

جودة أسئلة اللغة العربية في اختبار المدرسة الأخيرة للفصل التاسع في المدرسة المتوسطة منبع القرآن
بمدينة موجوكرتو: تحليل بنود الأسئلة

Kualitas Soal Bahasa Arab Ujian Madrasah Siswa Akhir Kelas IX di MTs Manba'ul Qur-an Kota Mojokerto: Analisis Butir Soal

Ahmad Nuruddin¹⁾, Elok Rufaiqoh²⁾, Achmad Ja'far Sodik³⁾, Muhammad Arif⁴⁾

^{1,4)}Institut Agama Islam Sunan Kalijogo Malang ²⁾, Institut Agama Islam Al-Qodiri, Jember, Indonesia, ³⁾Universitas Islam Negeri Fatmawati Sukarno, Bengkulu, Indonesia

¹⁾ahmadnuruddin123@gmail.com, ²⁾elokrufaiqoh90@gmail.com, ³⁾sodik@uinfasbengkulu.ac.id,
⁴⁾ariv.muh2614@gmail.com

Abstrak. Kegiatan penilaian merupakan salah satu cara untuk mengetahui keberhasilan pembelajaran. Untuk memastikan kualitas alat penilaian yang digunakan, perlu dilakukan analisis. Penelitian ini bertujuan untuk mengkaji kualitas soal-soal penilaian pada ujian akhir sekolah kelas IX di SMP Sumber Ilmu Kota Mojokerto tahun pelajaran 2022-2023 dalam bentuk analisis butir soal. Jenis penelitian ini adalah penelitian deskriptif kuantitatif. Data yang digunakan dalam penelitian ini adalah soal ujian akhir sekolah, kunci jawaban, dan hasil jawaban siswa yang diambil dengan teknik dokumen. Data yang dianalisis digunakan untuk menentukan tingkat kesukaran soal dan membedakan daya pembeda soal dengan bantuan software Anates versi 4.0.2. Hasil dari penelitian ini ada tiga. Pertama, hasil analisis tingkat kesukaran soal adalah sangat mudah (13,3%), mudah (13,3%), sedang (66,7%), sukar (3,3%), dan sangat sukar (3,3%). Kedua, hasil analisis daya beda soal memperoleh kategori sangat jelek sebesar 6,7%. Persentase soal pada kategori jelek sebesar 16,7%, soal pada kategori cukup sebesar 3,3%, soal pada kategori baik sebesar 16,7%, dan soal pada kategori sangat baik sebesar 56,7%. Ketiga, hasil analisis daya pembeda soal masuk dalam kategori sangat jelek dengan persentase sebesar 13,3%. Persentase soal pada kategori jelek sebesar 16,7%, soal pada kategori cukup sebesar 10%, soal pada kategori baik sebesar 16,7%, dan soal pada kategori sangat baik sebesar 43,3%. Dengan demikian, hasil meta-analisis butir soal yang dapat digunakan sebanyak 21 butir soal (70%) dan 9 butir soal (30%) tidak dapat digunakan.

Kata kunci: Kualitas soal, sekolah menengah, butir soal.

مستخلص البحث

تعتبر أنشطة التقييم إحدى طرق معرفة نجاح التعلم. ولضمان جودة أدوات التقييم المستخدمة، يجب إجراء التحليل. يهدف هذا البحث إلى فحص جودة أسئلة تقييم في اختبار المدرسة الأخيرة للفصل التاسع في المدرسة المتوسطة منبع القرآن بمدينة موجوكرتو سنة

دراسية 2022-2023 م في شكل تحليل للفقرات. ونوع هذا البحث هو البحث الوصفي الكمي. البيانات المستخدمة في هذا البحث هي أسئلة اختبار المدرسة الأخيرة، ومفاتيح الإجابة، ونتائج إجابات الطلاب المأخوذة باستخدام تقنيات المستندات. وتم استخدام البيانات التي تم تحليلها لتحديد مستوى صعوبة الأسئلة وتمييز القوة والمشتتات بمساعدة برنامج أناتيس (Anates) الإصدار 4-0-2. وتتضمن نتائج هذا البحث ثلاثة أشياء. أولاً، جاءت نتائج تحليل مستوى صعوبة الأسئلة في الفئة السهلة جداً بنسبة 13.3%، الفئة السهلة بنسبة 13.3%، الفئة المتوسطة بنسبة 66.7%، الفئة الصعبة بنسبة 3.3%، والفئة الصعبة جداً بنسبة 3.3% ثانياً، حصلت نتائج تحليل القوة التفاضلية للأسئلة على فئة ضعيفة جداً بنسبة 6.7%. وبلغت نسبة الأسئلة في الفئة السيئة 16.7%، والأسئلة في الفئة العادلة 3.3%، والأسئلة في الفئة الجيدة 16.7%، والأسئلة في الفئة الجيدة جداً 56.7%. ثالثاً، جاءت نتائج التحليل التشتيت في فئة الفقراء جداً بنسبة 13.3%. وبلغت نسبة الأسئلة في الفئة السيئة 16.7%، والأسئلة في الفئة العادلة 10%، والأسئلة في الفئة الجيدة 16.7%، والأسئلة في الفئة الجيدة جداً 43.3%. وبذلك كانت نتائج التحليل التلخيصي للفقرات التي يمكن استخدامها 21 سؤالاً أي 70% والأسئلة التي لا يمكن استخدامها 9 أسئلة أي 30%.

الكلمة المفتاحية: جودة أسئلة، المدرسة المتوسطة منبع القرآن، بنود الأسئلة.

المقدمة

إن من أحد المكونات التي تدعم العملية التعليمية هو التقييم، بل إنه من الأشياء المهمة فيها. يتطلب التعليم التقييم كوسيلة أو نشاط لمراقبة وضمان وتحديد جودة التعليم. يتم إجراء التقييم بشكل شامل في كل مسار أو مستوى تعليمي، مما يجعل التقييم أحد روح التعليم. بمعنى آخر، عندما نتحدث عن التعليم، لا يمكن فصلنا عن التقييم التربوي.

يُفترض دائماً أن التقييم التعليمي في إندونيسيا هو الخطوة الأخيرة في العملية التعليمية أو على وجه التحديد في التعلم. ومع ذلك، هذا ليس صحيحاً تماماً لأن هناك العديد من التقييمات التي يتم

إجراؤها على مراحل. نحو التقييم الذي يتم إجراؤه في منتصف الفصل الدراسي هو تقييم المدرسة (AM). وهنا يمكننا أن نرى أن استخدام التقييم المناسب سيكون قادرًا على قياس نتائج التعلم بدقة. وليس هذا فحسب، بل يمكن استخدام ذلك كوسيلة لتطوير الأسئلة المطبقة¹ ومقارنة جودة الأسئلة المطبقة.²

تؤكد عدة آراء على مفهوم التقييم. ويهدف التقييم الذي يتم إجراؤه إلى تحديد مدى تحقيق الأهداف التعليمية.³ وهذا ما أكد عليه كورنيوان حيث ذكر أن الأهداف التعليمية يمكن تحقيقها بمكونات متعددة، أحدها التقييم الذي يعد جزءًا لا يتجزأ من العملية التعليمية.⁴ فإن التقييم هو نشاط أو عملية تحديد قيمة التعليم، بحيث يمكن معرفة الجودة أو النتائج.⁵ وتمشيا مع ريشا⁶ يعد التقييم عنصرا هاما في التعليم، وخاصة تقييم نتائج التعلم. ومن الأشياء التي يتم قياسها في التقييم هو مستوى ذكاء الطلاب. وبذلك يصبح التقييم أحد المكونات التي تساعد المعلمين والمدارس على تحديد نتائج التعلم.

يمكن أن تختلف أشكال التقييم المستخدمة، على سبيل المثال في شكل اختبارات وغير اختبارات مع أدوات تقييم متنوعة،⁷ بما في ذلك الأسئلة أو الأسئلة والواجبات المصممة.⁸ وهكذا يمكننا أن نرى أن الاختبارات هي أداة التقييم الأكثر شعبية لأن جميع المعلمين تقريبًا يستخدمونها. يمكن أيضًا أن يختلف شكل الاختبار، على سبيل المثال في شكل اختبارات متعددة ومقالات أو أوصاف. من الناحية العملية، غالبًا ما يستخدم المعلمون الأسئلة المقالية أثناء الاختبارات اليومية وأسئلة الاختيار من متعدد أثناء تقييمات منتصف الفصل الدراسي أو الفصل الدراسي النهائي. يتم ذلك لجعل الاختبار أكثر كفاءة.

¹ Kiromim Muhson, Ali; Lestari, Barkah; Supriyanto, Supriyanto; Baroroh, "The Development of Practical Item Analysis Program for Indonesian Teachers," *International Journal of Instruction*, 2017, 199-210, doi:10.12973/iji.2017.10213a.

² N. S. Subali, B., Kumaidi. & Aminah, "The Comparison of Item Test Characteristics Viewed from Classic and Modern Test Theory," *International Journal of Instruction* 14 (2021), <https://doi.org/10.29333/iji.2021.14139a>.

³ Ali Hamzah, *Evaluasi Pembelajaran Matematika* (Jakarta: Rajawali Pers, 2014).

⁴ Tutut Kurniawan, "Analisis Butir Soal Ulangan Akhir Semester Gasal Mata Pelajaran IPS Sekolah Dasar," *Journal of Elementary Education* 4, no. 1 (2015), <http://journal.unnes.ac.id/sju/index.php/jee>.

⁵ M. & R Matsna, *Evaluasi Pembelajaran Bahasa Arab* (Ciputat: UIN Press, 2015).

⁶ U Ridho, "Evaluasi Dalam Pembelajaran Bahasa Arab," *An Nabighoh: Jurnal Pendidikan Dan Pembelajaran Bahasa Arab* 20, no. 01 (2018), <https://doi.org/https://doi.org/10.32332/annabighoh.v20i01.1124>.

⁷ M Purwati, "Analisis Butir Soal Ujian Akhir Mata Pelajaran Akuntansi Keuangan Menggunakan Microsoft Office Excel 2010," *Jurnal Pendidikan Akuntansi Indonesia*, n.d., <https://doi.org/10.21831/jpai.v12i2.2710>.

⁸ Kurniawan, "Analisis Butir Soal Ulangan Akhir Semester Gasal Mata Pelajaran IPS Sekolah Dasar."

تتطلب الأسئلة المستخدمة في التقييمات المدرسية أيضًا التحقق من الصحة أو التحليل الذي يوضح ما إذا كانت الأسئلة جيدة أم سيئة. إحدى الطرق هي إجراء تحليل للأسئلة التي تم تطبيقها. ويمكن التعرف على جودة الأسئلة التي تم تطبيقها باستخدام تحليل بنود الأسئلة⁹ وكذلك تحسين جودة النتائج هالادينا ورودريجز. وبصرف النظر عن ذلك، يمكن استخدامه أيضًا كمقارنة لتحديد عناصر الأسئلة المختارة.¹⁰ تشمل الأشياء التي يتم تحليلها مستوى الصعوبة، وقوة التمييز، وفعالية الخيارات أو عوامل التشتيت. شكل الأسئلة التي تم تحليلها هو أسئلة الاختيار من متعدد. ستقرر نتائج التحليل ما إذا كان السؤال سيتم تصنيفه على أنه جيد أم لا بناءً على استيفاء معايير مستوى الصعوبة والقوة التمييزية وفعالية الخيارات باستخدام التحليل الكمي. وهذا ما أكده جيرل وآخرون. الذي كشف أن استخدام أسئلة الاختيار من متعدد يتطلب أيضًا فهم خصائصها.¹¹

يعد تحليل العناصر مهمًا كوسيلة لتقييم استخدام أسئلة الاختيار من متعدد. ويمكن استخدام نتائج التحليل كوسيلة للتقييم لتحسين الأسئلة حتى تصبح أفضل. ستكون الأسئلة المحسنة أكثر فعالية لاستخدامها في تحديد جودة التعلم الذي تم تنفيذه، كما هو الحال في البحث الذي أجراه.¹² ومن ناحية أخرى، فإن التحليل الذي يتم إجراؤه سيوفر معلومات محددة فيما يتعلق بعنصر السؤال. تعد هذه المعلومات ونتائج التحليل المحددة بمثابة جهد لتحسين التعلم وتحسين جودة التعليم حتى لا يكون المعلمون مهملين في إجراء التقييمات لأنهم يحاولون قياسها باستخدام عدة فئات. عند التحليل، ما زلنا نعطي الأولوية للمعايير المحددة.¹³

⁹ M. Zuhdi Rachman, "Kajian Butir Soal Ujian Sekolah Matematika SMA Negeri 1 Gongdanglegi Tahun 2012," n.d., http://mulok.lib.um.ac.id/index.php?p=show_detail&id=58404.

¹⁰ Wojciech Malec and Małgorzata. Krzeminska-Adamek, "A Practical Comparison of Selected Methods of Evaluating Multiple-Choice Options through Classical Item Analysis," *Practical Assessment, Research, and Evaluation* 25, no. 7 (2020), <https://scholarworks.umass.edu/pare/vol25/iss1/7>.

¹¹ Mark J Gierl et al., "Evaluating the Psychometric Characteristics of Generated Multiple-Choice Test Items," *Applied Measurement in Education*, 2016, doi:10.1080/08957347.2016.1171768.

¹² Wiyasa et al., "Evaluating Quality of Teacher-Developed English Test in Vocational High School: Content Validity and Item Analysis," *Education Quarterly Reviews* 2, no. 2 (2019), doi: 10.31014/aior.1993.02.02.67.

¹³ Jari Metsämuuronen, "Dimension-Corrected Somers' D for the Item Analysis Settings," *International Journal of Educational Methodology* 6, no. 2 (2020): 297–317, doi: 10.12973/ijem.6.2.297.

واستند البحث الذي تم إجراؤه أيضًا إلى نتائج الأبحاث السابقة ذات الصلة، مثل البحث الذي أجراه راتنانينجسيه وإسفارودي (2013)،¹⁴ ويجايانتي وآخرون (2014)،¹⁵ وكورنياوان (2015)،¹⁶ وجيمو ونوغراهاني (2019)،¹⁷ وفاطمة وآخرون (2019)،¹⁸ ماجدالينا وآخرون. (2021)،¹⁹ سوكمافيترياني وآخرون. (2021)²⁰ الذي أجرى بحثًا حول تحليل العناصر بأنواع مختلفة من الأسئلة التي تتكون من عدة مواضيع مختلفة. هناك أيضًا أبحاث تركز على تطوير الأسئلة أو تطوير المكونات الداعمة للتقييم، مثل البحث الذي أجراه أوتومو (2012)،²¹ محصان وآخرون (2014)،²² أوتومو ويولياني وآخرون (2014)،²³ براتيوينينجتياس وآخرون. (2017)،²⁴ ماساي (2019)،²⁵ سوليسا ومانوبوتي (2019).²⁶ ويعتبر هذا البحث أساساً لتحسين جودة الأسئلة واختبارها بشكل جيد. نفذت العديد من الدراسات إعادة بناء الأسئلة والتحقق من صحتها، مثل البحث الذي أجراه رزقوه وواجيران (2020)،²⁷ وكافاه وآخرون. (2021)،²⁸

¹⁴ Ratnaningsih, Dewi Juliah, and Isfarudi, "Analisis Butir Tes Objektif Ujian Akhir Semester Mahasiswa Universitas Terbuka Berdasarkan Teori Tes Modern," *Jurnal Pendidikan Terbuka Dan Jarak Jauh* 14, no. 2 (2013): 98–109, <https://jurnal.ut.ac.id/index.php/jptj/article/view/402/379>.

¹⁵ Hani Wijayanti, Hari Bambang, and Mustika Ani Hety, "Analisis Butir Soal Objektif UAS Semester Genap Kelas VII Pada Mata Pelajaran Ilmu Pengetahuan Sosial (IPS) Tahun Pelajaran 2013/2014 Di SMP Negeri 3 Balung," *Artikel Universitas Jember*, 2014.

¹⁶ Kurniawan, "Analisis Butir Soal Ulangan Akhir Semester Gasal Mata Pelajaran IPS Sekolah Dasar."

¹⁷ Farida Nugrahani Gimo, "Analisis Butir Soal Ulangan Akhir Semester (UAS) Mata Pelajaran Bahasa Indonesia Kelas XI Sekolah Menengah Kejuruan Negeri 1 Wonogiri Tahun Pelajaran 2015/2016," *Stilistika*, 5, no. 1 (2019): 35–46.

¹⁸ Laela Umi Fatimah, "Analisis Kesukaran Soal, Daya Pembeda Dan Fungsi Distraktor," *Jurnal Komunikasi Dan Pendidikan Islam* 8, no. 2 (2019).

¹⁹ Ina Magdalena et al., "Analisis Butir Soal Ulangan Akhir Semester Ganjil Mata Pelajaran IPS Kelas IV SDS Sari Putra," *Bintang : Jurnal Pendidikan Dan Sains* 3, no. 1 (2021): 15–22, <https://ejournal.stitpn.ac.id/index.php/bintang>.

²⁰ Sukmafritriani et al., "Analisis Butir Soal Penilaian Keterampilan Apresiasi Sastra Pada Soal UAS Materi Puisi Kelas VII Tahun Pelajaran 2019/2020 Di SMPN 1 Kandanghaur," *Asas: Jurnal Sastra* 10, no. 2 (2021).

²¹ Asep Purwo Yudi Utomo, "Pengembangan Perangkat Evaluasi Kompetensi Menulis Kelas V SD Dengan Taksonomi The Structure Of Observed Learning Outcome," *Tesis. Universitas Negeri Semarang*, 2012.

²² Ali Muhson et al., "Pengembangan Software Analisis Butir Soal Yang Praktis Dan Aplikatif," *Jurnal Ilmu Pendidikan* 20, no. 2 (2014): 207–16.

²³ Asep Purwo Yudi Utomo and Uki Hares Yulianti, "Pengembangan Media Uji Kemahirann Berbahasa Indonesia Berbasis Teknologi Informasi Dan Komunikasi Bermuatan Nilai Konservasi," *Penelitian Pemula. Universitas Negeri Semarang*, 2014.

²⁴ Pratiwiningtyas et al., "Pengembangan Instrumen Penilaian Kognitif Untuk Mengukur Literasi Membaca Bahasa Indonesia Berbasis Model PIRLS Pada Siswa Kelas IV SD," *Journal of Educational Research and Evaluation, JERE* 6, no. 1 (2017).

²⁵ Aleeyah Masae, "Pengembangan Tes Membaca Bahasa Indonesia Berbasis Web Bagi Mahasiswa Thailand Di Universitas Negeri Yogyakarta," *LingTera* 6, no. 1 (2019): 91–99.

²⁶ Irwan Soulisa and Peter Manuputty, "Pengembangan Alat Penilaian Pembelajaran Bahasa Indonesia Berbasis Kompetensi Menggunakan Pendekatan Kontekstual Di SMA Papua Kota Sorong," *Jurnal Kredo* 4, no. 2 (2019).

²⁷ Rizqoh, Annisa Nur Ainun, and Wagiran, "Rekonstruksi Instrumen Penilaian Tengah Semester Kelas VIII Mata Pelajaran Bahasa Indonesia Bermuatan HOTS Dan Adiwiyata," *Jurnal Pendidikan Bahasa Dan Sastra Indonesia* 9, no. 2 (2020).

برامونو وآخرون. (2021).²⁹ وبصرف النظر عن ذلك، هناك أبحاث حول تحليل العناصر تستخدم التطبيقات أو برنامج Excel في عملية التحليل، مثل بحث نوريندرا وآخرون (2013)،³⁰ كورنياوان وآخرون. (2017).³¹ توضح العديد من موضوعات البحث هذه اختلافات حقيقية مع البحث الذي أجريناه على الرغم من قيام كلانا بإجراء بحث تقييمي.

منهج البحث

استخدم هذا البحث بالمنهج الكمي لوصف جودة أسئلة تقييم المدرسة (AM) من خلال تحليل العناصر من البيانات الكمية لمادة اللغة العربية للصف التاسع في المدرسة المتوسطة منبع القرآن بمدينة موجوكيرتو للعام الدراسي 2023/2022. البيانات المستخدمة في هذا البحث هي أسئلة تقييم المدرسة (AM) ونتائج إجابات الطلاب في مدرسة منبعول للقرآن، مدينة موجوكيرتو. أخذ الباحثون عينات من البيانات بناءً على جميع طلاب الصف التاسع في مدرسة منبعول القرآن بمدينة موجوكيرتو، العام الدراسي 2023/2022. وبلغ عدد العينات المأخوذة 61 طالباً. تم حساب العينات باستخدام صيغة سلوفين.

استخدمت عملية جمع البيانات التي أجراها الباحثون تحليل الوثائق التي تم جمعها بعد إجراء تصحيحات تقييم المدرسة (AM). كانت الخطوة الأولى التي تم اتخاذها هي جمع أسئلة تقييم المدرسة (AM) لمادة اللغة العربية للصف التاسع في مدرسة منبع القرآن بمدينة موجوكيرتو للعام الدراسي 2023/2022، ومفاتيح الإجابة، وأوراق الإجابة من الطلاب الذين تم استخدامهم كنماذج بحثية. تم تعديل اختيار أوراق الإجابة على العينة المحددة مبدئياً. تم تنفيذ تقنية التحليل باستخدام برنامج Anates الإصدار 4. وشملت الأشياء التي تم تحليلها: مستوى صعوبة الأسئلة، قوة تمييز الأسئلة،

²⁸ Rifada Robitha Silmi Kaaffah, Andi Wibowo, and Indah Rahmayanti, "Validitas Isi Pada Alat Evaluasi Buku Teks Bahasa Indonesia Kelas X SMA," *Imajeri: Jurnal Pendidikan Bahasa Dan Sastra Indonesia* 3, no. 2 (2021): 158–67.

²⁹ Didi Pramono et al., "Pelatihan Penyusunan Evaluasi Pembelajaran Berbentuk Portofolio Sebagai Implementasi Merdeka Belajar Di Sekolah Bagi Guru SMP 41 Semarang," *Jurnal Implementasi* 1, no. 1 (2021).

³⁰ Sindy Nurinda, Ely Rudyatmi, and Saiful Ridlo, "Analisis Butir Soal Olimpiade Biologi SMA Tingkat Kabupaten/Kota Tahun 2013," *Unnes Journal of Biology Education* 3, no. 1 (2014).

³¹ Riza Yonisa Kurniawan et al., "Pemberian Pelatihan Analisis Butir Soal Bagi Guru Di Kabupaten Jombang; Efektif?," *Jurnal Pemberdayaan Masyarakat Madani (JPMM)* 1, no. 2 (2017).

والمشتتات في أسئلة تقييم المدرسة (AM). وتم تحليل نتائج البيانات نوعياً وكشفت عن عدة أمثلة لتعزيز النتائج الكمية.

نتائج البحث

تم عرض النتائج والمناقشة في هذا البحث في عدة أقسام فرعية لتسهيل فهم محتوى البحث ونتائجه. يتم ترتيب العرض التقديمي حسب ترتيب تقديم البيانات الكمية والأمثلة والتحليل النوعي. ويتم تحليل جميع البيانات لوصف نتائج البحث. الأقسام الفرعية المقدمة هي مستوى صعوبة أسئلة تقييم المدرسة، والقوة المميزة لأسئلة تقييم المدرسة، وتحليل عوامل تشتت انتباه تقييم المدرسة.

مستوى صعوبة أسئلة تقييم في اختبار المدرسة الأخيرة للفصل التاسع في المدرسة المتوسطة منبع القرآن بمدينة موجوكرتو سنة دراسية 2022-2023 م

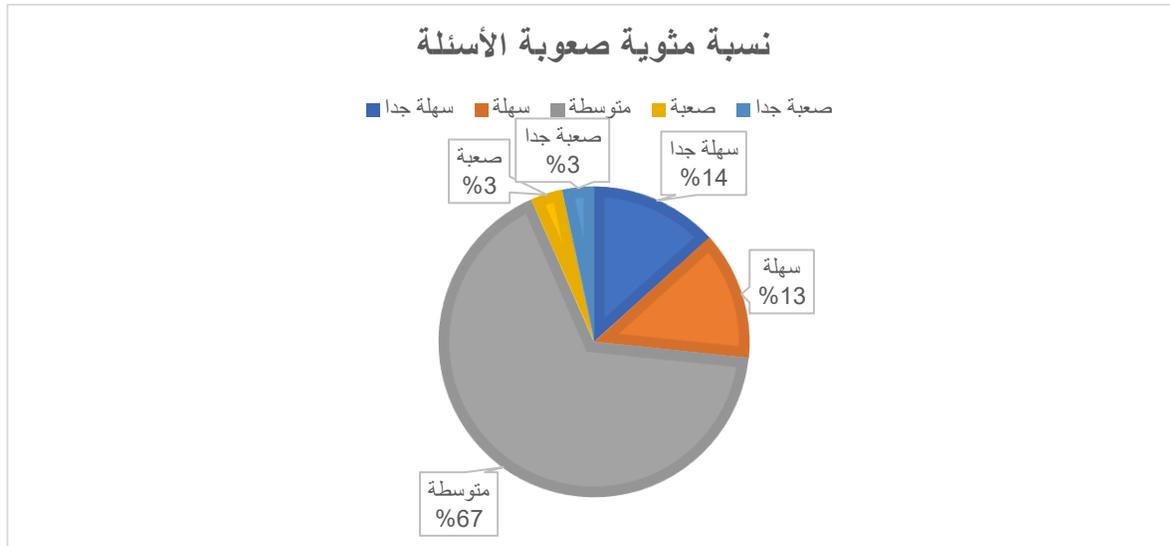
الأسئلة المقدمة في تقييم المدرسة (Asesmen Madrasah) في المدرسة المتوسطة منبع القرآن بمدينة موجوكرتو لديها مستويات متفاوتة من الصعوبة. حاول الباحثون تحليل الأسئلة مع الإشارة إلى مستويات الصعوبة: سهلة جداً، وسهلة، ومتوسطة، وصعبة، وصعبة جداً. ويهدف هذا إلى وصف تكافؤ مستوى صعوبة الأسئلة المستخدمة. ومع ذلك، تظهر نتائج التحليل أن الأمور متفاوتة لأن أسئلة تقييم المدرسة (AM) في مدرسة متوسطة منبع القرآن بمدينة موجوكرتو لا تحتوي إلا على مستوى من الصعوبة في شكل سهل جداً وسهل ومتوسط وفقاً للبيانات الواردة في الجدولين 1 والرسم البياني 1 أدناه.

الجدول 1: من نتائج تحليل صعوبة أسئلة

رقم	مستوى الصعوبة	بنود الأسئلة	العدد	نسبة مئوية
1	سهلة جداً	22 – 9 – 6 – 3	4	13,3%
2	سهلة	21 – 8 – 7 – 4	4	13,3%
3	متوسطة	1 – 2 – 5 – 10 – 11 – 12 – 13 – 15 – 16 – 17 – 18 – 19 – 20 – 24 – 25 – 26 – 27 – 28 – 29 – 30	20	66,7%

%3,3	1	14	صعبة	4
%3,3	1	23	صعبة جدًا	5

الرسم البياني: 1



يوضح الجدول 1 والرسم البياني 1 نفس البيانات فيما يتعلق بمستوى صعوبة أسئلة التقييم في اختبار المدرسة الأخيرة للفصل التاسع في المدرسة المتوسطة منبع القرآن بمدينة موجوكرتو سنة دراسية 2022-2023 م. وتشير البيانات إلى أن 13,3 بالمائة من الأسئلة تقع في الفئة السهلة جدًا، و13,3 بالمائة في الفئة السهلة، و66,7 بالمائة في الفئة المتوسطة، و3,3 بالمائة في الفئة الصعبة، و3,3 بالمائة في الفئة الصعبة جدًا. وتوضح نتائج هذا التحليل أيضًا أنه توجد أسئلة في كل فئات، ويكون هذا بمثابة تقييم لكاتب الأسئلة أو المعلم لتطوير أسئلة أكثر تنوعًا، على سبيل المثال تلك الموجودة في بحث سورباني (2017)،³² وهاليك وآخرين. (2019).³³

³² Yulinda Erma Suryani, "Pemetaan Kualitas Empirik Soal Ujian Akhir Semester Pada Mata Pelajaran Bahasa Indonesia SMA Di Kabupaten Klaten," *Jurnal Penelitian Dan Evaluasi Pendidikan* 21, no. 2 (2017): 142-52.

³³ Andi Surahma Halik, Sitti Mania, and Fitriani Nur, "Analisis Butir Soal Ujian Akhir Sekolah (UAS) Mata Pelajaran Matematika Pada Tahun Ajaran 2015/2016 SMP Negeri 36 Makassar," *Al-Asma: Journal of Islamic Education Vol. 1, No 1 1*, no. 1 (2019).

هناك سؤالاً في الفئة الصعبة جداً وهما في رقم 23. وتصنف هذا السؤال على أنه صعب جداً لأن كثير من الطلاب يخطؤون في الإجابة، وهي من عدد الطلاب 61 طالبا أنهم يخطؤون في الإجابة بين 92 بالمئة إلى 79 بالمئة، ويصوبون في الإجابة بين 8 بالمئة إلى 21 بالمئة. ومن مثال السؤال في الفئة الصعبة جداً فيما يلي:

إفهم هذا النص!

جاءت العُطلَةُ، ذهبَ حمَدانُ وعائِلَتُهُ إلى قَرْيَةٍ بَعِيدَةٍ عَنِ الْمَدِينَةِ، هُم ذَهَبُوا إِلَيْهَا بِالسَّيَّارَةِ، تَحَرَّكَتِ السَّيَّارَةُ مِنَ الْبَيْتِ فِي السَّاعَةِ السَّادِسَةِ صَبَاحًا. بَعْدَ سَاعَةٍ وَاحِدَةٍ مَرَّتِ السَّيَّارَةُ بِالْحُقُولِ وَبِغَابَةِ جَمِيلَةٍ. الْغَابَةُ رَائِعَةٌ وَمُنَظَّمَةٌ هُنَاكَ أَشْجَارٌ مُتَجَاوِرَةٌ، وَوَقَفَتِ الْحَافِلَةُ جَانِبَ الْغَابَةِ الْخَضِرَاءِ. ثُمَّ وَصَلَتِ السَّيَّارَةُ إِلَى بَيْتِ عَمِّهِ، يَقَعُ بَيْتُ الْعَمِّ فِي الْقَرْيَةِ فِي قِمَّةِ الْجَبَلِ. زَارَ حَمَدَانُ عَمَّهُ هُنَاكَ، إِسْمُ عَمِّهِ السَّيِّدُ هَاشِمٌ، هُوَ فَلَّاحٌ نَشِيطٌ وَمَعْرُوفٌ. لَهُ مَزْرَعَةٌ وَاسِعَةٌ. يَزْرَعُ عَمُّهُ وَأَهْلُ الْقَرْيَةِ الْأُرْزُ وَالْخَضِرَوَاتِ فِي مَزَارِعِهِمْ. نَزَلْنَا مِنَ السَّيَّارَةِ وَجَلَسْنَا تَحْتَ الشَّجَرَةِ قَرِيبًا مِنَ الْمَزْرَعَةِ لِنَتَمَتَّعَ بِجَمَالِ الطَّبِيعَةِ هُنَاكَ. وَالْهَوَاءُ هُنَاكَ طَرِيٌّ. هَذِهِ الْمَنَاطِرُ جَمِيلَةٌ جِدًّا، مَا أَعْظَمَ خَالِقَهَا، خَالِقُ الْكَوْنِ، اللَّهُ سُبْحَانَهُ وَتَعَالَى.

مَا الْفِكْرَةُ الْمُنَاسِبَةُ لِهَذَا النَّصِّ؟

أ. الْعُطْلَةُ

ب. جَمَالِ الطَّبِيعَةِ

ج. مَنَاطِرُ الْقَرْيَةِ

د. أَحْوَالُ الْقَرْيَةِ

هناك سؤالاً في الفئة الصعبة وهي في رقم 14. وتصنف هذا السؤال على أنه صعب لأن كثير من الطلاب يخطؤون في الإجابة، وهي من عدد الطلاب 61 طالبا أنهم يخطؤون في الإجابة بين 64 بالمئة إلى 61 بالمئة، ويصوبون في الإجابة بين 36 بالمئة إلى 39 بالمئة. ومن مثال السؤال في الفئة الصعبة جداً فيما يلي:

اقرأ هذا النص!

كَارًا فِي الْمَطْبَخِ لِصِنَاعَةِ الْأَطْعِمَةِ وَالْمَشْرُوبَاتِ، وَلَيْلَى فِي الْمُسْتَشْفَى لِمُعَالَجَةِ الْمَرْضَى، دِنْدَا فِي الْمُدْرَسَةِ لِبِنَاءِ

الْفَاعَةُ الْجَدِيدَةَ، وَلَا لَا تَحُوكُ ثِيَابَ الْمُدْرَسَةِ لِابْنَتِهَا.

مَا مِهْنَةٌ لِأَلَا ؟

أ. الطَّبِيبَةُ

ب. الْمُدْرَسَةُ

ج. الْخَيَّاطَةُ

د. الطَّبَّاخَةُ

وهناك 20 سؤالاً في الفئة المتوسطة وهي في الأرقام 1 – 2 – 5 – 10 – 11 – 12 – 13 – 15 – 16 – 17 – 18 – 19 – 20 – 24 – 25 – 26 – 27 – 28 – 29 – 30. وتصنف هذه الأسئلة على أنها متوسطة لأن بعض الطلاب تمكنوا من الإجابة على الأسئلة بشكل صحيح وهي بين 46 بالمئة إلى 64 بالمئة، ولكن بعضها أجابوا بشكل خاطئ وهي بين 64 بالمئة إلى 46 بالمئة. ومن مثال السؤال في الفئة المتوسطة فيما يلي:

"حروف النصب هي أحرف تسبق الفعل المضارع فتنصبه ؛ وهذه الأحرف هي : أن، لن، إذن، كي، لام التعليل، لام الجحود، فاء السببية، واو المعية، حتى، أو، ثم." "أَذْهَبُ إِلَى مَلْعَبِ الْمُدْرَسَةِ ... أَشَاهِدُ الْمُبَارَاةَ". وَمِنْ حَرْفِ النَّوَاصِبِ الْمُنَاسِبِ لِإِمْلَاءِ هَذِهِ الْجُمْلَةِ هِيَ ...

أ. أَنْ

ب. كَي

ج. لَنْ

د. لِ

هناك 4 أسئلة في الفئة السهلة وهي في الأرقام 4 – 7 – 8 – 21. وتصنف هذه الأسئلة على أنها سهلة لأن كثير من الطلاب يصوبون في الإجابة، وهي من عدد الطلاب 61 طالبا أنهم يصوبون في الإجابة بين 82 بالمئة إلى 67 بالمئة، ويخطؤون في الإجابة بين 18 بالمئة إلى 33 بالمئة. ومن مثال السؤال في الفئة الصعبة جدا فيما يلي:

اقْرَأْ هَذَا الْحَوَارِ!

س : أَيْنَ يَبْتَئِكُ؟

ج : بَيْتِي فِي شَارِعِ بَانْجَانِجِ

س : كَمْ رَقْمُ بَيْتِكَ؟

ج : رَقْمُ بَيْتِي 96

اُكْتُبِ الرَّقْمَ مِنَ الْحَوَارِ بِاللُّغَةِ الْعَرَبِيَّةِ!

أ. تِسْعَةٌ وَسِتُّونَ

ب. سِتَّةٌ وَسِعَةٌ

ج. تِسْعَةٌ وَسِتَّةٌ

د. سِتَّةٌ وَتِسْعُونَ

هناك 3 أسئلة في الفئة السهلة جدا وهما في الأرقام 3 – 6 – 9 – 22. وتصنف هذه الأسئلة على أنها سهلة جدا لأن كثير من الطلاب يصوبون في الإجابة، وهي من عدد الطلاب 61 طالبا أنهم يصوبون في الإجابة بين 92 بالمئة إلى 84 بالمئة، و يخطؤون في الإجابة بين 8 بالمئة إلى 16 بالمئة. ومن مثال السؤال في الفئة الصعبة جدا فيما يلي:

إفهم هذه النص!

نَزَلَ الْقُرْآنَ عَلَى النَّبِيِّ مُحَمَّدٍ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ فِي غَارِ حِزَاءٍ وَ كَانَ عُمُرُهُ أَرْبَعِينَ سَنَةً كَانَ النَّبِيُّ مُحَمَّدٌ يَخْلُو بِنَفْسِهِ فِي الْغَارِ وَيَتَعَبَّدُ اللَّهُ فِيهِ بِمَا بَقِيَ مِنْ دِينِ إِبْرَاهِيمَ عَلَيْهِ السَّلَامُ، ثُمَّ نَزَلَ عَلَيْهِ جِبْرِيلُ بِالْوَحْيِ وَأَمَرَهُ بِالْقِرَاءَةِ. وَكَانَ ذَلِكَ فِي اللَّيْلَةِ السَّابِعَةِ عَشْرَةَ مِنْ شَهْرِ رَمَضَانَ.

قال جبريل لمحمد صلى الله عليه وسلم: اقرأ، فقال محمد: ما أنا بقاري لأن محمد كان أمياً فقال جبريل:

اقرأ. فقال محمد: ما أنا بقاري ثم قال جبريل: ﴿ اِقْرَأْ بِاسْمِ رَبِّكَ الَّذِي خَلَقَ (١) خَلَقَ الْإِنْسَانَ مِنْ عَلَقٍ (٢)

اِقْرَأْ وَرَبُّكَ الْأَكْرَمُ (٣) الَّذِي عَلَّمَ بِالْقَلَمِ (٤) عَلَّمَ الْإِنْسَانَ مَا لَمْ يَعْلَمْ (٥)

وهذه الآيات هي أول ما نزل من القرآن على النبي محمد صلى الله عليه وسلم واستمر نزول القرآن الكريم على النبي ٢٣ سنة. سنة ١٣ سنة في مكة و ١٠ سنوات في المدينة.

مَا الْفِكْرَةُ الْمُنَاسِبَةُ لِهَذَا النَّصِّ؟

أ. نُزُولُ الْقُرْآنِ

ب. شَهْرُ رَمَضَانَ

ج. قِصَّةُ جِبْرِيلَ وَمُحَمَّدَ

د. دِينُ إِبْرَاهِيمَ عَلَيْهِ السَّلَامُ

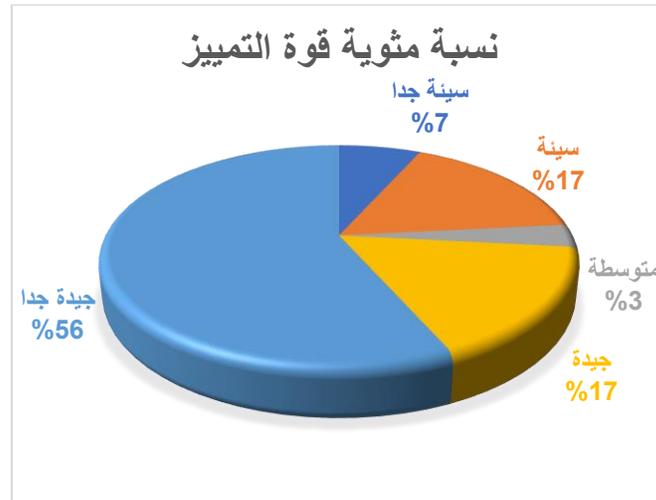
مستوى قوة التمييز في أسئلة التقييم في اختبار المدرسة الأخيرة للفصل التاسع في المدرسة المتوسطة منبع القرآن بمدينة موجوكرتو سنة دراسية 2023-2022 م

التعريف التالي هو قوة التمييز الموجودة في عناصر السؤال. يمكن تصنيف قوة التمييز الموجودة في الأسئلة على أنها سيئة جدًا، وسيئة، وعادلة، وجيدة، وجيدة جدًا. نعرض أدناه نتائج تحليل القوة التفاضلية لأسئلة تقييم اللغة العربية في اختبار المدرسة الأخيرة للفصل التاسع في المدرسة المتوسطة منبع القرآن بمدينة موجوكرتو سنة دراسية 2023-2022 م والتي تظهر في الجدول والرسم البياني التالي.

الجدول 2: من نتائج تحليل قوة التمييز

رقم	مستوى قوة التمييز	بنود الأسئلة	العدد	نسبة مئوية
1	سيئة جدًا	23 – 1	2	6,7%
2	سيئة	22 – 17 – 6 – 3 – 2	5	16,7%
3	متوسطة	14	1	3,3%
4	جيدة	21 – 20 – 9 – 8 – 4	5	16,7%
5	جيدة جدًا	– 19 – 18 – 16 – 15 – 13 – 12 – 11 – 10 – 7 – 5 30 – 29 – 28 – 27 – 26 – 25 – 24	17	56,7%

الرسم البياني: 2



يوضح الجدول 2 والرسم البياني 2 نفس البيانات فيما يتعلق بمستوى قوة تمييز التقييم في اختبار المدرسة الأخيرة للفصل التاسع في المدرسة المتوسطة منبع القرآن بمدينة موجوكرتو سنة دراسية 2023-2022 م. وتشير البيانات إلى أن 6,7 بالمئة من الأسئلة تقع في الفئة السيئة جدًا، و16,7 بالمئة في الفئة السيئة، و3,3 بالمئة في الفئة المتوسطة، و16,7 بالمئة في الفئة الجيدة، و56,7 بالمئة في الفئة الجيدة جدًا. وتوضح نتائج هذا التحليل أيضًا أنه توجد أسئلة في كل فئات، ويكون هذا بمثابة تقييم لكاتب الأسئلة أو المعلم لتطوير أسئلة أكثر تنوعًا ومطابقًا.

مستوى المشتتات في أسئلة التقييم في اختبار المدرسة الأخيرة للفصل التاسع في المدرسة المتوسطة منبع القرآن بمدينة موجوكرتو سنة دراسية 2023-2022 م

المشتتات هي اختيارات للإجابة وليست مفاتيح إجابة تعمل على خداع المتقدمين للاختبار الذين لا يتقنون الأسئلة/المادة التي يتم اختبارها. يجب أن تكون جميع المشتتات جذابة ليختارها المتقدم للاختبار كإجابة على عنصر السؤال. يتم اختيار المشتتات التي يُقال إنها فعالة من قبل ما لا يقل عن 5٪ من عدد المشاركين في الاختبار. بالإضافة إلى ذلك، يجب على المشتتات الجيدة أن يخدع المجموعة السفلية أكثر. وبعيداً عن ذلك، من المهم تحديد تشتت أسئلة الاختيار المتعدد لأنه يمكن استخدامه كأساس لاختيار الأسئلة بشكل موضوعي (Forthmann et al. 2020). الأسئلة الواردة في اختبار المدرسة الأخيرة للفصل التاسع في المدرسة المتوسطة منبع القرآن بمدينة موجوكرتو سنة دراسية 2022-2023 م بها أيضاً العديد من التشتيئات التي تحدث في الوصف التالي.

الجدول 3: من نتائج تحليل المشتتات

رقم	مستوى المشتتات	بنود الأسئلة	العدد	نسبة مئوية
1	سيئة جداً	26 – 25 – 22 – 4	4	13.3%
2	سيئة	24 – 23 – 21 – 3 – 2	5	16.7%
3	متوسطة	9 – 7 – 1	3	10%
4	جيدة	19 – 13 – 11 – 8 – 6	5	16.7%
5	جيدة جداً	30 – 29 – 28 – 27 – 20 – 18 – 17 – 16 – 15 – 14 – 12 – 10 – 5	13	43.3%

يوضح الجدول 3 نفس البيانات فيما يتعلق بمستوى مشتتات أسئلة التقييم في اختبار المدرسة الأخيرة للفصل التاسع في المدرسة المتوسطة منبع القرآن بمدينة موجوكرتو سنة دراسية 2022-2023 م. وتشير البيانات إلى أن 13.3 بالمائة من الأسئلة تقع في الفئة السيئة جداً، و16.7 بالمائة في الفئة السيئة، و10 بالمائة في الفئة المتوسطة، و16.7 بالمائة في الفئة الجيدة، و43.3 بالمائة في الفئة الجيدة جداً. وتوضح نتائج هذا التحليل أيضاً أنه توجد أسئلة في كل فئات، ويكون هذا بمثابة تقييم لكاتب الأسئلة أو المعلم لتطوير أسئلة أكثر تنوعاً ومطابقاً.

من التوضيحات أعلاه، يمكن ملاحظة أنه يمكن استخدام التقييمات (1) مستوى صعوبة أسئلة التقييم، (2) مستوى قوة التمييز في أسئلة التقييم، (3) مستوى المشتتات في أسئلة التقييم كتقييمات في تحديد قيمة التعليم، وهذا يتوافق مع نظرية ماتسنا القائلة بأن التقييم هو نشاط أو عملية تحديد قيمة التعليم، بحيث يمكن معرفة الجودة أو النتائج.³⁴ وتمشيا مع رضا³⁵ يعد التقييم عنصرا هاما في التعليم، وخاصة تقييم نتائج التعلم. ومن الأشياء التي يتم قياسها في التقييم هو مستوى ذكاء الطلاب. وبذلك يصبح التقييم أحد المكونات التي تساعد المعلمين والمدارس على تحديد نتائج التعلم.

الخاتمة

تظهر نتائج التحليل الذي تم إجراؤه باستخدام برنامج أناتيس (Anates) الإصدار 4.0 النقاط التالية. وأظهرت نتائج تحليل مستوى صعوبة الأسئلة أن هناك 13,3 بالمئة من الأسئلة تقع في الفئة السهلة جداً، و13,3 بالمئة في الفئة السهلة، و66,7 بالمئة في الفئة المتوسطة، و3,3 بالمئة في الفئة الصعبة، و3,3 بالمئة في الفئة الصعبة جداً. وهذا يعني أنه توجد أسئلة في جميع الفئات للغاية. وفي تحليل قوة التمييز للأسئلة كانت نسبة الأسئلة هي 6,7 بالمئة من الأسئلة تقع في الفئة السيئة جداً، و16,7 بالمئة في الفئة السيئة، و3,3 بالمئة في الفئة المتوسطة، و16,7 بالمئة في الفئة الجيدة، و56,7 بالمئة في الفئة الجيدة جداً. وكانت نتائج تحليل المشتتات هي 13,3 بالمئة من الأسئلة تقع في الفئة السهلة جداً، و13,3 بالمئة في الفئة السهلة، و66,7 بالمئة في الفئة المتوسطة، و3,3 بالمئة في الفئة الصعبة، و3,3 بالمئة في الفئة الصعبة جداً. وبذلك كانت نتائج تحليل خلاصة العناصر التي يمكن استخدامها 21 سؤالاً أي 70 بالمئة و9 أسئلة لا يمكن استخدامها أي 30 بالمئة. وبناء على نتائج التحليل، يمكن إعادة استخدام بعض الأسئلة لتقييم أنشطة التعلم ويجب مراجعة البعض الآخر حسب الحاجة.

³⁴ M. & R Matsna, *Evaluasi Pembelajaran Bahasa Arab* (Ciputat: UIN Press, 2015).

³⁵ U Ridho, "Evaluasi Dalam Pembelajaran Bahasa Arab," *An Nabighoh: Jurnal Pendidikan Dan Pembelajaran Bahasa Arab* 20, no. 01 (2018), <https://doi.org/https://doi.org/10.32332/annabighoh.v20i01.1124>.

المراجع

- Fatimah, Laela Umi. "Analisis Kesukaran Soal, Daya Pembeda Dan Fungsi Distraktor." *Jurnal Komunikasi Dan Pendidikan Islam* 8, no. 2 (2019).
- Gierl, Mark J, Hollis Lai, Debra Pugh, Claire Touchie, André Boulais, and André De Champlain. "Evaluating the Psychometric Characteristics of Generated Multiple-Choice Test Items." *Applied Measurement in Education*, 2016. doi:10.1080/08957347.2016.1171768.
- Gimo, Farida Nugrahani. "Analisis Butir Soal Ulangan Akhir Semester (UAS) Mata Pelajaran Bahasa Indonesia Kelas XI Sekolah Menengah Kejuruan Negeri 1 Wonogiri Tahun Pelajaran 2015/2016." *Stilistika*, 5, no. 1 (2019): 35–46.
- Halik, Andi Surahma, Sitti Mania, and Fitriani Nur. "Analisis Butir Soal Ujian Akhir Sekolah (UAS) Mata Pelajaran Matematika Pada Tahun Ajaran 2015/2016 SMP Negeri 36 Makassar." *Al-Asma: Journal of Islamic Education Vol. 1, No 1* 1, no. 1 (2019).
- Hamzah, Ali. *Evaluasi Pembelajaran Matematika*. Jakarta: Rajawali Pers, 2014.
- Kaaffah, Rifada Robitha Silmi, Andi Wibowo, and Indah Rahmayanti. "Validitas Isi Pada Alat Evaluasi Buku Teks Bahasa Indonesia Kelas X SMA." *Imajeri: Jurnal Pendidikan Bahasa Dan Sastra Indonesia* 3, no. 2 (2021): 158–67.
- Kurniawan, Riza Yonisa, Albrian Fiky Prakoso, Luqman Hakim, Retno Mustika Dewi, and Irin Widayanti. "Pemberian Pelatihan Analisis Butir Soal Bagi Guru Di Kabupaten Jombang; Efektif?" *Jurnal Pemberdayaan Masyarakat Madani (JPMM)* 1, no. 2 (2017).
- Kurniawan, Tutut. "Analisis Butir Soal Ulangan Akhir Semester Gasal Mata Pelajaran IPS Sekolah Dasar." *Journal of Elementary Education* 4, no. 1 (2015). <http://journal.unnes.ac.id/sju/index.php/jee>.
- Magdalena, Ina, Yanti Ardia Ningsih, Nesfi Berliana, and Alma Novianti Gunawan. "Analisis Butir Soal Ulangan Akhir Semester Ganjil Mata Pelajaran IPS Kelas IV SDS Sari Putra." *Bintang: Jurnal Pendidikan Dan Sains* 3, no. 1 (2021): 15–22. <https://ejournal.stitpn.ac.id/index.php/bintang>.
- Malec, Wojciech, and Małgorzata Krzeminska-Adamek. "A Practical Comparison of Selected Methods of Evaluating Multiple-Choice Options through Classical Item Analysis." *Practical Assessment, Research, and Evaluation* 25, no. 7 (2020). <https://scholarworks.umass.edu/pare/vol25/iss1/7>.

- Masae, Aleeyah. "Pengembangan Tes Membaca Bahasa Indonesia Berbasis Web Bagi Mahasiswa Thailand Di Universitas Negeri Yogyakarta." *LingTera* 6, no. 1 (2019): 91–99.
- Matsna, M. & R. *Evaluasi Pembelajaran Bahasa Arab*. Ciputat: UIN Press, 2015.
- Metsämuuronen, Jari. "Dimension-Corrected Somers' D for the Item Analysis Settings." *International Journal of Educational Methodology* 6, no. 2 (2020): 297–317. doi: 10.12973/ijem.6.2.297.
- Muhson, Ali; Lestari, Barkah; Supriyanto, Supriyanto; Baroroh, Kiromim. "The Development of Practical Item Analysis Program for Indonesian Teachers." *International Journal of Instruction*, 2017, 199–210. doi:10.12973/iji.2017.10213a.
- Muhson, Ali, Barkah Lestari, Supriyanto, and Kiromim Baroroh. "Pengembangan Software Analisis Butir Soal Yang Praktis Dan Aplikatif." *Jurnal Ilmu Pendidikan* 20, no. 2 (2014): 207–16.
- Nurinda, Sindy, Ely Rudyatmi, and Saiful Ridlo. "Analisis Butir Soal Olimpiade Biologi SMA Tingkat Kabupaten/Kota Tahun 2013." *Unnes Journal of Biology Education* 3, no. 1 (2014).
- Pramono, Didi, Ngabiyanto, Isnarto Utomo, wan Hardi Saputro, and Asep Purwo Yudi. "Pelatihan Penyusunan Evaluasi Pembelajaran Berbentuk Portofolio Sebagai Implementasi Merdeka Belajar Di Sekolah Bagi Guru SMP 41 Semarang." *Jurnal Implementasi* 1, no. 1 (2021).
- Pratiwiningsih, Bakti Nanda, Endang Susilaningsih, and I Made Sudana. "Pengembangan Instrumen Penilaian Kognitif Untuk Mengukur Literasi Membaca Bahasa Indonesia Berbasis Model PIRLS Pada Siswa Kelas IV SD." *Journal of Educational Research and Evaluation, JERE* 6, no. 1 (2017).
- Purwati, M. "Analisis Butir Soal Ujian Akhir Mata Pelajaran Akuntansi Keuangan Menggunakan Microsoft Office Excel 2010." *Jurnal Pendidikan Akuntansi Indonesia*, n.d. <https://doi.org/10.21831/jpai.v12i2.2710>.
- Rachman, M. Zuhdi. "Kajian Butir Soal Ujian Sekolah Matematika SMA Negeri 1 Gongdanglegi Tahun 2012," n.d. http://mulok.lib.um.ac.id/index.php?p=show_detail&id=58404.
- Ratnaningsih, Dewi Juliah, and Isfarudi. "Analisis Butir Tes Objektif Ujian Akhir Semester Mahasiswa Universitas Terbuka Berdasarkan Teori Tes Modern." *Jurnal Pendidikan Terbuka Dan Jarak Jauh* 14, no. 2 (2013): 98–109. <https://jurnal.ut.ac.id/index.php/jptjj/article/view/402/379>.
- Ridho, U. "Evaluasi Dalam Pembelajaran Bahasa Arab." *An Nabighoh: Jurnal Pendidikan Dan Pembelajaran Bahasa Arab* 20, no. 01 (2018).

<https://doi.org/https://doi.org/10.32332/annabighoh.v20i01.1124>.

- Rizqoh, Annisa Nur Ainun, and Wagiran. "Rekonstruksi Instrumen Penilaian Tengah Semester Kelas VIII Mata Pelajaran Bahasa Indonesia Bermuatan HOTS Dan Adiwiyata." *Jurnal Pendidikan Bahasa Dan Sastra Indonesia* 9, no. 2 (2020).
- Soulisa, Irwan, and Peter Manuputty. "Pengembangan Alat Penilaian Pembelajaran Bahasa Indonesia Berbasis Kompetensi Menggunakan Pendekatan Kontekstual Di SMA Papua Kota Sorong." *Jurnal Kredo* 4, no. 2 (2019).
- Subali, B., Kumaidi, & Aminah, N. S. "The Comparison of Item Test Characteristics Viewed from Classic and Modern Test Theory." *International Journal of Instruction* 14 (2021). <https://doi.org/10.29333/iji.2021.14139a>.
- Sukmafitriani, Adhetia, Finda Septina Munjariyati, Wagiran, and Deby Luriawati Naryatmojo. "Analisis Butir Soal Penilaianketerampilan Apresiasi Sastra Pada Soal UAS Materi Puisi Kelas VII Tahun Pelajaran 2019/2020 Di SMPN 1 Kandanghaur." *Asas: Jurnal Sastra* 10, no. 2 (2021).
- Suryani, Yulinda Erma. "Pemetaan Kualitas Empirik Soal Ujian Akhir Semester Pada Mata Pelajaran Bahasa Indonesia SMA Di Kabupaten Klaten." *Jurnal Penelitian Dan Evaluasi Pendidikan* 21, no. 2 (2017): 142–52.
- Utomo, Asep Purwo Yudi. "Pengembangan Perangkat Evaluasi Kompetensi Menulis Kelas V SD Dengan Taksonomi The Sructure Of Observed Learning Outcome." *Tesis. Universitas Negeri Semarang*, 2012.
- Utomo, Asep Purwo Yudi, and Uki Hares Yulianti. "Pengembangan Media Uji Kemahirann Berbahasa Indonesia Berbasis Teknologi Informasi Dan Komunikasi Bermuatan Nilai Konservasi." *Penelitian Pemula. Universitas Negeri Semarang*, 2014.
- Wijayanti, Hani, Hari Bambang, and Mustika Ani Hety. "Analisis Butir Soal Objektif UAS Semester Genap Kelas VII Pada Mata Pelajaran Ilmu Pengetahuan Sosial (IPS) Tahun Pelajaran 2013/2014 Di SMP Negeri 3 Balung." *Artikel Universitas Jember*, 2014.
- Wiyasa, Putu Irmayanti, Laksana, I Ketut Darma, Indrawati, and Ni Luh Ketut Mas. "Evaluating Quality of Teacher-Developed English Test in Vocational High School: Content Validity and Item Analysis." *Education Quarterly Reviews* 2, no. 2 (2019). doi: 10.31014/aior.1993.02.02.67.
- Fatimah, Laela Umi. "Analisis Kesukaran Soal, Daya Pembeda Dan Fungsi Distraktor." *Jurnal Komunikasi Dan Pendidikan Islam* 8, no. 2 (2019).
- Gierl, Mark J, Hollis Lai, Debra Pugh, Claire Touchie, André Boulais, and André De Champlain. "Evaluating the Psychometric Characteristics of Generated Multiple-Choice Test Items." *Applied Measurement in Education*, 2016. doi:10.1080/08957347.2016.1171768.

- Gimo, Farida Nugrahani. "Analisis Butir Soal Ulangan Akhir Semester (UAS) Mata Pelajaran Bahasa Indonesia Kelas XI Sekolah Menengah Kejuruan Negeri 1 Wonogiri Tahun Pelajaran 2015/2016." *Stilistika*, 5, no. 1 (2019): 35–46.
- Halik, Andi Surahma, Sitti Mania, and Fitriani Nur. "Analisis Butir Soal Ujian Akhir Sekolah (UAS) Mata Pelajaran Matematika Pada Tahun Ajaran 2015/2016 SMP Negeri 36 Makassar." *Al-Asma: Journal of Islamic Education Vol. 1, No 1* 1, no. 1 (2019).
- Hamzah, Ali. *Evaluasi Pembelajaran Matematika*. Jakarta: Rajawali Pers, 2014.
- Kaaffah, Rifada Robitha Silmi, Andi Wibowo, and Indah Rahmayanti. "Validitas Isi Pada Alat Evaluasi Buku Teks Bahasa Indonesia Kelas X SMA." *Imajeri: Jurnal Pendidikan Bahasa Dan Sastra Indonesia* 3, no. 2 (2021): 158–67.
- Kurniawan, Riza Yonisa, Albrian Fiky Prakoso, Luqman Hakim, Retno Mustika Dewi, and Irin Widayanti. "Pemberian Pelatihan Analisis Butir Soal Bagi Guru Di Kabupaten Jombang; Efektif?" *Jurnal Pemberdayaan Masyarakat Madani (JPMM)* 1, no. 2 (2017).
- Kurniawan, Tutut. "Analisis Butir Soal Ulangan Akhir Semester Gasal Mata Pelajaran IPS Sekolah Dasar." *Journal of Elementary Education* 4, no. 1 (2015). <http://journal.unnes.ac.id/sju/index.php/jee>.
- Magdalena, Ina, Yanti Ardia Ningsih, Nesfi Berliana, and Alma Novianti Gunawan. "Analisis Butir Soal Ulangan Akhir Semester Ganjil Mata Pelajaran IPS Kelas IV SDS Sari Putra." *Bintang: Jurnal Pendidikan Dan Sains* 3, no. 1 (2021): 15–22. <https://ejournal.stitpn.ac.id/index.php/bintang>.
- Malec, Wojciech, and Małgorzata Krzeminska-Adamek. "A Practical Comparison of Selected Methods of Evaluating Multiple-Choice Options through Classical Item Analysis." *Practical Assessment, Research, and Evaluation* 25, no. 7 (2020). <https://scholarworks.umass.edu/pare/vol25/iss1/7>.
- Masae, Aleeayah. "Pengembangan Tes Membaca Bahasa Indonesia Berbasis Web Bagi Mahasiswa Thailand Di Universitas Negeri Yogyakarta." *LingTera* 6, no. 1 (2019): 91–99.
- Matsna, M. & R. *Evaluasi Pembelajaran Bahasa Arab*. Ciputat: UIN Press, 2015.
- Metsämuuronen, Jari. "Dimension-Corrected Somers' D for the Item Analysis Settings." *International Journal of Educational Methodology* 6, no. 2 (2020): 297–317. doi:10.12973/ijem.6.2.297.
- Muhson, Ali; Lestari, Barkah; Supriyanto, Supriyanto; Baroroh, Kiromim. "The Development of Practical Item Analysis Program for Indonesian Teachers." *International Journal of Instruction*, 2017, 199–210. doi:10.12973/iji.2017.10213a.
- Muhson, Ali, Barkah Lestari, Supriyanto, and Kiromim Baroroh. "Pengembangan Software Analisis Butir Soal Yang Praktis Dan Aplikatif." *Jurnal Ilmu Pendidikan* 20, no. 2

(2014): 207–16.

Nurinda, Sindy, Ely Rudyatmi, and Saiful Ridlo. “Analisis Butir Soal Olimpiade Biologi SMA Tingkat Kabupaten/Kota Tahun 2013.” *Unnes Journal of Biology Education* 3, no. 1 (2014).

Pramono, Didi, Ngabiyanto, Isnarto Utomo, wan Hardi Saputro, and Asep Purwo Yudi. “Pelatihan Penyusunan Evaluasi Pembelajaran Berbentuk Portofolio Sebagai Implementasi Merdeka Belajar Di Sekolah Bagi Guru SMP 41 Semarang.” *Jurnal Implementasi* 1, no. 1 (2021).

Pratiwiningtyas, Bekti Nanda, Endang Susilaningih, and I Made Sudana. “Pengembangan Instrumen Penilaian Kognitif Untuk Mengukur Literasi Membaca Bahasa Indonesia Berbasis Model PIRLS Pada Siswa Kelas IV SD.” *Journal of Educational Research and Evaluation, JERE* 6, no. 1 (2017).

Purwati, M. “Analisis Butir Soal Ujian Akhir Mata Pelajaran Akuntansi Keuangan Menggunakan Microsoft Office Excel 2010.” *Jurnal Pendidikan Akuntansi Indonesia*, n.d. <https://doi.org/10.21831/jpai.v12i2.2710>.

Rachman, M. Zuhdi. “Kajian Butir Soal Ujian Sekolah Matematika SMA Negeri 1 Gongdanglegi Tahun 2012,” n.d. http://mulok.lib.um.ac.id/index.php?p=show_detail&id=58404.

Ratnaningsih, Dewi Juliah, and Isfarudi. “Analisis Butir Tes Objektif Ujian Akhir Semester Mahasiswa Universitas Terbuka Berdasarkan Teori Tes Modern.” *Jurnal Pendidikan Terbuka Dan Jarak Jauh* 14, no. 2 (2013): 98–109. <https://jurnal.ut.ac.id/index.php/jptjj/article/view/402/379>.

Ridho, U. “Evaluasi Dalam Pembelajaran Bahasa Arab.” *An Nabighoh: Jurnal Pendidikan Dan Pembelajaran Bahasa Arab* 20, no. 01 (2018). <https://doi.org/https://doi.org/10.32332/annabighoh.v20i01.1124>.

Rizqoh, Annisa Nur Ainun, and Wagiran. “Rekonstruksi Instrumen Penilaian Tengah Semester Kelas VIII Mata Pelajaran Bahasa Indonesia Bermuatan HOTS Dan Adiwiyata.” *Jurnal Pendidikan Bahasa Dan Sastra Indonesia* 9, no. 2 (2020).

Soulisa, Irwan, and Peter Manuputty. “Pengembangan Alat Penilaian Pembelajaran Bahasa Indonesia Berbasis Kompetensi Menggunakan Pendekatan Kontekstual Di SMA Papua Kota Sorong.” *Jurnal Kredo* 4, no. 2 (2019).

Subali, B., Kumaidi, & Aminah, N. S. “The Comparison of Item Test Characteristics Viewed from Classic and Modern Test Theory.” *International Journal of Instruction* 14 (2021). <https://doi.org/10.29333/iji.2021.14139a>.

Sukmafutriani, Adhetia, Finda Septina Munjariyati, Wagiran, and Deby Luriawati Naryatmojo. “Analisis Butir Soal Penilaianketerampilan Apresiasi Sastra Pada Soal UAS Materi Puisi

- Kelas VII Tahun Pelajaran 2019/2020 Di SMPN 1 Kandanghaur.” *Asas: Jurnal Sastra* 10, no. 2 (2021).
- Suryani, Yulinda Erma. “Pemetaan Kualitas Empirik Soal Ujian Akhir Semester Pada Mata Pelajaran Bahasa Indonesia SMA Di Kabupaten Klaten.” *Jurnal Penelitian Dan Evaluasi Pendidikan* 21, no. 2 (2017): 142–52.
- Utomo, Asep Purwo Yudi. “Pengembangan Perangkat Evaluasi Kompetensi Menulis Kelas V SD Dengan Taksonomi The Structure Of Observed Learning Outcome.” *Tesis. Universitas Negeri Semarang*, 2012.
- Utomo, Asep Purwo Yudi, and Uki Hares Yulianti. “Pengembangan Media Uji Kemahirann Berbahasa Indonesia Berbasis Teknologi Informasi Dan Komunikasi Bermuatan Nilai Konservasi.” *Penelitian Pemula. Universitas Negeri Semarang*, 2014.
- Wijayanti, Hani, Hari Bambang, and Mustika Ani Hety. “Analisis Butir Soal Objektif UAS Semester Genap Kelas VII Pada Mata Pelajaran Ilmu Pengetahuan Sosial (IPS) Tahun Pelajaran 2013/2014 Di SMP Negeri 3 Balung.” *Artikel Universitas Jember*, 2014.
- Wiyasa, Putu Irmayanti, Laksana, I Ketut Darma, Indrawati, and Ni Luh Ketut Mas. “Evaluating Quality of Teacher-Developed English Test in Vocational High School: Content Validity and Item Analysis.” *Education Quarterly Reviews* 2, no. 2 (2019). doi: 10.31014/aior.1993.02.02.67.