الباب الثاني

الأساس النظري

أ. الوصف المفاهيمي

١. جوهر تقويم التعلم

من الناحية اللغوية، يُعرف مصطلح "التقويم" في اللغة العربية بكلمات مثل :التقويم، التقدير، أو التثمين (Munawir, ٢٠٠٧). أما من الناحية الاصطلاحية، فإنّ مفهوم التقويم لم يعد غريبًا في الحياة المعاصرة، لا سيما في المجال التربوي، حيث يُعدّ من المفاهيم الأساسية المرتبطة ارتباطًا وثيقًا بعملية التعليم والتعلّم. وفي الحقيقة، فإنّ نشاط التقويم ليس جديدًا في سلوك الإنسان، بل هو جزء أصيل من طبيعة الإنسان منذ القدم، حين بدأ الإنسان يُعمل فكره وينظر في ما حوله ويحكم على الأشياء والأحداث وفق معاييره وقيمه. في اللغة الإندونيسية، يُعدّ مصطلح "التقويم "مرادفًا لكلمة "التقييم "أو "القياس التقديري (*penilaian*) "، كما أشار إلى ذلك (Sudaryono, ٢٠١٢).

يعود أصل مصطلح "التقويم" إلى الكلمة الإنجليزية"Evaluation" ، ويُقابله في اللغة العربية مصطلح "التقدير"، والذي يعني من الناحية اللغوية "التقويم" أو "التثمين". أما جذره اللغوي، فيرجع إلى كلمة "القيمة" التي تُعرف في العربية ب "القِيمة"، وهي المقابل لكلمة "nilai" في اللغة الإندونيسية. ومن هذا المنطلق، يُفهم "التقدير التربوي" (التقويم في المجال التربوي) بأنه عملية إصدار حكم أو تقدير ضمن نطاق الأنشطة التعليمية، أو تقييم ما يتعلق بعمليات التعليم والتعلّم. وتجدر الإشارة إلى أن نطاق "التقويم" أوسع بكثير من نطاق "القياس" أو "الاختبار"، إذ يشمل جوانب متعددة من العملية التعليمية، في حين يركّز "الاختبار" عادة على عناصر محددة ضمن هذا الإطار الواسع. ويُلاحظ أن كلاً من "التقويم" و"التقدير" يتّسمان بالطبيعة النوعية، إذ يركّزان على الحكم القيمي والتحليلي، بينما يُعد "القياس" ذا طبيعة كميّة، حيث يعتمد على نتائج عددية يتم الحصول عليها من خلال تطبيق اختبارات معيارية أو أدوات قياس محددة (Miladya, ٢٠١٦).

وفقًا لأريكُنتو (٢٠١٠)، فإنّ التقويم هو سلسلة من الأنشطة المنهجية والمنظّمة تهدف إلى قياس وتقدير نتائج تعلّم الطلاب، باستخدام أدوات تقويم مناسبة مثل الاختبارات، أو الملاحظة، أو الاستبيانات.

ويُنفّذ التقويم في مراحل مختلفة من العملية التعليمية، ويشمل ما يلي:

أ) الوظيفة التشخيصية

من خلال استخدام أدوات تقويم صادقة وثابتة، يتمكّن المربّون من تحديد نواحي القصور في تعلّم الطلاب، وتتبع الأسباب الكامنة وراء تلك النواقص. إنّ التقويم الدقيق والمُحكم يُعدّ أساسًا لاتخاذ قرارات تربوية مدروسة تهدف إلى تحسين العملية التعليمية، وضمان أن تكون الاستراتيجيات التدريسية أكثر توجيهًا وفعالية. ونتيجة لذلك، يمكن تحقيق الأهداف التعليمية على نحو أمثل وذي معنى.

ب)الوظيفة التكوينية

يُجرى التقويم التكويني في أثناء سير العملية التعليمية، ويهدف أساسًا إلى المتابعة المستمرة لتقدّم تعلم الطلاب، وتقديم

تغذية راجعة بنّاءة لكلٍّ من المتعلّمين والمعلّمين، وذلك بغرض تحسين جودة التعليم بشكل عام (Sudaryono, ٢٠١٢).

ج) الوظيفة الختامية

يُجرى التقويم الختامي في نهاية الفترة التعليمية، كختام الفصل الدراسي، أو السنة الدراسية، أو المرحلة التعليمية. ويهدف هذا النوع من التقويم أساسًا إلى تحديد مدى نجاح البرنامج التعليمي، وكذلك تقييم مستوى تحصيل الطلاب واستحقاقهم للترقية أو التخرّج.

يؤدّي تقويم التعلّم دورًا جوهريًا، إذ يهدف إلى قياس مدى فهم الطلاب للمحتوى الذي قُدّم إليهم، ومعرفة نسبة المتعلّمين الذين تمكّنوا من تحقيق الدرجات أو الكفاءات المتوقعة (Syukron,٢٠١٥).

ومن جهة أخرى، ترى وجهة نظر مختلفة أنّ أهداف تقويم نتائج التعلّم يمكن تصنيفها إلى فئتين: الأهداف العامة والأهداف الخاصة:

أ) الأهداف العامة

تشمل الأهداف العامة لتقويم نتائج التعلّم ثلاثة أهداف رئيسية، وهي:

١) جمع المعلومات المتعلّقة بتطوّر مستوى الطلاب ونموّهم.

٢) تقويم الأنشطة التعليمية والخبرات التعلّمية التي يقدّمها المعلّم.

٣) تقييم مدى فعالية أساليب التدريس المستخدمة.

ب) الأهداف الخاصة

تشمل الأهداف الخاصة لتقويم نتائج التعلّم النقاط الخمس الآتية:

١) تحديد العوامل التي تسهم في نجاح الطلاب أو إخفاقهم في إتمام البرامج التعليمية، وتحليلها.

٢) تحفيز مشاركة الطلاب الفعّالة وانخراطهم في عملية التعلّم.

٣) تقديم الإرشاد المناسب الذي يراعي حاجات كل متعلّم ومرحلته النمائية وميوله الفردية.

٤) عداد تقارير معلوماتية حول تقدّم الطلاب الأكاديمي، لفائدة أولياء الأمور والمؤسسات التعليمية.

UNIVERSITAS ISLAM NEGERI FATMAWATI SUKARNO BENGKULU

٥) تحسين جودة عملية التعلّم بشكل عام، بما في ذلك استراتيجيات التعلّم لدى الطلاب وأساليب التدريس التي يعتمدها المعلّمون.

٢. أنواع الاختبارات في تقويم التعليم

يُعدّ الاختبار من الأدوات الأساسية المستخدمة في عملية التقويم لقياس مدى تحقيق الطلاب للأهداف التعليمية المقصودة. ولا يقتصر دور الاختبار على كونه وسيلةً لتقويم التحصيل المعرفي للمتعلمين، بل يُعدّ أيضًا أساسًا مهمًّا لاتخاذ القرارات، لا سيّما في ما يتعلّق بتحسين استراتيجيات التدريس وتصنيف المتعلمين بحسب مستويات كفاءتهم.

وفقًا لأريكُنتو (٢٠١٢:٦٧)، يُعرَّف الاختبار بأنه "مجموعة من الأسئلة أو التمارين أو الوسائل الأخرى التي تهدف إلى قياس المهارات أو المعارف أو الذكاء أو القدرات أو المواهب التي يمتلكها فرد أو مجموعة من الأفراد ويتميّز الاختبار بطبيعته المنهجية، إذ يُبنى بطريقة منظّمة ومدروسة، ويهدف إلى تقديم صورة موضوعية عن كفاءة الطالب في مجال معرفي معيّن.

يقوم الخبراء بتصنيف أنواع الاختبارات وفقًا لعدة معايير، من بينها، وقت التطبيق، والغرض من التنفيذ، وصيغة العرض، والجوانب المحددة التي يتم قياسها. وفيما يلي شرحٌ تفصيلي لكلّ فئة من هذه الفئات:

أ) بناء على وقت التطبيق

١) الاختبار التشخيق

يُجرى هذا النوع من الاختبارات قبل بدء العملية التعليمية، ويهدف أساسًا إلى تقييم المعارف السابقة لدى الطلاب، واكتشاف صعوبات التعلّم، وقياس مدى جاهزيتهم للتعلّم. وتُساعد النتائج المستخلصة من الاختبار التشخيصي المعلّمين في تصميم استراتيجيات وأساليب تدريسية تتوافق مع احتياجات الطلاب الخاصة (Sudijono, ٢٠١١: ٢١٩).

٢) الاختبار التكويني

يُطبَّق هذا الاختبار أثناء سير العملية التعليمية، ويتمثّل دوره في تحديد مدى فهم الطلاب للمادة التعليمية الجارية. ويُعدّ الاختبار التكويني مؤقتًا في طبيعته، ويُستخدم كآلية للتغذية الراجعة التي

تساعد المعلّمين على تعديل وتحسين استراتيجياتهم التعليمية (Sudijono, ٢٠١١:٢١٩).

٣) الاختبار الختامي

يُجرى هذا النوع من الاختبارات في نهاية وحدة تعليمية أو فترة دراسية محددة، مثل نهاية الفصل الدراسي أو السنة الأكاديمية. ويهدف إلى تقييم مدى تحقيق الأهداف التعليمية بشكل عام. كما يُستخدم كمرجع أساسي في إصدار الأحكام النهائية بشأن ترقية الطلاب أو تخرّجهم (Sudijono, ٢٠١١:٢١٩).

ب) بناء على الغرض من التطبيق

١) اختبار التحصيل الدراسي

يصمَّم هذا الاختبار لقياس مدى تمكّن الطلاب من المادة التعليمية التي تم تدريسها. وغالبًا ما يكون هذا النوع من الاختبارات مرتبطًا بالمناهج الدراسية ويعكس مباشرة محتوى التعليم داخل الصف. ووفقًا لمرداني (٢٠١٢: ٩٥)، فإن

اختبارات التحصيل تهدف إلى تقييم مدى تحقيق الطلاب للأهداف التعليمية المحددة.

٢) اختبار الاستعداد

يهدف اختبار الاستعداد إلى قياس القدرات الفطرية أو الإمكانات الأساسية لدى الطلاب في مجال معين لم يتلقّوا فيه تدريبًا رسميًا من قبل. وغالبًا ما يُستخدم هذا النوع من الاختبارات في عمليات الانتقاء أو التوجيه الأكاديمي والتخصّص (Mardapi, ٢٠١٢:٩٥).

٣) اختبار القدرة العامة

تُستخدم اختبارات القدرة العامة لتقييم المستوى العقلي العام أو الكفاءة الفكرية الشاملة لدى الطلاب. وعلى عكس اختبارات التحصيل، لا ترتبط هذه الاختبارات بمنهج دراسي محدد، بل تركّز على قياس مهارات التفكير العليا، وقدرات حل المشكلات، والذكاء التحليلي.

٤) اختبار الميول

يُستخدم هذا الاختبار للكشف عن ميول الطلاب واتجاهاتهم نحو مجالات دراسية أو أنشطة معيّنة. وغالبًا ما تُستخدم نتائج هذه الاختبارات في الإرشاد المهني أو في تحديد البرامج الأكاديمية المناسبة للطلاب.

ج) أنواع الاختبارات من حيث طريقة العرض

١) الاختبار الشفهي

الاختبار الشفهي هو صيغة من صيَغ التقييم يتم فيها طرح الأسئلة وتقديم الإجابات شفويًا .ويُستخدم هذا النوع غالبًا في تقييم مهارات التحدث، وفهم الاستماع، والاستجابات الفورية لدى الطلاب. وقد أشار Nurgiyantoro (٢٠١١:٥٥) إلى أن الاختبارات الشفهية تُستخدم على نطاق واسع لتقييم الكفاءة التواصلية، خصوصًا في سياقات تعليم اللغة.

٢) الاختبار التحريري

يُعدّ الاختبار التحريري أكثر صِيَغ التقييم استخدامًا في البيئات التعليمية الرسمية. ووفقًا Nurgiyantoro (٢٠١١: ٥٢)، فإنّ الاختبارات التحريرية تمكّن المعلّمين من تقييم مجموعة واسعة من قدرات الطلاب بشكل فعّال وسريع. وقد تتخذ هذه الصيغة عدة أشكال، منها:

(أ) أسئلة الاختيار من متعدّد :وتتكوّن من فقرة سؤال يليها مجموعة من البدائل، يكون أحدها فقط صحيحًا بينما تُعدّ البقية مشتّتات.

(ب) أسئلة الإكمال أو الربط :تُستخدم لربط قائمتين من المعلومات، أو لاستكمال عبارة ناقصة بكلمة أو عبارة مناسبة.

(ج) الأسئلة المقالية :تتطلّب من الطالب تقديم إجابات موسّعة تشتمل على شرح أو تحليل أو رأي شخصي.

٣) اختبار الأداء

يستخدم اختبار الأداء لتقييم المهارات النفسحركية لدى الطلاب من خلال تنفيذ مهام معيّنة يمكن ملاحظتها بشكل مباشر وفي الوقت الحقيقي.

د) من حيث المجال المعرفي المقاس

١) الاختبار المعرفي

يصمَّم الاختبار المعرفي لقياس قدرات التفكير لدى الطلاب، ابتداءً من استرجاع المعلومات وفهمها، مرورًا بتطبيقها وتحليلها وتقييمها، وصولًا إلى إبداع أفكار أو حلول جديدة. ويتوافق هذا النوع من التقييم مع المجال المعرفي في تصنيف بلوم، الذي يُبرز تسلسلاً هرميًا للمهارات الذهنية اللازمة للتعلّم العميق والتفكير النقدي البنّاء.

٢) الاختبار الوجداني

يهدف الاختبار الوجداني إلى تقييم مواقف الطلاب وقِيَمهم ودوافعهم واستجاباتهم العاطفية تجاه موضوع معيّن

أو مادة دراسية أو نشاط تعليمي محدّد. يوفّر هذا النوع من الاختبارات فهمًا عميقًا للميول الداخلية للمتعلمين، والتي تلعب دورًا أساسيًا في تشكيل سلوكهم ودرجة تفاعلهم داخل العملية التعليمية.

٣) الاختبار النفسحركي

يُستخدم الاختبار النفسحركي لتقييم المهارات الجسدية أو القدرات الحركية لدى الطلاب، والمتعلّقة بالأنشطة التعليمية، مثل: كتابة النصوص بالحروف العربية، تركيب الجمل بشكل سليم، أداء الحوارات الشفهية.

UNIVERSITAS ISLAM NEGERI FATMAWATI SUKARNO BENGKULU

٣. اختبارات بصيغة الاختيار من متعدّد والأسئلة المقالية

تعدّ الاختبارات أدوات قياس منهجيّة البناء، صُمّمت لتقييم خصائص فردٍ أو كائنٍ ما وفقًا لمعايير أو إجراءات محددة مسبقًا .وتُجرى هذه الاختبارات لجمع معلومات أو بيانات من الفرد الخاضع للتقويم، ثمّ يتم تصحيح النتائج وتقويمها بناءً على ذلك.

وفي سياق تقويم التعلُّم، تلعب صيغة فقرات الاختبار دورًا بالغ الأهمية كوسيلة لقياس مدى تحقيق الطلاب للأهداف التعليمية المنشودة. وتُعدّ أسئلة الاختيار من متعدّد والأسئلة المقالية من بين أكثر الصيغ استخدامًا وانتشارًا. لكلٍّ من هذين النوعين خصائص ومزايا وقيود محددة، مما يستدعي اختيارهما بعناية وفقًا للهدف التعليمي المُراد قياسه.

أ) أسئلة الاختيار من متعدّد

يعدّ اختبار الاختيار من متعدّد أحد أشكال التقويم التي تتطلّب من الطالب تقديم إجابة محدّدة لكل فقرة من فقرات الاختبار. وبهذا، فإن نتائج التصحيح تبقى متّسقة بغض النظر عن الشخص الذي يقوم بتصحيح الإجابات. ومن منظور الطالب، فإن هذا النوع من الاختبارات يتطلّب إجابات دقيقة وموحَّدة لكي تُعدّ صحيحة. أمّا من منظور المعلم، فينبغي أن تكون عملية التصحيح موضوعية بالكامل، بحيث يمكن لأي مصحّح، أيًّا كان، أن يصل إلى النتيجة نفسها في تقويم إجابات الطلاب (Modul Evaluasi Pembelajaran PPG, ٢٠٢٣:٢٢).

تتكوّن فقرة الاختيار من متعدّد من جذع السؤال (النص الرئيس) وعددٍ من الخيارات المحتملة للإجابة، والتي تكون عادةً أربعة أو خمسة، من بينها إجابة صحيحة واحدة فقط، بينما تُعدّ الخيارات الأخرى مشتّتات (مُلهيات). وبحسب ما ذكرته Arikunto (٢٠١٢: ١٠٢) فإن سؤال الاختيار من متعدّد هو سؤالٌ يتبعه عددٌ من الخيارات، تكون واحدة منها فقط صحيحة .وتُعتبر هذه الصيغة فعّالة من حيث قدرتها على تغطية محتوى واسع، بالإضافة إلى أنها تُسهّل عملية التصحيح بشكل مباشر. ومع ذلك، فإنه إذا كانت **المشتّتات** ضعيفة الصياغة أو غير فعّالة، فقد تُصبح الفقرة سهلة للغاية، وبالتالي تفشل في التمييز بين الطلاب ذوي القدرات المختلفة. إنّ فعالية فقرة الاختيار من متعدّد تعتمد اعتمادًا كبيرًا على مدى واقعية وقوة المشتّتات، بحيث تكون منطقية وجذّابة للطالب الذي لم يتقن المادة، ولا يُمكن استبعادها بسهولة عن طريق التخمين.

١) مزايا وعيوب أسئلة الاختيار من متعدّد

(أ) مزايا أسئلة الاختيار من متعدّد

(١) تغطية شاملة للمحتوى

تتيح صيغة الاختيار من متعدِّد إدراج عدد كبير من الفقرات، مما يضمن تمثيلاً واسعًا وشاملاً لموضوعات المادة الدراسية.

(٢) التصحيح الموضوعي

يمكن تصحيح هذا النوع من الأسئلة بشكل موثوق ومتّسق، حتى إذا تم تصحيحه من قبل شخص غير واضع الاختبار.

(٣) سهولة التصحيح

يمكن تقييم الإجابات بسرعة دون الحاجة إلى إصدار أحكام شخصية أو تفسير ذاتي، مما يُسهِّل عملية التصحيح.

(٤) ارتفاع صدق الاختبار

UNIVERSITAS ISLAM NEGERI FATMAWATI SUKARNO BENGKULU

بفضل القدرة على تغطية نطاق واسع من المحتوى، تُعدُّ أسئلة الاختيار من متعدّد وسيلة دقيقة لقياس الأهداف التعليمية بشكل شامل.

(٥) رتفاع موثوقية الاختبار

نظرًا لكون التصحيح موضوعيًا ومتسقًا بين مختلف المصحّحين، فإن احتمالية التحيز أو التفاوت في النتائج تقلُّ بدرجة كبيرة.

(ب) عيوب أسئلة الاختيار من متعدّد

(١) تشجيع التخمين

نظرًا لأن المطلوب من الطالب هو اختيار خيار واحد فقط، فقد يميل إلى الاعتماد على التخمين بدلاً من الفهم الحقيقي للمادة.

(٢) إتاحة فرص الغش والتواطؤ

قد يسهل على الطلاب تبادل الإجابات أو تقليدها أثناء أداء الاختبار، مما يُضعف مصداقية النتائج.

(٣) تقييد قدرة الطالب على التعبير عن التفكير

لا تُتيح هذه الصيغة للطالب فرصة شرح أفكاره أو توضيح أسلوب تفكيره، مما يقلل من قدرة المعلم على تقييم مهارات التفكير العليا.

(٤) صعوبة في البناء

يحتاج إعداد اختبار متوازن يتضمن مشتتات (خيارات خاطئة) معقولة إلى وقت وجهد كبيرين، بالإضافة إلى خبرة تربوية عالية.

٢) أنواع اختبارات الاختيار من متعدّد

UNIVERSITAS ISLAM NEGERI FATMAWATI SUKARNO BENGKULU

(أ) اختبار الصواب والخطأ

هو نوع من الاختبارات يُطلب فيه من الطلاب تحديد ما إذا كانت العبارة صحيحة أو خاطئة من خلال وضع دائرة أو علامة ✓ أو ✗ على الحرف (ص) للصواب أو (خ) للخطأ.

(ب) اختبار الاختيار من متعدِّد القياسي

اختبار يُطلب فيه من الطلاب اختيار الإجابة الصحيحة من بين عدة بدائل معطاة، وذلك بوضع دائرة أو علامة على الخيار الذي يعتقدون أنه الأكثر دقة.

(ج) اختبار المطابقة

نمط من الاختبارات يُطلب فيه من الطلاب ربط العبارات أو المصطلحات الموجودة في العمود الأيسر بالتعريفات أو الردود المناسبة الموجودة في العمود الأيمن.

(د) اختبار الإكمال

اختبار يُطلب فيه من الطلاب إكمال الجملة بملء الفراغ بالكلمة أو العبارة المناسبة لتكوين جملة صحيحة نحويًا وسياقيًا.

(ه) اختبار الإجابة القصيرة

نمط يُطلب فيه من الطلاب تقديم إجابات مختصرة على الأسئلة من خلال كتابة المعلومات الناقصة، وغالبًا ما يكون ذلك في نهاية الجملة أو السؤال.

ب) أسئلة المقال

تتكوّن اختبارات المقال من بنود تتطلّب من الطلاب الرد كتابيًا باستخدام كلماتهم الخاصة، مما يسمح بتنوّع الإجابات بين الطلاب (Modul Evaluasi Pembelajaran PPG, ٢٠٢٣:١٩).

ووفقًا Nurgiyantoro (٢٠١١: ٧٥)، تهدف أسئلة المقال إلى تقييم مهارات التفكير العليا مثل التحليل، والتركيب، والتقويم. وقد

تأخذ هذه الأسئلة شكل مقالات مقيّدة الاستجابة (حيث تُحدّد التعليمات أو الحدود بشكل دقيق) أو مقالات ممتدة الاستجابة (حيث يُسمح بالإجابة بحرية دون قيود واضحة).

١) مزايا وعيوب أسئلة المقال

(أ) مزايا أسئلة المقال

(١) تشجع الطلاب على التعبير عن آرائهم، وطرح أفكارهم، وبناء جمل منظمة ومنهجية؛

(٢) سهلة البناء نسبيًا، وإن كان إعداد أسئلة مقالية عالية الجودة يتطلّب دراسة دقيقة؛

(٣) تحفّز الطلاب على الدراسة بجدّ وتقلّل من الميل إلى التخمين؛

(٤) تقلل من احتمالية التعاون أو الغش بين الطلاب أثناء الاختبار.

(ب) عيوب أسئلة المقال

(١) صعوبة في التصحيح، وغالبًا ما تكون عملية التقييم ذات طابع شخصي، مما يجعل من الصعب تفويضها إلى أشخاص آخرين؛

(٢) تغطية محدودة للمحتوى، نظرًا لأن اختبارات المقال تتضمّن عادة عددًا قليلاً من البنود، مما قد يؤدي إلى استبعاد بعض الموضوعات الهامة من التقييم؛

(٣) انخفاض مستوى الصدق (الصلاحية)، لأن نطاق المادة المقيمة ضيق، مما يقلل من دقة الاختبار في قياس الأهداف التعليمية المقصودة؛

(٤) انخفاض مستوى الثبات (الموثوقية)، بسبب التقييم الذاتي من قبل المعلمين، مما قد يؤدي إلى تباين في نتائج التصحيح بين طالب وآخر.

٢) أنواع أسئلة المقال

(أ) أسئلة المقال المفتوح

في هذا النوع من الأسئلة، تكون الإجابات غير متوقعة ولا يمكن الحكم عليها بشكل صارم على أنها صحيحة أو خاطئة، لأنها تعتمد على تفسير الطالب ومنطقه. ونتيجة لذلك، فإن عملية التصحيح تكون عرضة بدرجة كبيرة للتقدير الذاتي.

(ب) سئلة المقال المقيد

تتيح هذه الأسئلة للطلاب تقديم إجابات يمكن تحديد صحتها أو خطئها بشكل عام. ومع ذلك، نظرًا لأن الطلاب يعبرون عن أفكارهم بكلماتهم الخاصة، فإن هناك درجة من الذاتية في التقييم، وإن كانت قابلة للتقليل من خلال وجود معايير تصحيح واضحة.

(ت) أسئلة المقال المنظم

في هذا النوع، تكون كل من الإجابات المتوقعة ومعايير التصحيح محددة إلى حد كبير. وغالبًا ما تكون الأسئلة

مصحوبة بإرشادات أو نقاط رئيسية يجب على الطلاب تناولها، مما يسمح بإجراء تقييم أكثر اتساقًا وموضوعية.

٤. معايير جودة فقرات الاختبار

إنّ الفقرة الاختبارية الجيّدة هي تلك التي تقيس كفاءة الطالب بدقّة ووفقًا للأهداف التعليمية المقصودة. لضمان الصدق والثبات في أداة التقويم، لا بدّ من مراعاة عدة معايير أساسية عند إعداد فقرات الاختبار، وهي:

أ. درجة الصعوبة

تشير درجة الصعوبة إلى مدى قدرة المتعلمين على الإجابة عن الفقرة الاختبارية بشكل صحيح. وتتمثل وظيفتها الأساسية في ضمان تضمين مستويات متنوّعة من الصعوبة داخل أداة التقييم، بحيث يُسهم ذلك في قياس قدرات الطلاب بشكل أدقّ وشامل. نّ توزيع درجات الصعوبة بشكل متوازن يُتيح للمعلّمين رصد الأداء عبر مستويات الإتقان المختلفة، ويضمن أن يكون الاختبار لا سهلاً جدًا ولا صعبًا جدًا، بل مناسبًا لجميع فئات المتعلّمين.

وفقًا لما ذكره Ardapi (٢٠١٧:١١٢)، تُعَدُّ صعوبة الفقرة أحد المؤشرات الأساسية لجودة الفقرات، إذ تعكس قدرة السؤال على التمييز بين الطلاب ذوي المستويات المختلفة من القدرة. وبعبارة أخرى، فإن مستوى الصعوبة المُعايَر بشكل مناسب يُسهم بفعالية في الكشف عن الفروقات في مدى تمكّن الطلاب من المادة العلمية.

ذكر Nitko dan Brookhart (٢٠١١:٣٥٠) أن مستوى صعوبة الفقرة الاختبارية يُحدَّد بنسبة الذين يجيبون عنها بشكل صحيح من بين المختبَرين. ويُعتَبَر تحليل صعوبة الفقرات أمرًا ضروريًّا لتقويم مدى توازن توزيع مستويات الصعوبة في الاختبار، بما يضمن ألّا يكون الاختبار سهلًا للغاية أو صعبًا بصورة مفرطة بوجه عام.

وبالمعنى نفسه، أوضحت Arikunto (٢٠١٨:٢٢٣) أن صعوبة الفقرة تشير إلى نسبة الذين يجيبون عن سؤال معين إجابة صحيحة مقارنةً بعدد المشاركين الكلّي. فكلما زاد عدد الطلاب الذين أجابوا عنه بشكل صحيح، اعتُبِرت الفقرة أسهل؛ وعلى العكس، كلما قلّ عدد الذين أجابوا عنه بشكل صحيح، اعتُبِرت الفقرة أكثر صعوبة.

ويمكن تحليل درجة صعوبة الفقرة الاختبارية تحليلًا كيفيًا من خلال إجابات الطلاب، وملاحظات المعلم، وبيانات المقابلات. وتُعدّ الفقرة: سهلة إذا أجاب عليها معظم الطلاب بشكل صحيح, ومتوسطة إذا تباينت الإجابات بشكل ملحوظ بين الطلاب, وصعبة إذا عجز أغلب الطلاب عن تقديم الإجابة الصحيحة. تُسهم هذه التصنيفات في ضمان أن يُشكّل الاختبار تحدّيًا مناسبًا للطلاب باختلاف مستوياتهم، مع الحفاظ على العدالة والدقّة التشخيصية في التقييم.

ب.قوة التمييز

تشير قوة التمييز إلى مدى قدرة الفقرة الاختبارية على التفرقة بين الطلاب ذوي الأداء العالي والطلاب ذوي الأداء المنخفض. فالفقرة التي تتمتع بدرجة تمييز عالية تكون أكثر فاعلية في الكشف عن الطلاب الذين أتقنوا المحتوى فعلًا، مقابل الذين لم يتقنوه. عدّ مثل هذه الفقرات ضرورية لضمان أن تعكس نتائج التقييم الفروقات الحقيقية في مستوى كفاءة الطلاب، مما يُسهم في تحقيق عدالة التقييم ودقّته.

وفقًا لما ذكره Mardani (٢٠١٧:١١٦)، تُعدُّ القدرة التمييزية مؤشرًا يعكس مدى قدرة الفقرة الاختبارية على التمييز بين الطلاب ذوي التمكن العالي والمنخفض من المادة التي يُختبرون فيها. وبعبارة أخرى، يُساعد هذا المؤشر في تحديد ما إذا كانت الفقرة قادرة على التمييز بدقة بين قدرات الطلاب.

يشرح Sudijono (٢٠١٧:١٣٤) أن مؤشر التمييز للفقرة الاختبارية يُمثّل الفرق في أداء الطلاب بين المجموعة العليا والمجموعة السفلى عند الإجابة عن سؤال معين. وهذا يُبرز دور القدرة التمييزية في تحديد مدى فعالية الفقرة في التمييز بين الطلاب ذوي مستويات القدرة المختلفة.

يضيف Arikunto (٢٠١٨:٢٢٦) أن القدرة التمييزية هي قدرة الفقرة الاختبارية على التمييز بين الطلاب ذوي القدرة العالية والطلاب ذوي القدرة المنخفضة. والفقرة الجيدة هي التي تستطيع أن تميّز بين هاتين المجموعتين بوضوح وبشكل ملموس.

ويمكن تقييم قوة التمييز من خلال ملاحظة أداء الطلاب ذوي التحصيل العالي والمنخفض، ومقارنة إجاباتهم على نفس الفقرة. فإذا أجاب الطلاب المتفوّقون على السؤال بشكل صحيح، في حين أخطأ فيه الطلاب ذوو الأداء الضعيف، فإنّ الفقرة تُعدّ ذات قدرة تمييز عالية. وتُعدّ هذه الفقرات ذات قيمة تشخيصية عالية، حيث تُساعد على تمييز مستويات الإتقان بين المتعلّمين بشكل دقيق.

ج. فعالية المشتّتات

تشير فعالية المشتّتات إلى مدى قدرة البدائل الخاطئة في أسئلة الاختيار من متعدّد على جذب انتباه الطلاب الذين لا يعرفون الإجابة الصحيحة. فالمشتّت الفعّال هو الذي يبدو منطقيًّا ومعقولًا بدرجة كافية ليختاره الطالب الذي لا يمتلك فهمًا عميقًا للمحتوى, بينما يميل الطالب المتمكّن إلى اختيار الإجابة الصحيحة. عدّ المشتّتات المصاغة بعناية من تالعناصر الأساسية للحفاظ على صدق التقييم ودقّته التشخيصية في اختبارات الاختيار من متعدّد.

وفقًا لما ذكره Mardapi (٢٠١٧: ١٢٤) فإنّ فعالية المشتّت تكمن في قدرته على إرباك الطلاب الذين لم يفهموا المادة بشكل كامل، مما يدفعهم إلى اختيار البديل الخاطئ. ويُعدّ المشتّت وظيفيًّا إذا نجح في جذب إجابات عدد من الطلاب ذوي التحصيل المنخفض، حيث يدلّ ذلك على أنّه مُضلِّل منطقيًّا ويؤدي وظيفته التشخيصية ضمن الفقرة الاختبارية.

وأوضح Sudijono (٢٠١٧: ١٣٦) أن المشتّت يُعدّ فعّالًا إذا اختاره ما لا يقل عن ٥٪ من مجموع المُجتازين للاختبار. أمّا إذا لم يختره أي طالب، فيُعدّ غير وظيفي ويجب تعديله أو استبداله. يُعتبر هذا الأمر بالغ الأهمية، لأنّ المشتّتات غير الفعّالة تقلّل من مستوى صعوبة الفقرة، مما قد يُضعف من صدق أداة التقييم وكفاءتها التشخيصية.

ذكرَت Arikunto (٢٠١٨: ٢٣٠) أنّه في أسئلة الاختيار من متعدّد، يجب أن يكون المشتّت الفعّال جذّابًا للطلاب الذين لا يعرفون الإجابة الصحيحة. فإذا لم ينجح المشتّت في جذب هؤلاء الطلاب، قد ينجؤون إلى التخمين من بين الخيارات المتبقية، مما يؤدي إلى تشويه مستوى الصعوبة المقصود للفقرة.

ويُعتبر المشتّت فعّالًا إذا نجح في جذب انتباه الطلاب وخاصة الذين لم يتقنوا المادة بعد مما يدفعهم إلى اختيار البديل الخاطئ. وفي إطار التحليل النوعي، يمكن للمعلمين أو الباحثين ملاحظة اختيارات الطلاب للإجابات وإجراء مقابلات لفهم الأسباب الكامنة وراء اختياراتهم، مما يُسهم في تقييم مدى فعالية المشتّتات بدقّة.

٥. تصنيف بلوم (Taksonomi Bloom).

عند بناء أسئلة الامتحان، من الضروري مراعاة المستويات المعرفية التي يتم تقييمها. ويُعتبر تصنيف بلوم أحد الأطر المُعترف بها على نطاق واسع لهذا الغرض، حيث يُصنّف الأهداف التعليمية بشكل منهجي وفقًا لدرجة تعقيد العمليات الذهنية المتضمّنة. وقد تم تطوير هذا التصنيف في الأصل من قبل Benjamin Bloom، ثم تم تعديله لاحقًا من قبل Anderson & Krathwohl (٢٠٠١)، حيث حددا ستة مستويات هرمية للتفكير تتراوح من الأبسط إلى الأكثر تقدّمًا، وهي من المستوى C١ إلى المستوى C٦. ويُعدّ هذا النهج المنظم مرجعًا أساسيًا لضمان أن فقرات التقييم لا تتوافق فقط مع الأهداف التعليمية، بل تعكس أيضًا بدقة عمق الفهم

المطلوب من الطالب.

(١) C١(التذكر–Remembering)

يركّز هذا المستوى على قدرة الطلاب في استرجاع المعلومات التي تم تعلمها سابقًا. وتُقيِّم فقرات الاختبار في هذا المستوى قدرة المتعلّم على التعرّف على الحقائق الأساسية أو المفردات أو المفاهيم التمهيدية، أو سردها أو تحديدها، دون الحاجة إلى فهم عميق لها. ومن الأمثلة المناسبة على هذا النوع من الأسئلة: "*اذكر معنى كلمة (مكتبة) باللغة الإندونيسية*"، أو "اكتب حروف الهجاء مع حركات الفتحة". ووفقًا (Anderson & Krathwohl, ٢٠٠١)، فإن عملية التذكر تشمل كلًا من التعرف على المعلومات واسترجاعها من الذاكرة.

(٢) C٢(الفهم–Understanding)

يتمحور هذا المستوى حول قدرة الطالب على فهم المعلومات المُتلقاة والتعبير عن معناها. وفي هذا المستوى المعرفي، تهدف فقرات الاختبار إلى تقييم مهارة المتعلّم في

تفسير المعلومات أو شرحها أو استخلاص النتائج منها باستخدام تعبيراته الخاصة. ومن الأمثلة المناسبة على هذا النوع من الأسئلة: *"اقرأ الجملة الآتية ثم ترجمها إلى اللغة الإندونيسية: ذهب أحمد إلى المدرسة الساعة السابعة صباحًا"*. وذكر (Anderson & Krathwohl, ٢٠٠١) أن الفهم يشمل القدرة على الترجمة، والتلخيص، وتحديد الأفكار الرئيسة الكامنة في المعلومة.

(٣) C٣(التطبيق–Applying)

يركّز هذا المستوى على قدرة الطلاب على استخدام المعلومات التي تم اكتسابها سابقًا في مواقف أو سياقات جديدة. وتتطلب الأسئلة في هذا المستوى من المتعلّمين تطبيق معرفتهم بالقواعد النحوية أو المفردات أو تراكيب الجمل في مهام عملية. ومن الأمثلة على ذلك: "استخدم كلمة (قلم) في جملة مناسبة"، أو *"كوّن جملة اعتمادًا على الصورة الآتية"*. ووفقًا (Anderson & Krathwohl, ٢٠٠١)، فإن التطبيق يشمل القدرة على تنفيذ إجراءات أو ت.

(٤) c٤ (التحليل- Analyzing)

يركّز هذا المستوى على قدرة المتعلّم على تفكيك المعلومات إلى أجزائها المكوّنة، وفهم العلاقات بينها، والتعرّف على الأنماط والهياكل الكامنة فيها. وتُقيِّم الأسئلة في هذا المستوى قدرة الطالب على التمييز، أو التصنيف، أو تحديد الأخطاء داخل الجملة. ومن الأمثلة التوضيحية على ذلك: "*حلّل الخطأ النحوي في الجملة الآتية: الطالبات يذهب إلى المدرسة*". وفي تصنيف بنوم المعدَّل، يؤكّد (Anderson & Krathwohl) أن التحليل يتضمن القدرة على تنظيم وربط المكوّنات المختلفة داخل وحدة معرفية معينة.

(٥) C٥(التقويم–Evaluating)

يركّز هذا المستوى على قدرة الطالب على إصدار الأحكام استنادًا إلى معايير أو مقاييس محددة. وتتطلب فقرات الاختبار في هذا المستوى من المتعلّمين تقييم دقّة المعلومات أو اتساقها أو مدى صحتها، سواء من حيث المعنى أو بنية الجملة. ومن الأمثلة المناسبة على هذا النوع من الأسئلة:

"*قيّم استخدام الفعل في الجملة الآتية وبرّر رأيك*". ووفقًا (Anderson & Krathwohl, ٢٠٠١)، فإن التقويم يشمل الفحص المنهجي والنقدي لفكرة أو منتَج ما، بهدف تحديد جودته العامة أو مدى ترابطه واتساقه.

(٦) C٦ (الإبداع–Creating)

يتمحور هذا المستوى حول قدرة الطالب على إنتاج مخرجات جديدة وأصلية من خلال تركيب العناصر التي تم تعلمها مسبقًا. وفي هذا المستوى المعرفي، يُشجَّع المتعلّمون على إنشاء جمل أو فقرات أو حوارات باستخدام المفردات والتراكيب النحوية التي أتقنوها سابقًا. ومن الأمثلة المعبّرة عن هذا النوع من الأسئلة: "*اكتب فقرة قصيرة باللغة العربية عن أنشطتك يوم الأحد*". ووفقًا لتصنيف بلوم المعدَّل، يوضح (Anderson & Krathwohl) أن الإبداع يتضمن عمليات التصميم، والتنظيم، والبناء لشيء جديد وذي معنى. تم تطوير تصنيف بلوم ليُعبّر عن تسلسل هرمي لمستويات القدرات المعرفية لدى الطلاب، حيث يتدرج من المستويات

الأساسية إلى الأكثر تعقيدًا. يمثل كل مستوى من المستويات من (C١)إلى (C٦) عمقًا مميزًا في العمليات المعرفية، مما يوضح التطور التدريجي في مهارات التفكير لدى الطلاب، بدءًا من القدرة على استرجاع المعلومات(C١) ، وفهم المفاهيم(C٢) ، وتطبيق المعرفة (C٣)، وتحليلها(C٤) ، وتركيبها(C٥) ، وفي النهاية تقييمها .(C٦) ولكي يكون بناء بنود الاختبار فعالًا في قياس كفاءات الطلاب، يجب أن يشمل التوزيع المتوازن لهذه المستويات المعرفية، مما يضمن تقييماً شاملاً لمخرجات التعلم (Anderson & Krathwohl، ٢٠٠١).

بالفعل، يجب أن تُصمم أسئلة الامتحان المتوافقة مع هذا التصنيف بعناية فائقة لتكشف بدقة عن القدرات المعرفية للطلاب في كل مستوى من المستويات المذكورة. فعلى سبيل المثال، في مستوى (C١)ينبغي أن تطلب الأسئلة من الطلاب استدعاء المفردات أو التعريفات؛ أما في مستوى (C٢) فتتطلب إظهار الفهم لمعاني الجمل أو المحادثات؛ وفي مستوى (C٣) يجب أن يطبق الطلاب القواعد النحوية بشكل مناسب داخل سياق الجمل. ومع تقدم المستويات، من (C٤) إلى(C٦) ، يجب أن تتحدى الأسئلة الطلاب لتحليل

الأخطاء النحوية داخل النصوص، وتركيب فقرات متماسكة منطقياً، وفي النهاية تقييم الحجج في نقاشٍ باللغة العربية.

في سياق تعليم اللغة العربية، يُعدُّ استخدام تصنيف بلوم إطاراً فعّالاً لتصميم أسئلة امتحانات نهاية الفصل، بهدف التقييم الشامل لمجمل قدرات الطلاب. فلا تقتصر بنود الاختبار على قياس الفهم النظري، كمعرفة المفردات (المفردات) وقواعد اللغة (القواعد)، بل تمتد أيضاً لتقويم المهارات العملية لدى الطلاب في مهارات التحدث والكتابة والتفكير النقدي باللغة العربية. وبذلك، فإن الأسئلة التي تغطي المستويات من (C١) إلى (C٦) تتيح تصويراً أكثر دقة لمدى تمكُّن الطلاب من المادة العلمية، شاملاً المعرفة الأساسية وكذلك الكفاءة اللغوية التطبيقية.

ب. الدراسات السابقة

من الأساس، ينبغي على أي بحث علمي أن يأخذ بعين الاعتبار الدراسات السابقة ذات الصلة، وأن يستفيد منها كمراجع علمية تدعم توجهه. فالدراسات السابقة تُعدُّ أساساً مقارنياً، ودعماً نظرياً، وإطاراً توجيهياً في صياغة مسار البحث ومنهجيته. وفيما يلي عرضٌ لعدد من

الدراسات السابقة التي تشترك في بعض الجوانب أو تُعدُّ ذات صلة بهذا البحث:

١) أجرت Nurul Aini (٢٠١٧) دراسة بعنوان "*Analisis Soal Tes Pilihan Ganda Mata Pelajaran Bahasa Arab di MAN Purbalingga*"، وهدفت دراستها إلى تحليل جودة أسئلة الاختيار المتعدد لمادة اللغة العربية للصف الحادي عشر في الفصل الدراسي الفردي من السنة الدراسية ٢٠١٤/٢٠١٥م. استخدمت الباحثة المنهج النوعي، معتمدةً على المقابلات والوثائق كأدوات لجمع البيانات، وتم تحليل المعطيات باستخدام الأسلوب الوصفي النوعي. وقد أظهرت نتائج الدراسة أن جودة الأسئلة لم تكن مثالية بعد، ويُعتقد أن ذلك يرجع إلى انخفاض مستوى كفاءة الطلاب في اللغة العربية، مما جعل حتى الأسئلة السهلة نسبيًا تُعدُّ صعبة في نظرهم.

٢) أجرت Sagita Virginia dkk (٢٠٢٠) دراسة بعنوان "*Analisis Butir Soal Ulangan Harian IPA Terpadu Kelas VII SMPN ٠٥ Kota Bengkulu*"، هدف هذا البحث إلى تحليل الأسئلة بناءً على المجال المعرفي في تصنيف بلوم ومستوى الصعوبة. وقد استُخدم في هذه

الدراسة منهج وصفي نوعي. أظهرت النتائج أن الأسئلة تهيمن عليها المستويات C١ إلى C٣ بنسبة ١٠٠%، مما يعكس تركيزًا على المهارات المعرفية الدنيا. أما من حيث مستوى الصعوبة، فقد تبيّن أن ٤٤% من الأسئلة تُعد سهلة، و٣٦% متوسطة، و٢٠% صعبة. وبناءً وبناءً على هذه النتائج، يمكن تقييم جودة الأسئلة بأنها جيدة إلى حدٍّ ما، مع توصية بمزيد من التنوع في المستويات المعرفية لتعزيز التفكير النقدي لدى الطلاب.

٣) أجرى Sonia Meliasari dkk (٢٠٢٢) دراسة بعنوان "*Analisis Soal Ujian Akhir Semester Mata Pelajaran Biologi di SMA Negeri ٦ Seluma*"، هدف هذا البحث إلى تقييم جودة أسئلة اختبار نهاية الفصل الدراسي في مادة الأحياء للصف العاشر، وذلك باستخدام المنهج الوصفي النوعي. وقد أظهرت نتائج الدراسة أن معظم الأسئلة تندرج ضمن المستويين C١ (التذكر) و C٢(الفهم) من تصنيف بلوم، كما أن ما نسبته ٨٥% من الأسئلة تُعد سهلة من حيث مستوى الصعوبة. ومع ذلك، فإن قوة التمييز وفعالية البدائل الخاطئة (المُموِّهات) في الغالب لا تزال غير كافية. وبناءً على ذلك، فإن جودة الأسئلة تُعد

مقبولة بشكل عام، إلا أن هناك حاجة إلى تحسينات، لا سيما في جانب قوة التمييز وتوزيع المستويات المعرفية لتحقيق توازن أفضل في تقييم القدرات الفكرية للطلبة.

٤) . قامت Annisa Rakhmaningrum (٢٠٢٢) بدراسة بعنوان "*Analisis Butir Soal Penilaian Tengah Semester Mata Pelajaran Bahasa Arab Wajib Kelas XII Keagamaan Madrasah Aliyah Negeri ١ Banyumas Tahun Ajaran* ٢٠٢١/٢٠٢٢"، وهدفت الدراسة إلى وصف جودة فقرات الاختبار بناءً على مستوى الصعوبة، وقوة التمييز، وفعالية البدائل. اعتمدت الدراسة على المنهج الوصفي الكمي، واستخدمت التوثيق والمقابلات كوسائل لجمع البيانات. أظهرت النتائج أن ٥٥٪ من الفقرات صُنّفت سهلة، و٤٥٪ متوسطة، ولم توجد فقرات صعبة. أما بالنسبة لقوة التمييز، فقد صُنّفت ٦٠٪ من الفقرات ممتازة، و٥٪ بحاجة إلى مراجعة، و١٠٪ مرفوضة. وفيما يتعلق بفعالية البدائل، فقد حصلت ٣٠٪ من الفقرات على تقييم جيد جداً، و٤٠٪ جيد، و٢٠٪ أقل فعالية.

٥) Siti Ardianti Rukmana (٢٠٢١)، في بحثها المعنون "*Analisis Soal Latihan Buku Pendidikan Agama Islam dan Budi Pekerti Kurikulum*

"٢٠١٣ *kelas VIII Berdasarkan Taksonomi Bloom*، يهدف هذا البحث إلى وصف أبعاد العمليات المعرفية في أسئلة التمارين الواردة في كتاب "التربية الإسلامية" للصف الثامن وفقًا لمنهاج ٢٠١٣، وذلك بالاستناد بالاستناد إلى تصنيف بنوم للأهداف التربوية، كما يهدف إلى تحديد تصنيف هذه الأسئلة حسب المستويات المعرفية الستة من C١ إلى C٦. وقد أظهرت نتائج التحليل أن أسئلة التمارين شملت أربع فقرات في مستوى C١ (التذكر)، وسبع فقرات في مستوى C٢ (الفهم)، وفقرة واحدة فقط في مستوى C٣ (التطبيق)، وثلاث فقرات في مستوى C٤ (التحليل)، بينما لم تُدرج أي أسئلة تقيس المستويين C٥ (التقويم) و C٦(الابتكار). وتشير هذه النتائج إلى وجود فجوة في توزيع الأسئلة من حيث المستويات المعرفية، إذ إن الغالبية العظمى من الأسئلة تركز على قياس مهارات التفكير الدنيا(LOTS) ، مما يدل على غياب التوازن في تنمية مهارات التفكير العليا، الأمر الذي يتطلب إعادة النظر في بناء أسئلة التقييم لتشمل جميع أبعاد التفكير وفقاً لمتطلبات القرن الحادي والعشرين.

استنادًا إلى مراجعة العديد من الدراسات السابقة، تبيّن وجود أوجه تشابه واختلاف بينها وبين البحث الحالي. فمن حيث أوجه التشابه، تناولت الدراسات السابقة تحليل فقرات الاختبار، ولا سيّما تلك المستخدمة في امتحانات نهاية الفصل الدراسي أو غيرها من أشكال التقييم الختامي في مادة اللغة العربية. وقد ركّزت هذه الدراسات أساسًا على تقييم جودة فقرات الاختبار من حيث الصلاحية (الصدق)، والموثوقية، ومستوى الصعوبة، وقوة التمييز، وفعالية البدائل الخادعة (المُشتّتات).

ومع ذلك، توجد عدة فروق بارزة بين هذا البحث والدراسات السابقة المذكورة. فقد اعتمدت معظم الدراسات السابقة على المنهج الكمي مستخدمة التحليل الإحصائي والبرامج المتخصصة مثل برنامج (ANATES)، بينما يتبنى هذا البحث المنهج الوصفي النوعي، معتمدًا على المقابلات والملاحظة والوثائق كأدوات لجمع البيانات. إضافة إلى ذلك، لا يقتصر هذا البحث على تحليل الجوانب الفنية لفقرات الاختبار فحسب، بل يتناول أيضًا الإجراءات المتبعة في بناء الاختبار وعلاقتها في تحسين نتائج تعلم الطلاب في تعليم اللغة العربية.

وعليه، فعلى الرغم من أن هذا البحث يشترك في الموضوع مع عدد من الدراسات السابقة، إلا أنه يختلف عنها اختلافًا جوهريًا من حيث النهج والمنهجية وتركيز التحليل. وتُعدّ هذه الفروقات إسهامًا مميزًا، لا سيّما في السياق المحلي لمدرسة الثانوية فنجاسيلا بنجكولو.

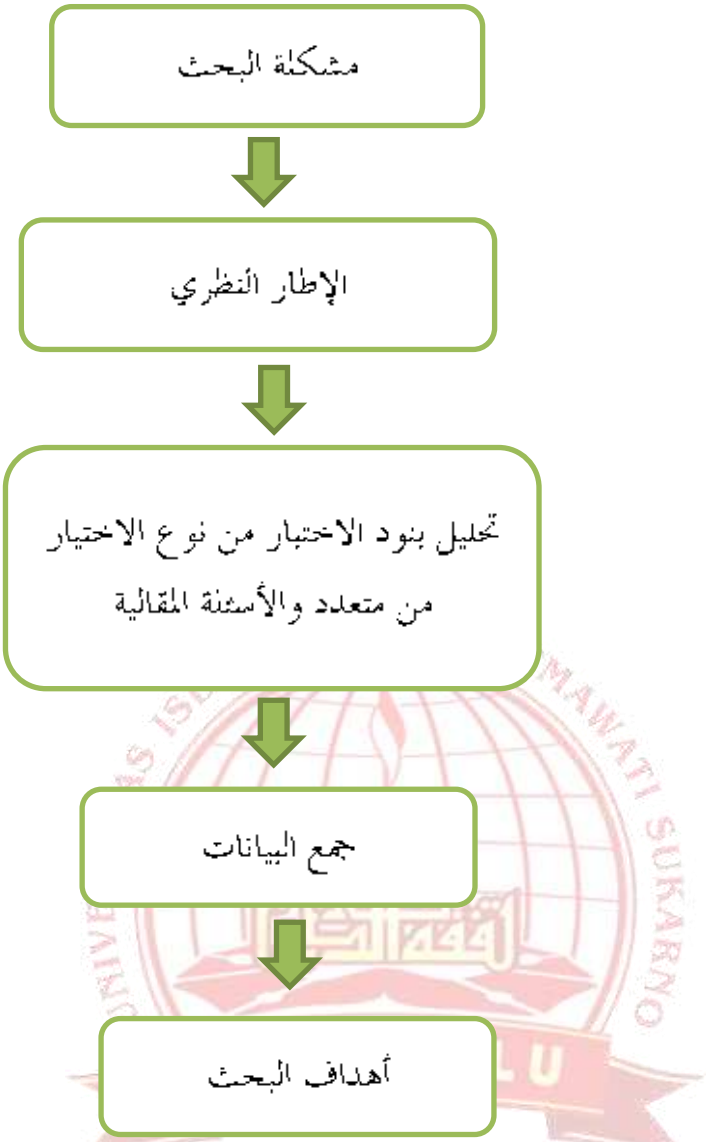
ج. الإطار النظري

الإطار النظري يُمثّل البنية المنطقية التي يتم بناؤها استنادًا إلى النظريات القائمة والنتائج الميدانية، بهدف تفسير العلاقات بين المفاهيم التي يتناولها البحث. وفي البحوث النوعية، يلعب الإطار النظري دورًا حيويًا باعتباره الأساس الذي يُوجّه الباحث في صياغة بؤرة البحث وأهدافه، كما يُسهم في توضيح السياق والمنهجية المُعتمدة في الدراسة.

يُعدّ الإطار النظري تركيبًا تكامليًا للنظريات ذات الصلة والقضايا التي تم تحديدها في الميدان، إذ يُشكّل الأساس المفاهيمي للبحث، ويُستخدم كمرجعية أساسية في تحليل البيانات وتفسيرها.

في هذه الدراسة، يُبنى الإطار النظري استنادًا إلى مفاهيم جودة بنود الاختبار، وإجراءات إعداد الاختبار، وتطبيقها في عملية تقييم التعلم. تهدف هذه الدراسة، التي تحمل عنوان "تحليل بنود أسئلة إمتحان نهاية الفصل الدراسي لمواد اللغة العربية بالمدرسة الثانوية فنجاسيلا بنجكولو"، إلى فحص جودة بنود الاختبار وفقًا للمعايير الفنية وكذلك الإجراءات التي يتبعها المعنمون في إعدادها.

تحليل البنود هو عملية منهجية تُستخدم لتقييم جودة بنود الاختبار، وتشمل جوانب مثل الصلاحية، والثبات، ومستوى الصعوبة، وقوة التمييز، وفعالية المُلهيات (أو المُربكات). يمكن الاستفادة من نتائج هذا التحليل في تحسين بنود الاختبار، لضمان أن تكون أكثر فعالية ومتوافقة مع أهداف تعلم اللغة العربية .استنادًا إلى ما سبق من شرح، يمكن بناء الإطار النظري كما يلي:


مشكلة البحث
الإطار النظري
تحليل بنود الاختبار من نوع الاختيار من متعدد والأسئلة المقالية
جمع البيانات
أهداف البحث