Kalamuna: Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban
P-ISSN: 2655-4267, E-ISSN: 2745-6943
Vol. 07, No. 2, July 2026
DOI: *https://doi.org/10.52593/klm.07.2.12*

Analysis of Arabic Language Madrasah Assessment Test Items Based on Educational Assessment Standards

تحليل أسئلة الامتحان النهائي لمادة اللغة العربية في ضوء معايير التقويم

Mohammad Zaelani Musonif[1*], Habib Maulana Maslahul Adi[2], Ahmad Nawirul Huda[3], Muhammad Yahya Abdullah[4]

[1]International Open University, Gambia
[2]Universitas Islam Negeri Sunan Kudus, Indonesia
[3]Universitas Negeri Malang, Indonesia
[4] Universitas Islam Negeri Maulana Malik Ibrahim, Malang, Indonesia
**Corresponding E-mail: zaelani@bahasa.iou.edu.gm*

Keywords:
Arabic language assessment; Difficulty index; Item analysis; Reliability; Validity.

Abstract

Analyzing test items is essential for improving the quality of the Arabic language curriculum, as it helps identify weaknesses in the evaluation process. This study aims to examine the validity, reliability, difficulty level, and discrimination power of Arabic language questions used in the Madrasah Assessment at MA Al-Ikhlas Nanggulan for the 2022–2023 academic year. A descriptive quantitative method was employed, with 40 multiple-choice questions as the research population. Data were collected through documentation of the answer sheets of 25 twelfth-grade students. Data analysis included the Product-Moment correlation for validity, Cronbach's Alpha for reliability, the difficulty index (P), and the discrimination index (D). The results showed that the test items were generally of good quality. Of the 40 items, 65% were valid, while 35% were invalid. The reliability coefficient reached 0.915, indicating very high reliability. In terms of difficulty, 17.5% of the items were easy, 70% were moderate, and 12.5% were difficult. Regarding discrimination power, 62.5% of the items were classified as good, 17.5% as moderate, 12.5% as poor, 5% as ineffective, and 2.5% as negative. These findings indicate that item analysis is an effective tool for evaluating and improving the quality of Arabic language assessments.

الكلمات المفتاحية:
تقويم اللغة العربية؛ معامل الصعوبة؛ تحليل أسئلة الاختبار؛ الثبات؛ الصدق.

ملخص البحث

يُسهم تحليل أسئلة الاختبار في تحسين جودة منهج اللغة العربية من خلال الكشف عن جوانب القوة والضعف في عملية التقويم. يهدف هذا البحث إلى تحليل صدق الأسئلة وثباتها ومعاملات الصعوبة والتمييز للكشف عن جودة أسئلة اللغة العربية في الامتحان النهائي بمدرسة الإخلاص الثانوية الإسلامية نانغولان للعام الدراسي ٢٠٢٢–٢٠٢٣. اعتمد البحث المنهج الوصفي الكمي، وتكوّنت العينة من ٤٠ سؤالًا من نوع الاختيار من متعدد، وجُمعت البيانات من إجابات ٢٥ طالبًا باستخدام أسلوب

التوثيق. وتم تحليل البيانات باستخدام معامل ارتباط بيرسون لقياس الصدق، ومعامل ألفا كرونباخ لقياس الثبات، ومؤشر الصعوبة (P)، ومعامل التمييز(D). أظهرت النتائج أن جودة الأسئلة كانت جيدة بوجه عام؛ إذ بلغت نسبة الأسئلة الصادقة ٦٥٪، في حين بلغت نسبة الأسئلة غير الصادقة ٣٥٪. كما حققت الأسئلة درجة ثبات مرتفعة جدًا بمعامل ألفا كرونباخ قدره (٠,٩١٥). ومن حيث الصعوبة، توزعت الأسئلة إلى ١٧,٥٪ سهلة، و٧٠٪ متوسطة، و١٢,٥٪ صعبة. أما من حيث التمييز، فبلغت نسبة الأسئلة ذات التمييز الجيد ٦٢,٥٪، والمتوسط ١٧,٥٪، والضعيف ١٢,٥٪، وغير المميزة ٥٪، في حين بلغت نسبة الأسئلة ذات التمييز السلبي ٢,٥%.

Article Information | **Submitted 2026-06-25. Received 2026-06-26. Revised 2026-07-30. Accepted 2026-07-30. Published 2026-07-31.**

المقدمة

التقويم الذي يحتوي على تمارين وأسئلة هو أحد المكونات الهامة في المنهج الدراسي (Agung, 2022). ولذا، فإن دراسة البحث المتعلقة بتحليل الأسئلة في مناهج اللغة العربية يمكن أن تساهم في تطوير مناهج اللغة العربية في إندونيسيا. على سبيل المثال، قد يساعد دراسة تحليل أسئلة الإختبار لدرس اللغة العربية لطلاب مدرسة روضة العلوم الثانوية الإسلامية باغاك مالانغ يساعد في تحسين جودة الأسئلة وأساليب التدريس المستخدمة في المدرسة (Fathiyah, 2019). يمكن استخدام نتائج تحليل البحث لتطوير مناهج اللغة العربية الأكثر فاعلية وكفاءة، والتي يمكن أن ترفع جودة تعليم اللغة العربية (Maulana & Sanusi, 2020, p. 22) .

تحليل الأسئلة هو جزء من عملية تقويم اللغة العربية وينطوي على تقويم وتحسين جودة منهج اللغة العربية (Nurbayan et al., 2021). من خلال تحليل الأسئلة المستخدمة في التدريس، يمكننا تقويم صحة الأسئلة، ومدى موثوقيتها، ومستوى صعوبتها، وقدرتها على تمييز الطلاب. ويمكن استخدام هذه النتائج كأساس لتحقيق التوافق مع أهداف التعلم والارتباط بسياق حياة الطلاب الواقعي (Gusmizain, 2022). و يمكن أن يساعد التحليل في تحديد النقاط الضعيفة أو النقائص في المنهج الحالي وتوفير الأساس لإجراء التحسينات

اللازمة. وكذلك هذه الدراسة توضح لنا جودة الاختبار التي يضعها المعلِّمُ مؤشرًا غيرَ مباشرٍ على مستوى كفاءته المهنية، ولا سيما فيما يتعلَّق بجانب تقويم التعلُّم (Alamsyah et al., 2025). وبالتالي، يمكن أن يسهم تحليل الأسئلة في تحسين مستدام لمنهج اللغة العربية في إندونيسيا.

بالإضافة إلى ذلك، يمكن أن يساهم تحليل الأسئلة في اللغة العربية في تطوير استراتيجيات التدريس الأكثر فاعلية. من خلال فهم أنواع الأسئلة الفعالة في قياس قدرات الطلاب، يمكن للمعلمين تطوير أساليب تدريس أكثر توجهًا وتركيزًا على جوانب مهمة في اللغة العربية كمثل قواعد النحو والمفردات وفهم النصوص. وبالتالي، فإن دراسة تحليل أسئلة اختبار اللغة العربية في مدرسة الإخلاص الثانوية الإسلامية نانغولان في السنة الدراسية ٢٠٢٢-٢٠٢٣ تعد خطوة هامة لضمان جودة الأسئلة التي يمكن أن تكون أساسًا لتدريس اللغة العربية. وله هدف يعني تحقيق أهداف التعلم المرجوة وتوفير الفوائد للطلاب.

الأبحاث السابقة المتعلقة بتحليل أسئلة اللغة العربية تشمل بحث جنة الذي قامت فيه بتحليل أسئلة اللغة العربية من جوانب المؤشرات والكفاءات، وبناء الأسئلة، والمحتوى (Jannah, 2022). ثم قامت قمرية وحاكم وأخرين أيضًا ببحث حول تحليل قوة التمييز وصعوبة أسئلة اللغة العربية في اختبار *TOAFL* (Hakiem et al., 2025; Qomariyah, 2022). أضافت أمبيا وآخرون بحثًا حول تحليل الأسئلة من خلال تحليل امتحان اللغة العربية في المدرسة الثانوية المتعلق بالعلاقة بين الأسئلة وقدرات استخدام اللغة، والصدق والثبات، وجوانب التفكير العالي (Ambiya et al., 2022). ثم قام رمضان وفردوس بدراسة تحليلية لأسئلة اللغة العربية بمنهج نوعي فيما يتعلق بأنواع الأسئلة وتنفيذها في عملية التعلم (Ramadhan & Firdaus, 2022). ثم قام فهمي وآخرون بتحليل أسئلة اللغة العربية في المدرسة الثانوية بنمط المدرسة في المعهد الإسلامي فيما يتعلق بأربعة جوانب هي الصدق والثبات وقوة التمييز ودرجة الصعوبة (Fahmi et al., 2022). ومخيصة وسروية قد قامتا

بتحليل أسئلة اختبار المدرسة في اللغة العربية باستخدام نظرية التصنيف الضريبي لبلوم (Makhisoh & Suroyya, 2022).

من البيان السابق علم على هذه الدرسة تختلف بالأبحاث المذكورة أعلاه. سيتم تركيز هذا البحث على أربعة جوانب وهي صدق الأسئلة وثبات الاختبار درجة صعوبة الأسئلة ودرجة التمييز لكل سؤال في اختبار اللغة العربية في مدرسة الإخلاص الثانوية الإسلامية نانغولان في السنة الدراسية ٢٠٢٢-٢٠٢٣. وعلى الرغم من أن الدراسة مشابهة لدراسة فهمي وآخرين. ولكنه يختلف موضوع هذا البحث ببحث فهمي وهو اختبار اللغة العربية في مدرسة الإخلاص الثانوية الإسلامية نانغولان.

منهج البحث

استخدم هذا البحث منهج البحث الوصفي بنهج كمّي الذي يتضمن استخدام الأرقام في مراحل مختلفة مثل جمع البيانات وتفسيرها وتقديم نتائج البحث (Arikunto, 2013, p. 12). البحث الوصفي هو شكل من أشكال البحث يهدف إلى فهم قيم المتغيرات المستقلة، سواء كانت متغيرًا واحدًا أو أكثر دون المقارنة أو الربط بمتغيرات أخرى (Sugiyono, 2012, p. 13).

يشمل عينة الدراسة في هذا البحث أسئلة اللغة العربية في اختبار اللغة العربية في مدرسة الإخلاص الثانوية الإسلامية نانغولان في السنة الدراسية ٢٠٢٢-٢٠٢٣ التي تبلغ ٤٠ سؤالًا. تتمثل البيانات المستخدمة في هذا البحث في ورقة الإجابة للطلاب التي يتم معالجتها بتحليل صدق الأسئلة وثبات الاختبار درجة صعوبة الأسئلة وقدرة الأسئلة سؤال على تفريق الطلاب. ويكون نتيجة البحث تقييما لجودة امتحان اللغة. وتم جمع البيانات بتقنية الوثيقة المتعلقة برقة الأسئلة وإجاباتها.

وأما تحليل البيانات فاستخدام تطبيق *Microsoft Excel* و *IBM SPSS 26* لحصول على تحليلات إحصائية للبحث. وتحليل صدق البيانات فيستخدم اختبار الارتباط الإنتاجي (*product moment*). وتحليل الثبات استخدم اختبار واختبار ألفا كرونباخ (*Cronbach's Alpha*). وتحليل درجة الصعوبة استخدم صيغة مؤشر الصعوبة (P). وتحليل القدرة على التفريق استخدم حسابات إحصائية وصيغة مؤشر القدرة على التفريق (P).

نتائج البحث

تحليل الأسئلة يستند إلى أربعة عناصر مهمة وهي صدق الأسئلة، وثبات الأسئلة، ودرجة الصعوبة، وقدرة الأسئلة على تمييز الطلاب في الفئة العليا والفئة السفلى. لا تقتصر أهميةُ دراسةِ تحليلِ أسئلة الاختبار على تقويم جودة أدوات التقييم فحسب، بل يمكن أن تُعَدَّ أيضًا أحدَ المؤشرات على كفاءة معلِّم اللغة العربية، ولا سيما في مجال تصميم أدوات التقويم وتنفيذها. فتعكس جودةُ أدوات التقويم التي يُعِدُّها المعلِّم، بما في ذلك الأسئلة في الاختبارات، مدى قدرتها على تطبيق مبادئ التقويم التربوي تطبيقًا سليمًا، مثل الصدق والثبات ومعامل الصعوبة ومعامل التمييز وفاعلية المشتتات. ومن ثمَّ، تؤدي كفاءةُ المعلِّم دورًا أساسيًّا في ضمان جودة عمليتَي التعليم والتقويم (Zahro' & Alfian, 2026).

صدق الأسئلة في الاختبار النهائي المدرسي لدرس اللغة العربية

يمكن معرفة قيمة صدق عنصر السؤال من خلال حساب قيمة الارتباط باستخدام صيغة ارتباط المنتج اللحظي. لحساب معامل الارتباط باستخدام ذلك الرمز، تم إجراء حسابات باستخدام الصيغة التالية (Syafril, 2019, p. 92):

$$r_{xy} = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{[N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2]}}$$

البيان:

r_{xy} : معامل الارتباط (ارتباط المنتج اللحظي)

X و Y : قيمتين من المتغيرات المرتبطة

N : عدد الطلاب

بعد معرفة قيمة ارتباط المنتج اللحظي (*product moment*)، يتم مقارنة نتائج حساب rxy مع قيمة المعامل في جدول r ارتباط المنتج اللحظي. إذا كانت قيمة r_{xy} أقل من قيمة المعامل في جدول r ارتباط المنتج اللحظي، فإن الارتباط في عنصر السؤال غير معنوي. وبالتالي، يعتبر عنصر الأسئلة غير صدق. وبالعكس إذا كانت قيمة r_{xy} أكبر من قيمة المعامل في جدول r أو تساوي به، فإن الارتباط في عنصر الأسئلة هو معنوي. وبالتالي يعتبر عنصر الأسئلة صدقا (Sumardi, 2020, p. 84).

بالإضافة إلى مراعاة قيمة معامل الارتباط، يمكن أيضًا معرفة صدق السؤال من خلال حساب قيمة الدلالة في اختبار ارتباط المنتج اللحظي باستخدام برنامج SPSS. تم تحديد معايير اتخاذ القرار على النحو التالي (Gunawan, 2020, p. 93):

أ. إذا كانت قيمة الدلالة أقل من ٠,٠٥، يعتبر السؤال صدقا.

ب. إذا كانت قيمة الدلالة أكبر من ٠,٠٥، يعتبر السؤال غير صدق.

بناءً على الوصف أعلاه، يتم الحصول على النتائج بعد إجراء اختبار ارتباط المنتج اللحظي باستخدام برنامج *IBM SPSS 26* على أسئلة الاختبار النهائي لدرس اللغة العربية بمدرسة الإخلاص الثانوية العالية على النحو التالي:

جدول ١. نتائج اختبار الارتباط بواسطة اختبار المنتج اللحظي

رقم السؤال	قيمة الارتباط	قيمة r في الجدول	قيمة الدلالة	تفسير
١	٠.٠٠٠	٠.٣٩٦	-	غير صدق
٢	٠.٦٩٠	٠.٣٩٦	٠.٠٠٠	صدق
٣	٠.١٨٤	٠.٣٩٦	٠.٣٧٨	غير صدق
٤	٠.٥٩٧	٠.٣٩٦	٠.٠٠٢	صدق

٥	٠.٢٧٢	٠.٣٩٦	٠.١٨٨	غير صدق
٦	٠.٣٩٤	٠.٣٩٦	٠.٠٥١	غير صدق
٧	٠.٧٢٣	٠.٣٩٦	٠.٠٠٠	صدق
٨	٠.١٨٤	٠.٣٩٦	٠.٣٧٧	غير صدق
٩	٠.٧٩٤	٠.٣٩٦	٠.٠٠٠	صدق
١٠	٠.٢٣٠	٠.٣٩٦	٠.٢٦٩	غير صدق
١١	٠.٤١٨	٠.٣٩٦	٠.٠٣٨	صدق
١٢	٠.٦٢٩	٠.٣٩٦	٠.٠٠١	صدق
١٣	٠.٥٤٥	٠.٣٩٦	٠.٠٠٥	صدق
١٤	٠.٤٧٩	٠.٣٩٦	٠.٠١٦	صدق
١٥	٠.٨٣٠	٠.٣٩٦	٠.٠٠٠	صدق
١٦	٠.٥٧١	٠.٣٩٦	٠.٠٠٣	صدق
١٧	٠.٥٠٣	٠.٣٩٦	٠.٠١٠	صدق
١٨	٠.٠٩٢	٠.٣٩٦	٠.٦٦٢	غير صدق
١٩	٠.٣٦٩	٠.٣٩٦	٠.٠٧٠	غير صدق
٢٠	٠.٦٩٥	٠.٣٩٦	٠.٠٠٠	صدق
٢١	٠.٥٢٥	٠.٣٩٦	٠.٠٠٧	صدق
٢٢	٠.٣٦٣	٠.٣٩٦	٠.٠٧٤	غير صدق
٢٣	٠.٢٧١	٠.٣٩٦	٠.١٩١	غير صدق
٢٤	٠.٠٤٥	٠.٣٩٦	٠.٨٣٠	غير صدق
٢٥	٠.٦٦٧	٠.٣٩٦	٠.٠٠٠	صدق
٢٦	٠.٣٧٦	٠.٣٩٦	٠.٠٦٤	غير صدق

٢٧	٠.٤٦٦	٠.٣٩٦	٠.٠٢٢	صدق
٢٨	٠.٧٠٤	٠.٣٩٦	٠.٠٠٠	صدق
٢٩	-٠.٠٩٧	٠.٣٩٦	٠.٦٤٤	غير صدق
٣٠	٠.٧٤٥	٠.٣٩٦	٠.٠٠٠	صدق
٣١	٠.٦٣٨	٠.٣٩٦	٠.٠١٠	صدق
٣٢	٠.٦٩٥	٠.٣٩٦	٠.٠٠٠	صدق
٣٣	٠.٧٢٨	٠.٣٩٦	٠.٠٠٠	صدق
٣٤	٠.٥٢٧	٠.٣٩٦	٠.٠٠٧	صدق
٣٥	٠.٦١٤	٠.٣٩٦	٠.٠١٦	صدق
٣٦	٠.٧١٨	٠.٣٩٦	٠.٠٠٠	صدق
٣٧	٠.٤٧٩	٠.٣٩٦	٠.٠١٥	صدق
٣٨	٠.٧١٤	٠.٣٩٦	٠.٠٠٠٦١٦	صدق
٣٩	٠.١٠٦	٠.٣٩٦	٠.٦١٦	غير صدق
٤٠	٠.٤٨٢	٠.٣٩٦	٠.١٧	صدق

بناءً على النتائج أعلاه، يمكن الاستنتاج أنه من أربعين سؤالًا في اختبار اللغة العربية في مدرسة الإخلاص الثانوية الإسلامية نانغولان في السنة الدراسية ٢٠٢٢-٢٠٢٣، هناك ٢٦ سؤالًا صديقا و ١٤ سؤالًا غير الصدق. الأسئلة الصديقة هي الأسئلة في رقم ٢ و ٤ و ٧ و ٩ و ١١ و ١٢ و ١٣ و ١٤ و ١٥ و ١٦ و ١٧ و ٢٠ و ٢١ و ٢٥ و ٢٧ و ٢٨ و ٣٠ و ٣١ و ٣٢ و ٣٣ و ٣٤ و ٣٥ و ٣٦ و ٣٧ و ٣٨ و ٤٠. وأما الأسئلة غير الصدق فهي الأسئلة في رقم ١ و ٣ و ٥ و ٦ و ٨ و ١٠ و ١٨ و ١٩ و ٢٢ و ٢٣ و ٢٤ و ٢٦ و ٢٩ و ٣٩.

ثبات الأسئلة في الاختبار النهائي المدرسي لدرس اللغة العربية

تم اختبار ثبات أسئلة اختبار اللغة العربية في مدرسة الإخلاص الثانوية الإسلامية نانغولان في السنة الدراسية ٢٠٢٢-٢٠٢٣ باستخدام تقنية كرونباخ ألفا (*Cronbach's Alpha*) بمساعدة برنامج *IBM SPSS 26*. يتم تحديد قرار القبول أو الرفض في هذا الاختبار عن طريق مقارنة قيمة كرونباخ ألفا مع مستوى الدلالة (٠,٦٠). إذا كانت قيمة كرونباخ ألفا أكبر من مستوى الدلالة (٠,٦٠)، فإن الأسئلة تعتبر ثباتة. إذا كانت قيمة كرونباخ ألفا أقل من مستوى الدلالة (٠,٦٠)، فإن الأسئلة غير ثباتة (Darma, 2021, p. 17). يمكن تصنيف قيمة الثبات المحصلة في الفئات التالية (Djiwandono, 1996):

جدول ٢. فئات الثبات/ الموثوقية

قيمة معامل الارتباط	فئة الموثوقية
٠,٩٠ – ١,٠٠	عالية جدًا
٠,٧٠ – ٠,٨٩	عالية
٠,٥٠ – ٠,٦٩	متوسطة
٠,٣٠ – ٠,٤٩	منخفضة
أقل من ٠,٣٠	منخفضة جدًا

اجريت اختبار الثبات أو الموثوقية للأسئلة مرتين. الاختبار الأول لمجموعة الأسئلة الصديقة فقط والاختبار الثاني لمجموعة الأسئلة سواء الصديقة وغير الصديقة. وقد تم استخدام تقنية كرونباخ ألفا (*Cronbach's Alpha*) وتطبيق برنامج *IBM SPSS 26* للحصول على نتائج الاختبار التالية:

Reliability Statistics

Cronbach's Alpha	N of Items
.932	26

Reliability Statistics

Cronbach's Alpha	N of Items
.915	40

الصورة ١. نتائج اختبار الثبات أو الموثوقية للأسئلة

بناءً على النتائج المذكورة أعلاه، يمكن ملاحظة أن قيمة كرونباخ ألفا لمجموعة الأسئلة الصديقة فقط هي ٠,٩٣٢، وللمجموعة الكاملة من الأسئلة هي ٠,٩١٥. كلا القيمتين أعلى من قيمة المستوى الدلالي (٠,٦٠). وبالتالي، يمكن الاستنتاج بأن أسئلة اللغة العربية في الإمتحان النهائي في مدرسة الإخلاص الثانوية الإسلامية نانغولان في السنة الدراسية ٢٠٢٢-٢٠٢٣ هي ثابتة أو موثوقة. ووفقًا لتصنيف جيواندونو (١٩٩٦)، يندرج هذا السؤال في فئة الأسئلة التي لديها ثابتة عالية جدًا لأنه يقع ضمن نطاق القيم بين ٠,٩ و ١,٠.

درجة صعوبة الأسئلة في الاختبار النهائي المدرسي لدرس اللغة العربية

درجة صعوبة الأسئلة هو قياس يهدف إلى تحديد مدى صعوبة السؤال. ووفقًا لما ذكره أريكونتو(2009)، يجب أن يكون الاختبار المثالي لا يكون سهلاً جداً ولا صعباً جداً. فالأسئلة السهلة جداً تجعل الطلاب لا يكتسبون مهارات التفكير العالي، بينما الأسئلة الصعبة جداً تجعل الطلاب يستسلمون ويشعرون باليأس (Hanim, 2018, p. 144). بعبارة أخرى، إذا تمت الإجابة على سؤال الاختبار بشكل صحيح من قبل جميع الطلاب أو تمت الإجابة على السؤال بشكل خاطئ من قبل جميع الطلاب، فإن الاختبار ليس جيدًا.

صيغة التي تُستخدم لحساب نسبة درجة الصعوبة هي على النحو التالي (Marudut :& Akbar, 2022, p. 258)

$$P = \frac{B}{JS}$$

التوضيح:

P: مستوى الصعوبة

B: عدد الطلاب الذين أجابوا بشكل صحيح

JS: عدد المشاركين في الاختبار

لتفسير درجة الصعوبة، قام أريكونتو (٢٠١١) بوضع المعايير التالية: (Fikriyah, 2021)

- $P > 0.70$ = سهل
- $P \geq 0.70 \geq 0.03$ = متوسط
- $P < 0.30$ = صعب

نتائج حساب درجة صعوبة كل سؤال في اختبار اللغة العربية في مدرسة الإخلاص الثانوية الإسلامية نانغولان في السنة الدراسية ٢٠٢٢-٢٠٢٣ هي كما يلي:

جدول ٣. نتائج حساب درجة صعوبة كل سؤال

رقم السؤال	B	JS	P	التفسير
١	٢٥	٢٥	١	سهل
٢	١٤	٢٥	٠.٥٦	متوسط
٣	٢٤	٢٥	٠.٩٦	سهل
٤	٢٠	٢٥	٠.٨	سهل
٥	١١	٢٥	٠.٤٤	متوسط
٦	٢٠	٢٥	٠.٨	سهل
٧	٧	٢٥	٠.٢٨	صعب
٨	٢١	٢٥	٠.٨٤	سهل
٩	١٢	٢٥	٠.٤٨	متوسط
١٠	١٠	٢٥	٠.٤	متوسط
١١	١٤	٢٥	٠.٥٦	متوسط
١٢	١٤	٢٥	٠.٥٦	متوسط
١٣	١٤	٢٥	٠.٥٦	متوسط

١٤	١٥	٢٥	٠.٦	متوسط
١٥	١٢	٢٥	٠.٤٨	متوسط
١٦	١٠	٢٥	٠.٤	متوسط
١٧	٨	٢٥	٠.٣٢	متوسط
١٨	٢٤	٢٥	٠.٩٦	سهل
١٩	٤	٢٥	٠.١٦	صعب
٢٠	١٦	٢٥	٠.٦٤	متوسط
٢١	١٥	٢٥	٠.٦	متوسط
٢٢	١١	٢٥	٠.٤٤	متوسط
٢٣	٨	٢٥	٠.٣٢	متوسط
٢٤	١٣	٢٥	٠.٥٢	متوسط
٢٥	٩	٢٥	٠.٣٦	متوسط
٢٦	١٦	٢٥	٠.٦٤	متوسط
٢٧	١٣	٢٥	٠.٥٢	متوسط
٢٨	٩	٢٥	٠.٣٦	متوسط
٢٩	٣	٢٥	٠.١٢	صعب
٣٠	١٤	٢٥	٠.٥٦	متوسط
٣١	١٧	٢٥	٠.٦٨	متوسط
٣٢	١٦	٢٥	٠.٦٤	متوسط
٣٣	٦	٢٥	٠.٢٤	صعب
٣٤	٢٢	٢٥	٠.٨٨	سهل
٣٥	١٢	٢٥	٠.٤٨	متوسط

٣٦	١٥	٢٥	..٦٠	متوسط
٣٧	١٦	٢٥	..٦٤	متوسط
٣٨	١٦	٢٥	..٦٤	متوسط
٣٩	٦	٢٥	..٢٤	صعب
٤٠	١٧	٢٥	..٦٨	متوسط

من الحسابات المذكورة أعلاه ، يمكننا معرفة أن الأسئلة في اختبار اللغة العربية في مدرسة الإخلاص الثانوية الإسلامية نانغولان في السنة الدراسية ٢٠٢٢-٢٠٢٣ تتكون من أسئلة في فئات السهولة والمتوسطة والصعوبة وفقًا للتفاصيل التالية:

أ. هناك سبعة أسئلة تندرج في فئة السهولة ، وهي الأسئلة رقم ١ و٣ و٤ و٦ و٨ و١٨ و ٣٤.

ب. هناك ٢٨ سؤالًا يندرج في فئة المتوسطة ، وهي الأسئلة رقم ٢ و٥ و٩ و١٠ و١١ و١٢ و١٣ و١٤ و١٥ و١٦ و١٧ و٢٠ و٢١ و٢٢ و٢٣ و٢٤ و٢٥ و٢٦ و٢٧ و٢٨ و٣٠ و ٣١ و٣٢ و٣٥ و٣٦ و٣٧ و٣٨ و٤٠.

ج. هناك خمسة أسئلة تندرج في فئة الصعوبة ، وهي الأسئلة رقم ٧ و١٩ و٢٩ و٣٣ و ٣٩.

درجة التمييز في الاختبار النهائي المدرسي لدرس اللغة العربية

درجة التمييز هي قدرة السؤال على التمييز بين الطلاب الماهرين أو ذوي القدرات العالية والطلاب الضعفاء أو ذوي القدرات المنخفضة (Arikunto, 2009, p. 211). بالنسبة لهذا الاختبار، يتم تقسيم الطلاب إلى مجموعة من طلاب الصف العلوي (الطبقة العليا) ومجموعة من طلاب الصف الوسطى ومجموعة من طلاب الصف السفلي. هناك ٢٥ طالبًا في الصف الثاني عشر العربية في مدرسة الإخلاص الثانوية الإسلامية نانغولان الذين شاركوا في اختبار اللغة المدرسة في السنة الدراسية ٢٠٢٢-٢٠٢٣. من ذلك العدد، يتم أخذ ٢٧٪

كمجموعة عليا و ٢٧٪ كمجموعة سفلي. ولذلك تأخذ سبعة طلاب (وهو ناتج تقريبي للرقم ٦,٧٥) ينتمون إلى المجموعة العليا والمجموعة السفلي. الصيغة المستخدمة لتحديد القدرة التمييزية (D) لاختبار سؤال هي كما يلي (Asrori et al., 2017):

$$D = \frac{\sum \text{الإجابات الصحيحة للمجموعة العليا} \; - \; \sum \text{الإجابات الصحيحة للمجموعة الدنيا}}{\sum \text{أفراد المجموعتين}} \times 100\%$$

في تحديد نطاق مؤشر قدرة التمييز (D) لاختبار، يمكن استخدام الصيغة التي اقترحها دجيواندونو على النحو التالي (Djiwandono, 1996, p. 144):

- D > 0.50 : جيدة
- D ≥ 0.50 ≥ ٠,٢٠ : متوسط
- D < 0.2 : ضعيفة
- D = 0 : لا يوجد تمييز
- D = - (سالب) : سلبي

أما نتائج حساب قدرة التمييز للأسئلة فيمكن معرفتها من الجدول التالي:

جدول ٤. نتائج حساب قدرة التمييز للأسئلة

رقم السؤال	الصحيحة فئة العلوي	الصحيحة فئة السفلي	الفارق	عدد الطلاب	درجة التمييز	التفسير
١	٧	٧	٠	٧	٠	لا يوجد تمييز
٢	٧	٢	٥	٧	٠.٧١	جيد
٣	٧	٦	١	٧	٠.١٤	ضعيف

٤	٧	٣	٤	٧	٠.٥٧	جيد
٥	٤	٣	١	٧	٠.١٤	ضعيف
٦	٧	٥	٢	٧	٠.٢٩	متوسط
٧	٥	٠	٥	٧	٠.٧١	جيد
٨	٧	٦	١	٧	٠.١٤	ضعيف
٩	٧	٠	٧	٧	١	جيد
١٠	٣	٢	١	٧	٠.١٤	ضعيف
١١	٦	٢	٤	٧	٠.٥٧	جيد
١٢	٦	١	٥	٧	٠.٧١	جيد
١٣	٧	١	٦	٧	٠.٨٦	جيد
١٤	٧	٢	٥	٧	٠.٧١	جيد
١٥	٧	٠	٧	٧	١	جيد
١٦	٥	٢	٣	٧	٠.٤٣	متوسط
١٧	٦	٢	٤	٧	٠.٥٧	جيد
١٨	٧	٧	٠	٧	٠	لا يوجد تمييز
١٩	٣	١	٢	٧	٠.٢٩	متوسط
٢٠	٧	١	٦	٧	٠.٨٦	جيد
٢١	٦	١	٥	٧	٠.٧١	جيد
٢٢	٦	٢	٤	٧	٠.٥٧	جيد
٢٣	٣	١	٢	٧	٠.٢٩	متوسط
٢٤	٤	٢	٢	٧	٠.٢٩	متوسط

٢٥	٥	٠	٥	٧	٠.٧١	جيد
٢٦	٧	٣	٤	٧	٠.٥٧	جيد
٢٧	٦	٢	٤	٧	٠.٥٧	جيد
٢٨	٦	٠	٦	٧	٠.٨٦	جيد
٢٩	٠	١	-١	٧	-٠.١٤	سلبي
٣٠	٧	١	٦	٧	٠.٨٦	جيد
٣١	٧	٣	٤	٧	٠.٥٧	جيد
٣٢	٧	١	٦	٧	٠.٨٦	جيد
٣٣	٥	٠	٥	٧	٠.٧١	جيد
٣٤	٧	٤	٣	٧	٠.٤٣	متوسط
٣٥	٦	١	٥	٧	٠.٧١	جيد
٣٦	٧	٢	٥	٧	٠.٧١	جيد
٣٧	٦	٣	٣	٧	٠.٤٣	متوسط
٣٨	٧	١	٦	٧	٠.٨٦	جيد
٣٩	٢	١	١	٧	٠.١٤	ضعيف
٤٠	٧	٢	٥	٧	٠.٧١	جيد

من البيانات أعلاه، علم أن هناك أسئلة في اختبار اللغة العربية في مدرسة الإخلاص الثانوية الإسلامية نانغولان في السنة الدراسية ٢٠٢٢-٢٠٢٣ تستطيع التمييز بين مجموعتي الطلاب في فئتي الصف العلوي والصف السفلي. وهناك أيضًا أسئلة لا تستطيع التمييز بين الطلاب في الفئتين. من بين الأسئلة التي يمكنها التمييز بين مجموعتي الطلاب العلوي والسفلي، هناك أسئلة تصنف على أنها لها قدرة تمييز جيدة ومتوسطة وضعيفة. وأما

بالنسبة للأسئلة التي لا يمكنها التمييز بين المجموعتين، فتنقسم إلى أسئلة بقدرة تمييز سلبية وأخرى لا توجد فيها تمييز. تفصيل درجة التمييز لكل سؤال على حدة هو كما يلي:

أ. يوجد ٢٥ سؤالًا لها قدرة تمييز جيدة، وهي الأسئلة رقم ٢ و٤ و٧ و٩ و١١ و١٢ و١٣ و١٤ و١٥ و١٧ و٢٠ و٢١ و٢٢ و٢٥ و٢٦ و٢٧ و٢٨ و٣٠ و٣١ و٣٢ و٣٣ و٣٥ و٣٦ و٣٨ و٤٠.

ب. يوجد سبعة أسئلة لها قدرة تمييز متوسطة، وهي الأسئلة رقم ٦ و١٦ و١٩ و٢٣ و٢٤ و٣٤ و٣٧.

ج. يوجد خمسة أسئلة لها قدرة تمييز ضعيفة، وهي الأسئلة رقم ٣ و٥ و٨ و١٠ و٣٩.

د. يوجد سؤالان ليس لديهما قدرة تمييز بسبب عدم وجود تمييز، وهما الأسئلة رقم ١ و١٨.

يوجد سؤال واحد له قدرة تمييز سلبية، وهو السؤال رقم ٢٩.

الخلاصة

بعد إجراء تحليل لصدق الاختبار، الثبات أو الموثوقية، درجة التمييز، ودرجة الصعوبة لأسئلة اللغة العربية في اختبار اللغة العربية في مدرسة الإخلاص الثانوية الإسلامية نانغولان في السنة الدراسية ٢٠٢٢-٢٠٢٣، أظهرت النتائج في ما يالي، الأول فيما يتعلق بصدق الاختبار، تم اعتبار ٢٦ سؤالًا أو ما يعادل ٦٥٪ من الأسئلة صديقة وتم اعتبار ١٤ سؤالًا أو ٣٥٪ من الأسئلة غير صديقة. والثاني فيما يتعلق بالثبات، فإن أسئلة اللغة العربية تدخل في فئة الأسئلة ذات الثبات العالي جدًا بسبب وجود قيمة Cronbach's Alpha بقيمة ٠,٩١٥. والثالث فيما يتعلق بدرجة الصعوبة، توجد سبعة أسئلة أو ما يعادل ١٧,٥٪ من الأسئلة تدخل في فئة السهل، ٢٨ سؤالًا أو ٧٠٪ من الأسئلة تدخل في فئة المتوسطة، وخمسة أسئلة أو ١٢,٥٪ تدخل في فئة الصعب. والرابع فيما يتعلق بقدرة التمييز، هناك ٢٥ سؤالًا أو ٦٢,٥٪ من الأسئلة تمتلك قدرة تمييز جيدة، وهناك سبعة

أسئلة أو ١٧,٥٪ تمتلك قدرة تمييز متوسطة، وخمسة أسئلة أو ١٢,٥٪ تمتلك قدرة تمييز ضعيفة، وهناك سؤالان أو ٥٪ ليس لديهما قدرة تمييز بسبب عدم وجود تمييز، وهناك سؤال واحد أو ٢,٥٪ لديه قدرة تمييز سلبية.

المراجع

Agung, N. (2022). Kurikulum Pembelajaran Bahasa Arab Perguruan Tinggi: Analisis Komparatif Kurikulum Pembelajaran Bahasa Arab. *AL Waraqah: Jurnal Pendidikan Bahasa Arab*, *3*(2), 1–18. https://doi.org/http://dx.doi.org/10.30863/awrq.v3i2.2945

Alamsyah, T., Arifin, M., Romdoni, M. F., & F.A., M. H. (2025). Competency Analysis of Arabic Language Teachers at Madrasah Aliyah Level in Cileunyi District. *Kalamuna: Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban*, *6*(1), 1–16. https://doi.org/10.52593/klm.06.1.01

Ambiya, I. Z., Mulyawan, S., & Saefuloh, H. (2022). Analisis Soal Ujian Mata Pelajaran Bahasa Arab Kelas 12 di Madrasah Aliyyah. *EL-IBTIKAR: Jurnal Pendidikan Bahasa Arab*, *11*(1), 70–87. https://doi.org/http://dx.doi.org/10.24235/ibtikar.v11i1.10521

Arikunto, S. (2009). *Dasar-Dasar Evaluasi Pendidikan*. Bumi Aksara.

Arikunto, S. (2013). *Prosedur Penelitian: Suatu Pendekatan Praktik*. Rineka Cipta.

Asrori, I., Thohir, M., & Ainin, M. (2017). Evaluasi Pembelajaran Bahasa Arab. *Malang: Misykat*.

Darma, B. (2021). *Statistik Penelitian Menggunakan SPSS (Uji Validitas, Uji Reabilitas,Regresi Linier Sederhana, Regresi Linier Berganda, Uji t, Uji F, R2)*. Guepedia.

Djiwandono, M. S. (1996). *Tes Bahasa dalam Pengajaran*. ITB Bandung.

Fahmi, B., Rizqi, S., & Harmeilinda, N. E. (2022). Analisis Butir Soal Bahasa Arab MAS Pondok Pesantren Assalam Kampar Riau. *Ta'lim Al-'Arabiyyah: Jurnal Pendidikan Bahasa Arab & Kebahasaaraban*, *6*(1), 95–105. https://doi.org/https://doi.org/10.15575/jpba.v6i1.16193

Fathiyah, S. F. Al. (2019). Analisis Butir Soal Pelajaran Bahasa Arab Di Ma Roudlotul Ulum Pagak Malang. *Tarbiyatuna: Jurnal Pendidikan Ilmiah*, *4*(1), 76–100. https://ejournal.stitibnusina.ac.id/index.php/tarbiyatuna/article/view/357

Fikriyah, N. (2021). Analisis Butir Soal Ulangan Tengah Semester Mata Pelajaran Bahasa Arab Kelas VII Semester Ganjil SMP Muhammadiyah 1 Yogyakarta Tahun Ajaran

2019/2020. *Maharaat: Jurnal Pendidikan Bahasa Arab*, *3*(2), 128–140. https://doi.org/10.18196/mht.v3i2.10501

Gunawan, C. (2020). *Mahir Menguasai SPSS Panduan Praktis Mengolah Data Penelitian New Edition Buku Untuk Orang Yang (Merasa) Tidak Bisa Dan Tidak Suka Statistika*. Deepublish.

Gusmizain, A. (2022). Karakteristik Butir Soal Tes Mata Kuliah Matriks & Ruang Vektor Mahasiswa Matematika. *Jurnal Evaluasi Pendidikan*, *13*(2), 127–130. https://doi.org/https://doi.org/10.21009/jep.v13i2.28328

Hakiem, M. R., Abdurrohman, M. U., & Faturrahman, M. I. (2025). The Feasibility of TOAFL Instrument as a Measuring Tool for Reading Skills of Students Faculty of Education. *Kalamuna: Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban*, *6*(1), 173–185. https://doi.org/https://doi.org/10.52593/klm.06.1.11

Hanim, N. (2018). Penerapan Model Pembelajaran Aktif Tipe Index Card Match untuk Meningkatkan Prestasi Belajar Siswa Pada Materi Sistem Gerak. *BIOTIK: Jurnal Ilmiah Biologi Teknologi Dan Kependidikan*, *5*(2), 141–148. https://doi.org/http://dx.doi.org/10.22373/biotik.v5i2.3023

Jannah, R. (2022). Analisis Butir Soal Ujian Sekolah Berstandar Nasional (USBN) Mata Pelajaran Bahasa Arab Madrasah Tsanawiyah Kota Parepare. *INTERNATIONAL CONFERENCE ON ISLAM, LAW, AND SOCIETY (INCOILS) 2021*, *1*(1), 165–172. http://www.incoilsfdpdiktis2021.iaipd-nganjuk.ac.id/index.php/incoils/article/view/37

Makhisoh, I., & Suroyya, E. N. (2022). Analisa Butir Soal Ujian Madrasah Bahasa Arab Menurut Taksonomi Bloom. *Muhadasah: Jurnal Pendidikan Bahasa Arab*, *4*(1), 40–52. https://ejournal.iaiskjmalang.ac.id/index.php/muhad/article/view/501

Marudut, J., & Akbar, R. (2022). Analisis Butir Soal Ujian Akhir Sekolah Mata Pelajaran Bahasa Indonesia Kelas VIII SMP Negeri 3 Kutacane Tahun Pelajaran 2020/2021. *Tuwah Pande: Jurnal Ilmu Pendidikan Dan Pengajaran*, *1*(2), 254–269. http://tuwah-pande.com/index.php/tuwahpande/article/view/25

Maulana, D., & Sanusi, A. (2020). Analisis Butir Soal Bahasa Arab Ujian Akhir Madrasah Bersama Daerah (Uambd) Madrasah Ibtidaiyah Tahun 2017-2018. *Ta'lim Al-'Arabiyyah: Jurnal Pendidikan Bahasa Arab & Kebahasaaraban*, *4*(1), 12–24. https://doi.org/10.15575/jpba.v4i1.8054

Nurbayan, Y., Sauri, S., & Sanusi, A. (2021). Developing an International Standardized Arabic Language Education Curriculum: Introducing a Conception-Focused Design and Outcome. *Al-Ta'rib : Jurnal Ilmiah Program Studi Pendidikan Bahasa Arab IAIN Palangka Raya*, *9*(2), 155–172. https://doi.org/10.23971/altarib.v9i2.3264

Qomariyah, L. (2022). Analisis Tingkat Kesukaran dan Daya Pembeda Butir Soal TOAFL Universitas Hasyim Asy'ari Tebuireng Jombang. *Lisanan Arabiya: Jurnal Pendidikan*

Bahasa Arab, *6*(1), 1–18. https://doi.org/https://doi.org/10.32699/liar.v6i1.2549

Ramadhan, R., & Firdaus, F. N. (2022). Analisis Butir Soal Ujian Tengah Semester Bahasa Arab Kelas XII Di SMA Al-Izzah IIBS Malang. *Tsaqofiya: Jurnal Pendidikan Bahasa Dan Sastra Arab*, *4*(1), 126–135. https://doi.org/https://doi.org/10.21154/tsaqofiya.v4i1.49R

Sugiyono. (2012). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Alfabeta.

Sumardi. (2020). *Teknik Pengukuran Dan Penilaian Hasil Belajar*. Deepublish.

Syafril. (2019). *Statistik Pendidikan* (Edisi I). Kencana.

Zahro', K., & Alfian, M. I. (2026). Digital Competence of Arabic Language Teachers and Its Role in Developing Technology-Based Arabic Education. *Kalamuna: Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban*, *7*(1), 25–35. https://doi.org/10.52593/klm.07.1.03

Copyright holder :
©Mohammad Zaelani Musonif, Habib Maulana Maslahul Adi, Ahmad Nawirul Huda, Muhammad Yahya Abdullah. (2026)

First publication right:
Kalamuna: Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban

This is an open access article under the CC BY-SA license