

الباب الثاني

الأسس النظرية

أ. إطار النظري

١. التقويم

أ. تعريف التقويم

التقويم يأتي من كلمة قَوْم-يَقَوْم-تقويم. التقويم مرتبط بالقياس والتقييم.^١ القياس (measurement) هو نشاط أو عملية تحديد الحقائق الكميّة من خلال مقارنة كائن ليتمّ قياسه بوحدة قياس قياسيةّ محدّدة مسبقاً. التقييم (assessment) هو اتخاذ المقرّرات لتحديد شيء نوعيّ بناءً على معايير معيّنة. في التعليم، القياس هو عملية جمع البيانات من خلال الملاحظة التحريية لجمع المعلومات ذات الصلة بهدف محدد مسبقاً. التقييم هو تطبيق طرق مختلفة واستخدام أدوات التقويم المختلفة للحصول على معلومات حول مدى نتائج تعلم الطلاب أو تحقيق كفاءات الطلاب.^٢ لذلك، التقويم هو عملية لاتخاذ المقررات باستخدام مجموعة من نتائج القياس واستناداً إلى الأهداف التي تمت صياغتها. بهذه الطريقة، يمكن إجراء التقويم بعد القياس.

1 Muhammad Rapono, Safrial Safrial, dan Candra Wijaya, "Urgensi Penyusunan Tes Hasil Belajar: Upaya Menemukan Formulasi Tes Yang Baik dan Benar," *Jupiis: Jurnal Pendidikan Ilmu-Ilmu Sosial* 11, no. 1 (2019): 95, <https://doi.org/10.24114/jupiis.v11i1.12227>.

2 Arief Aulia Rahman dan Cut Eva Nasryah, *Evaluasi Pembelajaran, Uwais Inspirasi Indonesia* (Ponorogo: Uwais Inspirasi Indonesia, 2019).

التقويم التربوي وفقا لبلوم نقل عن هاندريياوان (٢٠٢١)، أي الجمع المنهجي للحقائق لتحديد ما إذا كان هناك تغيير في الطلاب في الواقع وتحديد مستوى التغيير في الطلاب.^٣ التقويم التربوي هو عملية تحديد قيمة النشاط المخطط والمنظم لتحليل وتفسير عمليات ونتائج التعلم. يستخدم القياس في التعليم أدوات قياس في شكل اختبارات وغير اختبارات، في حين أن التقييم هو تحديد نتائج درجات الاختبارات التي تم إجراؤها.^٤ ترتبط القياس والتقييم والتقويم بعضها البعض. القياس هو المرحلة الأولى في عملية التقييم حيث يكون التقييم أحد جوانب التقويم التربوي.^٥

ب. أهداف التقويم

التقويم التربوي هو عامل مهم كمياري لنجاح العملية التعليمية لكل المعلمين والطلاب. التقويم بشكل عام عدة أهداف منها:

١. للحصول على معلومات دقيقة وموضوعية حول برنامج يستخدم كمرجع لتحديد كفاءة وفعالية أنشطة التقويم. يمكن أن تكون المعلومات في شكل عملية تنفيذ

3 Dony Handriawan dan Muhammad Nurman, *Evaluasi Pembelajaran Bahasa Arab I*, ed. oleh MA Sultan, 1 ed. (Mataram: Sanabil, 2021).

4 Komarudin dan Sarkadi, *Evaluasi Pembelajaran*, ed. oleh Sugeng Miryanto Yuyus Kardiman, Andika, 2 ed. (Yogyakarta: RizQita Publishing & Printing, 2017)

5 Roviin, "EVALUASI PEMBELAJARAN BAHASA ARAB," n.d., 195–213.

البرنامج، الأثر أو النتائج المحققة، كفاءة واستخدام نتائج
التقويم.^٦

٢. بحث عن العوامل التي تسبب نجاح وفشل الطلاب في
المشاركة برامج التعليمية لإيجاد الحلول لتحسينها.

٣. اخبر عن مدى تحصيل وكفاءته كل طالب في نهاية
الفترة.^٧

٤. لأغراض تطوير وتحسين المناهج الدراسية للمدرسة
المعنية.^٨

وفي الوقت نفسه، أوضح سوهارسمي اريكونتو (٢٠١٨)،^٩
أهداف ووظائف تقويم نتائج التعلم على النحو التالي:
١. تقويم الوظيفة الانتقائية

يسهل هذا التقويم على المعلمين إجراء اختيارات أو
تقويمات للطلاب ذوي القدرات العالية أو التي لا تزال منخفضة.
يمكن استخدام هذا الانتقائي في اختيار الطلاب الذين ينتقلون
في الصف التالي والطلاب الذين يحصلون على منحة دراسية

6 Dony Handriawan dan Muhammad Nurman, *Evaluasi Pembelajaran Bahasa Arab I*, ed. oleh MA Sultan, 1 ed. (Mataram: Sanabil, 2021).

٧فؤادحسن حسين ابوالهيحاج، اساليب وطرق تدريس اللغة العربية واعداد دروسها

اليومية (عمان: دارالمناهج, ٢٠٠٧)

8 Ubaid Ridho, "Evaluasi Dalam Pembelajaran Bahasa Arab," *An Nabighoh Jurnal Pendidikan dan Pembelajaran Bahasa Arab* 20, no. 01 (2018): 19, <https://doi.org/10.32332/an-nabighoh.v20i01.1124>.

9 Suharsimi Arikunto, *DASAR-DASAR EVALUASI PENDIDIKAN*, 3 ed. (Jakarta: Bumi Aksara, 2018). hlm10

والطلاب الذين تم الإعلان عن تخرجهم من المدرسة أو نجاحهم من المواد.

٢. تقويم الوظيفة التشخيصية

يعمل هذا التقويم على معرفة نقاط القوة والضعف في عملية التعليم والتعلم. عندما تعرف أسباب هذا الضعف، سيكون من الأسهل إيجاد طرق للتغلب على الضعف.

٣. يعمل التقويم كمحدد لمجموعات الطلاب

يمكن وضع الطلاب وفقا لقدراتهم. يتم وضع الطلاب عن طريق تجميع الطلاب الذين لديهم نفس القدرات.

٤. التقويم كمقياس للنجاح

لمعرفة مدى نجاح البرامج الذي تم تنفيذه في عملية التعليم والتعلم.

ج. نوع التقويم

هناك نوعان من التقويم في التعليم، وهما الاختبار وغير الاختبار. التفسير هو كما يلي:

١. غير الاختبار

يتم تجميع التقويم غير الاختبار في التعليم على النحو التالي^{١٠}:

10 Suharsimi Arikunto, *DASAR-DASAR EVALUASI PENDIDIKAN*, 3 ed. (Jakarta: Bumi Aksara, 2018) hlm 41-44

أ. الاستبيان

الاستبيان هو قائمة من الأسئلة التي يجب أن يملأها المستجيبون. وهدف هو معرفة البيانات الشخصية والخبرة ومعرفة الموقف وغيرها.

ب. قائمة تدقيق

هي سلسلة من التصريحات، حيث وضع المستجيبون علامة اختيار في العمود المقدم.

ج. المقابلة

المقابلة هي طريقة للحصول على المعلومات في شكل إجابات من المستجيبين عن طريق سؤال وجواب من جانب واحد.

د. الملاحظات

إنها تقنية يتم تنفيذها عن طريق إجراء ملاحظات دقيقة وتسجيل منهجي. هناك عدة أنواع من الملاحظات، وهي:

١) الملاحظة المشتركة

هي الملاحظات التي أدلى بها المراقبون مباشرة.

٢) الملاحظة المنهجية

هي الملاحظات التي تم فيها إدراج العوامل التي يجب ملاحظتها بشكل منهجي وفقا لفئاتها.

٣) الملاحظة التجريبية

هي الملاحظات التي لم يشارك فيها المراقبون في
المجموعة التجريبية

٢. الاختبار

أ. تعريف الاختبار

الإختبار لغويا يأتي من كلمة اختبر-يختبر أو امتحن
-يمتحن، صيغة الجمع هي الاختبارات يرادف معنى
الامتحان.^{١١} هناك العديد من المصطلحات المتعلقة
بالاختبار. *Tes* هو اختبار أو تجربة. *Testing* هو حدث
قياس وتقييم. *Tester* هو الطرف الذي أنشأ الاختبار. و
Teste هو الطرف الذي يجري الاختبار (المشارك
للاختبار). يعرف أريكونتو في كتابها بعنوان أساسيات
التقويم التربوي، الاختبار هو تقنية أو طريقة تستخدم من
أجل تنفيذ أنشطة القياس التي توجد فيها أسئلة أو
عبارات أو سلسلة من الوظيفة المختلفة التي يجب أن
يقوم بها الطلاب لقياس جوانب سلوك الطالب.^{١٢}
الاختبار هو طريقة تستخدم لتحديد قدرة الطلاب
على إكمال وظيفة معينة مما يدل على إتقان المهارات أو

١١ المعاني عربي، " الإختبار" زمن الوصول في ١٥ نوفمبر ٢٠٢٣،

<https://play.google.com/store/apps/details?id=com.almaany.arar>

12 Suharsimi Arikunto, *DASAR-DASAR EVALUASI PENDIDIKAN*, 3
ed. (Jakarta: Bumi Aksara, 2018) hlm 90

معرفة المحتوى.^{١٣} إذا ارتبطت بالتقويم التربوي، فإنّ الاختبارات هي طرق أو إجراءات تستخدم في القياس والتقييم في مجال التعليم في شكل وظيفة أو سلسلة من الوظيفة في شكل أسئلة أو أوامر يجب القيام بها حتى يتم الحصول على الدرجات التي ترمز إلى السلوك.^{١٤} أما بالنسبة لاختبارات اللغة وفقا لسوناردي، فإن تعريفها واستخدامها يرتبطان أكثر بالمهارات اللغوية، وليس بالمعرفة اللغوية. لذلك، لا يمكن القول إن الاختبار الذي يهدف إلى قياس معرفة الطلاب اللغوية هو اختبار لغة.^{١٥} وبالتالي فإن المقصود باختبارات اللغة بشكل عام هي عدد من الإجراءات والأدوات التي يتم ترتيبها بشكل منهجي لتحديد مستوى إتقان الطلاب لمهارة أو كفاءة لغوية أو ككل. للاختبارات وظيفة مزدوجة، وهي قياس قدرات الطلاب وقياس نجاح برامج التعليم.

ب. وظائف الاختبار لقياس نجاح الطلاب

من حيث الفائدة في قياس نجاح الطلاب، ينقسم الاختبار إلى ثلاثة، وهي:

13Awaluddin Fajar, *EVALUASI PEMBELAJARAN BAHASA ARAB BERBASIS INFORMASI, KOMUNIKASI, DAN TEKNOLOGI*, ed. oleh Sultan Muslihin, 1 ed. (Yogyakarta: Arti Bumi Intara, 2023). hlm 11

14 Dony Handriawan dan Muhammad Nurman, *Evaluasi Pembelajaran Bahasa Arab I*, ed. oleh MA Sultan, 1 ed. (Mataram: Sanabil, 2021) hlm.46

15 Islakhul Anam dkk, "Bentuk – Bentuk Tes Bahasa" (Purwokerto, 2022).

١. الاختبارات التشخيصية

تعرف سوهارسيمي أريكونتو (٢٠١٨) الاختبارات التشخيصية هي اختبار تستخدم لتحديد نقاط ضعف الطلاب في إتقان المادة.^{١٦} من خلال الاختبارات التشخيصية، يكون المعلم قادرا على إيجاد الحلول المناسبة لحل الصعوبات التي يواجهها المتعلمون. خصائص الاختبارات التشخيصية هي كما يلي:

- أ. للكشف عن صعوبات التعلم
 - ب. تم تطويره بناء على مصدر الصعوبة
 - ج. استخدام نماذج أسئلة الإجابة القصيرة
 - د. مصحوبة بخطة متابعة حسب مصدر الصعوبة.^{١٧}
- ## ٢. الاختبارات التكوينية

الاختبارات التكوينية هي اختبار تستخدم لمراقبة تقدم الطلاب أثناء عملية التعليم والتعلم التي يتم إجراؤها في نهاية الدرس والذي يسمى عادة الإختبارات اليومية الذي يتم باستمرار.^{١٨} من خلال الاختبارات التكوينية، يتوقع من المعلمين

16 Suharsimi Arikunto, *DASAR-DASAR EVALUASI PENDIDIKAN*, 3 ed. (Jakarta: Bumi Aksara, 2018) hlm 47

17 Ellen Sutopo Putri dan Rinaningsih Rinaningsih, "Review: Tes Diagnostik Sebagai Tes Formatif Dalam Pembelajaran Kimia," *UNESA Journal of Chemical Education* 10, no. 1 (2021): 20–27, <https://doi.org/10.26740/ujced.v10n1.p20-27>.

18 Hardianti Okta, "Penerapan Tes Formatif dalam Pembelajaran Pendidikan Agama Islam di SMPN 1 Pulau Panggung Kabupaten Tanggamus," *Angewandte Chemie International Edition*, 6 (11), 2019, 951–52.

تحسين برنامج التدريس واستراتيجيات التنفيذ. يتم إجراء الاختبارات التكوينية في كل مرة تنتهي فيها وحدة الدراسة أو الموضوع الفرعي.

تهدف الاختبارات التكوينية إلى مراقبة تقدم التعلم للطلاب في عملية التعليم من خلال تقديم التغذية الراجعة كمكمل لبرنامج التعليم ومعرفة نقاط الضعف التي تتطلب التحسين حتى تصبح نتائج التعلم المتوقعة أفضل.

٣. الاختبارات التلخيصية

الاختبارات التلخيصية هي الاختبار التي يتم إجراؤها بعد انتهاء توفير مجموعة معينة من البرامج. يمكن تسمية الاختبارات التلخيصية باختبارات الامتحانات النهائية لأن الاختبارات التلخيصية يتم إجراؤها بعد الانتهاء من جميع المواد التعليمية، أي في نهاية الفصل الدراسي أو نهاية التعليم، مثل تقويم نهاية الفصل الدراسي. تؤكد الاختبارات التلخيصية على السلوك المعرفي.

وظيفة الاختبارات التلخيصية:

١. تحديد الدرجات المتعلقة بتحصيل على تعلم الطلاب في المدرسة

٢. لتحديد ما إذا كان الطفل قادراً على حضور البرنامج في الفصل التالي أم لا

٣. الإشارة للطلاب إلى أنهم شاركوا في أحد البرامج، وتحديد موقع قدرات الطلاب أمقارنة بأصدقائه في المجموعة.

٤. لتعبئة الملاحظات الدراسية للطلاب التي تفيد في:

أ. الوالدا الطلاب

ب. أطراف التوجيه والإرشاد في المدرسة

ج. أطراف أخرى إذا كان الطالب سيستمر في الدراسة أو حتى يدخل مجال العمل.

ج. نماذج الاختبار

بناء على شكل السؤال، قسم سوهارسيمي أريكونتو الاختبار إلى شكلين، وهما الاختبارات شخصيات والاختبارات الموضوعية.

١. الاختبارات الشخصيات

الاختبارات شخصيات وفقا لسوهارسيمي أريكونتو هي عموما في شكل مقالات (أوصاف). اختبار نموذج المقال هو الاختبارات التي يتطلب إجابات عبارة مناقشة أو وصف كلمة.^{١٩} تبدأ خصائص الأسئلة كلمات مثل وصف، شرح، لماذا، كيف، قارن، استنتاج، وغير ذلك. تحتوي أسئلة نموذج المقال على عدد

19 Suharsimi Arikunto, *DASAR-DASAR EVALUASI PENDIDIKAN*, 3 ed. (Jakarta: Bumi Aksara, 2018). hlm 108

من الأسئلة التي ليست كثيرة فقط حوالي ٥-١٠ أسئلة في وقت المعالجة من حوالي ٩٠ إلى ١٢٠ دقيقة. تتطلب أسئلة النموذج المقالي قدرة الطلاب على أن يكونوا قادرين على تنظيم وتفسير وربط التفاهات التي لديهم بالفعل. باختصار يمكن القول أن الاختبارات المقالية تتطلب من الطلاب أن يكونوا قادرين على التذكر والتعرف مرة أخرى، خاصة يجب أن يكون لديهم إبداع عالي.

تتمتع الاختبارات شخصيات بالعديد من المزايا، وهي إعداد الأسئلة بسهولة، وتشجع الطلاب على التعبير عن آرائهم وتجميعها في شكل جمل مناسبة، وتمنح الطلاب الفرصة للإجابة على الأسئلة وفقا لأسلوب لغتهم وطريقتهم، وقادر على معرفة مدى فهم الطلاب للمشكلة التي يتم اختبارها. بالإضافة إلى وجود مزايا، فإن الاختبارات شخصيات لها أيضا العديد من العيوب، بما في ذلك انخفاض مستويات الصدق والثبات، وتتأثر طريقة تصحيح الإجابات بالعناصر شخصيات، وأوقات التصحيح الطويلة، وأقل تمثيلا في تمثيل كل نطاق من مواد الدراسة بسبب العدد المحدود من الأسئلة المقدمة.^{٢٠}

20 Suharsimi Arikunto, *DASAR-DASAR EVALUASI PENDIDIKAN*, 3 ed. (Jakarta: Bumi Aksara, 2018) hlm 109

٢. الاختبارات الموضوعية

الاختبارات الموضوعية هي اختبار كتابي يتطلب من الطلاب اختيار الإجابات التي تم تقديمها أو تقديم إجابات قصيرة محدودة. تتمتع الاختبارات الموضوعية بالعديد من المزايا، بما في ذلك التصحيح الأسهل نظراً لوجود مفتاح إجابة. يمكن ترك تصحيح الاختبارات الموضوعية للآخرين لأنه لا يوجد عنصر شخصي في التصحيحها.

هناك عدة أنواع من الاختبارات الموضوعية، وهي: ^{٢١}

أ. اختبار صحيح خطيء

أسئلة هذا الاختبار هي في شكل أسئلة بين صواب و خطأ. ثم يضع الطالب علامة على السؤال بالرمزين B إذا كان صواباً و S إذا كان خطأ.

الفضل	النقص
أخذ عينات كافية	سهل التخمين (٥٠٪ من المتقدمين للاختبار يمكنهم التخمين بشكل صحيح)
يمكن إجراء التقويم بكفاءة وموضوعية	يمكن فقط قياس ترتيب المهارات العليا بشكل غير مباشر
سهل التحضير و سهل	يميل إلى تشجيع المتقدمين للاختبار على

21 Lumbantobing Linda Ria, "Kelebihan dan Keterbatasan dari Bentuk-Bentuk Penilaian Tes Tertulis," *Kemenkeu Learning Center*, 2021, 1–2, <https://klc2.kemenkeu.go.id/kms/knowledge/kelebihan-keterbatasan-penilaian-tes-tertulis-0f81178c/detail/>.

الفهم	حفظ الحقائق بدلا من دراسة المحتوى
-------	-----------------------------------

ب. اختبار المطابقة

هي اختبار يتم إجراؤها عن طريق مطابقة أو إقران سؤال واحد بسلسلة من الإجابات المتاحة. ومفتاح استخدام هذا النموذج هو وجود علاقة منطقية بين العناصر المراد ربطها. يجب أن تكون العناصر في كل عمود متجانسة، أي مرتبطة بنفس الموضوع.

الفضل	النقص
التنسيق المستخدم أكثر كفاءة وفعالية لاختبار المتقدمين للاختبار فيما يتعلق بالارتباطات الواقعية الأساسية في مجال المحتوى	أقل قدرة على تقييم المهارات المعرفية الأخرى بما يتجاوز مجرد الارتباطات الواقعية التي تؤكد على الحفظ
التقييم السريع والموضوعي والموثوق	مادة السؤال محدودة بالذاكرة البسيطة أو عوامل المعرفة
أوقات القراءة والاستجابة قصيرة نسبيا	من الصعب بناء سؤال يحتوي على عدد من الإجابات المتجانسة

ج. اختبار الاختيار من متعدد

اختبار الاختيار من متعدد هو اختبار يتكون من معلومات أو إخطار بفهم غير كامل. لإكماله، يجب على الشخص الذي يأخذ للاختبار اختيار أحد خيارات الإجابة العديدة التي تم توفيرها. تتكون الإجابات البديلة المقدمة من أربعة إلى خمسة عناصر مع إجابة واحدة صحيحة كمفتاح إجابة، والباقي في شكل إجابات مشتتة.^{٢٢} في هذه الدراسة، ركز على تحليل اختبارات نموذج الاختيار من متعدد.

الفضل	النقص
يمكن إجراء التقويم بسهولة وسرعة وله موضوعية عالية	يتطلب وقتا طويلا لتجميعها
قياس المستويات المعرفية المختلفة ويمكن أن تغطي نطاقا واسعا من المواد في الاختبار	من الصعب عمل مشتت متجانس وعامل
مناسبة لعدد كبير من المتقدمين للاختبار	يتمتع المقدمون للاختبار فرصة تخمين الإجابة على شيء ما
من المناسب استخدامه للامتحانات واسعة النطاق مثل الامتحانات الوطنية والامتحانات النهائية للمدرسة وما إلى	أقل فعالية لقياس أنواع معينة من حل المشكلات، والقدرة على تنظيم الأفكار والتعبير

22 Muhibbatul Laili, “Ketepatan Kontruksi Butir Pilihan Ganda Bahasa Arab” 3, no. 2 (2020): 111–24.

ذلك.	عنها
------	------

بعض الأشياء المتعلقة باختبارات خيارات متعددة هي كما يلي:

٢٣

- أ. يجب أن تكون تعليمات العمل واضحة
- ب. في اختبارات الاختيار من متعدد، هناك إجابة واحدة صحيحة فقط
- ج. يجب أن تكون الجملة في كل بند من الأسئلة قصيرة قدر الإمكان
- د. الإجابة المقدمة لها توافق نحوي مع الجملة الرئيسية
- هـ. من الناحية اللغوية، لا ينبغي أن تكون بنود الأسئلة صعبة للغاية.

٢. تحليل بنود الأسئلة

أ. تعريف تحليل بنود الأسئلة

كلمة تحليل مشتقة من حلل - يحلل - تحليل، بمعن فكاك.^{٢٤}

وكلمة بنود هي جمع من بند بمعنى نقطة من نقاط البحث.^{٢٥} بينما الأسئلة

23 Suharsimi Arikunto, *DASAR-DASAR EVALUASI PENDIDIKAN*, (Jakarta: Bumi Aksara, 2018), hlm. 170-171

24 Ahmad Warsono Munawir, *Kamus Al-Munawir Arab-Indonesia*, (Surabaya: Pustaka Progresif, 1997), hlm.292

هي جمع من كلمة سؤال، (سؤال-سؤالين- أسئلة) بمعنى ما يطرح من استفساراتٍ عن موضوع مالمالإجابة عليها، وذلك لاختيار تحصيل الطلاب.^{٢٦}

تحليل بنود الأسئلة هو نشاط يقوم به المعلم لتحديد جودة الأسئلة التي تم اختبارها. تحليل بنود الأسئلة هو نشاط لفحص جودة السؤال بعد اختبارها للطلاب. وفقا لسوهارسيمي أريكونطو، فإن الغرض من تحليل بنود الأسئلة في الاختبار الذي يجريه المعلم هو تحديد الأسئلة الجيدة وليس جيدة والسيئة. بتحليل الأسئلة يمكن الحصول على معلومات حول ضعف السؤال وتعليمات لإجراء تحسينات. بناء على هذه الأهداف، فإن تحليل بنود الأسئلة له فوائد عديدة، بما في ذلك:

- أ. يمكن أن يساعد مستخدمي الاختبار في تقويم الاختبارات المستخدمة.
- ب. يدعم كتابة بنود الأسئلة الفعالية
- ج. تحسين الاختبار في الفصل.
- د. تحسين صدق الأسئلة وثباتها.

٢٥ معجم المحيط (قاموس اليكترونيك عرب-ايندونيسيا) " بنود" زمن الوصول في

٢٠٢٣،

نوفمبر

١٥

<https://play.google.com/store/apps/details?id=com.ristekmuslim.kamusarabindo>

٢٦معجم الغني " الأسئلة " زمن الوصول في ١٥ نوفمبر ٢٠٢٣،

<https://play.google.com/store/apps/details?id=com.almaany.arar>

ذ. تحسين بنود الأسئلة من خلال ثلاثة مكونات هي تحليل الصعوبة والقوة التمييزية والمشتت، وتحسين التعليم من خلال غموض السؤال وبعض المهارات التي تسبب صعوبة للطلاب.

ب. خصائص الاختبار الجيد

سيكون الاختبار أداة قياس جيدة عندما يفي بمتطلبات الاختبار. ذكر سوهارسمي أريكونتو في كتابها بعنوان أساسيات التقويم التربوي، أن متطلبات الاختبار يقال جيدة عندما يكون لها صدق وثبات وصعوبة، وتمييز.²⁷ للتوضيح على النحو التالي:

١. الصدق

كلمة صدق له معنى الدقيقة والحقيقة والصلاحية. يقال إن الاختبار صدق عندما يتمكن الاختبار من قياس ما يريد قياسه. يتم التأكيد على هذا الصدق على نتائج الاختبار أو الدرجات، وليس على الاختبار. وبالتالي، تصبح نتائج الاختبار صحيح عندما تصف نتائج الاختبار بالفعل القدرة التي يتم قياسها أو اختبارها. على سبيل المثال، سيقس المعلم قدرة الطلاب على القراءة من خلال إجراء اختبار يتكون من القراءة والأسئلة المتعلقة بالقدرة على القراءة، كما حول الفكرة الرئيسية للقراءة والحقائق التي تتم قراءتها وما إلى ذلك. بهذه الطريقة سيتم الحصول على نتائج الاختبار بمستوى عال نسبياً من الصدق. بشكل عام، يقترح

27 Suharsimi Arikunto, *DASAR-DASAR EVALUASI PENDIDIKAN*, (Jakarta: Bumi Aksara, 2018), hlm. 232

سوهارسيمي أريكونتوأن هناك صدقين، وهما الصدق المنطقي والصدق التجريبي.

تستخدم صدق المحتوى بشكل شائع لاختبار مكونات جهاز القياس.^{٢٨} صدق المحتوى هي الصدق التي تقيس أهدافا معينة وفقا لمادة أو محتوى الدرس (المنهج) المعطى.^{٢٩} يجب أن تكون بنود الأسئلة المقدمة عينة تمثيلية من مجتمع المواد التعليمية التي تم تدريسها حتى يتمكنوا من تمثيل جميع الأداء الذي تم الحصول عليه أثناء الطلاب في وحدة التعليم.^{٣٠} في صدق المحتوى، يتم إجراء القياس من خلال مقارنة الملاءمة بين محتوى الأسئلة بشكل شامل مع شبكة الاختبار التي تعد مرجعا في إعداد الاختبار. بنود الاختبارات، في إعدادها مناسبة وتمثل محتوى المادة التي يجب قياسها أم لا.^{٣١} في الصدق المنطقي، يجب أن يكون مستوى صدق بنود الأسئلة وفقا للمنهج الدراسي.^{٣٢}

28 Mina Raeisi, Mohammad Ali Nadi, dan Majid Sharifi Ghoortani, "Validity, reliability and factor analysis of Reid's English language learning styles for high school students: A case study on Iranian students," *Cogent Education* 9, no. 1 (2022), <https://doi.org/10.1080/2331186X.2022.2107305>.

29 Suharsimi Arikunto, *DASAR-DASAR EVALUASI PENDIDIKAN*, (Jakarta: Bumi Aksara, 2018), hlm. 184

30 Melisa Rezi dan Annisa Aulia, "Tahapan Penyusunan dan Analisis Tes Bahasa Arab" 5 (2020): 52–68.

31 Muhammad Lukman Arifianto, Mohammad Ahsanuddin, dan Nurul Fitria, *Evaluasi Pembelajaran dan Pengembangan Tes Interaktif Bahasa Arab*, 1 ed. (Yogyakarta: TONGGAK MEDIA, 2022). hlm 3

32 Nurul Fikriyah, "Analisis Butir Soal Ulangan Tengah Semester Mata Pelajaran Bahasa Arab Kelas VII Semester Ganjil SMP Muhammadiyah 1 Yogyakarta Tahun Ajaran 2019 / 2020," *Maharat : Jurnal Pendidikan Bahasa Arab* 3, no. 2 (2021): 128–40, <https://doi.org/10.18196/mht.v3i2.10501>.

تركز هذه الدراسة على تحليل بنود الأسئلة. تعتمد الصدق الإجمالية للاختبار على صدق كل بنود الأسئلة فيه.^{٣٣} يقال إن الاختبار يكون صحيحا إذا كانت النتائج تتطابق مع المعايير، مما يعني أن هناك توازيا بين نتائج الاختبار والمعايير. بعد معرفة صدق الاختبار، يتم إجراء صدق البند لمعرفة بنود الأسئلة التي تتسبب في أن يكون للاختبار صدق منخفض أو عالية. يقال إن السؤال صدق إذا قام r بحساب $r < 0.5$ من جدول بمستوى دلالة ٥٪ أو إذا كانت قيمة الدلالة > 0.05 . يقال إن البند يتمتع بصدق عالية إذا كانت درجة البند لها محاذاة (ارتباط) مع النتيجة الإجمالية. لتحديد مستوى صدق، يستخدم سوهارسيمي أريكونتو الصيغة التالية:

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

البيانات:

 r_{xy} : معامل الارتباط بين المتغيرات X و Y

N : عدد العينات

X : درجة البند

Y : مجموع الدرجات

المعاملات الضعيفة المستخدمة في تحليل الصدق،^{٣٤} هي كما يلي:

بين ٠.٨٠٠ و ١.٠٠٠ : مرتفع جدا

٣٣ محمد عبد الخالق، اختبارات اللغة، (الرياض: جامعة الملك سعود، ١٩٩٦)

³⁴ Suharsimi Arikunto, *DASAR-DASAR EVALUASI PENDIDIKAN*, (Jakarta : Bumi Aksara, 2018), hlm 21

مرتفع :	بين ٠.٦٠٠ و ٠.٨٠٠
متوسط :	بين ٠.٤٠٠ و ٠.٦٠٠
منخفض :	بين ٠.٢٠٠ و ٠.٤٠٠
منخفض جدا :	بين ٠.٠٠٠ و ٠.٢٠٠

٢. الثبات

الثبات هو تحديد أو استهزاء بمدى استقرار نتائج الاختبار عند اختباره على نفس الموضوع. يشار إلى ثبات مرتفع أو منخفض تجريبيا برقم يسمى قيمة معامل الثبات. يمكن القول أن الاختبار يتمتع بمستوى عال من الثبات عند ما يتم إجراء الاختبار في فترات زمنية مختلفة يمكن أن يعطي نتائج ثابتة. عند تحديد ثبات بنود أسئلة الاختبار من متعدد في أسئلة الاختبار النهائي للفصل الدراسي الأول، حيث يتم اختبار الأسئلة مرة واحدة فقط، يمكن البحث عنها باستخدام الصيغة *Kuder and Richard* (K-R.20) ³⁵ كما يلي:

$$R_{11} = \left(\frac{n}{n-1} \right) \left(\frac{S^2 - \sum pq}{S^2} \right)$$

البيانات:

R_{11} : معامل ثبات الاختبار

P : نسبة الطلاب الذين أجابوا على البند الصحيح

Q : نسبة الطلاب الذين أجابوا على البند غير صحيح

$$(q=1-p)$$

$\sum pq$: عدد نتائج الضرب بين p و q

³⁵ Suharsimi Arikunto, *DASAR-DASAR EVALUASI PENDIDIKAN*, (Jakarta : Bumi Aksara, 2018), hlm 42

n : عدد بنود الاختبار

S : الانحراف المعياري للاختبار

يستخدم أن تفسير نتائج حساب معامل ثبات الاختبار بشكل عام على المعايير التالية:³⁶

منخفض جدا : بين ٠,٢٠ — ٠,٤٠

منخفض : بين ٠,٤٠ — ٠,٦٠

متوسط : بين ٠,٦٠ — ٠,٨٠

مرتفع : بين ٠,٨٠ — ٠,٩٠

مرتفع جدا : بين ٠,٩٠ — ٠,٩٩

١. التمييز

التمييز هو قدرة المشكلة على التمييز بين الطلاب الأذكياء (القدرة العالية) والطلاب الأغبياء (القدرة المنخفضة).³⁷ يرمز إلى التمييز بـ "D" (Distractor). تم تقسيم جميع المتقدمين للاختبار إلى مجموعتين، وهي المجموعة المرتفعة والمجموعة المنخفضة. إذا كان بإمكان جميع المتقدمين لاختبار المجموعة العالية الإجابة على سؤال اختبار بشكل صحيح، بينما يجب جميع المتقدمين لاختبار المجموعة المنخفضة بشكل غير صحيح، فإن

³⁶ Suharsimi Arikunto, *DASAR-DASAR EVALUASI PENDIDIKAN*, (Jakarta : Bumi Aksara, 2018), hlm 42

³⁷ Suharsimi Arikunto, *DASAR-DASAR EVALUASI PENDIDIKAN*, (Jakarta : Bumi Aksara, 2018), hlm 42

السؤال له أكبر فهرس قوة تمييز وهو + ١.٠٠٠. ومع ذلك، إذا أجاب جميع المتقدمين لاختبار المجموعة العالية بشكل غير صحيح، لكن جميع المتقدمين لاختبار المجموعة المنخفضة أجابوا بشكل صحيح، فإن السؤال يحتوي على أصغر فهرس قوة تمييز، وهو - ١.٠٠٠. إذا كانت المجموعة المرتفعة والمجموعة المنخفضة كلاهما أجابتا بشكل صحيح أو خاطئ، فإن السؤال له تمييز في فهرس القدرة يساوي ٠.٠٠٠ لأنه لا يوجد تمييز على الإطلاق.

يمكن تحديد التمييز للسؤال بالطريقة التالية:

١. فرز درجات جميع المتقدمين للاختبار أولاً من أعلى إلى أسفل، ثم تقسيمهم إلى مجموعتين، المجموعة العالية والمجموعة المنخفضة وفقاً للدرجات التي تم حصولوا عليهم
٢. مجموعات صغيرة

في المجموعة الصغيرة، تم تقسيم جميع المتقدمين للاختبار إلى مجموعتين كبيرتين، وهما ٥٠٪ من مجموعة الدرجات المرتفعة و ٥٠٪ من مجموعة الدرجات المنخفضة.

٣. مجموعات كبيرة

للمجموعات الكبيرة، يكفي حساب ٢٧٪ للدرجات المرتفعة و ٢٧٪ للدرجات المنخفضة. يتم ذلك لتوفير التكاليف والوقت في التحليل.

٤. يقارن نسبة الإجابات الصحيحة لكلا المجموعتين

لحساب فهرس التمييز يمكن استخدام الصيغة:^{٣٨}

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B} = P_A - P_B$$

البيانات:

D	:	عدد المتقدمين للاختبار
J _A	:	عدد المشاركين في المجموعة المرتفعة
J _B	:	عدد المشاركين في المجموعة المنخفضة
B _A	:	عدد طلاب المجموعة المرتفعة الذين أجابوا بصحيح
B _B	:	عدد طلاب المجموعة المنخفضة الذين أجابوا بصحيح
$P_A(\frac{B_A}{J_A})$	:	نسبة المشاركين في المجموعة المرتفعة الذين أجابوا بصحيح
$P_B(\frac{B_B}{J_B})$	:	نسبة المشاركين في المجموعة المنخفضة الذين أجابوا بصحيح

(P كفهرس صعوبة)

نسبة المشاركين في المجموعة المنخفضة الذين أجابوا بصحيح

بصحيح

يمكن تصنيف مستوى التمييز للأسئلة على النحو التالي:^{٣٩}

D : ٠,٠٠ - ٠,٢٠ (منخفض)

D : ٠,٢٠ - ٠,٤٠ (متوسط)

D : ٠,٤٠ - ٠,٧٠ (جيد)

D : ٠,٧٠ - ١,٠٠ (مرتفع)

D : سلبى (يجب عدم استخدام الأسئلة)

³⁸ Suharsimi Arikunto, *DASAR-DASAR EVALUASI PENDIDIKAN*, (Jakarta : Bumi Aksara, 2018), hlm 45

³⁹ Suharsimi Arikunto, *DASAR-DASAR EVALUASI PENDIDIKAN*, (Jakarta : Bumi Aksara, 2018), hlm 47

٤. الصعوبة

يعرف سوهارسيمي أريكوننتو مستوى الصعوبة على أنها الأرقام التي تشير إلى صعوبة أو سهولة المشكلة.^{٤٠} الرقم الذي يوضح صعوبة المشكلة يسمى فهرس الصعوبة. يرمز فهرس الصعوبة ب P الذي يرمز إلى "Proporsi". السؤال الجيد هو السؤال الذي ليس سهلاً جداً وليس صعباً للغاية. الأسئلة السهلة تسبب انخفاض في دافعية الطلاب، من ناحية أخرى، فإن الأسئلة الصعبة تصيب الطلاب باليأس لأن الأسئلة تفوق قدراتهم.^{٤١} يهدف مستوى الصعوبة إلى تحديد الأسئلة المتحيزة، وتوفير تدخلات لإعادة التعليم، وتوفير تدخلات في المناهج الدراسية، وتجميع الاختبارات التي لها دقة البيانات.

هناك عدة أشياء في الكشف عن مستوى الصعوبة، وهي:^{٤٢}

١. لماذا الأسئلة سهلة؟

أ. أداة مشتت بنود الأسئلة لا تعمل

ب. الطلاب يفهمون بسهولة

٢. لماذا الأسئلة صعبة؟

أ. مفتاح الإجابة المحتمل خاطيء

ب. لديه أكثر من إجابتين صحيحتين

ج. لم يتم تحقيق الحد الأدنى من كفاءة الطالب

40 Suharsimi Arikunto, *DASAR-DASAR EVALUASI PENDIDIKAN*, (Jakarta : Bumi Aksara, 2018), hlm 232

41 Fina Aunul Kafi dan Zidni Ilma, "Analisis Soal Tematik Bahasa Arab berbasis HOTS pada Kelas XII MAN 3 Jember" 4, no. 1 (2022): 1–9.

42 Muhson Ali, "Analisis Butir Soal," *Journal of Chemical Information and Modeling* 53, no. 9 (2019): 1689–99.

ح. المواد المطلوبة لا تتوافق مع شكل السؤال

خ. أسئلة غير واضحة

يصنف أريكوننتو سوهارسيمي مؤشر الصعوبة بالقيم التالية:^{٤٣}

السؤال مع $p = 0,00 - 0,30$ هو سؤال صعب

السؤال مع $p = 0,30 - 0,70$ هو سؤال متوسط

السؤال مع $p = 0,70 - 1,00$ هو سؤال سهل

هناك علاقة عكسية بين درجة صعوبة البند ورقم الفهرس نفسه.

كلما انخفض عدد فهرس صعوبة البند الذي يملكه عنصر ما، زادت درجة صعوبة البند. على العكس من ذلك، كلما ارتفع رقم فهرس الصعوبة الذي يملكه عنصر ما، انخفضت درجة صعوبة البند.

لمعرفة حجم فهرس الصعوبة، يمكن العثور عليه باستخدام الصيغة

التالية^{٤٤}:

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

البيانات:

$P =$ فهرس الصعوبة

$B =$ عدد الطلاب الذين أجابوا على السؤال بصحيح

$JS =$ العدد الإجمالي لطلاب الاختبار

⁴³ Suharsimi Arikunto, *DASAR-DASAR EVALUASI PENDIDIKAN*, (Jakarta : Bumi Aksara, 2018), hlm 50

⁴⁴ Suharsimi Arikunto, *DASAR-DASAR EVALUASI PENDIDIKAN*, (Jakarta : Bumi Aksara, 2018), hlm 50

٣. تعليم اللغة العربية

أ. فهم تعليم اللغة العربية

تعليم اللغة العربية هي الأنشطة التعليمية يقوم به المعلم على النحو الأمثل بحيث يقوم الطلاب الذين يعلمهم اللغة العربية بأنشطة التعلم بشكل جيد، بحيث يفضي إلى تحقيق هدف تعليم اللغة العربية.^{٤٥} الغرض من تعليم اللغة هو إتقان اللغويات ومهارات اللغة العربية مثل المطالعة والمحادثة والإنشاء والنحو والصرف وذلك لاكتساب المهارات اللغوية التي تشمل أربعة جوانب من الكفاءة، وهي مهارة الاستماع ومهارة القراءة ومهارة الكتابة ومهارة الكلام.

ب. مجال اختبار اللغة العربية

في سياق اللغة العربية، ينقسم الاختبار إلى قسمين، وهما:

١. اختبار عناصر اللغة

يتكون اختبار عناصر اللغة من اختبار صوت اللغة واختبار المفردات واختبار القواعد. يتكون اختبار المفردات من اختبار المفردات السلبية، وهو فهم معنى المفردات عند سماع صوت أو قراءة نص. ثم المفردات النشطة وهو اختبار لاستخدام المفردات في الكلام والكتابة. وفي الوقت نفسه، يتم تجميع اختبارات القواعد في اختبارات الفهم واختبارات لاستخدام قواعد نحو وصرف في سياق الجمل أو الفقرات.

٢. اختبار مهارات اللغة العربية

45 Salamah, Rifqi Aulia Rahman, dan M. Elfan Kaukab, "PENILAIAN AUTENTIK DALAM PEMBELAJARAN BAHASA ARAB," *Nivedana : Jurnal Komunikasi & Bahasa 2* (2021).

يتم تصنيف المهارات اللغوية إلى أربعة وهي مهارة الاستماع ومهارة الكلام ومهارة الكتابة ومهارات القراءة. تصنف هذه المهارات إلى مجموعتين، المهارة الأولى هي منتجة تعرف باللغة العربية باسم مهارات التحصيل، أو تسمى أيضا كفاية بيان (القدرة على الشرح الشفهي والكتابي)، أي مهارة الاستماع ومهارة الكام. والمهارة الثانية هي التقبلية، والذي يسمى في المصطلحات العربية مهارات الاستقبالية، أو يسمى أيضا كفاية التباين (القدرة على التقاط وتلقي التفسيرات شفها وكتابيا)، أي مهارة الكتابة ومهارة القراءة.

ترتبط هذه المهارات اللغوية الأربعة ارتباطا وثيقا ببعضها البعض. على سبيل المثال، عند مناقشة مهارة كلام بالطبع، لن ينجح إذا لم يكن متوازنا مع مهارة الاستماع. وبالمثل في مهارة الكتابة، حيث لن يتمكن الشخص من الكتابة بشكل جيد إذا لم يكن مدعوما بالقدرة على القراءة. مهارة الكتابة هي أنشطة تواصلية يقوم بها الكاتب مع القراء بناء على مجموعة من الأسس والمبادئ العامة التي تمثل بشكل أساسي الهدف النهائي لاستخدام اللغة.^{٤٦} تسمح القراءة للشخص بتطوير عقله وتحسين فهمه وأسلوبه في الكتابة.^{٤٧} هذه المهارات اللغوية الأربعة مهمة جدا لإتقانها في تعليم اللغة العربية.

^{٤٦} إبراهيم خليل و إمتنان الصمادي، فن الكتابة والتعبير. (عمان: دار المسيرة،

٢٠٠٨).

^{٤٧} د. عبد اللطيف الصوفي، فن القراءة أهميتها، مستوياتها، مهاراتها، أنواعها. (دمشق:

دار الفكر، ٢٠٠٧).

ب. الدراسات السابقة

الدراسات السابقة ذات الصلة بهذا البحث هي كما يلي:

١. تحليل أسئلة اللغة العربية للامتحان النهائي للمدرسة الجهوية المشتركة (UAMBD) المدرسة الابتدائية سنة دراسية ٢٠١٧-٢٠١٨ تأليف مولانا والسنوسي (٢٠٢١).
- صياغة المشكلة في هذه الدراسة هي: هل الأسئلة العربية التي تم اختبارها في UAMBN تفي بمعايير التقويم؟ لا تدع الطلاب أيعتبرون فاشلين ليس لأنهم لا يتقنون المواد العربية، ولكن لأنه من الصعب فهم الأسئلة المقدمة. الطريقة المستخدمة وصفية نوعية مع تقنيات أخذ العينات الهادفة. مجتمع البحث هم الطلاب المدرسة الابتدائية الذين هم مشاركون في UAMBD من KKM المدرسة الابتدائية، منطقة جيامبيا مدينة بوجور. أدوات البحث الداعمة هي برنامج anates الإصدار ٤.٠.٠.٢. والاستنتاج من ذلك البحث : (١) صدق محتوى سؤال UAMBD مدرسة ابتدائية اللغة العربية في عام ٢٠١٨ تظهر عالية جدا. يظهر ذلك من خلال مستوى مطابقة ١٠٠٪ لجميع بنود الأسئلة مع الشبكة المحددة مسبقا، (٢) تتمتع أسئلة UAMBD مدرسة ابتدائية اللغة العربية في عام ٢٠١٨ بمستوى عال من الثبات، مع درجة اختبار ٠.٦٨ (في الفترة من ٠.٦٠ إلى ٠.٨٠). درجة الصعوبة ٦٣.٢٣٪ (٠.٦٣)، فيمكن

القول أنه يمكن استخدامها كاختبار قياسي، (٣) درجة التمييز مقبولة تصل إلى ٣١.٤٪، ولكنها تحتاج إلى تحسين تصل إلى ٢٥.٧٪، (٤) ١١ سؤالاً باللغة العربية UAMBD مدرسة ابتدائية اللغة العربية في عام ٢٠١٨ لديه توزيع جيد جدا لخيارات الإجابة (المشتتات)، خمسة أسئلة جيدة، ستة أسئلة ليست جيدة، تسعة أسئلة سيئة، وأربعة أسئلة السيئة للغاية.^{٤٨}

٢. تحليل الأسئلة الموضوعية اللغة العربية القائمة على *HOTS* في الفصل الثاني عشر من مدرسة الثانوية الحكومية ٣ جمبير تالف كافي وإلما (٢٠٢٢). تحليل هذه الدراسة أسئلة الاختبار المواضيعي للغة العربية يستند إلى *HOTS* في الفصل الثاني عشر من مدرسة الثانوية الحكومية ٣ جمبير للعام الدراسي ٢٠٢٠/٢٠٢١ في مواد الرياضة والشباب. لرسم لنتائج التعلم الموضوعية، والمواد اللازمة للتخطيط للدرس القادم وقياس جودة التعليم وتوقع التعليم التالي، للتدريب على مهارات التفكير النقدي وقياس قدرات الطلاب وفقا للقاعدة المعرفية لـ *HOTS*. استخدمت هذه الدراسة الطريقة الكمية مع تطبيق *anates* الإصدار ٤.٠.٢. أخذ العينات

48 Deni Maulana dan Anwar Sanusi, "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, no. 1 (2020): 12–24, <https://doi.org/10.15575/jpba.v4i1.8054>.

عن طريق أخذ فصل واحد من كل قسم، وهي IAGA (٣٠)، الدراسات الاجتماعية (٣٥)، الدراسات الطبيعية (٣٦). إذن المجموع هو ١٠١. يقوم الطلاب بالأسئلة من خلال التعلم الإلكتروني الذي تيسره المدرسة. يتم استخراج نتائج الإجابة إلى excel والتي تتم معالجتها بعد ذلك في تطبيق anates. والاستنتاج من حيث الجودة، فإن الفئة كافية بناء على أداة قياس anates التي تظهر ارتباط XY بمقدار ٠.٦٣، وثبات الاختبار ٠.٧٨. ومع ذلك، من حيث الكمية، لا يزال هناك تسعة أسئلة لم تكن HOTS لذا لا يزالون بحاجة إلى الحصول على تصحيحات وتحسينات حتى يمكن اختبارها مرة أخرى في الفترة القادمة.^{٤٩}

٣. الصدق والثبات في تحليل بنود الأسئلة بقلم فريدة ومشاركة (٢٠٢١). يحلل هذه الدراسة أسئلة اختبار النهائي للفصل الدراسي مادة اللغة العربية التي أعدها المعلم في مدرسة الثانوية الإسلام كيموجا التي لم يتم تحليلها من حيث النوعية والكمية، لذلك تم إجراء تحليل لتحديد مستوى الصدق والثبات ومستوى الصعوبة والتمييز في أسئلة اختبار النهائي للفصل الدراسي لمواد اللغة العربية للصف

49 Kafi dan Ilma, "Analisis Soal Tematik Bahasa Arab berbasis HOTS pada Kelas XII MAN 3 Jember."

العاشر في مدرسة الثانوية الإسلام كيموجا للعام الدراسي ٢٠٢٠/٢٠١٩. استخدم البحث الطريقة الكمية ذات المنهج الوصفي. جمع البيانات عن طريق التوثيق. الاستنتاج من ذلك البحث، هناك ١١ سؤالاً صحيح و ٢٤ سؤالاً غير صحيح. تظهر ثبات السؤال فئة منخفضة مع معامل ارتباط $r = 0.669$. كانت ١٠ أسئلة صعبة (٢٨٪)، ٢٣ متوسطة (٦٦٪)، ٢ سهلة (٦٪). التمييز غير الجيد هي ١٩ سؤالاً (٥٤٪)، جيد جداً ٣ أسئلة (٩٪)، جيد ١٣ سؤالاً (٣٧٪).^{٥٠}

٤. تحليل أسئلة امتحان مادة اللغة العربية للصف ١٢ في المدرسة الثانوية بقلم أميبا وآخرون (٢٠٢٢). المشكلة في هذه الدراسة هي عدد درجات الطلاب تحت *KKM*. يقوم المعلم فقط بتصحيح الصحيح والخطأ دون أي تحليل للأسئلة التي أعدها المعلم أولاً حتى لا تعرف الأسئلة ما إذا كانت الأسئلة جيدة أم لا. باستخدام الطريقة الوصفية الكمية، يكون المجتمع في شكل أسئلة الامتحان النهائي باللغة العربية الصف الثاني عشر في المدرسة الثانوية الحكومية ٣ ماجالنجكا سنة دراسية ٢٠٢٠/٢٠٢١. عينته من الصف ١٢ *MIPA 1* (٣٠ طالباً) و ١٢ *MIPA 2* (٢٧

50 Ana Musyarofah Farida, "Validitas dan Reliabilitas dalam Analisis Butir Soal," *Al-Mu'arrib: Jurnal Pendidikan Bahasa Arab I*, no. 1 (2021): 34-44.

طالباً). تشمل تقنيات جمع البيانات المقابلات (معلمي اللغة العربية المتعلقة بحالة طلاب الصف ١٢) والتوثيق (مفاتيح الإجابة والشبكات). وجد الباحثون أن العلاقة بين الأسئلة والمهارات اللغوية شملت ٢ مهارات لغوية، وهما: القدرة على "القراءة" ٥٢.٥٪ و "الكتابة" ٤٧.٥٪، المدرجة في فئة كافية بنسبة ٥٠٪. هناك ثلاثة نطاقات من الأسئلة التي تتوافق مع الكفايات الأساسية وهي KD "٣.١" ٥٧.٥٪. KD "٣.٢" ٢٠٪ و KD "٣.٣" ٢٢.٥٪ مشمولة في الفئة الجيدة بنسبة ٧٥٪. جودة أسئلة الامتحان من ناحية الصدق والثبات. الأسئلة الصدق هو ٢٨ سؤالاً. وفي الوقت نفسه، أسئلة غير الصحيحة هي ١٢ سؤالاً. الأسئلة التي أعدها المعلم لها ثبات عالية مع الرقم $> ٠,٧١١$ ر الجدول ٢٥٤,٠٠. تتم مراجعة بنود الأسئلة الاختبار من مستوى أعلى من التفكير. أظهرت بنود الاختبار مستويات المجال المعرفي، وحصلت $C1$ "متذكرة" على ٣٧.٥٪، و $C2$ "مفهومة" ٢٠٪، و $C3$ "مطبقة" ٢٢.٥٪، و $C4$ "محللة" ٢٠٪. السؤال الذي يوضح معدل $HOTS$ هو ٢٠٪، بينما السؤال الذي يوضح معدل $LOTS$ هو ٨.٠٪^{٥١}.

⁵¹ Iza Zainal Ambiya, "Analisis Soal Ujian Mata Pelajaran Bahasa Arab Kelas 12 di Madrasah Aliyyah" 11, no. 1 (2022): 70–87.

فيما يلي جدول بالاختلافات والتشابه مع الأبحاث السابقة:

الرقم	العنوان والمؤلف	التشابه	الاختلاف
١	تحليل أسئلة اللغة العربية للامتحان النهائي للمدرسة الجهوية المشتركة (UAMBD) المدرسة الابتدائية - ٢٠١٧	كلاهما يحللان الأسئلة المتعلقة بالصدق والثبات والتمايز ومستوى الصعوبة وأنماط الإجابات	يستخدم مولانا منهج وصفي نوعي مع تقنيات أخذ العينات الهادفة. بينما تستخدم المؤلفة نهجا كميًا مع طرق الوصفي. استخدمت ادوات البحث إصدار تطبيق <i>ANATES</i> ٤.٠.٠٩ بينما استخدم المؤلف الإصدار ٢٥ من تطبيق <i>spss</i> . أدى تحليل الأسئلة إلى الامتحان النهائي للمدرسة الجهوية المشتركة (UAMBD) على مستوى المدرسة الابتدائية في ٢٠١٧/٢٠١٨ في منطقة بوغور. وفي الوقت نفسه، أجرت الباحثة بحثًا حول اختبار الفصل الدراسي النهائي الأول في الصف

الثامن بمدرسة نهضة العلماء مفتاح الفلاح المتوسطة الإسلامية قدس السنة الدراسية ٢٠٢٣/٢٠٢٤ م.			
يستخدم كافي وعلماء إصدار تطبيق الأنايس ٤٠٠.٢، بينما تستخدم المؤلفة الإصدار ٢٥ من تطبيق spss. حللت الدراسة الأسئلة المواضيعية القائمة على HOTS في الفصل الثاني عشر من مدرسة الثانوية الحكومية ٣ جمبير سنة دراسية ٢٠٢٠/٢٠٢١، بينما حللت الباحثة أسئلة اختبار الفصل الدراسي النهائي في الفصل الثامن بمدرسة نهضة العلماء مفتاح الفلاح المتوسطة الإسلامية قدس	كلاهما يستخدم الطريقة الكمية	تحليل الأسئلة الموضوعية اللغة العربية القائمة على HOTS في الفصل الثاني عشر من مدرسة الثانوية الحكومية ٣ جمبير تالف كافي وعلماء (٢٠٢٢).	٢

السنة الدراسية ٢٠٢٣/٢٠٢٤ م.			
٣	الصدق والثبات في تحليل بنود الأسئلة بقلم فريدة ومشاركة (٢٠٢١).	يحلل كلاهما أسئلة اختبار الفصل النهائي. ويستخدم نمجا كيميا	أجرت فريدة ومشاركة بحثا في الفصل العاشر في مدرسة الثانوية الإسلامية كيموجا للعام الدراسي ٢٠١٩/٢٠٢٠. وفي الوقت نفسه، أجرت المؤلفة بحثا حول الفصل الثامن بمدرسة نهمضة العلماء مفتاح الفلاح المتوسطة الإسلامية قدس السنة الدراسية ٢٠٢٣/٢٠٢٤ م. في الدراسة، لا تحلل المشتتات بينما سيقوم المؤلف بإجراء تحليل حول أنماط الإجابات.
٤	تحليل أسئلة امتحان مادة اللغة العربية للفصف ١٢ في	يستخدم كلاهما الطريقة الكمية	أجرى أمبيا وآخرون بحثا حول أسئلة امتحان مادة اللغة العربية للصف ١٢ في المدرسة الثانوية الحكومية ٣

ماجاليينجكا سنة دراسية ٢٠٢٠/٢٠٢١. تجرى المؤلفه بحثا حول أسئلة اختبار الفصل الدراسي النهائي في الصف الثامن بمدرسة نهضة العلماء مفتاح الفلاح المتوسطة الإسلامية قدس السنة الدراسية ٢٠٢٣/٢٠٢٤ م. تقوم أمبيا بتحليل مستوى الصدق والثبات فقط. وفي الوقت النفسه، تقوم المؤلفه بتحليل مستوى الصدق، والثبات والتمييز ومستوى الصعوبة وانماط الإجابات المشتتة للإنتباه. في البحث الذي سيتم تنفيذه، لا تقوم المؤلفه بتحليل أسئلة المتعلقة بمستوى المجال المعرفي كما فعل أمبيا.		المدرسة الثانوية بقلم أمبيا وآخرون (٢٠٢٢).	
--	--	---	--

ج. إطار التفكير

وفقا لسوغيونو (٢٠١٩) نقل عن أديني، فإن إطار التفكير هو نموذج مفاهيمي لكيفية ارتباط النظرية بالجوانب المختلفة التي تم تحديدها.^{٥٢} يحتوي إطار التفكير في البحث على الأفكار الأساسية أو النظريات أو الافتراضات من البحث التي تم تجميعها من الحقائق والملاحظات ومراجعة الأدبيات. يشرح هذا الإطار التفكير العلاقات والمتبادلة بين المتغيرات.

يمكن معرفة مستوى جودة السؤال من خلال تحليل بنود الأسئلة. من خلال تحليل بنود الأسئلة، نحن نعرف الأسئلة التي تحتاج إلى تحسين أو إزالتها أو تخزينها في بنك الأسئلة. من خلال تحليل بنود الأسئلة، من المأمول أن تكون الأسئلة التي سيتم اختبارها في الاختبارات المستقبلية ذات جودة جيدة حتى تتمكن من قياس نتائج التعلم للطلاب بشكل صحيح ودقيق.

يمكن ملاحظة جودة الأسئلة الجيدة من عدة أشياء، بما في ذلك مدى ملائمة المنهج المستخدم، وتلبية متطلبات الصدق والثبات، قوة تمييزية عالية، مستوى الصعوبة، وجود مفتاح إجابة ووظيفة تشتيت جيدة في أسئلة اختيار من متعدد.

تحليل الأسئلة هو نشاط لمراجعة الأسئلة في الاختبار سواء كانت تستوفي المؤهلات كاختبار جودة بحيث يمكن تحديد الأسئلة الجيدة أو

52 Addini Zahra Syahputri, Fay Della Fallenia, dan Ramadani Syafitri, "Kerangka berfikir penelitian kuantitatif," *Tarbiyah: Jurnal Ilmu Pendidikan dan Pengajaran* 2, no. 1 (2023): 160–66.

المتوسطة أو الجيدة جدا أو غير الجيدة أو حتى السيئة جدا. يؤدي تحليل بنود الأسئلة إلى اختبار نهاية الفصل الدراسي الأول في مادة اللغة العربية في الصف الثامن في مدرسة المتوسطة الإسلامية نهضة العلماء مفتاح فلاح قدس السنة الدراسية ٢٠٢٣/٢٠٢٤م. لمعرفة جودة الأسئلة وقدرة الطلاب على الإجابة عن الأسئلة، يتم ذلك من خلال تحليل أوراق الأسئلة وأوراق الإجابة الناتجة عن اختبارات الفصل النهائي الأول في طلاب الصف الثامن فيما يلي مخطط للإطار التفكير:



د. فرضية البحث

الفرض هو إجاب مؤقت على صياغ مشكل البحث، بعد أن طرح الباحثون إطار النظري وهيكل التفكير.^{٥٣} ليست كل الباحثون الكمية تتطلب أو تحتاج فروض البحث. في البحث الكمي الوصفي والاستكشافي ليست هناك تحتاج فروض البحث لأن هذا لبحث يستخدم منهج الكمي الوصفي فإن هذا البحث لا يحتاج ولا يطلب فروض البحث.



53 Sugiyono, *Metode Penelitian Kualitatif, Kuantitatif, dan R&D* (Bandung: ALFABETA, 2014).