



Bunūd As'ilah Al-Lughah Al-'Arabiyyah Li-Ikhtibār Al-Taḡyīm Al-Nihā'ī Li-Al-Faṣl Al-Dirāsī Binā'an 'alā Al-Manhaj Al-Mustaḡill Min Mudarrisī Al-Lughah Al-'Arabiyyah Bi-Al-Madrasah Al-Thānawīyyah Al-Islāmiyyah Al-Ḥukūmiyyah Pekanbaru

تحليل بنود أسئلة اللغة العربية لاختبار التقييم النهائي للفصل الدراسي بناء على المنهج
المستقل من مدرسي اللغة العربية بالمدرسة الثانوية الإسلامية الحكومية بكنبارو

Aathifah Al Farros*, Agustiar, Dasman Yahya Ma'ali
Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia
aathifahalfarros19@gmail.com

ملخص البحث

تهدف هذا البحث إلى تحليل درجة الصعوبة ومستوى التمييز للأسئلة اللغة العربية لاختبار التقييم النهائي للفصل الدراسي بناء على المنهج المستقل من مدرسي اللغة العربية بالمدرسة الثانوية الإسلامية الحكومية بكنبارو. تصميم البحث المستخدم هو وصفي كمي. شمل مجتمع الدراسة جميع أسئلة اللغة العربية، والعينة هي أسئلة اختبار التقييم النهائي للفصل الدراسي للعام الدراسي ٢٠٢٣ للصف العاشر. أظهرت نتائج التحليل لدرجة صعوبة الأسئلة بالمدرسة الثانوية الإسلامية الحكومية الواحدة بكنبارو كان غير متوازن، ٤٤٪ من الأسئلة على أنها سهلة، و ٥٠٪ متوسطة، و ٦٪ تم صعبة. وفي المدرسة الثانوية الإسلامية الحكومية الثانية بكنبارو، ٦٠٪ من الأسئلة على أنها سهلة، و ٩٠٪ متوسطة، و ٣٠٪ صعبة. وبناءً على مستوى التمييز للأسئلة، فإن ٤٤ سؤالاً بالمدرسة الثانوية الإسلامية الحكومية الواحدة بكنبارو لها مستوى التمييز تزيد عن ٠,٣٠، مما يشير إلى أن الأسئلة فعالة في التمييز بين الطلاب ذوي التحصيل العالي والمنخفض. ومع ذلك، فإن ٦ أسئلة لها مستوى التمييز أقل من ٠,٢٠، لذلك يوصى بمراجعتها أو حذفها. وفي المدرسة الثانوية الإسلامية الحكومية الثانية بكنبارو، فإن ٨٣,٣٣٪ من الأسئلة لها مستوى التمييز تزيد عن ٠,٣٠، مما يشير إلى فعالية الأسئلة في قياس الاختلافات في قدرات الطلاب. ومع ذلك، فإن ٥ أسئلة بقوة تمييز أقل من ٠,٣٠ تتطلب المراجعة أو الاستبدال.

الكلمات المفتاحية: تحليل بنود، درجة الصعوبة، مستوى التمييز، اللغة العربية

Abstrak

Penelitian ini bertujuan untuk menganalisis tingkat kesukaran dan tingkat perbedaan soal bahasa arab untuk tes penilaian akhir semester berdasarkan kurikulum mandiri guru bahasa arab di Madrasah Ibtidaiyah Negeri Kinbaru. Desain penelitian yang digunakan adalah deskriptif kuantitatif. Populasi penelitian meliputi seluruh soal bahasa Arab, dan sampelnya adalah soal ulangan penilaian akhir semester tahun pelajaran 2023 untuk kelas X. Hasil analisis tingkat kesukaran soal di MAN 1 Pekanbaru menunjukkan tidak seimbang, yaitu 44% soal dinilai mudah, 50% sedang, dan 6% sulit. Di MAN 2 Pekanbaru 6,70% soal tergolong mudah, 90% sedang, dan 3,30% sulit. Berdasarkan tingkat diskriminasi soal, 44 soal di MAN 1 Pekanbaru memiliki tingkat diskriminasi lebih dari 0,30 yang menunjukkan bahwa soal tersebut efektif dalam membedakan siswa yang berprestasi tinggi dan yang berprestasi rendah. Namun terdapat 6 soal yang tingkat diskriminasinya kurang dari 0,20 sehingga disarankan untuk direvisi atau dihapus. Di Di MAN 2 Pekanbaru 83,33% soal memiliki tingkat diskriminasi lebih dari 0,30 yang menunjukkan efektifitas soal dalam mengukur perbedaan kemampuan siswa. Namun, 5 soal dengan daya diskriminasi kurang dari 0,30 memerlukan revisi atau penggantian.

Kata kunci : Analisis Butir Soal, Tingkat Kesukaran, Daya Pembeda, Bahasa Arab

Abstract

This study aims to analyze the level of difficulty and level of discrimination of Arabic language questions for the final semester assessment test based on the independent curriculum of Arabic language teachers at Madrasah Ibtidaiyah Negeri Kinbaru. The research design used is descriptive quantitative. The study population includes all Arabic language questions, and the sample is the final semester assessment test questions for the 2023 academic year for class X. The results of the analysis of the level of difficulty of questions at MAN 1 Pekanbaru showed an imbalance, namely 44% of questions were considered easy, 50% moderate, and 6% difficult. At MAN 2 Pekanbaru 6.70% of questions were classified as easy, 90% moderate, and 3.30% difficult. Based on the level of discrimination of questions, 44 questions at MAN 1 Pekanbaru had a discrimination level of more than 0.30 which indicated that the questions were effective in distinguishing high-achieving and low-achieving students. However, there were 6 questions whose discrimination level was less than 0.20 so it was recommended to be revised or deleted. In MAN 2 Pekanbaru, 83.33% of questions have a discrimination level of more than 0.30, which indicates the effectiveness of the questions in measuring differences in student abilities. However, 5 questions with a discrimination power of less than 0.30 require revision or replacement.

Keywords: Item Analysis, Level of Difficulty, Discriminating Power, Arabic Language

المقدمة

يلعب التعليم دورًا مهمًا جدًا في إنشاء موارد بشرية متفوقة وقادرة على المنافسة. من خلال التعليم، يتم توجيه الأفراد إلى تطوير إمكاناتهم الكاملة إلى الحد الأقصى، سواء في قدرات التفكير (المعرفية)، أو المواقف والقيم (الوجدانية)، أو المهارات العملية (الحركية النفسية). ولضمان تحقيق هذا الهدف، هناك حاجة إلى نظام منظم لرصد

وتقييم فعالية عملية التعلم المستمرة. وفي هذه الحالة، يعد تقييم نتائج التعلم أحد الأدوات الرئيسية المستخدمة لتقييم مدى تحقيق الطلاب للكفاءات المستهدفة، فضلاً عن توفير مدخلات لتحسين جودة التعلم. (Helean, Kandowangko, and Goni 2021)

من أجل إجراء تقييم نتائج التعلم بشكل فعال، من المهم للغاية التأكد من أن أدوات التقييم، مثل أسئلة الامتحانات أو الاختبارات، تلبي معايير الجودة الواضحة. لا يمكن للأدوات التي يتم إعدادها بعناية ودقة قياس إنجازات كفاءة الطلاب بشكل موضوعي فحسب، بل تغطي أيضاً الجوانب المختلفة التي تريد قياسها، مثل القدرات المعرفية والعاطفية والحركية النفسية. (Rizky Ananda Setiyawan and Palupi Sri Wijayanti 2020) ستوفر الأسئلة عالية الجودة معلومات أكثر دقة فيما يتعلق بفهم الطلاب للمادة التي تم تدريسها. وهذا يسمح لعملية التقييم بتقديم صورة أكثر شمولاً لقدرات الطلاب، وبالتالي يساعد المعلمين على اتخاذ قرارات أكثر دقة وتركيزاً في تصميم استراتيجيات التعلم المستقبلية. في السياق التعليمي، يكون لتقييم نتائج التعلم وظيفة أوسع من مجرد توفير الأرقام أو القيمة طلاب. يلعب هذا التقييم دوراً مهماً في ضمان قدرة الطلاب على تحقيق أهداف التعلم التي تم تصميمها بشكل فعال. من خلال عملية التقييم، يمكن للمدرسين الحصول على فكرة عن مدى قدرة الطلاب على فهم المواد المقدمة وتقييم ما إذا كانت طرق التدريس المستخدمة تعمل بشكل جيد. وبغض النظر عن ذلك، فإن المعلومات التي يتم الحصول عليها من نتائج التقييم تصبح أساساً قيماً للمعلمين لعكس وتحسين جودة عملية التعلم في المستقبل، بحيث تصبح تجارب التعلم للطلاب أكثر مثالية وذات معنى

في السياق التعليمي، يكون لتقييم نتائج التعلم وظيفة أوسع من مجرد توفير الأرقام أو الدرجات للطلاب. يلعب هذا التقييم دوراً مهماً في ضمان إمكانية تحقيق الطلاب لأهداف التعلم التي تم تصميمها بفعالية. من خلال عملية التقييم، يمكن للمعلمين الحصول على فكرة عن مدى قدرة الطلاب على فهم المواد المقدمة وتقييم ما إذا كانت طرق التدريس المستخدمة تعمل بشكل جيد. بالإضافة إلى ذلك، تصبح المعلومات التي تم الحصول عليها من نتائج التقييم أساساً قيماً للمعلمين للتفكير في جودة عملية التعلم وتحسينها في المستقبل، بحيث تصبح تجارب التعلم للطلاب أكثر مثالية وذات معنى. (Idrus 2019)

أدوات التقييم، وخاصة التي تكون على شكل أسئلة، لها دور مهم جداً في دعم عملية تقييم التعلم. إن الأسئلة التي يتم إعدادها بشكل صحيح لا يمكنها فقط قياس قدرات الطلاب بدقة، ولكنها أيضاً تحقق مبادئ الصحة والموثوقية والموضوعية. ومع ذلك، من الناحية العملية، غالباً ما يمثل تصميم أدوات تقييم الجودة تحدياً في حد ذاته. إحدى العقبات التي تظهر غالباً هي التأكد من توافق الأسئلة مع معايير التقييم التربوي. إذا حدثت أخطاء في الإعداد، مثل مستوى غير متناسب من صعوبة الأسئلة أو عدم وجود قوة تمييز كافية، فقد يكون لذلك تأثير سلبي على دقة نتائج التقييم، مما يجعلها أقل قدرة على تمثيل الكفاءة الشاملة للطلاب. (Nurhairiyah 2021)

ولذلك، يعد تحليل العناصر خطوة مهمة في ضمان جودة أداة التقييم. ومن خلال إجراء هذا التحليل، يمكن للمعلمين تقييم صحة وموثوقية الأسئلة، والتأكد من أن الأسئلة قادرة على قياس قدرات الطلاب بدقة. يوفر تحليل العناصر أيضًا معلومات حول مستوى الصعوبة، وقوة التمييز، وفعالية عوامل التشبث، وكلها تساهم في جودة التقييم. (Siregar, Remiswal, and Khadijah 2024)

إن التقييم الجوده دور يتجاوز مجرد عكس مدى إنجاز الطلاب في عملية التعلم. التقييم الجيد هو أيضًا وسيلة لتحسين وتعزيز الجودة الشاملة للتعليم. من خلال الفهم العميق لأهمية تقييم نتائج التعلم، يمكن للمعلمين اتخاذ قرارات أكثر ملاءمة في تصميم وتنفيذ عملية التعلم. وبهذه الطريقة، يمكن للمعلمين إنشاء تجارب تعليمية أكثر فعالية وملاءمة ولها تأثير إيجابي على تطور الطلاب، مع ضمان تحقيق أهداف التعلم على النحو الأمثل. (Budiyanto dan Titik Haryat 2023)

تقييم نتائج التعلم هو عملية منهجية تهدف إلى قياس وتقييم مستوى تحصيل كفاءة الطلاب بعد متابعة عملية التعلم. (Prastiwi et al. 2023) ومن خلال هذا التقييم، يمكن للمعلمين الحصول على بيانات دقيقة فيما يتعلق بفهم الطلاب ومهاراتهم ومواقفهم تجاه المواد التي تم تدريسها. وبالتالي، يصبح تقييم نتائج التعلم أداة مهمة في دعم تحسين الجودة الشاملة للتعليم، وضمان تحقيق أهداف التعلم، ومساعدة الطلاب في تحقيق أقصى إمكاناتهم الأكاديمية.

يعد تقييم نتائج التعلم عنصرًا أساسيًا في العملية التعليمية التي تهدف إلى قياس مدى تحقيق الطلاب للكفايات المتوقعة بعد المشاركة في برنامج التعلم. (Magdalena et al. 2023) ومع ذلك، على الرغم من أن تقييم نتائج التعلم يوفر صورة عامة عن تحصيل الطلاب، إلا أنه غالبًا ما تكون هناك فجوة في الفهم فيما يتعلق بجودة أداة التقييم نفسها، أي عناصر الاختبار. (Maulana 2023) يعد تحليل العناصر، والذي يتضمن مستوى التمييز درجة الصعوبة، طريقة مهمة لتقييم فعالية وموثوقية أسئلة الاختبار. تتمثل الفجوة الحالية في عدم تطبيق تحليل العناصر في العديد من السياقات التعليمية، بما في المدرسة الثانوية الإسلامية الحكومية الواحدة و الثانية بكنبارو، مما يؤدي إلى عدم دقة محتملة في تقييم كفاءة الطلاب. وبدون تحليل متعمق لعناصر الاختبار، قد لا تعكس تقييمات نتائج التعلم بشكل كامل قدرات الطلاب وفهمهم، الأمر الذي يمكن أن يؤثر في النهاية على عملية التعلم وصنع القرار التعليمي. ولذلك، من المهم دمج تحليل العناصر كجزء لا يتجزأ من عملية التقييم لتحسين جودة وصلاحية تقييمات نتائج التعلم.

منذ تطبيق منهج مستقل في عام ٢٠٢٢ في إندونيسيا، كان هناك تركيز قوي على تطوير كفاءات الطلاب ككل، بما في ذلك جوانب المعرفة والمهارات والشخصية. (Tuerah and Tuerah 2023) ومع ذلك، هناك

فجوة واضحة بين الهدف من المنهج المستقل وجودة الأسئلة المستخدمة في عملية التقييم. العديد من الأسئلة التي تم طرحها لا تعكس بشكل كامل المبادئ الأساسية للمنهج المستقل، مثل تشجيع التفكير النقدي والإبداع وقدرات حل المشكلات. ولا يزال عدد كبير من الأسئلة يركز أكثر على الحفظ ولا يقيس قدرات الطلاب التحليلية والتطبيقية. ونتيجة لذلك، فإن التقييم الناتج يكون أقل فعالية في وصف الكفاءات الفعلية المتوقعة من منهج ميرديكا. ولهذا السبب، هناك حاجة إلى بذل المزيد من الجهود الجادة في إعداد عناصر الاختبار التي تتماشى مع جوهر المنهج المستقل، بما في ذلك التحليل الدقيق لعناصر الاختبار للتأكد من أنها قادرة على قياس الكفاءات الأوسع والأعمق وفقاً للاتجاه المنشود بواسطة المنهج.

لقد اتخذت المدرسة الثانوية الإسلامية الحكومية الواحدة و الثانية بكنبارو خطوة إلى الأمام من خلال تجميع أسئلة اللغة العربية التي تتوافق مع مبادئ المنهج المستقل. ومنذ عام ٢٠٢٣، تم استخدام هذه الأسئلة في امتحان التقييم نهاية الفصل الدراسي. تم اتخاذ هذه الخطوة للتأكد من أن التقييمات التي يتم إجراؤها يمكن أن تعكس قدرات الطلاب بشكل أكثر شمولاً، وفقاً لأهداف منهج ميرديكا الذي يؤكد على تنمية مهارات التفكير النقدي والإبداع وقدرات حل المشكلات. من خلال اعتماد الأسئلة التي تركز بشكل أكبر على الجوانب المعرفية العليا، تسعى المدرسة الثانوية الإسلامية الحكومية الواحدة و الثانية بكنبارو إلى تحسين جودة التعليم وتقديم تعليقات أكثر فائدة للطلاب والمعلمين. يُظهر هذا التنفيذ التزام المدرسة بدعم سياسات التعليم الوطنية وتطوير إمكانات الطلاب إلى الحد الأقصى. على الرغم من نجاح المدرسة الثانوية الإسلامية الحكومية الواحدة و الثانية بكنبارو في تجميع أسئلة اللغة العربية المتوافقة مع المنهج المستقل واستخدامها في اختبارات التقييم نهاية الفصل الدراسي منذ عام ٢٠٢٣، إلا أنه لا تزال هناك ثغرات تحتاج إلى معالجة، خاصة فيما يتعلق بتحليل الأسئلة التي تشمل التمييز بين مستوى التمييز ودرجة الصعوبة. على الرغم من أن الأسئلة مصممة لتعكس الكفاءات المطلوبة في منهج ميرديكا، إلا أنه لم يتم تنفيذ تحليل متعمق مستوى التمييز ومستوى الصعوبة بشكل كامل. وهذا أمر مهم لأن هناك حاجة إلى مستوى التمييز جيدة لضمان أن الأسئلة يمكن أن تميز بين الطلاب الذين يفهمون المادة حقاً وأولئك الذين لا يفهمونها، في حين أن هناك حاجة إلى مستوى مناسب من الصعوبة لضمان توزيع متوازن للأسئلة بين السهل والمتوسط والصعب. وبدون هذا التحليل الشامل، هناك خطر من أن الأسئلة قد لا تكون فعالة بشكل كامل في قياس قدرات الطلاب بدقة ونزاهة، وبالتالي فإن نتائج التقييم قد لا تعكس بشكل كامل الكفاءات الحقيقية للطلاب. لذلك، من الضروري تحليل بنود الأسئلة بانتظام لتحسين جودة التقييمات وضمان توافقها مع أهداف منهج مستقل.

و لاختبار التقييم النهائي للفصل الدراسي للفصل الدراسي النهائي ٢٠٢٣ في المدرستين، لم تقتصر الأسئلة المطروحة على الطلاب على الأسئلة الموضوعية. وفقاً لتنفيذ المنهج المستقل، فإن مجموعة الأسئلة المقدمة تتنوع بشكل متزايد. تقدم المدرسة أسئلة الامتحان على شكل أسئلة اختيار من متعدد، أسئلة الصواب والخطأ وأسئلة

المطابقة. يتم تضمين أسئلة الاختيار من متعدد في فئة الأسئلة الموضوعية، حيث يتعين على الطلاب اختيار إجابة واحدة من بين عدة خيارات متاحة. من ناحية أخرى، تتضمن أسئلة الصواب والخطأ وأسئلة المطابقة أسئلة وصفية مع نظام تسجيل مختلف لكل عنصر. لذلك، سيقوم الباحثون بتحليل هذين النوعين من الأسئلة لتحديد درجة الصعوبة ومستوى التمييز للأسئلة بناءً على المنهج المستقل المطبق.

تحليل مستوى التمييز لعناصر الأسئلة هو عملية تقييم تهدف إلى قياس مدى قدرة السؤال على التمييز بين الطلاب الذين لديهم فهم عالٍ للمادة والطلاب الذين لديهم فهم منخفض. (Pradita, Megawanti, and Indraprasta PGRI 2023) تشير مستوى التمييز العالية إلى أن الأسئلة فعالة في تحديد الفروق في القدرة بين الطلاب، حيث يميل الطلاب الذين يتقنون المادة إلى الإجابة بشكل صحيح، بينما يميل الطلاب الذين لا يتقنون المادة إلى الإجابة بشكل غير صحيح. (Elviana 2020) ويتم هذا التحليل عادة عن طريق حساب مؤشر مستوى التمييز، وهو عبارة عن قيمة عددية توضح قدرة الأسئلة على التمييز بين مجموعتي الطلاب. الأسئلة ذات مستوى التمييز الجيدة لها مؤشر مستوى التمييز إيجابي مرتفع، في حين أن الأسئلة ذات مستوى التمييز الضعيفة قد يكون لها مؤشر منخفض أو حتى سلبي. من خلال تحليل مستوى التمييز، يمكن للمعلمين وصانعي الأسئلة تحسين أسئلتهم وتحسينها، مما يضمن أن كل سؤال مستخدم في التقييم قادر على توفير معلومات دقيقة حول قدرات الطلاب، وزيادة صحة وموثوقية أداة التقييم الشاملة. (Muzayyanah 2020)

وفقا لعارفين، فإن مستوى صعوبة السؤال هو مقياس يوضح مدى صعوبة السؤال. تحليل مستوى صعوبة الأسئلة يعني تقييم الأسئلة لتحديد أي منها سهل أو متوسط أو صعب. الهدف هو ضمان التوازن في مستوى صعوبة الأسئلة. يجب أن يكون السؤال الجيد بمستوى متوسط من الصعوبة، أي أنه ليس سهلاً للغاية بحيث لا يشكل تحدياً لقدرات الطلاب، ولكنه أيضاً ليس صعباً للغاية بحيث لا يتمكن معظم الطلاب من الإجابة عليه. بمعنى آخر، يجب أن يقع السؤال الجيد في منتصف نطاق الصعوبة لقياس قدرة الطالب بشكل أكثر دقة وعدالة. (Rismaulhijjah and Kuswanti 2022) تحليل مستوى صعوبة الأسئلة هو عملية تقييم تستخدم لتقييم مدى صعوبة أو سهولة السؤال بالنسبة للطلاب. يتم التعبير عن هذا المستوى من الصعوبة عادة على شكل نسبة مئوية أو رقم معين يصف نسبة الطلاب الذين أجابوا على السؤال بشكل صحيح. (Ratni 2023) درجة الصعوبة الصحيح هو المفتاح لإعداد أسئلة فعالة، حيث يجب أن تحتوي الأسئلة على مجموعة متوازنة من مستويات الصعوبة، بما في ذلك الأسئلة السهلة والمتوسطة والصعبة. يمكن للأسئلة السهلة للغاية أن تنتج درجات عالية لجميع الطلاب، في حين أن الأسئلة الصعبة للغاية يمكن أن تنتج درجات منخفضة لجميع الطلاب، لذلك فهي لا توفر معلومات ذات معنى حول قدرات الطلاب. (Fatimah and Alfath 2019) ولذلك فمن

خلال تحليل درجة الصعوبة يستطيع المعلمون وواضعو الأسئلة تقييم درجة صعوبة أسئلتهم وتحسينه، مما يضمن أن الأسئلة المستخدمة في التقييم يمكن أن تقدم صورة دقيقة عن قدرات الطلاب، فضلاً عن زيادة الصدق والثبات من أداة التقييم الشامل.

من خلال إجراء هذا التحليل، من المأمول أن يتمكن أيضاً من تحفيز معلمي اللغة العربية في المدرسة الثانوية الإسلامية الحكومية الواحدة و الثانية بكنبارو لمواصلة تحسين جودة التعلم والتقييم. هذا البحث ليس مفيداً للمعلمين والطلاب في المدرسة فحسب، بل أيضاً لصانعي السياسات التعليمية في المستويات العليا. نتائج هذا البحث يمكن أن تكون الأساس لاتخاذ قرارات أفضل في تطوير المناهج والتقييم التربوي في المستقبل.

منهج البحث

يصف هذا القسم طريقة البحث باختصار. اكتب موقع البحث بالكامل، وعدد المستجيبين، وكيفية معالجة نتائج الملاحظات أو المقابلات أو الاستبيانات، وكيفية قياس معايير الأداء. الطريقة العامة لا تحتاج إلى كتابتها بالتفصيل، لكنها كافية للإشارة إلى الكتاب المرجعي. يجب كتابة الإجراء التجريبي من حيث الجملة تكون سياقها على شكل خبر وليست على أسلوب أمر. وتكون حدود مقترح خطة البحث حوالي ٣٠٠-٦٠٠ كلمة

تستخدم "خطة البحث" Traditional Arabic بحجم ١٦ بخط الأحرف الكبيرة العريضة والصغيرة مع تباعد ١٨ نقطة في الأعلى والأسفل ٠ نقطة. تبدأ مقدمة البحث باستخدام Traditional Arabic بحجم ١٦ بخط هوامش عادية ومستقيمة يسار ويمين بمسافة في الأعلى ٠ نقطة والأسفل ٤ نقطة.

النتائج والمناقشة

الطلاب المشاركون في الامتحان هم طلاب الصف العاشر، وكان عددهم ٥٨٦ طالبا، يعني ٣١٨ طالبا من المدرسة الثانوية الإسلامية الحكومية الواحدة بكنبارو و ٢٦٨ طالبا من المدرسة الثانوية الإسلامية الحكومية الثانية بكنبارو. ونظراً للكمية الكبيرة من البيانات المتاحة، استخدم الباحث في هذه الدراسة صيغة سلوفين Slovin للحد من كمية البيانات المطلوب تحليلها. مع وصف الحساب على النحو التالي:

$$n = \frac{318}{1 + 318(0,05^2)} = \frac{318}{1 + 318(0,025^2)} = \frac{318}{1.795} = 177.2$$

$$n = \frac{268}{1 + 268(0,05^2)} = \frac{268}{1 + 268(0,025^2)} = \frac{318}{1.67} = 160.5$$

وبذلك بلغ عدد الطلاب الذين تم تحليل نتائج امتحاناتهم ١٧٧ طالباً من المدرسة الثانوية الإسلامية الحكومية الواحدة بكنبارو و ١٦٠ طالباً من المدرسة الثانوية الإسلامية الحكومية الثانية بكنبارو.

١. تحليل درجة الصعوبة بنود أسئلة اللغة العربية لاختبار التقييم النهائي للفصل الدراسي بناء على المنهج المستقل من مدرسي اللغة العربية بالمدرسة الثانوية الإسلامية الحكومية الواحدة بكنبارو

درجة صعوبة السؤال هو قياس مدى حجمه درجة صعوبة السؤال. ويقال للسؤال أنه جيد إذا كان لديه - أن يكون مستوى صعوبة الأسئلة متوازناً من حيث الأسئلة إنها ليست سهلة للغاية ولا صعبة للغاية. (Ardhani 2020) أسئلة السؤال الجيد هو السؤال الذي ليس سهلاً للغاية ولا صعباً للغاية. الأسئلة السهلة للغاية لا تحفز الطلاب على التحسن محاولة حلها. ومن ناحية أخرى، فإن الأسئلة التي ستكون صعبة للغاية مما يتسبب في إحباط الطلاب. (Latifah et al. 2024)

وفقاً لسورابراناتا Surapranata التوزيع الجيد لمستوى أو درجة صعوبة عناصر الاختبار هو أن النسبة بين عناصر الأسئلة الصعبة - السهلة - المتوسطة تكون ٣-٤-٣، أي ٣٠٪ صعبة، ٤٠٪ متوسطة، و ٣٠٪ سهلة مع توقع أسئلة الاختبار يمكن أن يكون إلى جانب جميع الطلاب، سواء الطلاب الأذكاء أو الطلاب الأقل ذكاءً. (Restiyawati, Suama, and Sappaile 2023)

أ) درجة الصعوبة بنود أسئلة اللغة العربية لاختبار التقييم النهائي للفصل الدراسي بناء على المنهج المستقل من مدرسي اللغة العربية بالمدرسة الثانوية الإسلامية الحكومية الواحدة بكنبارو

جدول نتائج التحليل لدرجة الصعوبة للأسئلة الموضوعية و غير الموضوعية

رقم	نوع السؤال	سهل	متوسط	صعب
١	الموضوعية	١٣	٢٤	٣
٢	غير الموضوعية	٦	٤	-
	مجموع	١٩	٢٨	٣

ومن خلال الجدول أعلاه يتبين أن أسئلة امتحان اللغة العربية لاختبار التقييم النهائي للفصل الدراسي بناء على المنهج المستقل تتطلب تحليل نوعين من الأسئلة، وهما أسئلة الاختيار من متعدد أو الأسئلة الموضوعية، والأسئلة غير الموضوعية. تتضمن الأسئلة الموضوعية أنواع

الصواب والخطأ والمطابقة. بالإضافة إلى ذلك، تظهر أسئلة اللغة العربية في المدرسة الثانوية الإسلامية الحكومية الواحدة بكنبارو تبايناً في درجة الصعوبة، حيث ضمت ١٩ سؤالاً (٣٨٪) في الفئة السهلة، و ٢٨ سؤالاً (٥٦٪) في الفئة المتوسطة، و ٣ أسئلة (٦٪) في الفئة المتوسطة. فئة صعبة.

ب) درجة الصعوبة بنود أسئلة اللغة العربية لاختبار التقييم النهائي للفصل الدراسي بناء على المنهج المستقل من مدرسي اللغة العربية بالمدرسة الثانوية الإسلامية الحكومية الثانية بكنبارو

جدول نتائج التحليل لدرجة الصعوبة للأسئلة الموضوعية و غير الموضوعية

رقم	نوع السؤال	سهل	متوسط	صعب
١	الموضوعية	-	١١	١
٢	غير الموضوعية	٢	١٦	-
	مجموع	٢	٢٧	١

ومن خلال الجدول أعلاه يتبين أن أسئلة امتحان اللغة العربية في نهاية التقييم الفصلي في المنهج المستقل تتطلب تحليل نوعين من الأسئلة، وهما أسئلة الاختيار من متعدد أو الأسئلة الموضوعية، والأسئلة غير الموضوعية. تتضمن الأسئلة الموضوعية أنواع الصواب والخطأ والمطابقة. بالإضافة إلى ذلك، تظهر أسئلة اللغة العربية في المدرسة الثانوية الإسلامية الحكومية الثانية بكنبارو تبايناً في درجة الصعوبة، حيث ضمت ٣ أسئلة (٦٦٪) في الفئة السهلة، و ٢٧ سؤالاً (٩٠٪) في الفئة المتوسطة، وسؤال واحد (٣،٣٪) في الفئة المتوسطة. فئة صعبة.

٢. تحليل مستوى التمييز بنود أسئلة اللغة العربية لاختبار التقييم النهائي للفصل الدراسي بناء على المنهج

المستقل من مدرسي اللغة العربية بالمدرسة الثانوية الإسلامية الحكومية الواحدة بكنبارو

مستوى التمييز هي قدرة عنصر السؤال على التمييز بين الطلاب لديهم قدرات عالية (المجموعة العليا) ومنخفضة (المجموعة السفلى أو الأدنى). إتقان المواد التي تم اختبارها. وبصرف النظر عن ذلك، فإن لديها أيضاً مستوى التمييز مؤشر خاص يوضح مستوى قدرة عناصر الاختبار على التفريق المجموعات ذات التحصيل المرتفع والمنخفض بين الطلاب الذين يشتركون بالامتحان. ويمكن القول أن الأسئلة جيدة في تحديد الطلاب الذين يفهمون المادة مع أولئك الذين لم يفعلوا ذلك إذا كان يتمشى مع قوة التمييز الأعلى. (Pradita, Megawanti, and Indraprasta PGRI 2023)

وفقا لكيلى Kelley و سوجيونو Sudjiono لعدد المشاركين في الاختبار كبيرة، ١٠٠

أو أكثر، ففي تحديد القيمة المثلى يتم استخدام حجم المجموعة $MT = MR = 27\%$ ، أي

بالنسبة لاختيار المجموعة العليا أو العليا تكون ٢٧٪ وبالنسبة كانت المجموعة المنخفضة أو الأقل ٢٧٪.

وكل منهم تنقسم الباحثة إلى مجموعتين يعني أدنى درجات وأعلى الدرجات، بعدد ٢٧٪ لأدنى الدرجات و ٢٧٪ لأعلى الدرجات. وأما الرموز في تحليل مستوى التمييز لأسئلة المقالية وهو يبحث الفرق بين المعدلين يعني المعدل من الفرق الأعلى والمعدل من الفرق

الأدنى. في هذا البحث سوف تنقسم الباحثة نوع تحليل الأسئلة إلى قسمين؛ تحليل الأسئلة الموضوعية وتحليل الأسئلة غير الموضوعية.

أ) تحليل مستوى التمييز بنود أسئلة اللغة العربية لاختبار التقييم النهائي للفصل الدراسي بناء على المنهج المستقل من مدرسي اللغة العربية بالمدرسة الثانوية الإسلامية الحكومية الواحدة بكنبارو

جدول نتائج التحليل لمستوى التمييز للأسئلة الموضوعية و غير الموضوعية					
رقم	نوع السؤال	ضعيف	متوسط	جيد	جيد جدا
١	الموضوعية	٣	-	٣	٣٤
٢	غير الموضوعية	٢	١	٢	٥
مجموع		٥	١	٥	٣٩

ويبين الجدول أعلاه القوة التفاضلية لأسئلة اللغة العربية في المدرسة الثانوية الإسلامية الحكومية الواحدة بكنبارو تضم ٣٩ سؤالاً مصنفة على أنها جيدة جداً.

ب) تحليل مستوى التمييز بنود أسئلة اللغة العربية لاختبار التقييم النهائي للفصل الدراسي بناء على المنهج المستقل من مدرسي اللغة العربية بالمدرسة الثانوية الإسلامية الحكومية الثانية بكنبارو

جدول نتائج التحليل لمستوى التمييز للأسئلة الموضوعية و غير الموضوعية					
رقم	نوع السؤال	ضعيف	متوسط	جيد	جيد جدا
١	الموضوعية	١	-	١	١٠

٩	٥	٢	٢	غير الموضوعية	٢
١٩	٦	٢	٣	مجموع	

ويبين الجدول أعلاه القوة التفاضلية لأسئلة اللغة العربية في المدرسة الثانوية الإسلامية الحكومية الثانية بكنبارو تضم ١٩ سؤالاً مصنفة على أنها جيدة جداً.

خلاصة البحث

ومن خلال تحليل مستوى التمييز لأسئلة اللغة العربية للصف العاشر بالمدرسة الثانوية الإسلامية الحكومية الواحدة بكنبارو ، تبين أنه من أصل ٥٠ سؤالاً، كان ٤٤ سؤالاً تتمتع فيها مستوى التمييز أعلى من ٠,٣٠، والتي تم تصنيفها على أنها جيد وجيد جداً. وهذا يوضح أن معظم الأسئلة يمكنها التمييز بين الطلاب ذوي التحصيل العالي والمنخفض بشكل جيد. ولكن هناك ٦ أسئلة لها قوة تمييز أقل من ٠,٢٠، حيث تكون هذه الأسئلة ذات قدرة منخفضة على التمييز بين الطلاب ذوي القدرة العالية والمنخفضة. يجب مراجعة الأسئلة ذات مستوى التمييز المنخفض أو حذفه حتى يمكن تحسين الجودة الإجمالية للاختبار.

وأما بتحليل مستوى التمييز لأسئلة اللغة العربية للصف العاشر بالمدرسة الثانوية الإسلامية الحكومية الثانية بكنبارو، وقد وجد أن ما يصل إلى ٢٥ سؤالاً كانت مستوى التمييز أعلى من ٠,٣٠. وهذا يدل على أن ٨٣,٣٣٪ من إجمالي الأسئلة التي تم اختبارها صُنفت على أنها جيدة في تمييز الطلاب ذوي القدرات العالية والمنخفضة. هذه الأسئلة قادرة على قياس الاختلافات في مستويات الفهم لدى الطلاب بشكل فعال. وفي الوقت نفسه، هناك ٥ أسئلة ذات مستوى التمييز أقل من ٠,٣٠، مما يدل على أن هذه الأسئلة أقل من الأمثل في التفريق بين الطلاب ذوي مستويات القدرة المختلفة، لذا فهي تتطلب المراجعة أو الاستبدال.

المراجع

- Ardhani, Yusuf. 2020. "Kualitas Butir Soal Penilaian Akhir Tahun Mata Pelajaran Teknologi Dasar Otomotif Kelas X Teknik Kendaraan Ringan Otomotif Di Smk Muhammadiyah Gamping." *Jurnal Pendidikan Vokasi Otomotif* 3 (1): 85-94. <https://doi.org/10.21831/jpvo.v3i1.34917>.
- Budiyanto dan Titik Haryat. 2023. "Faktor-Faktor Yang Mempengaruhi Peningkatan Mutu Pendidikan." *Pendidikan Terintegrasi* 4 (1): 31-36.
- Elviana. 2020. "Analisis Butir Soal Evaluasi Pembelajaran Pendidikan Agama Islam Menggunakan Program Anates." *Analisis Butir Soal Evaluasi Pembelajaran Pendidikan Agama Islam Menggunakan Program Anates* 10 (2): 58-74.

- Fatimah, Laela Umi, and Khairuddin Alfath. 2019. "Analisis Kesukaran Soal, Daya Pembeda Dan Fungsi Distraktor." *Analytical Biochemistry* 8 (2): 1–5. <http://link.springer.com/10.1007/978-3-319-59379-1%0Ahttp://dx.doi.org/10.1016/B978-0-12-420070-8.00002-7%0Ahttp://dx.doi.org/10.1016/j.ab.2015.03.024%0Ahttps://doi.org/10.1080/07352689.2018.1441103%0Ahttp://www.chile.bmw-motorrad.cl/sync/showroom/lam/es/>.
- Helean, Stevi, Nicholas Kandowangko, and Shirley Goni. 2021. "Peranan Pendidikan Dalam Meningkatkan Sumber Daya Manusia Di SMA Negeri 1 Tampam Amma Di Talaud." *Journal Holistik* 14 (2): 1–17. <https://ejournal.unsrat.ac.id/index.php/holistik/article/download/34453/32350>.
- Idrus. 2019. "EVALUASI DALAM PROSES PEMBELAJARAN." *ADAARA: Jurnal Manajemen Pendidikan* 9 (2): 920–35.
- Latifah, Inayatul, Martalia Ardiyaningrum, Mufida Awalia Putri, and Ahmad Syamsul Arifin. 2024. "Analisis Hasil Ujian Sekolah Berstandar Nasional Di Sekolah Dasar Kecamatan Bulu Pada Muatan Matematika." *Indonesian Journal of Elementary Education and Teaching Innovation* 3 (1): 52. [https://doi.org/10.21927/ijeeti.2024.3\(1\).52-63](https://doi.org/10.21927/ijeeti.2024.3(1).52-63).
- Magdalena, Ina, Nurul Hidayati, Ratri Hersita Dewi, Sabgi Wulan Septiara, and Zahra Maulida. 2023. "Pentingnya Evaluasi Dalam Proses Pembelajaran Dan Akibat Memanipulasinya." *Masaliq* 3 (5): 810–23. <https://doi.org/10.58578/masaliq.v3i5.1379>.
- Maulana, Habib. 2023. "Analisis Kualitas Instrumen Evaluasi Pembelajaran Menggunakan Media Digitalisasi Untuk Memotivasi Hasil Belajar Peserta Didik." *Jurnal Pendidikan Bhinneka Tunggal Ika* 1 (4): 1–12. <https://doi.org/10.51903/bersatu.v1i4.255>.
- Muzayyanah, Infi'atin. 2020. "Analisis Butir Soal Penilaian Tengah Semester Multiple Choice PAI & BP Buatan Guru (Perbandingan SMPN 18 Dan SMPN 44 Semarang)." *Eprints.Walisongo.Ac.Id*. Universitas Islam Negeri Walisongo Semarang. https://eprints.walisongo.ac.id/15352/1/SKRIPSI_1603016019_INFI%27ATIN_MUZAYYANAH.pdf.
- Nurhairiyah, Siti. 2021. "Pengembangan Instrumen Tes Untuk Mengukur Kemampuan Penalaran Statistik." *Jurnal Pembelajaran Matematika* 1 (1): 1–19.
- Pradita, Eliza, Priarti Megawanti, and Universitas Indraprasta PGRI. 2023. "Analisis Tingkat Kesukaran, Daya Pembeda, Dan Fungsi Distraktor PTS Matematika SMPN Jakarta." *HIMPUNAN: Jurnal Ilmiah Mahasiswa Pendidikan Matematika* 3 (80): 109–

18.

- Prastiwi, Yunita Eka Nur, Arba'iyah, Afifah Amatullah Al Barru, and Achmad Syarif Hidayatullah. 2023. "Penilaian Dan Pengukuran Hasil Belajar Pada Peserta Didik Berbasis Analisis Psikologi." *Bersatu: Jurnal Pendidikan Bhineka Tunggal Ika* 1 (4): 218–31. <https://journal.politeknik-pratama.ac.id/index.php/bersatu/article/view/293>.
- Ratni. 2023. "Analisis Tingkat Kesukaran, Daya Pembeda Dan Efektifitas Pengecoh Soal Ulangan Bahasa Indonesia Di Sekolah Dasar." *Jurnal Penelitian Pendidikan Guru Sekolah Dasar* 1 (Januari): 117–23.
- Restiyawati, I Wayan Suama, and Pallawagau Sappaile. 2023. "Analisis Tingkat Kesukaran Dan Daya Pembeda Soal Ulangan Biologi Jurusan MIA MAN 1 Buton Selatan." *Suparyanto Dan Rosad* (2015 7 (4): 248–53.
- Rismaulhijjah, Wikha, and Nur Kuswanti. 2022. "Analysis Of Daily Exam Test Items Developed by Teacher on the Human Motion System Material for Class XI Natural Science." *Berkala Ilmiah Pendidikan Biologi (BioEdu)* 11 (3): 643–61. <https://doi.org/10.26740/bioedu.v11n3.p643-661>.
- Rizky Ananda Setiyawan, and Palupi Sri Wijayanti. 2020. "Analisis Kualitas Instrumen Untuk Mengukur Kemampuan Pemecahan Masalah Siswa Selama Pembelajaran Daring Di Masa Pandemi." *Jurnal Lebesgue : Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika* 1 (2): 130–39. <https://doi.org/10.46306/lb.v1i2.26>.
- Siregar, Nurul Hidayah, Remiswal Remiswal, and Khadijah Khadijah. 2024. "Analisis Butir Soal Ujian Tengah Semester Pada Mata Pelajaran Pendidikan Agama Islam." *Urwatul Wutsqo: Jurnal Studi Kependidikan Dan Keislaman* 13 (2): 179–89. <https://doi.org/10.54437/urwatulwutsqo.v13i2.1637>.
- Tuerah, M S Roos, and Jeanne M Tuerah. 2023. "Kurikulum Merdeka Dalam Perspektif Kajian Teori: Analisis Kebijakan." *Jurnal Ilmiah Wahana Pendidikan, Oktober* 9 (19): 982. <https://doi.org/10.5281/zenodo.10047903>.