu aspek (indikator) penilaiannya, yakni *mura`at al-kalimat wa al-ayat*. Peniadaan indikator ini dilakukan karena penilaian dalam model pengembangan ini berbasis pada huruf, harakat dan elemen-elemen *maqra`* (ayat) lainnya secara detail, bukan berbasis pada kalimat atau ayat yang terdapat dalam *maqra`* secara global. Model pengembangan ini memandang bahwa aspek *mura`at al-kalimat wa al-ayat* sudah tercakup dalam aspek *mura`at al-huruf wa al-barakat*. Karena setiap kesalahan pada *mura`at al-kalimat wa al-ayat* dihitung berdasarkan jumlah huruf dan harakat yang terdapat di dalam kalimat atau ayat tersebut.

Adaptasi indikator penilaian bidang suara dilakukan dengan cara menambah dan menyelaraskan sebagian indikator yang ada pada bidang tersebut. Dalam “Pedoman Musabaqah al-Quran 2010” tidak disebutkan bahwa kenyaringan suara merupakan indikator penilaian bidang suara. Aspek kenyaringan suara ini sangat penting, sehingga ia dijadikan sebagai salah satu indikator penilaian suara dalam model penilaian yang dikembangkan dalam penelitian ini. Sedangkan aspek keserasian nada yang dalam buku “Pedoman Musabaqah al-Quran 2010” dijadikan sebagai salah satu indikator penilaian suara (Tim LPTQ, 2011, p.54) diubah menjadi salah satu indikator penilaian bidang lagu. Peneliti menilai bahwa keserasian nada lebih cocok menjadi salah satu indikator penilaian bidang lagu, karena dalam penilaian bidang lagu terdapat aspek peralihan lagu serta irama dan gaya sebagai bagian dari

indikatornya. Penilaian terhadap kedua indikator itu tidak bisa dilepaskan dari unsur keserasian nada.

Adaptasi indikator penilaian bidang lagu dilakukan dengan memilah indikator ke-3, yakni: peralihan, keutuhan dan tempo lagu (Tim LPTQ, 2011, pp.53-54) menjadi tiga indikator yang terpisah dari yang lain. Alasan pemilahan itu adalah bahwa indikator-indikator yang tersebut dalam point ke-3 merupakan substansi yang berbeda, yang masing-masing berdiri sendiri. Peralihan lagu menyangkut kemampuan peserta dalam melakukan perpindahan dari lagu yang satu ke lagu yang lain. Kemudian keutuhan lagu menyangkut kemampuan peserta dalam membawakan setiap lagu secara utuh atau lengkap, yaitu: dengan nada rendah, sedang dan tinggi. Sedangkan tempo lagu menyangkut kemampuan peserta dalam mengatur ritme lagu apakah cepat, sedang ataukah lambat. Dengan demikian, indikator penilaian bidang lagu dalam model penilaian hasil pengembangan berjumlah tujuh, yaitu: (1) lagu pertama dan penutup; (2) jumlah lagu; (3) peralihan lagu; (4) keutuhan lagu; (5) tempo lagu; (6) irama dan gaya; dan (7) variasi lagu. Ketujuh indikator tersebut merupakan pengembangan dari lima indikator penilaian lagu yang selama ini digunakan dalam MTQ Cabang Tilawah.

Pengembangan Item-Item Penskoran dari Setiap Indikator Penilaian Tajwid, Fashahah, Suara, dan Lagu

Penilaian dalam MTQ Cabang Tilawah di samping dilakukan terhadap ketepatan bacaan juga dilakukan terhadap keindahan bacaan. Penilaian terhadap ketepatan bacaan al-Quran diaktualisasikan dalam bentuk penilaian tajwid dan *fashahah*. Sedangkan penilaian terhadap keindahan bacaan diaktualisasikan dalam bentuk penilaian suara dan lagu.

Penilaian dalam MTQ Cabang Tilawah terutama yang menyangkut ketepatan bacaan dilakukan terhadap semua elemen bacaan yang termuat dalam *maqra`*. Dengan demikian, item-item pengukuran dalam bidang tajwid dan *fashahah* adalah keseluruhan hu-

ruf, harakat, tanda waqf, dan semua elemen bacaan lainnya yang secara riil terhimpun dalam ayat-ayat yang dibaca oleh peserta lomba. Sedangkan item-item penskoran bidang suara dan lagu tentu bukan berupa elemen bacaan yang termuat dalam *maqra'*, tetapi berupa sejumlah kriteria yang dikembangkan berdasarkan prosedur pembuatan angket berskala empat untuk setiap indikator penilaiannya.

Perbedaan strategi pengembangan item-item penilaian bidang tajwid dan *fasabab* di satu sisi serta item-item penilaian bidang suara dan lagu di sisi lain dilakukan karena adanya perbedaan karakter penilaian di kedua kelompok tersebut. Penilaian

bidang tajwid dan *fasabab* dilakukan berdasarkan kriteria benar dan salah, karena penilaian di kedua bidang ini memang dilakukan untuk menilai benar atau tidaknya para peserta MTQ Cabang Tilawah dalam membaca *maqra'* dalam lomba. Sedangkan penilaian bidang suara dan lagu yang menilai aspek keindahan bacaan tidak bisa atau tidak cocok menggunakan kriteria benar dan salah. Penilaian di kedua bidang ini lebih cocok didasarkan pada kriteria bagus atau kurang bagus, indah atau kurang indah dan sebagainya. Untuk mendapatkan gambaran yang lebih jelas, berikut ini disajikan kisi-kisi penilaian dari keempat bidang penilaian MTQ Cabang Tilawah.

Tabel 1. Kisi-kisi Instrumen Penilaian MTQ Cabang Tilawah Hasil Pengembangan

Bidang Penilaian	Indikator Penilaian	Butir Penilaian
1. Tajwid	a. <i>Makbarij al-huruf</i>	Ketepatan membaca semua huruf berharakat yang terhimpun dalam <i>maqra'</i> menurut tempat keluarnya
	b. <i>Sifat al-huruf</i>	Ketepatan membaca semua huruf berharakat yang terhimpun dalam <i>maqra'</i> menurut sifat keluarnya
	c. <i>Abkam al-huruf</i>	1) Ketepatan membaca hukum nun sukun dan tanwin 2) Ketepatan membaca hukum mim sukun 3) Ketepatan membaca <i>mufabbham</i> dan <i>muraqqiq</i> 4) Ketepatan membaca <i>alif lam qamariyah</i> dan <i>syamsiyah</i> 5) Ketepatan membaca <i>imalah</i> 6) Ketepatan membaca <i>isyamam</i>
	d. <i>Hukum al-mad wa al-qasr</i>	1) Ketepatan membaca huruf-huruf ber- <i>madd</i> 2) Ketepatan membaca huruf-huruf <i>qasr</i> (tidak ber- <i>madd</i>)
2. <i>Fasabab</i>	a. <i>Al-Waqf wa al-ibtida'</i>	1) Ketepatan melakukan <i>waqf</i> 2) Ketepatan melakukan <i>ibtida'</i> (memulai bacaan setelah <i>waqf</i>)
	b. <i>Mura'at al-huruf wa al-harakat</i>	1) Ketepatan membaca semua huruf yang ada dalam <i>maqra'</i> 2) Ketepatan membaca semua harakat yang ada dalam <i>maqra'</i> 3) Keterhindaran dari penambahan/pengurangan huruf/harakat
	3. Suara	a. Kejernihan suara Kemampuan bersuara jernih dan tingkat kestabilannya b. Kehalusan suara Kemampuan bersuara halus dan tingkat kestabilannya c. Keutuhan suara Kemampuan bersuara utuh dan tingkat kestabilannya d. Kenyaringan suara Kemampuan bersuara nyaring pada nada-nada tinggi dan tingkat kestabilannya e. Pengaturan nafas Kemampuan mengatur nafas dan tingkat kestabilannya
4. Lagu	a. Lagu pertama dan penutup	Kemampuan memulai dan menutup tilawah dengan lagu bayati/husaini
	b. Jumlah lagu	Kemampuan tilawah minimal dengan 5 buah lagu
	c. Peralihan lagu	Kemampuan melakukan peralihan lagu dengan serasi
	d. Keutuhan lagu	Kemampuan membawakan lagu secara utuh
	e. Tempo lagu	Kemampuan dalam mengatur tempo lagu
	f. Irama dan gaya	Kemampuan tilawah dengan irama dan gaya yang indah
	g. Variasi lagu	Kemampuan dalam melakukan variasi lagu

Pengembangan Teknik Penskoran

Penilaian dalam MTQ Cabang Tilawah tidak bisa dilepaskan dari kegiatan penskoran. Penskoran dilakukan untuk menyatakan kemampuan tilawah setiap peserta. Peserta yang mendapatkan skor besar menunjukkan bahwa kemampuan tilawahnya sangat baik. Sebaliknya, peserta yang mendapatkan skor kecil menunjukkan bahwa kemampuan tilawahnya tidak baik. Menyatakan kemampuan tilawah dengan angka-angka (skor) lebih jelas dan meyakinkan daripada menyatakan secara kualitatif.

Penskoran dalam penilaian MTQ Cabang Tilawah dilakukan untuk setiap indikatornya. Oleh karena itu, penelitian ini mengembangkan formula yang dapat digunakan untuk mengkomputasi skor-skor dari indikator tersebut. Skor-skor untuk indikator-indikator penilaian tajwid dan *fashahah* dikomputasi dengan menggunakan formula sebagai berikut:

$$x_i = \frac{\sum m_i - \sum m_e}{\sum m_i}$$

di mana x_i adalah notasi untuk skor indikator i , $\sum m_i$ adalah jumlah seluruh item di setiap indikator, dan $\sum m_e$ adalah jumlah item di setiap indikator yang dibaca salah.

Adapun skor untuk indikator-indikator penilaian bidang suara dan lagu tidak dikomputasi dengan menggunakan formula khusus. Skor untuk setiap indikator di kedua bidang penilaian tersebut dapat diperoleh secara langsung dari instrumen penilaian yang menyerupai angket berskala empat.

Validitas dan Reliabilitas

Uji coba dalam penelitian ini dilakukan terutama untuk memperoleh informasi tentang kelayakan semua instrumen hasil pengembangan untuk digunakan dalam penilaian MTQ Cabang Tilawah. Kelayakan instrumen dapat diketahui melalui pembuktian validitas dan estimasi reliabilitas instrumen tersebut. Jika pembuktian validitas dan estimasi reliabilitas menghasilkan nilai koefisien yang lebih besar atau minimal sama dengan kriteria koefisien yang ditetapkan,

maka instrumen penilaian ter-sebut dapat dikatakan valid dan reliabel, sehingga layak untuk digunakan. Tetapi jika nilai koefisien yang dihasilkan lebih kecil dari kriteria koefisien yang ditetapkan, maka instrumen tersebut dapat dikatakan tidak valid dan tidak reliabel, sehingga tidak layak untuk digunakan.

Penetapan validitas sebuah instrumen di samping bisa dilakukan melalui komputasi skor-skor hasil uji coba dengan menggunakan persamaan yang telah dirumuskan oleh para pakar pengukuran, juga bisa dilakukan melalui *judgment* pakar. Dalam penelitian ini, penetapan validitas instrumen hasil pengembangan dilakukan melalui *judgment* pakar. Penetapan validitas seperti ini dilakukan melalui proses analisis rasional mengenai isi instrumen sebuah tes, yang penentuannya didasarkan pada pendapat individu pakar, pendapat subjektif (Allen & Yen, 1979, p.95). Penetapan validitas yang demikian juga disebut dengan istilah *professional judgment* (Azwar, 2009, p.45).

Alasan penggunaan prosedur *judgment* pakar adalah bahwa penetapan indikator-indikator dan item-item penilaian dalam MTQ Cabang Tilawah memerlukan pendapat atau penilaian orang-orang yang ahli di bidangnya. Dalam hal ini adalah para pakar al-Quran atau para hakim penilai MTQ Cabang Tilawah di berbagai bidangnya.

Dalam penelitian ini, *judgment* terhadap konten dari instrumen-instrumen hasil pengembangan dilakukan oleh sejumlah pakar al-Quran, hakim penilai dan praktisi MTQ yang ada di lingkungan UIN Sunan Gunung Djati Bandung dan LPTQ Provinsi Jawa Barat. Adapun para pakar dimaksud adalah: (1) H. Abdul Mujib, M.Ag. (hakim dan praktisi MTQ), (2) Abdul Rozak, M.Ag. (hafizh dan praktisi MTQ), (3) H. Wawan Kusnawan, S.Pd.I (praktisi MTQ), (4) Dra. Hj. Farida (hakim dan praktisi MTQ), (5) Drs. Ending Syarifuddin, M.Ag (pakar al-Quran), (6) Drs. H. Dadang Suherman (hakim dan praktisi MTQ), (7) Salim Slamet, S.Pd.I (praktisi MTQ), dan (8) Drs. Saiful Anwar, M.Ag. (hafizh dan pakar al-

Quran). Keseluruhan pakar ini menilai bahwa instrumen penilaian tajwid, *fashahah*, suara, dan lagu yang dikembangkan dalam penelitian ini layak untuk digunakan.

Berbeda dengan estimasi validitas di atas, estimasi reliabilitas instrumen hasil pengembangan dilakukan melalui komputasi skor-skor hasil uji coba dengan menggunakan aplikasi program SPSS. Estimasi reliabilitas instrumen yang dikembangkan dalam penelitian ini menggunakan pendekatan konsistensi internal. Pendekatan konsistensi internal bisa dikatakan paling praktis di antara pendekatan-pendekatan yang lain, karena hanya menggunakan tes tunggal. Meskipun demikian, prosedur yang ditempuh bukan dengan cara membelah skor hasil uji coba menjadi beberapa bagian (*split-half method*), sebagaimana yang lazim ditempuh dalam pendekatan ini, tetapi dengan melalui analisis varians (ANOVA).

Estimasi reliabilitas melalui analisis varians dapat dikatakan logis terutama jika mengingat bahwa konsepsi reliabilitas sendiri memang merupakan rasio dari berbagai varians distribusi (Azwar, 2009, p.92). Adapun prosedur analisis varians dalam penelitian ini didasarkan pada teori generalisabilitas, atau yang lebih dikenal dengan istilah *generalizability theory* (GT). Sedangkan desain yang digunakan adalah model multifacet umum (*the general multifacet model*).

Varians-variens analisis dalam penelitian ini bersumber dari tiga variasi facet yang terlibat dalam penilaian MTQ Cabang Tilawah, yaitu: peserta, rater, dan butir. Menurut Thorndike (1982, p.161), variasi dari tiga facet ini dapat menghasilkan estimasi tujuh komponen varians yang berbeda. Dalam konteks penelitian ini, ketujuh komponen varians tersebut adalah σ_{pbr}^2 (variens peserta x butir x rater), σ_{pb}^2 varians peserta x butir, σ_{pr}^2 (variens peserta x rater), σ_{br}^2 (variens butir x rater), σ_p^2 varians peserta, σ_b^2 (variens butir), σ_r^2 (variens rater) (Thorndike, 1982, pp.161-163).

Prosedur estimasi reliabilitas instrumen penilaian MTQ Cabang Tilawah dilakukan dengan cara mencari harga dari varians-variens di atas. Harga-harga varians

tersebut dapat dikomputasi melalui kuadrat rata-rata (*mean square*) dari setiap variasi facetnya. Untuk mendapatkan hasil komputasi yang lebih akurat, penentuan kuadrat rata-rata (*mean square*) dimaksud dalam penelitian ini menggunakan aplikasi program SPSS.

Kegiatan yang dilakukan sebelum meng-*input* semua skor penilaian MTQ Cabang Tilawah ke dalam aplikasi program SPSS adalah menyamakan skala penilaian dari keempat bidang penilaian MTQ (tajwid, *fashahah*, suara dan lagu), karena penilaian terhadap keempatnya menggunakan skala yang berbeda. Penyamaan skala dilakukan dengan mengkonversi skor-skor hasil penilaian di keempat bidang tersebut menjadi *z-score*. Pengkonversian skor tersebut di samping untuk menyamakan skala, juga untuk memperoleh varians skor komposit (skor gabungan tajwid, *fashahah*, suara, dan lagu) yang sama.

Skor-skor yang telah dikonversi tersebut kemudian diinput ke dalam aplikasi program SPSS. Dengan melalui prosedur analisis *General Linear Model Univariate*, maka diperoleh *mean square* dari penilaian MTQ Cabang Tilawah disajikan pada Tabel 2.

Tabel 2. *Mean Square* Skor Gabungan Hasil Uji Coba Model Penilaian MTQ

<i>Source</i>	<i>Sum of Squares (SS)</i>	<i>Degrees of Freedom (DF)</i>	<i>Mean Square (MS)</i>
Peserta	494.500	29	17.052
Butir	0.940	14	0.067
Rater	0.317	8	0.040
P x B	224.386	406	0.553
P x R	97.359	232	0.420
B x R	1.879	28	0.067
P x B x R	207.145	812	0.255

Nilai-nilai *mean square* yang terdapat dalam Tabel 2 tersebut dapat digunakan untuk menghitung harga dari ketujuh macam varians yang terdapat dalam pengukuran ini. Adapun harga varians-variens tersebut sebagai berikut:

$$\begin{aligned}\sigma_{pqr}^2 &= MS_{pqr} \\ \sigma_{pq}^2 &= \frac{1}{n_r} (MS_{pq} - MS_{pqr}) \\ \sigma_{pr}^2 &= \frac{1}{n_q} (MS_{pr} - MS_{pqr}) \\ \sigma_{qr}^2 &= \frac{1}{n_p} (MS_{qr} - MS_{pqr}) \\ \sigma_p^2 &= \frac{1}{n_q n_r} (MS_p - MS_{pq} - MS_{pr} + MS_{pqr}) \\ \sigma_q^2 &= \frac{1}{n_p n_r} (MS_q - MS_{pq} - MS_{qr} + MS_{pqr}) \\ \sigma_r^2 &= \frac{1}{n_p n_q} (MS_r - MS_{pr} - MS_{qr} + MS_{pqr})\end{aligned}$$

Jika nilai setiap *mean square* yang terdapat dalam Tabel 2 dimasukkan ke dalam persamaan di atas, maka akan diperoleh harga setiap varians skor gabungan hasil uji coba model penilaian MTQ sebagai berikut:

$$\begin{aligned}\sigma_{pqr}^2 &= 0,255 \\ \sigma_{pq}^2 &= \frac{1}{12} (0,553 - 0,255) = 0,02483 \\ \sigma_{pr}^2 &= \frac{1}{18} (0,420 - 0,255) = 0,0092 \\ \sigma_{qr}^2 &= \frac{1}{30} (0,067 - 0,255) = -0,0063 \\ \sigma_p^2 &= \frac{1}{216} (17,052 - 0,553 - 0,420 + 0,255) = 0,07562 \\ \sigma_q^2 &= \frac{1}{360} (0,067 - 0,553 - 0,067 + 0,255) = -0,00083 \\ \sigma_r^2 &= \frac{1}{540} (0,040 - 0,420 - 0,067 + 0,255) = -0,000356\end{aligned}$$

Sebagaimana telah disinggung di atas, bahwa estimasi reliabilitas melalui analisis varians (ANOVA) merupakan rasio antar berbagai varians distribusi. Dalam konteks ini, reliabilitas adalah rasio atau perbandingan antara varians skor murni dan varians skor amatan (Thorndike, 1982, pp.163-164). Dengan demikian, koefisien reliabilitas dalam penelitian ini adalah besarnya harga koefisien yang diperoleh dari perbandingan antara varians skor murni dan varians skor amatan dari sebuah tes. Berdasarkan formula ini, maka koefisien reliabilitas skor gabungan adalah harga yang diperoleh dari perbandingan antara varians skor murni ga-

bungan dan varians skor amatan gabungan dari empat bidang penilaian MTQ Cabang Tilawah, yakni tajwid, *fashahah*, suara dan lagu.

Bertolak dari sini, reliabilitas skor gabungan hasil uji coba model penilaian MTQ Cabang Tilawah dapat diestimasi dengan menggunakan persamaan sebagai berikut:

$$r_{xx'} = \frac{\sigma_{true}^2}{\sigma_{obs}^2}$$

Komponen varians skor murni dan varians skor amatan dimaksud dapat diuraikan sebagai berikut:

$$\frac{\sigma_{true}^2}{\sigma_{obs}^2} = \frac{\sigma_p^2}{\sigma_p^2 + \frac{\sigma_q^2}{n_q} + \frac{\sigma_r^2}{n_r} + \frac{\sigma_{pq}^2}{n_q} + \frac{\sigma_{pr}^2}{n_r} + \frac{\sigma_{qr}^2}{n_q n_r} + \frac{\sigma_{pqr}^2}{n_q n_r}}$$

Dengan memasukkan harga setiap varians skor gabungan di atas ke dalam persamaan ini diperoleh koefisien reliabilitas gabungan sebagai berikut:

$$r_{xx'} = \frac{0,0756}{0,789} = 0,96$$

Dari perbandingan antara varians skor murni dan varians skor amatan di atas diperoleh harga koefisien reliabilitas skor gabungan hasil uji coba model penilaian MTQ Cabang Tilawah sebesar **0,96**. Dengan berpegang pada patokan konvensional tentang kriteria reliabilitas sebuah tes yang ditetapkan sebesar $r \geq 0,7$, maka dapat disimpulkan bahwa instrumen penilaian yang dikembangkan dalam penelitian secara keseluruhan adalah reliabel. Dengan demikian, instrumen-instrumen tersebut layak untuk digunakan dalam penilaian MTQ Cabang Tilawah, baik yang diselenggarakan oleh pemerintah melalui LPTQ maupun oleh institusi lainnya.

Simpulan

Pengembangan model asesmen MTQ Cabang Tilawah memerlukan pengkajian yang mendalam terhadap dua aspek, yaitu: (1) pedoman penilaian MTQ Cabang Tila-

wah yang biasa digunakan selama ini, dan (2) teori pengukuran atau asesmen sebagai landasan pengembangan model. Pengkajian terhadap pedoman penilaian yang biasa digunakan selama ini dilakukan untuk mengetahui karakteristik penilaian MTQ Cabang Tilawah dan kelemahan-kelemahan yang terkandung di dalamnya. Sedangkan pengkajian terhadap teori pengukuran dilakukan untuk mengetahui berbagai prinsip pengukuran dan penilaian yang dapat digunakan sebagai dasar untuk mengembangkan model penilaian MTQ Cabang Tilawah.

Pengembangan model asesmen MTQ Cabang Tilawah ini dilakukan melalui beberapa langkah, yaitu: (1) menetapkan *maqra'* tunggal dalam penilaian MTQ Cabang Tilawah; (2) melakukan perubahan terhadap instrumen penilaian MTQ Cabang Tilawah dengan melakukan perincian setiap indikator menjadi item-item yang secara operasional dapat diukur; (3) melakukan perubahan terhadap teknik penilaian MTQ Cabang Tilawah melalui tiga langkah, yaitu: melakukan penilaian terhadap seluruh elemen bacaan (item) untuk mengganti penilaian yang sebelumnya dilakukan terhadap sebagai elemen bacaan saja, menyesuaikan besaran pengurangan skor dengan jumlah elemen bacaan yang menjadi item-item penskoran agar penilaian MTQ Cabang Tilawah menjadi proporsional, dan merubah skor maksimal di setiap bidang penilaiannya menjadi bobot skor; (4) meniadakan konversi nilai dengan menjadikan skor yang diberikan para hakim penilai sebagai skor final.

Saran

Berdasarkan simpulan di atas, maka peneliti merekomendasikan kepada pembuat kebijakan (LPTQ Pusat) agar melakukan beberapa langkah, yaitu: (1) mengkaji ulang model penilaian MTQ yang digunakan selama ini dengan mendatangkan para pakar pengukuran dan penilaian, (2) membarui model penilaian MTQ dengan menerapkan teori pengukuran dan penilaian, (3) melakukan pelatihan secara berkala untuk para

hakim atau para calon hakim MTQ mengenai teknik-teknik penilaian yang benar, (4) menjadikan para hakim MTQ sebagai bagian dari pembaruan model penilaian MTQ agar proses dan hasil penilaian yang dilakukannya lebih berkualitas dari sebelumnya.

Daftar Pustaka

- Al-Baqi, Muhammad Fuad Abd. (1996). *Al-lu'lu' wal marjan: himpunan hadis shahih yang disepakati oleh Bukhari dan Muslim*. (Terjemahan Salim Bahreisy). Surabaya: PT. Bina Ilmu
- Annuri, Ahmad. (2010). *Panduan tafsir tilawah al-Quran & ilmu tajwid*. Jakarta: Pustaka al-Kautsar.
- Allen, M.J., & Yen, W.M. (1979). *Introduction to measurement theory*. Monterey, California: Brooks/Cole Publishing Company.
- Al-Nawawi, Yahya ibn Syarifuddin. (1984). *Al-tibyan fi adab hamalat al-Quran*. Bairut: Dar al-Nafais.
- Azwar, Saifuddin. (2009). *Reliabilitas dan validitas*. Yogyakarta: Pustaka Pelajar.
- Baker, E.L. (1998). *Model-based performance assessment*. Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Borg, W.R. & Gall, M.D. (1983). *Educational research: An introduction (4th Ed.)*. New York: Longman.
- Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. Holt, Rinehart & Winston.
- Departemen Agama RI. (2005). *Al-Quran dan terjemahnya*. Jakarta: P.T. Arga Publishing.
- Ebel, R.L. & Frisbie, D.A. (1986). *Essentials of educational measurement*. New Jersey: Prentice-Hall, Inc., Englewood Cliffs.
- Frishkopf, Michael. (2007). *Mediated qur'anic recitation and the contestation of Islam*

- contemporary Egypt*. Diakses tanggal 5 April 2011 dari <http://www.islamworld.net/historic>.
- Griffin, P. & Nix, P. (1991). *Educational assessment and reporting*. London: Harcourt Brace Javanovich Publisher.
- Linn, R.L. (1990). *Measurement and evaluation in teaching*. New York: Macmillan Publishing Company.
- Nashr, Al-Syaikh Muhammad Makki. (1349). *Nihayat al-qaul al-mufid fi 'ilm al-tajwid*. Surabaya: Dar al-'Ulum al-Islamiyyah.
- Shihab, Quraish M. (2002). *Wawasan al-Quran* Jakarta: Lentera Hati.
- Stiggins, R.J. (1994). *Student-centered classroom assessment*. New York: Macmillan Publishing Company.
- Thorndike, R.L. (1982). *Applied psychometrics*. Boston: Houghton Mifflin Company.
- Tim LPTQ. (2011). *Pedoman musabaqah al-Quran 2010*. Jakarta: LPTQ.
- Utsman, Husni Syaikh. (1994). *Haq al-tilawah*. Jeddah-Mekah: Dar al-Manarah li al-Nasyr wa al-Tauzi`.