

تحليل اختبارات العربية كلغة أجنبية

TAHLIL IKHTIBARAT AL-ARABIYAH KA LUGHAH AJNABIYAH

Nawang Wulandari

Institut Agama Islam Negeri Metro

Article Info

Article History:

Received: January 2023

Revised: June 2023

Accepted: June 2023

Published: June 2023

*Corresponding Author:

Name:

Nawang Wulandari

Email:

nawangwulandari@metrouniv.ac.id

Abstract

This study aims to determine the feasibility of TOAFL (Test of English as a Foreign Language) items at Institut Agama Islam Negeri Metro, especially on reading tests, by calculating the questions' validity, reliability, difficulty level and distractors. This study uses a quantitative approach. Data collection techniques in this study were carried out in two ways: documentation and interview techniques. The results of the study found that there were 24 valid category questions and 21 invalid category questions; the reliability of the item was 0.8, which was interpreted as having a high or reliable correlation. The difficulty level could be concluded that were 3 items included in the difficult category, 39 items with moderate difficulty and 8 items in the easy category; the discriminating power produced was 2 questions in the very bad category, so they had to be replaced.

Copyright © 2023, Nawang Wulandari

This is an open-access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Keywords:

Item Analysis; Test; TOAFL.

مستخلص البحث

تهدف هذه الدراسة إلى معرفة جودة عناصر اختبار اللغة العربية كلغة أجنبية وخاصة في أسئلة مهارة القراءة، وتحديدًا من خلال حساب صلاحية الأسئلة، والموثوقية، ومستوى الصعوبة، والمشتتات. تستخدم هذه الدراسة نهجًا كميًا. تم تنفيذ تقنيات جمع البيانات في هذه الدراسة بطريقتين، وهما تقنيات التوثيق والمقابلة. نتائج الدراسة، وجد أن هناك 24 سؤال طبقة صالح و 21 سؤال طبقة غير صالح، كانت موثوقية العنصر 0.8، والتي تم تفسيرها على أنها ذات علاقة عالية أو موثوقة، ويمكن استنتاج مستوى الصعوبة أن هناك 3 العناصر المدرجة في الطبقة الصعبة، 39 عنصرًا بصعوبة متوسطة و 8 عناصر في الطبقة السهلة، كانت القوة التمييزية الناتجة لسؤالين في طبقة سيئة جدًا أو سيئة للغاية لذلك كان لا بد من استبدالها.

كلمات أساسية: تحليل عناصر الاختبار؛ الاختبار؛ اختبار الكفاءة العربية.

المقدمة

كانت المنافسة العالمية في الثورة الصناعية ٥٠٠ تتطلب منا أن يكون لدينا قدرة تنافسية وكفاءة عالية، بما فيها كفاءة اللغة الأجنبية. لترقية كفاءة اللغة الأجنبية لدى طلابها، بالإضافة إلى عقد اختبار الكفاءة الإنجليزية، تعقد جامعة ميترو الإسلامية الحكومية اختبار الكفاءة العربية. يستخدم اختبار الكفاءة العربية لقياس مهارات الطلاب ومعرفتهم وذكائهم وكفاءتهم في اللغة العربية بشكل الكتابة. يهدف توفير دورات اللغة العربية في الحرم الديني الإسلامي إلى جعل جميع خريجي الجامعة يعرفون ويفهمون اللغة العربية. لأنه لا يمكن إنكاره ، فإن تعلم اللغة العربية سيساعد الطلاب (جميعهم مسلمون) في دراسة القرآن والحديث كدليل في تنفيذ حياتهم كخدام لله في المجتمع.¹ وجود دورات اللغة العربية واختبارات اللغة العربية هو محاولة لإدخال وتحسين كفاءة اللغة الأجنبية. بعيدا عن ذلك لأن المصادر الأصلية للشريعة الإسلامية واردة باللغة العربية.²

الاختبارات هي أدوات لقياس التحصيل العلمي.³ تتكون الاختبارات عادة من سلسلة من الأسئلة المستخدمة لقياس المهارات أو المعرفة أو الذكاء أو القدرات أو المواهب التي يمتلكها فرد أو مجموعة.⁴ يمكن تعريف اختبار اللغة العربية على أنه أداة مصممة بطريقة تحتوي على عدد من الأسئلة أو العبارات المتعلقة بالمواد العربية ، ويهدف اختبار اللغة العربية إلى قياس تحصيل الشخص أو تقدمه في التعلم في برنامج تعلم اللغة العربية.⁵

إذا كان اختبار TOEFL (اختبار اللغة الإنجليزية كلغة أجنبية) هو اختبار من المتوقع أن يقيس كفاءة اللغة الإنجليزية في الكتابة.⁶ إذن TOAFL هو اختبار من المتوقع أن يقيس كفاءة اللغة العربية في الكتابة. أهداف TOAFL هم الأشخاص الذين لا يتحدثون العربية كلغة أم. هذا يتفق مع الطلاب في إندونيسيا،

¹ Dian Ekawati, "Ragam Pemahaman Mahasiswa Pada Pembelajaran Bahasa Arab," *An Nabighoh* 21, no. 02 (2019): 183–202, <https://doi.org/10.32332/an-nabighoh.v21i02.1679>.

² Nawang Wulandari, "Rekonstruksi Pembelajaran Bahasa Arab," *An Nabighoh* 21, no. 01 (2019): 114, <https://doi.org/10.32332/an-nabighoh.v21i01.1631>.

³ Muhammad Faris Imaduddin, Hikmah Maulani, and Ijlal Haziq Taufik, "Test The Validity and Reliability of Arabic Learning Questions," *Arabi: Journal of Arabic Studies* 7, no. 2 (2022): 198–207, <https://doi.org/10.24865/ajas.v7i2.523>.

⁴ Purwanto, *Evaluasi Hasil Belajar* (Yogyakarta: Pustaka Pelajar, 2016), 64.

⁵ Rusydi Ahmad Thu'aimah, *Al-Marja' Fi Manhaj At-Ta'lim Al-Lughah Al-Arabiyyah Li An-Natiqin Bi Lughat Al-Ukhra* (Al-Qahirah: Dar Al-Fikr Al-Arabi, n.d.), 635.

⁶ Sulistyoto Seti Utami and Bayu Pirmansyah, "Peningkatan Skor Test Bahasa Inggris (TOEFL) Melalui Pelatihan Secara Intensif," *Lingua Franca: Jurnal Bahasa, Sastra, dan Pengajarannya* 2, no. 2 (2018): 36–46, <http://dx.doi.org/10.30651/lf.v2i2.1477>.

الذين يضعون اللغة العربية كلغة أجنبية أو لغة ثانية.⁷ يمكن تصنيف أهداف TOAFL إلى قسمين ، وهما أكاديمي وعام. الأهداف الأكاديمية لاختبار TOAFL هي لأغراض تعليمية أو بحثية أو أغراض أخرى تتعلق بالأنشطة الأكاديمية. تستخدم الأهداف بشكل عام في مجال العمل أو الترقية أو واجبات العمل. تتفق الجامعات الدينية الإسلامية الحكومية على أن اختبار اللغة العربية كلغة مسبقة هو أحد أدوات الاختبار لقياس المهارات اللغوية لطلابها.⁸ يتم تصنيف TOAFL في الجامعة ميترو الإسلامية الحكومية كنوع من التنبؤ باختبار اللغة العربية كلغة أجنبية على نطاق محلي، وليس اختبار اللغة العربية كلغة أجنبية مرخصا يمكن استخدام شهادته على نطاق دولي. الغرض من اختبار اللغة العربية كلغة أجنبية في الجامعة ميترو الإسلامية الحكومية هو قياس مهارات اللغة العربية الكتابية للطلاب، ثم تصبح الشهادة أحد متطلبات التسجيل في امتحان الأطروحة. يتم إجراء أسئلة اختبار اللغة العربية كلغة أجنبية في الجامعة ميترو الإسلامية الحكومية من قبل فريق من المحاضرين الخبراء في مجالهم. خيارات الإجابة هي خيارات متعددة يبلغ مجموعها ١٤٠ سؤالاً مع تفاصيل استيما ٥٠ سؤالاً و القواعد ٤٠ سؤالاً و القراءة ٥٠ سؤالاً.⁹ مهارة الاستملاء هي مهارة فهم الكلمات المسموعة من الحوار.¹⁰ بحيث يهدف اختبار الاستملاء على TOAFL إلى قياس دقة الطلاب في فهم الأسئلة العربية المسموعة. يمكن تفسير القواعد في تعلم اللغة العربية على أنها قواعد اللغة العربية.¹¹ في قواعد اللغة العربية ينقسم تعلم القويد إلى تعلم نحو وشرف.¹²

⁷ Nia Himatul Ulya, Chairani Astina, and Ashief El Qorny, "Implementasi Bi'ah Lughawiyah Dalam Peningkatan Maharah Kalam Di Pondok Modern Az-Zahra al-Gontory Purwokerto," *Mantiq Taysr: Journal of Arabic Language* 2, no. 2 (2022): 174, <https://doi.org/10.25217/mantiqtaysr.v2i2.2511>.

⁸ Syaiful Mustofa, "Analisis butir tes TOAFL Universitas Islam Negeri Maulana Malik Ibrahim Malang" (The International Conference on Language Teaching and Assessment, Jakarta, 2017), <http://repository.uin-malang.ac.id/6624/>.

⁹ Hasan Hasan, "Keterampilan Mengajar Bahasa Arab Materi Istima Menggunakan Media Lagu," *Al Qalam: Jurnal Ilmiah Keagamaan Dan Kemasyarakatan* 11, no. 23 (2018): 127, <https://doi.org/10.35931/aq.v0i0.7>.

¹⁰ Aim Matun Nihayati Al Muchith and Nely Rahmawati Zaimah, "Relevansi Pembelajaran Bahasa Asing Untuk Anak Usia Dini," *Mantiq Taysr: Journal of Arabic Language* 3, no. 1 (2023): 25–35, <https://doi.org/10.25217/mantiqtaysr.v3i1.3117>.

¹¹ Saiful Anwar, Guntur Cahaya Kesuma, and Koderi, "Pengembangan Modul Pembelajaran Al-Qawaid an-Nahwiyah Berbasis Metode Qiyasyah Untuk Mahasiswa Jurusan Pendidikan Bahasa Arab," *Mantiq Taysr: Journal of Arabic Language* 3, no. 1 (2023): 11–24, <https://doi.org/10.25217/mantiqtaysr.v3i1.2830>.

¹² Muhammad Abdul Qadir Ahmad, *Thuruq At-Ta'lim Al-Lughah Al-Arabiyyah* (Mesir: Maktabah Annahdhoh, 1979), 16.

مهارة القراءة أو مهارة القراءة هي القدرة على رؤية وفهم المعنى الوارد في القراءة.¹³ تنقسم مهارات القراءة إلى قسمين ، وهما مهارة القراءة بصوت عال (القراءة الجهرية) ومهارة القراءة والفهم (القراءة الشامطة). تهدف مهارات القراءة بصوت عال إلى قياس دقة نطق القراءة، بينما تهدف مهارات فهم القراءة إلى قياس فهم الطلاب لمعنى النص. سؤال القراءة على TOAFL في الجامعة ميترو الإسلامية الحكومية هو اختبار فهم القراءة. يتماشى المكون المادي مع أهداف تعلم اللغة العربية للناطقين بالأجنبية ، وهي فهم اللغة العربية بشكل صحيح ، ووسيلة التواصل والتعبير عن الروح ، والقراءة وإيجاد معنى مكة والتعبير عن الحالة والذات الوظيفية مع الكتابة.¹⁴ اللغة العربية كما نعرفها هي لغة أجنبية ، لذلك يختلف التدريس عن تدريس العلوم الأخرى. لأن تعليم اللغة يعطي الأولوية للعديد من المهارات اللغوية ، وهي مهارات الاستماع (الاستماع) ، ومهارات التحدث (الكلام) ، ومهارات القراءة (القراءة) ، ومهارات الكتابة (الكتابة).¹⁵

صممت أسئلة اختبار كفاءة العربية بجامعة ميترو الإسلامية الحكومية من قبل فريق من المحاضرين الخبراء في مجالاتهم. تحتوي الأسئلة على ثلاثة أجزاء من المادة، وهي الاستماع والقواعد (النحوية) والقراءة.¹⁶ والمشاركون في اختبار الكفاءة العربية هم الطلاب الذين انتهوا من مادة اللغة العربية الأولى و الثانية وشاركوا في تكثيف اللغة العربية، لمدة فصل دراسي واحد. أما درجة النجاح المحددة فهي ٤٢٥ خصيصا لطلاب شعبة تعليم اللغة العربية، و ٤٠٠ لطلاب الشعبة الأخرى. أما نجاح اختبار الكفاءة العربية شرط أساسي لمناقشة الرسالة العلمية. فأما المشكلة هي بالقرب من ٥٠٪ من المشاركين لاختبار الكفاءة العربية لم يحققوا النتيجة المستهدفة.

منذ بداية تنفيذ اختبار الكفاءة العربية، أي في عام ٢٠١٢، لم يوجد التحليل لعناصر الاختبار في أسئلة اختبار الكفاءة العربية ولم يوجد التحديث للأسئلة. كان تحليل عناصر الاختبار هو عملية جمع

¹³ Baiq Tuhfatul Unsi and Siti Robiatun Muniroh, "Penerapan Metode Jigsaw Dalam Pembelajaran Qiro'ah Di MTs Darul Ulum Kepuhdoko Tembelang Jombang," *Mantiqutayr: Journal of Arabic Language* 2, no. 2 (2022): 161, <https://doi.org/10.25217/mantiqutayr.v2i2.2474>.

¹⁴ Mahmud Kamil An-Naqoh, *Ta'lim Al-Lughah Ittishaliyan Baina Al-Manahij Wa Al-Istiratijiyat* (Rabath: Issesco, 2006), 123–124.

¹⁵ Muhammad Arif Mustofa, "Interferensi Bahasa Indonesia Terhadap Bahasa Arab," *An Nabighoh* 20, no. 02 (2018): 139, <https://doi.org/10.32332/an-nabighoh.v20i02.1275>.

¹⁶ Yuniar Yuniar and Atikah Marwa, "Al-'Alaqah Baina al-Kafaah al-Mihniyyah Li Mu'allim Wa Raghbat al-Talamidz Fi Ta'allum al-Lughah al-'Arabiyyah Bi Maharah Qiro'ah al-Nash al-'Arabiyyah Bi al-Madrasah al-'Aliyyah al-Muhajirin Muwas Rawas Sumatra al-Junubiyyah," *Jurnal Al Bayan: Jurnal Jurusan Pendidikan Bahasa Arab* 10, no. 1 (2018): 139, <https://doi.org/10.24042/albayan.v10i01.2600>.

وتلخيص واستخدام المعلومات من إجابات الطلاب لاتخاذ قرارات بشأن كل تقييم.¹⁷ للحصول على معايير الاختبار الجيدة والمؤهلة، يجب إجراء التحليل للأسئلة.¹⁸ فالهدف هو تحسين الاختبار من خلال المراجعة أو تجاهل الأسئلة غير الفعالة.¹⁹ أما تحليل عناصر الاختبار، يمكن إجراؤه نوعياً، فيما يتعلق بمحتواه وشكله، وكمياً فيما يتعلق بخصائصه الإحصائية.²⁰

يستخدم هذا البحث تطوير نموذج أدي. في البداية، كان أدي نموذجاً للتنمية في تصميم التعلم ونظرية جيش الولايات المتحدة في خمسينيات القرن العشرين وإنتاج جنود مؤهلين. في منتصف ثمانينيات القرن العشرين، قدم ممارسو التعليم نموذج المطابقة بحيث يمكن تطبيقه في عالم التعليم.²¹ أدي هو اختصار لخطوات تنفيذ تطويره، وهي التحليل والتصميم والتطوير والتنفيذ والتقييم.²² بحيث تبدأ المراحل في هذا البحث بإجراء تحليل، ثم عمل تصميم منتج، ثم يبدأ تطوير تصميم المنتج، ثم التحقق من صحته من قبل الخبراء، ستصبح نتائج هذا التحقق فيما بعد سجلاً لتحسين أو تقييم المنتج الذي تم تطويره.

قد قامت الدراسات السابقة المناسبة، منها دراسة إسابل سينجكاتون²³ ودراسة زكية²⁴ ودراسة أغس ميلو سوسيتيو²⁵ ودراسة فينا نور إنداه ساري²⁶ هناك التشابه بين الدراسات السابقة المناسبة مع هذه الدراسة، فهو لمعرفة جودة عناصر الأسئلة، من خلال معرفة الصلاحية والموثوقية ومستوى الصعوبة

¹⁷ Anthony J. Nitko, *Educational Assessment of Students* (Ohio: Merrill, 1996), 308.

¹⁸ Budi Utomo, "Analisis Validitas Isi Butir Soal Sebagai Salah Satu Upaya Peningkatan Kualitas Pembelajaran Di Madrasah Berbasis Nilai-Nilai Islam," *Jurnal Pendidikan Matematika (Kudus)* 1, no. 2 (2019), <https://doi.org/10.21043/jpm.v1i2.4883>.

¹⁹ Lewis Aiken, *Psychological Testing And Assessment* (Boston: Allyn And Baccon, 1994).

²⁰ Anne Anastasi and Susana Urbina, *Psychological Testing* (New Jersey: Prentice Hall, 1997), 172.

²¹ Fitria Hidayat and Muhamad Nizar, "Model ADDIE (Analysis, Design, Development, Implementation and Evaluation) dalam Pembelajaran Pendidikan Agama Islam," *Jurnal Inovasi Pendidikan Agama Islam (JIPAI)* 1, no. 1 (2021): 28–38, <https://doi.org/10.15575/jipai.v1i1.11042>.

²² Robert A. Reiser and John V. Dempsey, *Trends and Issues in Instructional Design and Technology* (New Jersey: Pearson Education, 2011), 19.

²³ Isabell Sengkaton, Muslih Hambali, and Soni Mirizon, "Teacher-Made Summative Test: An Analysis of Test Format, Index of Difficulty, Discrimination, and Distractors," *LINGUA: Jurnal Bahasa Dan Sastra* 20, no. 2 (2020), <https://ejournal.unsri.ac.id/index.php/lingua/article/view/13343>.

²⁴ Zakiyah Zakiyah and Jamillah Jamillah, "Analisis Butir, Pilihan, dan Reliabilitas Tes Prestasi Bahasa Inggris selama Covid-19," *JURNAL PENDIDIKAN* 30, no. 3 (2021): 547, <https://doi.org/10.32585/jp.v30i3.1911>.

²⁵ Agus Milu Susetyo, "Analisis Butir Soal Ujian Semester Kelas VIII Mata Pelajaran Bahasa Indonesia," *BELAJAR BAHASA: Jurnal Ilmiah Program Studi Pendidikan Bahasa dan Sastra Indonesia* 5, no. 2 (2020): 187, <https://doi.org/10.32528/bb.v5i2.2789>.

²⁶ Vina Nur Indah Sari, Asep Purwo Yudi Utomo, and Sumarwati Sumarwati, "Kualitas Soal Bahasa Indonesia Di SMP Muhammadiyah 1 Pontianak: Analisis Butir Soal," *Jurnal Pendidikan Bahasa Dan Sastra Indonesia* 11, no. 2 (2022): 112, <https://journal.unnes.ac.id/sju/index.php/jpbsi/article/view/58091>.

والقوة التمييزية. أما الفرق بين الدراسات السابقة المناسبة مع هذه الدراسة هو حللت الدراسات السابقة المناسبة عناصر الأسئلة للغة الإنجليزية والإندونيسية، بينما ستحلل هذه الدراسة عناصر أسئلة اختبار الكفاءة العربية للطلاب والتي يتم تلخيصها في أسئلة اختبار الكفاءة العربية.

والتالي، كانت الدراسة التي قامت بها محبة الليلي²⁷ وسيتي فاطمة الفتحي²⁸ فإن التشابه تلك الدراسة السابقة مع هذه الدراسة هو أن كلاهما يحلل عناصر أسئلة الاختبار العربية، أما الفرق بينهما هو أن عناصر الأسئلة التي تم تحليلها هي أسئلة امتحان المواد العربية بمدرسة الثانوية ومدرسة العالية. في حين أن بحث ليلة القمرية²⁹ على الرغم من أن كلاهما حلل الأسئلة العربية، إلا أن تلك الدراسة حللت فقط مستوى الصعوبة والقوة التمييزية، بينما حللت هذه الدراسة صلاحية الأسئلة وموثوقيتها وقوة تمييزيتها ومستوى صعوبتها.

كان تحليل عناصر أسئلة اختبار الكفاءة العربية أمرا مهما للغاية، بالنظر إلى أن اختبار الكفاءة العربية هو أداة تستخدمها جامعة ميترو الإسلامية الحكومية لقياس الكفاءة اللغوية لدى طلابها. بالإضافة إلى ذلك، كان تحليل عناصر أسئلة اختبار الكفاءة العربية كتنظيم لنشاط اختبار الكفاءة العربية من قبل وحدة تطوير اللغة بجامعة ميترو الإسلامية الحكومية. يتم إجراء التقييم لمعرفة كفاءة وفعالية تنفيذ البرنامج في تحقيق أهدافه، وكذلك شكل من أشكال مراقبة جودة التعليم.³⁰

منهج البحث

استخدمت هذه الدراسة نهجا كميا. تهدف هذه الدراسة إلى معرفة عناصر أسئلة اختبار الكفاءة العربية التي تم إعدادها من قبل فريق من محاضري اللغة العربية بجامعة ميترو الإسلامية الحكومية. أما أسلوب جمع البيانات مستخدمة لهذا الدراسة فهما تقنيات التوثيق والمقابلة. استخدمت تقنيات التوثيق للحصول

²⁷ Muhibbatul Laili, "Ketepatan Kontruksi Butir Pilihan Ganda Bahasa Arab," *ALSUNIYAT: Jurnal Penelitian Bahasa, Sastra, Dan Budaya Arab* 3, no. 2 (2020): 111, <https://doi.org/10.17509/alsuniyat.v3i2.25272>.

²⁸ Siti Fathimah Al Fathiyah, "Analisis Butir Soal Pelajaran Bahasa Arab Di MA Roudlotul Ulum Pagak Malang," *Tarbiyatuna: Jurnal Pendidikan Ilmiah* 4, no. 1 (2019): 76–100.

²⁹ Lailatul Qomariyah, "Analisis Tingkat Kesukaran Dan Daya Pembeda Butir Soal TOAFL Universitas Hasyim Asy'ari Tebuireng Jombang," *Lisanan Arabiya: Jurnal Pendidikan Bahasa Arab* 6, no. 1 (2022): 1–18, <https://doi.org/10.32699/liar.v6i1.2549>.

³⁰ Pemerintah Pusat, "Undang-Undang (UU) Nomor 20 Tahun 2003 Tentang Sistem Pendidikan Nasional," accessed May 17, 2023, <http://peraturan.bpk.go.id/Details/43920/uu-no-20-tahun-2003>.

على البيانات في شكل إجابات مخطوطة اختبار الكفاءة العربية لدى طلاب بجامعة ميترو الإسلامية الحكومية في عام ٢٠٢٠. واستخدمت المقابلة للحصول على بيانات في شكل معلومات حول عملية إعداد أسئلة اختبار الكفاءة العربية.

نتائج البحث

الصلاحية

كانت مرحلة إجراء تحليل عناصر أسئلة اختبار الكفاءة العربية بجامعة ميترو الإسلامية الحكومية من خلال مراجعة بعناصر الأسئلة من خلال معلومات من إجابات المشتركين في اختبار الكفاءة العربية. بعد جمع ورقة إجابة اختبار الكفاءة العربية، فالخطوة الأولى هي إجراء تحليل كمي لطبيعة السؤال. تقوم الباحثة أولاً بإنشاء جدول في تطبيق مايكروسوفت إكسل ويكتب الإجابة الصحيحة بقيمة ١ وغير الصحيحة بقيمة ٠ لكل عنصر من عناصر الأسئلة التي أجاب عليها مشترك اختبار. بعد كملت كتابة البيانات المتعلقة بالإجابات الصحيحة والخاطئة، فإن المرحلة التالية هي تحليل مستوى الصعوبة و وقوة التمييزية وفعالية المخادعين باستخدام الصيغ الواردة في تطبيق مايكروسوفت إكسل.

يمكن القيام باختبار صلاحية السؤال كميّة. يتم إجراء التحليل الكيفي قبل اختبار أداة الاختبار. في هذه الحالة، يتم إجراء التحليل الكيفي بناء على قواعد كتابة الأسئلة (الاختبار الكتابي والفعلي والموقفي). عادة ما تتم هذه المراجعة قبل استخدام السؤال. يهتم التحليل الكيفي بعدة جوانب، بما في ذلك المواد والبناء واللغة والثقافة ودلائل التقييم. تشمل الأسلوب المستخدم في التحليل الكيفي منها أساليب الوسيط واللوحة. في تنفيذه، يتم تنفيذ أسلوب الوسيط من خلال جمع خبراء المواد واللغويين وخبراء المناهج واللغويين، ثم مناقشة الأسئلة معا وهناك وسيط واحد.

في محاولة لمعرفة التحليل الكيفي لأسئلة اختبار الكفاءة العربية، قامت الباحثة بمقابلة مع فريق مكون أسئلة اختبار الكفاءة العربية لمعرفة تحليل الصلاحية النوعي الذي قام به فريق مكون الأسئلة من اللغويين. من نتائج المقابلة إلى رئيس فريق مكون أسئلة اختبار الكفاءة العربية الذي كان أيضا رئيس مركز اللغة عند إجراء أداة الاختبار، وجد أنه لم يتم إجراء أي تحليل نوعي أو كمي لاختبار كفاءة اللغة العربية. تبين أنه لم

يتم إجراء أي تحليل كيمي أو كمي لأداة اختبار الكفاءة العربية.³¹ لمعرفة اختبار الصلاحية الكمية، قامت الباحثة باختبار التحليلي باستخدام تطبيق إس بي إس إس. وحصلت بعد التحليل يعرف أن صلاحية أسئلة اختبار الكفاءة العربية بوحدة تطوير اللغة جامعة ميترو الإسلامية الحكومية في عام ٢٠٢٠ هي ٢٤ (٤٨٪) أسئلة في طبقة صالحة و ٢١ (٤٢٪) أسئلة في طبقة غير صالحة.

الموثوقية

رأى باسوكي وهاريانتو عن مستوى موثوقية عنصر السؤال.³² كجدوال التالي:

الجدول ١. مستوى موثوقية عناصر السؤال

مستوى الموثوقية	معامل الموثوقية
الموثوقية سافلة جدا	٠,١٩ - ٠,٤٠
الموثوقية سافلة	٠,٣٩ - ٠,٤٠
الموثوقية مقبولة	٠,٦٩ - ٠,٤٠
الموثوقية عالية	٠,٨٩ - ٠,٧٠
الموثوقية عالية جدا	٠,١٠٠ - ٠,٩٠

يستخدم حساب الموثوقية في هذه الدراسة تطبيق مايكروسوفت إكسل. وبعد التحليل عرف أن موثوقية عناصر أسئلة اختبار الكفاءة العربية وقعت في ٠,٨. بالرجوع إلى الجدول الثاني السابق يمكن الاستنتاج منه على أن أسئلة اختبار الكفاءة العربية بجامعة ميترو الإسلامية الحكومية في عام ٢٠٢٠ ذات علاقة عالية أو موثوقة.

مستوى الصعوبة

عند تحديد تصنيف مستوى صعوبة عنصر السؤال، صاغ سوهارسيمي أريكسينتو على النحو التالي:

الجدول ٢. معايير صعوبة عنصر السؤال

النمرة	مستوى الصعوبة	الطبقة	التقرير
١	١,٠٠ - ٠,٧١	سهلة	مردود
٢	٠,٧٠ - ٠,٣١	متوسطة	مقبول
٣	٠,٣٠ - ٠,٠٠	صعبة	مردود

³¹ Hasil wawancara kepada Bapak Jono Sutarjo pada 6 Juli 2020

³² Ismet Basuki, Hariyanto, and Nita Nur Muliawati, *Asesmen Pembelajaran* (Bandung: Remaja Rosda Karya, 2014), 119.

من الجدول الثالث السابق، فإذا وقعت نتيجة الحساب علي مستوى الصعوبة بين ٠،٣٠ و ٠،٣٠، فعنصر السؤال التي تم اختبارها في الطبقة الصعبة، وكان عنصر السؤال مردود أو يحتاج إلى الاصلاح. والتالي إذا وقعت نتيجة الحساب علي مستوى الصعوبة بين ٠،٣١ و ٠،٧٠، فعنصر السؤال التي تم اختبارها في الطبقة المتوسطة، وكان عنصر السؤال مقبول. أما إذا وقعت نتيجة الحساب علي مستوى الصعوبة بين ٠،٧١ و ١،٠٠، فعنصر السؤال التي تم اختبارها في الطبقة السهلة، وكان عنصر السؤال مردود أو يحتاج إلى الاصلاح. من الوصف السابق، يمكن الاستنتاج أن معيار مستوى الصعوبة الجيد هو ما إذا كان عنصر السؤال التي تم اختبارها في الطبقة المتوسطة، والتي وقع بين ٠،٣١ و ٠،٧٠. أما إذا كان عنصر السؤال يحتوي على طبقة سهلة أو صعبة، يكون عنصر السؤال مردود أو يحتاج إلى الاصلاح.

نتائج التحليل الكمي لمستوى صعوبة اختبار الكفاءة العربية معروضة في الجدول التالي:

الجدول ٣. نتائج التحليل الكمي لمستوى صعوبة اختبار الكفاءة العربية

النمرة	الطبقة	عناصر الأسئلة	المجموعة	النسبة المئوية
١	صعبة جدا	-	-	-
٢	صعبة	٤٨، ٤٤، ٢٥ ١٥، ١٤، ١١، ١٠، ٩، ٦، ٤، ٣، ١ ٢٣، ٢٢، ٢١، ٢٠، ١٩، ١٨، ١٧، ١٦ ٣٣، ٣٢، ٣١، ٣٠، ٢٨، ٢٧، ٢٦، ٢٤ ٤٣، ٤١، ٤٠، ٣٩، ٣٨، ٣٧، ٣٦، ٣٤ ٥٠، ٤٩، ٤٧، ٤٦، ٤٥	ثلاثة عناصر	٦٪
٣	متوسطة	٤٢، ٢٩، ١٣، ١٢، ٨، ٧، ٥، ٢	تسعة وثلاثون عنصرا	٧٨٪
٤	سهلة	-	ثمانية عناصر	١٦٪

بالرجوع إلى جدول مستوى الصعوبة لعناصر الأسئلة السابق، فنتائج تحليل مستوى الصعوبة في أسئلة اختبار الكفاءة العربية بجامعة ميترو الإسلامية الحكومية هي أن ثلاثة عناصر من الأسئلة مردودة، وهي عناصر الأسئلة رقم ٢٥ و ٤٤ و ٤٨ يحتاج إلى الاصلاح أو كانت مردودة لأنها من الطبقة الصعبة وعناصر الأسئلة رقم ٢، ٥، ٧، ٨، ١٢، ١٣، ٢٩ و ٤٢ يحتاج أيضا إلى الاصلاح أو كانت مردودة لأنها تعتبر سهلة. والتالي، من الممكن استخدام الأسئلة التي في الطبقة المتوسطة كبنك أسئلة اختبار الكفاءة العربية.

قوة التمييزية

في تفسير قوة التمييزية، صاغ علي حمزة المعيار³³ في الجدول التالي:

الجدول ٤. تفسير التمييزية

الطبقة	نتائج قوة التمييزية
سيئة جدا (يحتاج الاستبدال)	$Dp \leq 0,000$
سيئة	$0,000 < Dp \leq 0,20$
كافية	$0,20 < Dp \leq 0,40$
جيدة	$0,40 < Dp \leq 0,70$
جيدة جدا	$0,70 < Dp \leq 1,00$

من نتائج تحليل قوة التمييزية في أسئلة اختبار الكفاءة العربية بوحدة تطوير اللغة جامعة ميترو الإسلامية الحكومية حصلت النتائج التالية:

الجدول ٥. جدول تفسير قوة الفرق

النسبة المئوية	المجموع	عناصر الأسئلة	الطبقة	التمرة
44%	عنصران اثنان	٤٥، ٢	سيئة جدا	١
58%	تسعة وعشرون عنصرا	١٦، ١٥، ١٤، ١٣، ٩، ٨، ٧، ٦، ٥، ٤، ٣، ٣١، ٣٠، ٢٩، ٢٧، ٢٥، ٢٠، ١٩، ١٨، ١٧، ٤٧، ٤٦، ٤٣، ٣٦، ٣٥، ٣٤، ٣٣، ٣٢	سيئة	٢
24%	اثنا عشر عنصرا	٤٢، ٣٧، ٢٨، ٢٦، ٢٣، ٢١، ١٢، ١١، ١، ٥٠، ٤٩، ٤٤	كافية	٣
14%	سبعة عناصر	٤١، ٤٠، ٣٩، ٣٨، ٢٤، ٢٢، ١٠	جيدة	٤
0%	-	-	جيدة جدا	٥

بناء على جدول تفسير قوة التمييزية ونتائج حساب قوة التمييزية لأسئلة اختبار الكفاءة العربية السابقة، هناك عنصران من الأسئلة يجب استبدالهما لأنهما من الطبقة السيئة جدا السيئة للغاية، وهما السؤال رقم ٢ ورقم ٤٥.

لمعرفة جودة عناصر أسئلة اختبار الكفاءة العربية، خاصة في أسئلة مهارة القراءة، سيتم عرض ملخص لنتائج الحساب في الجدول التالي:

³³ Ali Hamzah, *Evaluasi Pembelajaran Matematika* (Jakarta: Rajawali Pers, 2014), 243.

الجدول ٦. عرض ملخص لنتائج الحساب

نمرة العنصر	الصلاحية	الموثوقية	مستوى الصعوبة	قوة التمييزية	التقرير
١	صالح	موثوقة	متوسطة	كافية	مقبول
٢	غير صالح	موثوقة	سهلة	سيئة جدا	مردود
٣	غير صالح	موثوقة	متوسطة	سيئة	مردود
٤	غير صالح	موثوقة	متوسطة	سيئة	مردود
٥	غير صالح	موثوقة	سهلة	سيئة	مردود
٦	صالح	موثوقة	متوسطة	سيئة	مقبول
٧	غير صالح	موثوقة	سهلة	سيئة	مردود
٨	غير صالح	موثوقة	سهلة	سيئة	مردود
٩	غير صالح	موثوقة	متوسطة	سيئة	مردود
١٠	صالح	موثوقة	متوسطة	جيدة	مقبول
١١	صالح	موثوقة	متوسطة	كافية	مقبول
١٢	غير صالح	موثوقة	سهلة	كافية	مردود
١٣	غير صالح	موثوقة	سهلة	سيئة	مردود
١٤	غير صالح	موثوقة	متوسطة	سيئة	مردود
١٥	غير صالح	موثوقة	متوسطة	سيئة	مردود
١٦	صالح	موثوقة	متوسطة	سيئة	مقبول
١٧	غير صالح	موثوقة	متوسطة	سيئة	مردود
١٨	غير صالح	موثوقة	متوسطة	سيئة	مردود
١٩	غير صالح	موثوقة	متوسطة	سيئة	مردود
٢٠	صالح	موثوقة	متوسطة	سيئة	مقبول
٢١	صالح	موثوقة	متوسطة	كافية	مقبول
٢٢	صالح	موثوقة	متوسطة	جيدة	مقبول
٢٣	صالح	موثوقة	متوسطة	كافية	مقبول
٢٤	صالح	موثوقة	متوسطة	جيدة	مقبول
٢٥	صالح	موثوقة	صعبة	سيئة	مردود/الاصلاح
٢٦	صالح	موثوقة	متوسطة	كافية	مقبول
٢٧	صالح	موثوقة	متوسطة	سيئة	مقبول
٢٨	صالح	موثوقة	متوسطة	كافية	مقبول
٢٩	صالح	موثوقة	سهلة	سيئة	مردود/الاصلاح
٣٠	صالح	موثوقة	متوسطة	سيئة	مقبول
٣١	صالح	موثوقة	متوسطة	سيئة	مقبول
٣٢	صالح	موثوقة	متوسطة	سيئة	مقبول
٣٣	صالح	موثوقة	متوسطة	سيئة	مقبول

مردود	سيئة	متوسطة	موثوقة	غير صالح	٣٤
مردود	سيئة	متوسطة	موثوقة	غير صالح	٣٥
مقبول	سيئة	متوسطة	موثوقة	صالح	٣٦
مقبول	كافية	متوسطة	موثوقة	صالح	٣٧
مقبول	جيدة	متوسطة	موثوقة	صالح	٣٨
مقبول	جيدة	متوسطة	موثوقة	صالح	٣٩
مقبول	جيدة	متوسطة	موثوقة	صالح	٤٠
مقبول	جيدة	متوسطة	موثوقة	صالح	٤١
مردود/الاصلاح	كافية	سهلة	موثوقة	صالح	٤٢
مردود	سيئة	متوسطة	موثوقة	غير صالح	٤٣
مردود	كافية	صعبة	موثوقة	غير صالح	٤٤
مردود	سيئة جدا	متوسطة	موثوقة	غير صالح	٤٥
مقبول	سيئة	متوسطة	موثوقة	صالح	٤٦
مقبول	سيئة	متوسطة	موثوقة	صالح	٤٧
مردود	سيئة	صعبة	موثوقة	غير صالح	٤٨
مقبول	كافية	متوسطة	موثوقة	صالح	٤٩
مردود	كافية	متوسطة	موثوقة	غير صالح	٥٠

الخلاصة

كان تحليل عناصر أسئلة اختبار اللغة العربية يحلل على وجه التحديد مهارة القراءة، والتحليل الكمي الذي يتم إجراؤه هو تحليل الصلاحية والموثوقية ومستوى الصعوبة والقوة التمييزية. استخدم حساب الموثوقية في هذه الدراسة تطبيق مايكروسوفت إكسل. وبعد التحليل عرف أن موثوقية عناصر أسئلة اختبار الكفاءة العربية وقعت في ٠,٨٠. فيمكن الاستنتاج أن اختبار أسئلة اختبار الكفاءة العربية بجامعة ميترو الإسلامية الحكومية في عام ٢٠٢٠ ذات علاقة عالية أو موثوقة. فنتائج تحليل مستوى الصعوبة في أسئلة اختبار الكفاءة العربية بجامعة ميترو الإسلامية الحكومية هي أن ثلاثة عناصر من الأسئلة مردودة، يحتاج إلى الاصلاح أو كانت مردودة لأنها من الطبقة الصعبة. وثمانية عناصر يحتاج أيضا إلى الاصلاح أو كانت مردودة لأنها تعتبر سهلة. والتالي، من الممكن استخدام الأسئلة التي في الطبقة المتوسطة كبنك أسئلة اختبار الكفاءة العربية. بناء على جدول تفسير قوة التمييزية ونتائج حساب قوة التمييزية لأسئلة اختبار الكفاءة العربية، هناك عنصران من الأسئلة يجب استبدالهما لأنهما من الطبقة السيئة جدا السيئة للغاية.

والتالي، بناء على تلخيص تحليل عناصر اختبار الكفاءة العربية والتي تشمل مدى ملائمة صلاحية الأسئلة وموثوقيتها ومستوى صعوبتها وقوة تمييزيتها، فحصلت أسئلة اختبار الكفاءة العربية بجامعة ميترو الإسلامية الحكومية في هذه الحالة خصيصا لأسئلة مهارة القراءة التي بلغت مجموعتها خمسين سؤالاً على ثلاثة قرارات: (١) هناك ستة وعشرون سؤالاً لا تزال مناسبة أو مقبولة للاستخدام. (٢) هناك ثلاثة أسئلة يجب اصلاحها لإعادة استخدامها أو يمكن أيضاً التخلص منها. (٣) هناك واحد وعشرون سؤالاً غير مناسب أو مردود للاستخدام ويجب التخلص منها.

المراجع

- Ahmad, Muhammad Abdul Qadir. *Thuruq At-Ta'lim Al-Lughah Al-Arabiyyah*. Mesir: Maktabah Annahdhoh, 1979.
- Aiken, Lewis. *Psychological Testing And Assessment*. Boston: Allyn And Baccon, 1994.
- Anastasi, Anne, and Susana Urbina. *Psychological Testing*. New Jersey: Prentice Hall, 1997.
- An-Naqoh, Mahmud Kamil. *Ta'lim Al-Lughah Ittishaliyan Baina Al-Manahij Wa Al-Istiratijiyyat*. Rabath: Issesco, 2006.
- Anwar, Saiful, Guntur Cahaya Kesuma, and Koderi. "Pengembangan Modul Pembelajaran Al-Qawaid an-Nahwiyah Berbasis Metode Qiyasiyah Untuk Mahasiswa Jurusan Pendidikan Bahasa Arab." *Mantiq Tayr: Journal of Arabic Language* 3, no. 1 (2023): 11–24. <https://doi.org/10.25217/mantiqutayr.v3i1.2830>.
- Basuki, Ismet, Hariyanto, and Nita Nur Muliawati. *Asesmen Pembelajaran*. Bandung: Remaja Rosda Karya, 2014.
- Ekawati, Dian. "Ragam Pemahaman Mahasiswa Pada Pembelajaran Bahasa Arab." *An Nabighoh* 21, no. 02 (2019): 183–202. <https://doi.org/10.32332/an-nabighoh.v21i02.1679>.
- Fathiyah, Siti Fathimah Al. "Analisis Butir Soal Pelajaran Bahasa Arab Di MA Roudlotul Ulum Pagak Malang." *Tarbiyatuna: Jurnal Pendidikan Ilmiah* 4, no. 1 (2019): 76–100.
- Hamzah, Ali. *Evaluasi Pembelajaran Matematika*. Jakarta: Rajawali Pers, 2014.
- Hasan, Hasan. "Keterampilan Mengajar Bahasa Arab Materi Istima Menggunakan Media Lagu." *Al Qalam: Jurnal Ilmiah Keagamaan Dan Kemasyarakatan* 11, no. 23 (2018): 127. <https://doi.org/10.35931/aq.v0i0.7>.
- Hidayat, Fitria, and Muhamad Nizar. "Model ADDIE (Analysis, Design, Development, Implementation and Evaluation) dalam Pembelajaran

- Pendidikan Agama Islam." *Jurnal Inovasi Pendidikan Agama Islam (JIPAI)* 1, no. 1 (2021): 28–38. <https://doi.org/10.15575/jipai.v1i1.11042>.
- Imaduddin, Muhammad Faris, Hikmah Maulani, and Ijlal Haziell Taufik. "Test The Validity and Reliability of Arabic Learning Questions." *Arabi: Journal of Arabic Studies* 7, no. 2 (2022): 198–207. <https://doi.org/10.24865/ajas.v7i2.523>.
- Laili, Muhibbatul. "Ketepatan Kontruksi Butir Pilihan Ganda Bahasa Arab." *ALSUNYAT: Jurnal Penelitian Bahasa, Sastra, Dan Budaya Arab* 3, no. 2 (2020): 111. <https://doi.org/10.17509/alsuniyat.v3i2.25272>.
- Muchith, Aim Matun Nihayati Al, and Nely Rahmawati Zaimah. "Relevansi Pembelajaran Bahasa Asing Untuk Anak Usia Dini." *Mantiqutayr: Journal of Arabic Language* 3, no. 1 (2023): 25–35. <https://doi.org/10.25217/mantiqutayr.v3i1.3117>.
- Mustofa, Muhammad Arif. "Interferensi Bahasa Indonesia Terhadap Bahasa Arab." *An Nabighoh* 20, no. 02 (2018): 139. <https://doi.org/10.32332/an-nabighoh.v20i02.1275>.
- Mustofa, Syaiful. "Analisis butir tes TOAFL Universitas Islam Negeri Maulana Malik Ibrahim Malang." Jakarta, 2017. <http://repository.uin-malang.ac.id/6624/>.
- Nitko, Anthony J. *Educational Assessment of Students*. Ohio: Merrill, 1996.
- Pemerintah Pusat. Undang-undang (UU) Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional. Accessed May 17, 2023. <http://peraturan.bpk.go.id/Details/43920/uu-no-20-tahun-2003>.
- Purwanto. *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Pelajar, 2016.
- Qomariyah, Lailatul. "Analisis Tingkat Kesukaran Dan Daya Pembeda Butir Soal TOAFL Universitas Hasyim Asy'ari Tebuireng Jombang." *Lisanan Arabiya: Jurnal Pendidikan Bahasa Arab* 6, no. 1 (2022): 1–18. <https://doi.org/10.32699/liar.v6i1.2549>.
- Reiser, Robert A., and John V. Dempsey. *Trends and Issues in Instructional Design and Technology*. New Jersey: Pearson Education, 2011.
- Sari, Vina Nur Indah, Asep Purwo Yudi Utomo, and Sumarwati Sumarwati. "Kualitas Soal Bahasa Indonesia Di SMP Muhammadiyah 1 Pontianak: Analisis Butir Soal." *Jurnal Pendidikan Bahasa Dan Sastra Indonesia* 11, no. 2 (2022): 112. <https://journal.unnes.ac.id/sju/index.php/jpbsi/article/view/58091>.
- Sengkaton, Isabell, Muslih Hambali, and Soni Mirizon. "Teacher-Made Summative Test: An Analysis of Test Format, Index of Difficulty, Discrimination, and Distractors." *LINGUA: Jurnal Bahasa Dan Sastra* 20, no. 2 (2020). <https://ejournal.unsri.ac.id/index.php/lingua/article/view/13343>.
- Susetyo, Agus Milu. "Analisis Butir Soal Ujian Semester Kelas VIII Mata Pelajaran Bahasa Indonesia." *BELAJAR BAHASA: Jurnal Ilmiah Program Studi Pendidikan*

- Bahasa dan Sastra Indonesia* 5, no. 2 (2020): 187. <https://doi.org/10.32528/bb.v5i2.2789>.
- Thu'aimah, Rusydi Ahmad. *Al-Marja' Fi Manhaj At-Ta'lim Al-Lughah Al-Arabiyyah Li An-Natiqin Bi Lughat Al-Ukhra*. Al-Qahirah: Dar Al-Fikr Al-Arabi, n.d.
- Ulya, Nia Himatul, Chairani Astina, and Ashief El Qorny. "Implementasi Bi'ah Lughawiyah Dalam Peningkatan Maharah Kalam Di Pondok Modern Az-Zahra al-Gontory Purwokerto." *Mantiqutayr: Journal of Arabic Language* 2, no. 2 (2022): 174. <https://doi.org/10.25217/mantiqutayr.v2i2.2511>.
- Unsi, Baiq Tuhfatul, and Siti Robiatun Muniroh. "Penerapan Metode Jigsaw Dalam Pembelajaran Qiro'ah Di MTs Darul Ulum Kepuhdoko Tembelang Jombang." *Mantiqutayr: Journal of Arabic Language* 2, no. 2 (2022): 161. <https://doi.org/10.25217/mantiqutayr.v2i2.2474>.
- Utami, Sulisty Seti, and Bayu Pirmansyah. "Peningkatan Skor Test Bahasa Inggris (TOEFL) Melalui Pelatihan Secara Intensif." *Lingua Franca: Jurnal Bahasa, Sastra, dan Pengajarannya* 2, no. 2 (2018): 36-46. <http://dx.doi.org/10.30651/lf.v2i2.1477>.
- Utomo, Budi. "Analisis Validitas Isi Butir Soal Sebagai Salah Satu Upaya Peningkatan Kualitas Pembelajaran Di Madrasah Berbasis Nilai-Nilai Islam." *Jurnal Pendidikan Matematika (Kudus)* 1, no. 2 (2019). <https://doi.org/10.21043/jpm.v1i2.4883>.
- Wulandari, Nawang. "Rekonstruksi Pembelajaran Bahasa Arab." *An Nabighoh* 21, no. 01 (2019): 114. <https://doi.org/10.32332/an-nabighoh.v21i01.1631>.
- Yuniar, Yuniar, and Atikah Marwa. "Al-'Alaqah Baina al-Kafaah al-Mihniyyah Li Mu'allim Wa Raghbat al-Talamidz Fi Ta'allum al-Lughah al-'Arabiyyah Bi Maharah Qiro'ah al-Nash al-'Arabiyyah Bi al-Madrasah al-'Aliyyah al-Muhajirin Muwas Rawas Sumatra al-Junubiyyah." *Jurnal Al Bayan: Jurnal Jurusan Pendidikan Bahasa Arab* 10, no. 1 (2018): 139. <https://doi.org/10.24042/albayan.v10i01.2600>.
- Zakiyah, Zakiyah, and Jamillah Jamillah. "Analisis Butir, Pilihan, dan Reliabilitas Tes Prestasi Bahasa Inggris selama Covid-19." *JURNAL PENDIDIKAN* 30, no. 3 (2021): 547. <https://doi.org/10.32585/jp.v30i3.1911>.

