الجمعية الأردنية للعلوم التربوية، المجلة التربوية الأردنية، المجلد العاشر، العدد الثاني، ملحق (1)، 2025
DOI: https://doi.org/10.46515/jaes.v10i2.1506

Constructing a Criterion-Referenced Test in General Physics for Tenth Grade Students According to the British Curriculum Using the Three-Parameter Logistic Model

Mohammad Zohdi Mohammad Al-Fagawi*
Prof. Ferial Mohammad Abu Awwad**

Received 14/10/2023 Accepted 25/11/2023

Abstract:

The study aimed to build a reference-based test in general physics for tenth grade (British Curriculum), based on IRT using the three-parameter logistic model. The sample consisted of (700) students, who were selected randomly by method. A test was constructed to achieve the objectives of the study. The number of items in the initial form reached (85) items, The characteristics of the test were verified through various methods of validity and reliability, in addition to the characteristics of the items considering CCT and the IRT. The results related to the characteristics of the test and the items showed that the test has good psychometric characteristics, which indicates its suitability for detecting students' achievement in Physics, as the reliability value by the Carver method was (0.90) and by the Kappa method (0.74). The cut-off score for the test was (22.8), and when the test was matched according to the item response theory models, it was found that (40) items matched the model. The study ended with a set of recommendations.

Keywords: Criterion-Referenced Test, Item Response Theory, General Physics, Three-Parameter Logistic Model.

* Jordan\ mohzohdi25@hotmail.com
** https://orcid.org/0000-0002-0975-7208
School of Educational Sciences\ The University of Jordan\ Jordan\ f.abuawwad@ju.edu.jo

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

بناء اختبار محكي المرجع في الفيزياء العامة لطلبة الصف العاشر وفق المنهاج البريطاني باستخدام الأنموذج اللوجستي ثلاثي المعلمة

محمد زهدي محمد الفقعاوي *

أ.د. فريال محمد أبو عواد **

ملخص:

هدفت الدراسة إلى بناء اختبار محكي المرجع في الفيزياء العامة للصف العاشر وفق المنهاج البريطاني، استنادًا إلى نظرية الاستجابة للفقرة باستخدام الأنموذج اللوجستي ثلاثي المعلمة. تكونت عينة الدراسة من (700) طالب وطالبة، جرى اختيارهم بالطريقة العشوائية العنقودية. ولتحقيق أهداف الدراسة تم بناء الاختبار، وبلغ عدد الفقرات في الصورة الأولية (85) فقرة، وتم التحقق من خصائص الاختبار من خلال طرق متنوعة من الصدق والثبات، فضلًا عن خصائص الفقرات في ضوء كل من النظرية الكلاسيكية، ونظرية الاستجابة للفقرة باستخدام الأنموذج اللوجستي ثلاثي المعلمة، أظهرت النتائج المتعلقة بخصائص الاختبار والفقرات، تمتع الاختبار بخصائص سيكومترية جيدة، مما يدلل على مدى صلاحيتها للكشف عن تحصيل الطلبة في الفيزياء لدى عينة الدراسة، إذ بلغت قيمة الثبات بطريقة كارفر (0.90) وبطريقة كابا (0.74)، كما وتم تحديد درجة القطع للاختبار باستخدام طريقة بوكمارك حيث بلغت قيمتها (22.8) ولدى مطابقة الاختبار وفق نماذج نظرية الاستجابة للفقرة تبين مطابقة (40) فقرة للأنموذج، وبناء عليه استخرجت معالم الأفراد والفقرات المطابقين للأنموذج. وانتهت الدراسة بجملة من التوصيات.

الكلمات المفتاحية: اختبار محكي المرجع، نظرية الاستجابة للفقرة، الفيزياء العامة، النموذج اللوجستي ثلاثي المعلمة.

* الأردن / mohzohdi25@hotmail.com

** كلية العلوم التربوية/ الجامعة الأردنية/ الأردن/ f.abuawwad@ju.edu.jo

المقدمة:

في كل عام، يستعد ما يقرب من مليون طالب وطالبة من طلبة كامبردج Cambridge من 10000 مدرسة في 160 دولة لمستقبلهم ولدخول الجامعات بعد اجتيازهم لاختبارات كامبردج الدولية Cambridge International. إذ تُعد الاختبارات الدولية International General Certificate of Secondary Education (IGCSE) طلبة المدارس للحياة الجامعية، وتمنحهم الشهادة العامة الدولية للتعليم الثانوي والتي هي عبارة عن امتحان يعتمد على اللغة الإنجليزية يشبه شهادة الثانوية العامة، ويُعترف به في المملكة المتحدة بصفته مكافئًا لامتحان شهادة الثانوية العامة لأغراض الاعتراف بالتحصيل السابق؛ تم تطويره بواسطة جامعة كامبردج الدولية، وذلك لكونهم يركزون على امتلاك الطلبة لمهارات القرن الحادي والعشرين التي تساعدهم على تنمية فضول مستنير وشغف دائم بالتعلم. إذ أن المؤهلات الدولية من قبل كامبردج معترف بها من قبل أفضل الجامعات والشركات على مستوى العالم، مما يمنح الطلبة مجموعة واسعة من الخيارات في تعليمهم ومهنتهم المستقبلية. كما أن كامبردج تقدم برامج دولية تم إنشاؤها وتحريرها من قبل خبراء في كل موضوع، متجذرة في الصرامة الأكاديمية ومحتوية لأحدث البحوث التربوية لأنها توفر منصة قوية مدعومة جيدًا بمصادر التعليم والتعلم للمتعلمين للتقدم من مرحلة إلى أخرى وتطويرهم لكي يتمتعوا بالثقة والمسؤولية والتفكير والابتكار والمشاركة والجاهزية للنجاح ومواكبة العالم الحديث (Cambridge Assessment International Education, 2021).

تم إنشاء مناهج كامبردج خصيصًا لطلبة البرامج الدولية لأكثر من 25 عامًا من خلال العمل مع المدارس والمعلمين والمشرفين التربويين في جميع أنحاء العالم، لتطوير مناهج تناسب مختلف البلدان والثقافات وأنواعًا مختلفة من المدارس والمتعلمين الذين لديهم مقدرات مختلفة، فمادة الفيزياء على سبيل المثال، تهدف إلى تمكين المتعلمين من زيادة فهمهم للعالم التكنولوجي، والاهتمام بالمسائل العلمية، والتعرف إلى مزايا المنهج العلمي ومحدداته، وكيفية تطبيق ذلك على التخصصات الأخرى وعلى كثير من أمور الحياة اليومية، كما أنها تمكنهم من تطوير الاتجاهات العلمية ذات الصلة، مثل الاهتمام بالدقة، والموضوعية، والنزاهة، والمبادرة، والابتكار، وطرح الأسئلة، وتنمية الاهتمام بالبيئة والعناية بها، وفهم أفضل للتأثيرات والحدود الموضوعة على الدراسة العلمية من قبل المجتمع والاقتصاد والتكنولوجيا والأخلاق والبيئة، وتطوير فهم المهارات العلمية الأساسية لكل من الدراسة الإضافية والحياة اليومية، من خلال توازن البرامج المقدمة بين

المعرفة والفهم الشامل لموضوع ما، وتطوير المهارات التي يحتاجها المتعلمون لخطواتهم اللاحقة في التعليم (Cambridge IGCSE (9–1) Physics 0625/0972 (2023; Cambridge International Examinations, 2017).

وتسعى برامج كامبريدج الدولية إلى إعداد طلبة المدارس للحياة الجامعية، من خلال مساعدتهم على تنمية فضول مستنير وشغف دائم بالتعلم، إذ تسهم برامج كامبريدج ومؤهلاته في إعداد الطلبة إعدادًا ملائماً للجامعة والتوظيف والحياة. ويمكن للمتعلمين أن يكونوا على ثقة من أن النتائج التي يحصلون عليها يتم منحها بشكل عادل وفقًا للمعايير العالمية المتسقة. وفي كل عام، يحقق الآلاف من متعلمي كامبريدج الدرجات التي يحتاجون إليها للتقدم إلى الجامعة التي اختاروها. كما تقدر الجامعات وأرباب العمل الرائدة في جميع أنحاء العالم مؤهلات كامبريدج وتعترف بها كدليل على الكفاية الأكاديمية. وتفتح الأبواب أمام متعلمي كامبريدج جنبًا إلى جنب مع المدارس، للقيام بتطوير المتعلمين الذين يتمتعون بالثقة والمسؤولية والتفكير والابتكار والمشاركة، والمجهزين للنجاح في العالم الحديث سريع التغير (Cambridge International Examinations, 2017).

ويتمثل غرض التقييم في كامبردج في قياس التعلم والتحصيل. إذ أن التقييم يؤكد التحصيل والأداء فيما يتعلق بالمعرفة والفهم والمهارات المحددة في المنهاج، إلى المستويات الموضحة في توصيف الدرجات من أجل إظهار احتمالية النجاح في المستقبل، كما أن النتائج تساعد في التنبؤ بالطلبة المستعدين جيدًا لمادة دراسية أو مهنة معينة، ومساعدة الطلبة في اختيار المسار أو المهنة الأنسب؛ لذلك يأخذ جميع المتقدمين لاختبار الفيزياء ثلاث أوراق. بحيث يتم إدخال المتقدمين الذين درسوا محتوى الموضوع الأساسي (Core) أو الذين من المتوقع أن يحصلوا على درجة D أو أقل إلى الورقة 1 والورقة 3. أما الورقة 5 أو الورقة 6. فيكون المتقدمون لها مؤهلين لنيل الدرجات من C إلى G. كما يجب إدخال المتقدمين الذين درسوا محتوى الموضوع الموسع Extended (أساسي وملحق)، والذين من المتوقع أن يحصلوا على درجة C أو أعلى، في ورقة 2، وورقة 4 أو ورقة 5 وورقة 6. إذ سيكون هؤلاء المتقدمون مؤهلين للحصول على مدى درجات A * إلى C. أي أنه يجب تعليم جميع المتقدمين محتوى المادة الأساسية Core. بحيث يمكن للمتقدمين الذين يتم تعليمهم محتوى المادة الأساسية فقط تحقيق الحد الأقصى من الدرجة "C". كما يجب تعليم المتقدمين الذين يهدفون للحصول على الدرجات من "A*" إلى "C" محتوى

الموضوع الموسع Extended. بحيث يتضمن محتوى الموضوع الموسع كلًا من المحتوى الأساسي والملحق Supplement (Cambridge International Examinations, 2017). وقد أشارت نتائج درجات المرشحين الذين تقدموا لامتحان الفيزياء في دورة شهر حزيران 2021 على مقياس تتراوح علاماته من *A إلى G أن 38.3% من الطلبة حصلوا على *A و8.9% حصلوا على A، و15.4% حصلوا على B، و14.5% حصلوا على C. عمومًا، لا توجد اختبارات تجريبية كافية يمكن أن تطبق على الطلبة قبل تعرضهم للاختبارات الدولية، فقط يتم تزويد الطلبة باختبار واحد عند تعديل الخطة تظهر فيه التعديلات والإضافات الجديدة.

ولعل الاتجاهات المعاصرة في التقويم تتجه إلى الاختبارات محكية المرجع Criterion Referenced Testالتي يسعى مطور الاختبار من خلالها إلى تعظيم وزيادتها معلومات الاختبار لتقترب من موقع درجة القطع، إذ يُعتقد أن هذا ينتج أعلى دقة تصنيفية للطلبة حسب مستويات الإتقان مقارنة بالاختبارات معيارية المرجع التي تحدد المركز النسبي للفرد بين جماعة مرجعية معينة دون الاهتمام بما يمكن أن يفعله أو يؤديه الفرد (Allam, 2006)، مما يجعلها عرضة للشك والتساؤلات في مقدرتها على تجويد التعليم نظرًا لما تقدمه من تفسير تقليدي حول أداء الطلبة (Ababneh, 2009). وهذا ما أدركه التربويون وأوضحوا أن تلك الاختبارات ربما لا تنجح في تشخيص الصعوبات التحصيلية التي يواجهها الطلبة، ولا تتمكن من قياس مستوى الإتقان بالنسبة لمجال محدد. فالنجاح والتحصيل له معان ليست تربوية والنتائج تتحدد بالنجاح والرسوب فقط، فالفجوة موجودة بين نتيجة الطالب ومستوى الأداء الفعلي له (Airasian & Madaus, 1972). ومع ذلك، فإن الموقع الأمثل للقيمة العظمى لمعلومات الاختبار يعتمد أيضًا على عوامل أخرى، مثل نوع الأنموذج اللوجستي المستخدم وموقع درجة القطع، وتوزيع مقدرة المفحوص، وأنموذج الاختبار المطبق في تقدير السمات الكامنة. ولعل تفسير نقاط القطع في الاختبار محكي المرجع يقوم على ما يعرفه الأفراد أو ما يمكنهم فعله ضمن مستوى معين من الأداء، وهي تصف المستوى المطلق للأداء لدى المفحوص، إذ يوفر الاختبار معلومات عمّا إذا كان المفحوص قد حقق مستوى معينًا من إتقان مجال المعرفة أو المهارات أم لا. ويشتمل الاختبار محكي المرجع على درجة قطع محددة مسبقًا لتصنيف الأفراد إلى متقنين وغير متقنين (Hu, 2018).

يتم في الاختبار محكي المرجع تحديد المجال السلوكي تحديدًا جيدًا حتى يكون للتحقق من

موقع الطالب معنى ملائمًا، فيتم في معظم الحالات التعامل مع المجال السلوكي المقاس بحيث يشير إلى مجموعة من المعارف والمهارات والكفايات المتضمنة في هذا المجال، والتي يتم استخدامها لتقييم درجة معرفة الطالب ومهاراته وكفايته في ذلك المجال، بشكل يخدم عملية التدريس اللاحق، فالأفراد الذين تظهر نتائجهم على الاختبار تحقيقهم لأحد جوانب المجال السلوكي المقاس بالاختبار بدرجة ملائمة فإنهم يتقدمون إلى المجال التدريسي اللاحق، والأفراد الذين يظهرون كفاية أقل يتلقون تدريسًا لاحقا من نوع آخر (Walsh & Betz, 1995). وذكر برين (Brain, 2009) إن الاختبار محكي المرجع هو ذلك الاختبار الذي يتناول موضوعات وكفايات محددة تتم مقارنتها بما يسمى درجة القطع، وأن أكثر أنماط التقويم إسهامًا في التخطيط وبناء العمليات الخاصة بعمليتي التعلم والتعليم الصفي هو التقييم محكي المرجع (Murphy & Davidshfer, 2001; Al-Hamouri, H. & Al-Kahlot, 2009).

عمومًا، لا يوجد اختبار مثالي، فكل الدرجات تحتوي على خطأ في القياس لا علاقة له بما يتم قياسه. ومع ذلك تعد نماذج نظرية الاختبار ضرورية لمساعدة المهتمين على فهم عملية القياس بشكل أفضل وكيف تتأثر بمصادر الخطأ. بالإضافة إلى ذلك، نظرًا لأن القدرات التحصيلية لا يتم ملاحظتها بشكل عام، تبرز الحاجة إلى نماذج نظرية الاختبار لتوضيح العلاقة الموجودة بين درجات الاختبار المشاهدة والأداء المقدر في المجال، أي أن المقدرة الحقيقية للطالب يتم استنتاجها من درجاته في الاختبار. وعادة ما يتم تصميم أدوات التقييم استنادًا إلى نماذج نظرية الاستجابة للفقرة أحادية البعد، وبالتالي فهي تهدف إلى استخدام أنموذج الاستجابة لإنشاء مقياس وتقديم تقديرات للطلبة على طول سلسلة متصلة. إذ توفر نماذج نظرية الاستجابة للفقرة Item Response Theory (IRT) أحادية البعد تقديرات مقدرة الأفراد على سمة كامنة أحادية البعد مستمرة؛ فتقدير مقدرة أعلى يساوي احتمالية أعلى للإجابة على الفقرة بنجاح (Conrad, 2021). وتتمتع نظرية الاستجابة للفقرة بعدة مزايا، فهي لا تعتمد في تقدير مقدرة المفحوصين على عينة الفقرات المقدمة لهم وهو ما يعبر عنه بالتحرر من الفقرات، كما لا تعتمد خصائص الفقرات (معامل الصعوبة والتمييز) على عينة المفحوصين الذين خضعوا للاختبار، وهو ما يعبر عنه بالتحرر من الأشخاص (Hambelton & Swaminathan, 1985).

ونظرًا لأهمية الاختبارات محكية المرجع في العلوم بشكل عام، وفي الفيزياء بشكل خاص، إذ تركز الاختبارات محكية المرجع على توفير معايير الكفايات، وتعليم المادة وتحسين المستوى

لدى الفرد نفسه، وليس التركيز على التقويم معياري المرجع القائم على المعدل الرقمي أو مقارنة الطلبة بعضهم ببعض (Allam, 1995)؛ فقد اهتمت المؤسسات التربوية ببناء مثل هذه الاختبارات، ونظرًا لأن أعداد الطلبة في البرنامج البريطاني كبيرة (Cambridge Assessment International Education, 2021) والتحديث الجديد لخطة مادة الفيزياء من حيث إضافة موضوعات جديدة، وعدم توفر اختبارات تجريبية مجانية تحاكي نمط الاختبار؛ هدفت هذه الدراسة إلى بناء اختبار محكي المرجع في وحدة الفيزياء العامة لطلبة الصف العاشر وفق المنهاج البريطاني باستخدام الأنموذج اللوجستي ثلاثي المَعْلَمَة.

مشكلة الدراسة وأسئلتها:

على الرغم من الاهتمام المتزايد في الأردن بالاختبارات بمختلف أنواعها، والدعم المقدم من قبل مؤسسات التعليم لجميع الجهود العلمية التي تهدف إلى التحسين النوعي للطلبة في جميع المناهج واكتساب مهارات التفكير العليا بشكل عام، والمناهج العلمية بشكل خاص، وترجمة كتب المناهج البريطانية وتكييفها واعتمادها في جميع المراحل الدراسية إلا أن الجهود لا زالت بحاجة إلى توجيه نحو تقديم اختبارات محكية المرجع لقياس مدى تحقق النتاجات التعلمية في الفيزياء للصف العاشر بشكل خاص، إذ إن مجال الاختبارات محكية المرجع يستوعب مزيدًا من الدراسات والأبحاث. لذا تأتي هذه الدراسة لبناء اختبار محكي المرجع في مادة الفيزياء للصف العاشر للمنهاج البريطاني وفق نظرية الاستجابة للفقرة باستخدام النموذج اللوجستي ثلاثي المعلمة.

وبالتحديد، تحاول الدراسة الإجابة عن السؤالين الآتيين:

1. ما تقديرات معالم الفقرات والأفراد للاختبار محكي المرجع في الفيزياء العامة الذي تم بناؤه وفق المنهاج البريطاني للصف العاشر في الأردن باستخدام النموذج اللوجستي ثلاثي المَعْلَمَة؟
2. ما الخصائص السيكومترية للاختبار محكي المرجع في الفيزياء العامة الذي تم بناؤه وفق المنهاج البريطاني للصف العاشر في الأردن باستخدام النموذج اللوجستي ثلاثي المَعْلَمَة؟

أهداف الدراسة

تهدف هذه الدراسة إلى بناء اختبار محكي المرجع في وحدة الفيزياء العامة لطلبة الصف العاشر وفق المنهاج البريطاني واستخراج خصائصه السيكومترية ودقة تقديرات معالم الفقرات والأفراد باستخدام نماذج نظرية الاستجابة للفقرة أحادية البعد.

أهمية الدراسة

تستمد هذه الدراسة أهميتها من أنها تسعى إلى توفير اختبار محكي المرجع في موضوع الفيزياء العامة للصف العاشر وفق المنهاج البريطاني، نظرًا لما لهذه المرحلة من أهمية في قبول الطالب الجامعي واختياره للمسار والتخصص المطلوب، وتأسيسه وتزويده بالمفاهيم والمهارات التي يبنى عليه تعلمه اللاحق، إذ تعد مرحلة حرجة مصيرية بالنسبة للطلبة، إذ أن ما يقارب 70% من معدل القبول الجامعي يعتمد عليها، وفي تعلم الفيزياء بشكل خاص، لما يترتب على دراسة الفيزياء من قبول للدراسة في تخصصات جامعية معينة دون غيرها.

وعلى الصعيد العملي يمكن أن يسهم الاختبار المعد في هذه الدراسة في تنمية التفكير وطرح الأسئلة الاستقصائية والربط مع بقية العلوم من أجل التوسع فيها وامتلاك المقدرة على حل المشكلات، وتطبيق مهارات القرن الحادي والعشرين في مواقف حياتية مختلفة، كما أنها توظف نمط الاختبار محكي المرجع، ونماذج الاستجابة للفقرة، مما يجعلها تستخدم كدليل للباحثين الآخرين، وتمكن التربويين من العمل على مراحل أخرى كالمرحلة الابتدائية والثانوية وفق النمط ذاته مع الأخذ بعين الاعتبار اختلاف الخطط والمحتوى.

مصطلحات الدراسة والتعريفات الإجرائية:

تتضمن الدراسة عددًا من المفاهيم والمصطلحات يمكن تعريفها كالآتي:

- **الاختبارات محكية المرجع:** هي اختبارات يتم فيها تحديد أداء الفرد بالنسبة لمجال معرفي محدد (Sax & Newton, 1997) كما أنها تعرف إجرائيًا بأنها مجموعة الفقرات التي تم بناؤها وفق الخطوات المتعارف عليها لدى علماء القياس والتي تقيس النتاجات المعرفية في موضوع الفيزياء العامة للصف العاشر، وتم فيها تحديد درجة قطع لتصنيف الطلبة إلى متقنين وغير متقنين.
- **الخصائص السيكومترية:** خصائص الاختبارات وصفاتها والتقييمات النفسية التي تحدد جودتها وثباتها وصدقها، وتساعد هذه الخصائص على التأكد من أن القياسات التي يتم الحصول عليها دقيقة وذات معنى. ويقصد بها في هذه الدراسة دلالات الصدق والثبات للاختبار التحصيلي محكي المرجع في وحدة الفيزياء العامة للصف العاشر.
- **معالم الفقرات والأفراد:** قيم معالم الصعوبة في الأنموذج أحادي المعلم، ومعلمتي الصعوبة والتمييز في الأنموذج ثنائي المعلم، ومعالم الصعوبة والتمييز والتخمين في الأنموذج ثلاثي

المعلم، فضلًا عن دالة معلومات الفقرات والاختبار، ومَعْلَمَة المقدرة للأفراد كما تم حسابها في الدراسة الحالية.

حدود الدراسة:

تم تنفيذ هذه الدراسة في العام الدراسي 2022/ 2023 في محافظة عمان في الأردن في المدارس الدولية والتابعة للمعهد البريطاني، إذ تم تطبيق الدراسة على عينة عشوائية من طلبة الصف العاشر في البرنامج البريطاني.

الدراسات السابقة

فيما يأتي عرض لأبرز الدراسات ذات الصلة التي أمكن للباحث الاطلاع عليها:

في دراسة قام بها حمادنة (Hamadneh, 2009) هدفت إلى استخدام نظرية الاستجابة للفقرة في بناء اختبار محكي المرجع في الرياضيات، وفق الأنموذج اللوجستي ثلاثي المعلمة، ولتحقيق هذا الهدف تم بناء اختبار في الرياضيات، يقيس تحصيل الطالب في الإحصاء، مكون من (28) فقرة من نوع الاختيار من أربعة بدائل. تكونت عينة الدراسة من 411 طالبًا وطالبة، من طلبة الصف الأول الثانوي العلمي. أظهرت نتائج الدراسة ما يأتي: تحقق افتراضات نظرية الاستجابة للفقرة في بيانات الدراسة، وكذلك مطابقة الاستجابات عن 24 فقرة من فقرات الاختبار لتوقعات الأنموذج اللوجستي ثلاثية المعلمة، وحذف 4 فقرات لم تطابق الأنموذج. كما بينت النتائج المتعلقة بتقديرات معالم الفقرات (الصعوبة، والتمييز، والتخمين) أنها كانت مقبولة ضمن المحكات التي أوردتها أدبيات القياس التربوي.

وأجرت البلوي (Al-Blewey, 2015) دراسة هدفت إلى بناء اختبارات تشخيصية محكية المرجع في العلوم للصفوف الرابع والخامس والسادس وفق نظرية الاستجابة للفقرة، واستخدم المنهج الوصفي التحليلي في هذه الدراسة، وصممت ثلاثة اختبارات تشخيصية محكية المرجع في العلوم للصفوف الرابع والخامس والسادس الابتدائي، وقد تكونت الاختبارات في صورتها النهائية من (98) فقرة لاختبار العلوم للصف الرابع، و(136) فقرة لاختبار العلوم للصف الخامس، و(142) فقرة لاختبار العلوم للصف السادس. وطبقت هذه الاختبارات في صورتها النهائية على عينة الدراسة التي تكونت من (942) طالبًا وطالبة من طلبة الصفوف الرابع والخامس والسادس الابتدائي في مدينة تبوك، والتي اختيرت بصورة عشوائية، وتم استقصاء خصائصها السيكومترية بطرق متعددة، كما تم استخراج معالم الصعوبة، والتمييز، والتخمين للفقرات، ومعلمة المقدرة،

وكذلك تم تحديد درجة القطع لكل اختبار باستخدام طريقة أنجوف.

وفي دراسة أجراها راكابو وآخرون (Rakkapao, Prasitpong & Arayathanitkul, 2016) التي عنيت باختبار الاختيار من متعدد لفهم المتجهات في الفيزياء، من خلال تطبيق نظرية استجابة الفقرة. كانت معلمات الصعوبة والتمييز والتخمين متوافقة مع الأنموذج اللوجستي ثلاثي المعلمات باستخدام برنامج PARSCALE. وتم جمع البيانات عن 2392 طالبًا في العلوم والهندسة من ثلاث جامعات في تايلاند. كشفت النتائج أن نظرية IRT مفيدة في تقييم الاختبار نظرًا لأن معلمات الفقرة الخاصة به مستقلة عن معلمات المقدرة، وأشار إلى نطاقات المقدرة المناسبة للاختبار. فضلًا عن، أمكن استخدام النتائج لتقييم فعالية عوامل تشتيت الانتباه في الاختبار.

وفي دراسة أجراها المطيري (Al-Mutairi, 2018) والتي هدفت إلى بناء اختبار مرجعي المحك لقياس الكفاءة والمهارة البحثية لدى طلبة كلية التربية باستخدام نظرية استجابة الفقرة. وتم بناء اختبار محكي المرجع تتمحور فقراته حول مجالات البحث العلمي مثل: أنواع البحث حسب المنهج والتصميم، عناصر البحث، مصادر المعلومات للبحوث التربوية، المتغيرات، أنواع العينات، أدوات البحث التربوي، الخصائص السيكومترية للبحث التربوي والإحصاء. طبق الاختبار المكون من (40) فقرة على عينة مكونة من (362) طالبًا وطالبة من كلية التربية بجامعة الكويت وكلية التربية الأساسية بدولة الكويت. توصلت هذه الدراسة من خلال استخدام الأنموذج الأحادي إلى اختبار يتكون من (37) فقرة لقياس المهارات البحثية، إذ يتمتع بالصدق من خلال تطابقه مع فرضيات أنموذج راش، بالإضافة إلى أنه يتمتع بدرجة ثبات عالية لكل من صعوبة الفقرة ومقدرة الفرد.

وفي دراسة أجراها سوهارتو وروزميادي (Soeharto & Rosmaiyadi, 2018) هدفت إلى تحديد وصف لمهارات التفكير العليا لطلبة الفيزياء في إندونيسيا. أستخدمت في هذه الدراسة اختبارات الصدق والثبات وفق أنموذج Rasch بمساعدة برنامج Winsteps. وتم الحصول على بيانات الطلبة في مهارات التفكير العليا في الفيزياء لموضوع الموجات والبصريات، وكانت النتائج تتمثل في أن مدى المقدرة قد تراوح بين +2.56 إلى −2.42 لوجيت بمستوى صعوبة قياسي بقيمة 0.48 لوجيت، كما تبين أن الطلبة يواجهون صعوبات عند صياغة المعادلة الفيزيائية وعملية الحساب، بينما كانت مهاراتهم في التحليل هي الأعلى، متبوعة بالمقدرة الإبداعية، والمقدرة على

التقييم.

وأجرى استينو ويديهاستوتي وحمدي (Istiyono, Widihastuti & Hamdi, 2020) دراسة هدفت إلى قياس أداء الطلبة في مهارات التفكير الإبداعي في الفيزياء، ومقارنة مستوى هذه المهارات بين الطلاب والطالبات، وتحديد المهارات الفرعية العليا والدنيا لديهم. كان المشاركون في البحث 534 طالبًا وطالبة من المدرسة الثانوية الحكومية في بانتول ريجنسي، تم اختيارها باستخدام أسلوب العينات العشوائية الطبقية. طبق عليهم اختبار وفق نظرية الاستجابة للفقرة وتم تحليل البيانات باستخدام أنموذج التقدير الجزئي. أظهرت النتائج أن أداء الطلبة وقع في الفئة المتوسطة بنسبة (48٪)، وتبين وجود فروق في المهارات الفرعية بين الطلاب والطالبات. بالنسبة للفئة العليا، يهيمن الطلاب على الجوانب الفرعية للتخطيط والتعبير والصياغة بينما تهيمن الطالبات على التخطيط والصياغة والتعبير. بالنسبة للفئة الدنيا، يهيمن الطلاب على إيجاد البدائل والنقد والاختبار بينما تهيمن الطالبات على إيجاد البدائل والتطوير والاختبار.

وفي دراسة آم وهادي (Am & Hadi, 2020) بينت أن الاختبار الجيد هو الذي تم التحقق من صحته واختبار دقته بهدف تقديم دراسة حول تنفيذ نظرية استجابة الفقرة باستخدام أنموذج Rasch لتحليل جودة فقرات الاختبار النهائي في الفيزياء. كانت العينة 131 طالبًا في الصف الحادي عشر. تم استخدام اختبار الامتحان النهائي في الفيزياء كتقنيات لجمع البيانات. أظهرت النتائج أن قيمة الثبات باستخدام معادلة كرونباخ ألفا بلغت 0.47. وقيمة ثبات الشخص 0.48 وقيمة ثبات الفقرة 0.88. وأن اتساق الإجابات من المتقدمين للاختبار لا يزال ضعيفًا، لكن جودة الفقرات في جانب الثبات للأداة جيدة جدًا. بحيث لا يزال الاختبار النهائي للأسئلة في مادة الفيزياء بحاجة إلى تحسين ومراعاة لاستخدامه مرة أخرى في المستقبل.

وفي دراسة قام بها المصاروة والبطوش (Al-Massarweh & Al-Btoosh, 2021) كان الهدف منها هو بناء اختبار تحصيلي مرجعي باستخدام نظرية استجابة الفقرة في مقرر القياس النفسي وفق الأنموذج اللوغاريتمي ثلاثي المعلمة، إذ اشتمل الاختبار على (36) فقرة من نوع الاختيار من متعدد مع أربعة بدائل. تم تطبيق الاختبار على عينة الدراسة وعدد أفرادها (115) طالبًا وطالبة بجامعة الإسراء، وقد توصلت نتائج الدراسة إلى استيفاء افتراضات نظرية استجابة الفقرة في بيانات الدراسة، ومطابقة استجاباتهم. وكانت هناك (34) فقرة اختبار مطابقة لافتراضات الأنموذج اللوغاريتمي ثلاثي المعلمة وتم حذف (فقرتين) لا تتطابقان مع الأنموذج، وأظهرت نتائج

تقييمات معلمات الفقرات (الصعوبة والتمييز والتخمين) أنها كانت مقبولة ضمن المعايير المذكورة في أدبيات القياس والتقويم التربوي والنفسي.

ومن خلال استعراض الدراسات السابقة يتضح ما يأتي:

هناك دراسات أوصت باستخدام نماذج الاستجابة للفقرة في تحليل نتائج الاختبارات محكية المرجع مثل دراسة حمادنة (Hamadneh, 2009) ودراسة (Rakkapao, 2016). وأكدت دراسات أخرى على مدى فاعلية نماذج نظرية الاستجابة للفقرة في تحليل نتائج الاختبارات محكية المرجع مثل دراسة البلوي (Al-Blewey, 2015) ودراسة المطيري (Al-Mutairy, 2018) ودراسة المصاروة والبطوش (Al-Btoosh & Al-Masarweh, 2021).

يلاحظ عدم وجود دراسات عربية استهدفت بناء اختبار محكي المرجع في مادة الفيزياء للصف العاشر وتحديدًا لطلبة البرنامج البريطاني، باستخدام الأنموذج اللوجستي ثلاثي المعلمة مما يعد مبررًا لإجراء الدراسة الحالية، ويميزها عن الدراسات الأخرى التي وظفت على مناهج وطنية. وقد أفادت الدراسة الحالية من الدراسات السابقة في الإطار النظري والمنهجية البحثية المتبعة.

منهج الدراسة:

تم في هذه الدراسة بناء اختبار محكي المرجع في موضوع الفيزياء العامة، ومن ثم تطبيقه وفق المنهج الوصفي المسحي على عينة الطلبة واستخراج البيانات اللازمة باستخدام برنامج الرزمة الإحصائية في العلوم الاجتماعية SPSS وبرنامج BILOG-MG للإجابة عن أسئلة الدراسة.

مجتمع الدراسة وعينتها:

تكون مجتمع الدراسة من جميع طلبة الصف العاشر الأساسي في الأردن الدارسين لمادة الفيزياء Physics 0625 في المنهاج البريطاني والذين هم على مقاعد الدراسة. ويقدر عدد المدارس التي توفر المنهاج البريطاني في الأردن بـ 96 مدرسة حسب إحصائيات 2019 الصادرة من موقع كامبردج. أما عينة الدراسة فتم اختيارها بالطريقة العشوائية العنقودية على اعتبار أن المدرسة وحدة الاختيار مع مراعاة الجنس من حيث البرامج التي تدرس بها والمعدل الأكاديمي للطلبة. وبلغ حجم العينة (500) طالب وطالبة، اختيروا بالطريقة العنقودية موزعين على ما لا يقل عن 40 شعبة.

أداة الدراسة:

تضمنت الدراسة الحالية اختبارًا محكي المرجع في المهارات الأساسية في وحدة الفيزياء العامة للصف العاشر. وقد اتبع في بنائه منهجية بناء هذه الاختبارات بعد الاطلاع على الأدب التربوي المتعلق ببناء الاختبارات محكية المرجع بشكل عام، وكذلك الاطلاع على الدراسات السابقة التي هدفت إلى تطوير اختبارات محكية المرجع في العلوم والفيزياء بشكل خاص، كما تم الاطلاع على الاختبارات والأدوات المستخدمة للكشف عن النتاجات التعلمية في موضوع الفيزياء العامة للصف العاشر فضلًا عن الاطلاع على المناهج وأدلة المعلم التي تعدها وتعتمدها كامبردج من دور نشر مختلفة مثل (Cambridge, Collins, Oxford, Pearson, Hodder) وغيرها، وقد أفاد الباحث من الأدب السابق في تحديد الخطوات الأساسية لتطوير الاختبار محكي المرجع، وعليه تم السير في خطوات منهجية لتطوير الاختبارات المعدة في الدراسة الحالية للصف العاشر، وفيما يأتي عرض للمراحل المتبعة في إعدادها وتطويرها:

تم تطوير الاختبار محكي المرجع في وحدة الفيزياء العامة وفق الخطوات الآتية:

1. تحديد الغرض من الاختبار، ويتمثل في قياس تحصيل طلبة الصف العاشر الملتحقين بالمنهاج الخاص بالبرنامج الدولي في موضوع الفيزياء العامة، من خلال اختبار محكي المرجع، واشتقاق خصائصه السيكومترية في ضوء نظرية الاستجابة للفقرة.
2. تحديد المجال السلوكي الذي يقيسه الاختبار، وذلك بتحليل محتوى الوحدة الدراسية الخاصة بالفيزياء العامة في منهاج البرنامج الدولي للصف العاشر، من خلال الاستعانة بالكتب المدرسية وأدلة المعلم، فضلًا عن الحصول على النتاجات التعلمية الخاصة بالطلبة وفق النظام الموسع (Extended). إذ جرى كتابة مسودة بالموضوعات والعناوين المدرجة في الخطة التفصيلية للمادة وتم ترجمة النتاجات في جدول وضع أسفل منه الأسئلة المتعلقة به وذلك من أجل تحديد أدق للمجال السلوكي، وتم عرض جدول النتاجات على عشرة محكمين من حملة الدكتوراة في تخصص الفيزياء وتخصص أساليب تدريس العلوم؛ للتأكد من احتواء الجدول على جميع النتاجات التعلمية المتضمنة في الخطوة والأخذ بآرائهم وتوصياتهم فيما يتعلق بحذف أو إضافة مهارة أو سؤال بدون أن يؤثر في تمثيل المجال السلوكي. وبناء على ما سبق تكون لدى الباحث جدول بالنتاجات التعلمية لجميع المهارات والأسئلة للصف العاشر فيما يتعلق بموضوع الفيزياء العامة.

3. بناء مواصفات الاختبار، إذ تم في هذه الخطوة استخدام طريقة الأهداف المكبّرة لتوصيف الاختبار عن طريق تكوين قائمة تضمّ جميع النّتاجات في خطة الفيزياء Physics 0625 الّتي يشتمل عليها المجال السلوكي للصف العاشر في موضوع الفيزياء العامة، وتحديد نوع الفقرة ومكوّناتها لكل نتاج، وكتابة توصيف لكيفيّة بناء كل فقرة باستخدام أسلوب صيغ الفقرات، ومحكّات الإجابة الصّحيحة عنها.
4. كتابة فقرات الاختبار ضمن قواعد كتابة الفقرات، وبما يتفق مع مواصفات الاختبارات، إذ تم بناء تجمعات من الفقرات لنتاجات التعلم للمرحلة من نوع الاختيار من متعدد حسب الاختبارات التي تطرحها كامبردج (Paper 2)، كما تم بناء عدد لا يقل عن فقرتين على كل نتاج تعلمي، وبذلك بلغ عدد فقرات هذا الاختبار 85 فقرة.
5. تم عرض الصورة الأولية من المواصفات والفقرات على محكمين متخصصين في مبحث العلوم والفيزياء (أعضاء هيئة تدريس في مجال أساليب ومناهج العلوم، ومعلّمين ذوي خبرة في الفيزياء للصف العاشر، ومشرفين تربويين على البرنامج البريطاني في الفيزياء)، للتحقّق من صدق المحتوى للفقرات الاختبارية، والطلب منهم إبداء الرأي في مدى وضوح الفقرات، واتساقها مع مواصفات الفقرات، وارتباطها أو تعلقها مع الأهداف ونتاجات التعلم التي تنطوي تحتها، ومدى ملاءمتها للفئة المستهدفة. ومناسبة هذه الصياغة مع مستوى الطلبة، وأن إجابة الفقرة لا تعتمد على إجابة فقرة أخرى في الاختبار، وأنه لا توجد أي دلالات لفظية أو نحوية أو في الصور تدل على الإجابة، وكذلك مناسبة المموهات لكل فقرة من الفقرات، وتم إجراء ما يلزم من تعديل في صياغة الفقرات في ضوء آراء المحكمين.
6. تم إعداد الصورة الأولية للاختبار مع التعليمات وكيفية الإجابة على الأسئلة بحيث يكون مشابهًا لما يقدمه المعهد البريطاني.
7. تم مخاطبة إدارات المدارس المعنية بكتاب موجه من إدارة الجامعة الأردنية بهدف تسهيل مهمّة تطبيق الاختبارات.
8. التجريب الأولي للاختبار على عينة استطلاعية بلغ حجمها (200) طالب وطالبةً من الصف العاشر، من خارج عينة الدراسة ومن داخل مجتمع الدراسة، وذلك للتحقّق من وضوح الفقرات والتعليمات، وتحديد الزمن اللازم للإجابة، واستخراج المؤشرات الأولية للثبات. وصعوبة فقرات الاختبار وتمييزها وثباتها.

9. تم استخراج معامل الثبات لاختبار الفيزياء العامة محكي المرجع للصف العاشر بطريقة الاتساق الداخلي باستخدام معادلة كرونباخ ألفا، وقد بلغ (0.970)، وبطريقة الإعادة، إذ بلغ (0.89)، وهي قيم ملائمة للبحث العلمي.

كما وتم استخراج معاملات الصعوبة والتمييز لكل فقرة من فقرات الاختبار وفق النظرية الكلاسيكية. وقد تراوحت معاملات الصعوبة بين (0.35–0.95)، بمتوسط مقداره (0.66)، مع العلم أن المعايير السيكومترية المتعلقة بصعوبة الفقرات يمكن تجاهلها في الاختبارات محكية المرجع (Allam, 2006).

وفيما يتعلق بمعاملات التمييز للفقرات فقد أوضح إيبل (Ebel, 1972) وصفًا لكل مدى من معاملات التمييز فإذا كانت قيمة معامل التمييز أقل من صفر (سالبة)، فتحذف الفقرة ولا داعي للاحتفاظ بها، وإذا كانت من (صفر إلى 0.19) فتكون الفقرة ضعيفة التمييز وينصح بحذفها، وإذا كانت من (0.20 إلى 0.39) فتمييز الفقرة مقبول وينصح بتحسينها، وإذا كان تمييز الفقرة أعلى من (0.39) فالفقرة جيدة ويحتفظ بها. إذ تراوحت معاملات التمييز بين (0.19 حتى 0.74)، مع وجود فقرة ذات تمييز منخفض، وهي الفقرة (85)، إذ قام الباحث بمراجعة تلك الفقرة من خلال عرضها على المتخصصين وإجراء التعديلات بما يتفق مع آرائهم وبما يخدم غرض الاختبار، أما بقية فقرات الاختبار والبالغ عددها (84) فقرة، أي ما يعادل معظم فقرات الاختبار (99٪) اتصفت بدلالة تمييزية ملائمة بحكم بناء الفقرات القائم على التحليل الدقيق للمهارات والنتاجات التي اشتمل عليها الاختبار وفقًا لنماذج نظرية نمائية.

نتائج الدراسة ومناقشتها:

للإجابة عن السؤال الأول: **ما تقديرات معالم الفقرات والأفراد للاختبار محكي المرجع في الفيزياء العامة الذي تم بناؤه وفق المنهاج البريطاني للصف العاشر في الأردن باستخدام النموذج اللوجستي ثلاثي المَعْلَمَة؟**

تم التحقق من افتراضات الأنموذج اللوجستي ثلاثي المعلمة باستخدام استجابات طلبة الصف العاشر، وهي:

أحادية البعد

تم التحقق من أحادية البعد لاختبار الفيزياء العامة محكي المرجع من خلال إجراء التحليل العاملي الاستكشافي (Exploratory Factor Analysis)، بطريقة المكونات الرئيسة

(Principal Component Analysis)، وذلك لبحث العوامل المسؤولة عن الأداء على الاختبار، والتأكد من أن الأداء على الاختبارات يعود لعامل المقدرة، كما وتم استخدام التدوير المتعامد (Varimax Rotation) للعوامل التي كان لها جذر كامن أكبر من الواحد، وقبل ذلك تم التحقق من كفاية حجم العينة باستخدام مؤشر جودة القياس Kaiser-Meyer-Olkin (KMO index)، وبلغت قيمته (0.730)، إذ تعد هذه القيمة كبيرة وتدل على كفاية حجم العينة، فتعتبر قيمة المؤشر مقبولة إذا تراوحت بين (0.5 – 0.7)، وجيدة إذا تراوحت بين (0.7 – 0.8)، وكبيرة إذا تراوحت بين (0.8 – 0.9)، وتكون كبيرة جدًا إذا تعدت (0.9) كما أشار إليه كايزر (Kaiser, 1974).

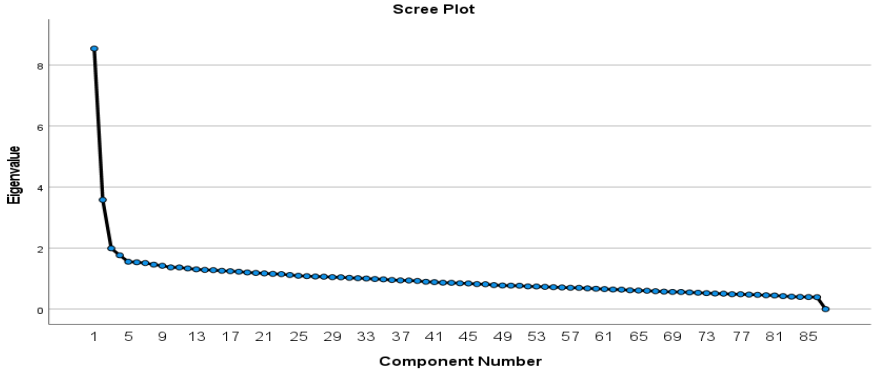
ويبين الجدول (1) قيم الجذور الكامنة (Eigenvalues) ونسبة التباين المفسر، ونسبة التباين التراكمي للعوامل التي لها جذر كامن أكبر من الواحد الصحيح والمستخلصة من اختبار الفيزياء محكي المرجع للصف العاشر للمنهاج البريطاني.

الجدول (1): قيم الجذور الكامنة ونسبة التباين المفسر للعوامل المكونة لاختبار الفيزياء العامة محكي المرجع للصف العاشر للمنهاج البريطاني

العامل	الجذر الكامن	نسبة التباين المفسر ٪	نسبة التباين المفسر التراكمي ٪
1	8.539	9.814	9.814
2	3.581	4.117	13.931
3	1.991	2.288	16.219
4	1.763	2.027	18.246
5	1.553	1.785	20.031
6	1.535	1.764	21.795
7	1.510	1.736	23.531
8	1.458	1.676	25.207
9	1.421	1.633	26.841
10	1.368	1.573	28.413
11	1.367	1.571	29.984
12	1.333	1.532	31.516
13	1.304	1.498	33.015
14	1.285	1.478	34.492
15	1.278	1.469	35.961
16	1.256	1.444	37.405
17	1.244	1.430	38.835
18	1.226	1.409	40.244
19	1.202	1.381	41.625
20	1.186	1.363	42.988

العامل	الجذر الكامن	نسبة التباين المفسر ٪	نسبة التباين المفسر التراكمي ٪
21	1.171	1.346	44.334
22	1.155	1.328	45.662
23	1.147	1.318	46.981
24	1.119	1.286	48.267
25	1.093	1.256	49.523
26	1.080	1.242	50.765
27	1.069	1.228	51.993
28	1.063	1.221	53.214
29	1.047	1.203	54.418
30	1.043	1.199	55.617
31	1.026	1.179	56.796
32	1.014	1.166	57.962
33	1.003	1.153	59.114

يتضح من الجدول (1) وجود (33) عاملًا لها جذور كامنة أكبر من الواحد الصحيح، فسرت ما يقارب (59%) من التباين الكلي، وكانت نسبة قيمة الجذر الكامن للعامل الأول إلى العامل الثاني (2.5)، وهي أكبر من (2)، مما يدل على وجود عامل مسيطر واحد أو سائد على أداء الأفراد على الاختبار إذ فسر ما يقارب (20٪) من التباين الكلي، وبناء على المحكات التي أشار إليها ريكاس (Reckase, 1997) لقبول أحادية البعد والتي من أهمها أن نسبة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني أكبر من (2)، فإن الاختبار يمتاز بأحادية البعد، والشكل (1) يوضح أن المنحنى ينحدر بشدة بعد العامل الأول مما يشكل دليلًا على وجود العامل المسيطر.



الشكل (1): التمثيل البياني لقيم الجذور الكامنة للعوامل المكونة لاختبار الفيزياء العامة محكي المرجع للصف العاشر للمنهاج البريطاني

الاستقلال الموضعي:

تم التحقق من الاستقلال الموضعي بحسب ما أشار إليه (Hambleton & Swaminathan, 1985) أن تحقق أحادية البعد يعد دليلًا على تحقق افتراض الاستقلال الموضعي لفقرات الاختبار.

الاختبارات متحررة من السرعة:

بناء على استجابات الطلبة لوحظ أن جميع الطلبة أجابوا على جميع الفقرات دون ترك أي فقرة بدون إجابة، مما يدل على أن عامل السرعة لم يكن مهمًا في الإجابة على فقرات الاختبار، فكما أوضح (Hambleton & Jones, 1993) أن إجابة (75٪) من الطلبة على (80٪) من فقرات الاختبار فلا يعد اختبارًا يقيس السرعة.

بعد ذلك تم استخلاص خصائص الفقرات (معالم الصعوبة، والتمييز والتخمين ودالة المعلومات) في ضوء نظرية الاستجابة للفقرة، فقد خضع الاختبار للتحليل وفق النماذج اللوجستية الثلاثة: أحادي المعلمة، وثنائي المعلمة، وثلاثي المعلمة، وتم تبني الأنموذج اللوجستي ثلاثي المعلمة كونه الأكثر ملاءمة من إذ عدد الفقرات والأفراد المطابقين له، كما وتم استخلاص نتائج مطابقة البيانات من خلال الاختبار الإحصائي مربع كاي (Chi Square) لفحص مؤشرات المطابقة الخاصة بالفقرات، حيث استخرجت معالم الاختبار محكي المرجع ونتائج مطابقة الفقرات للأنموذج ثلاثي المعلمة. وبلغ متوسط قيم معلمة الصعوبة 1.19 لوجيت بانحراف معياري (0.52)، وكان متوسط قيم معلمة التمييز 2.82 لوجيت بانحراف معياري (0.99)، ومتوسط قيم معلمة التخمين 0.44 لوجيت بانحراف معياري (0.03). وبلغ متوسط دوال معلومات الفقرات 0.895، كما بينت النتائج أن هناك (45) فقرة من أصل (85) فقرة لم تتطابق مع الأنموذج ثلاثي المعلمة. ولدى حذفها أصبح الاختبار مكونًا من (40) فقرة وأصبحت معالم الفقرات على النحو الآتي: متوسط معلمة الصعوبة (1.22) لوجيت بانحراف معياري (0.45)، ومتوسط معلمة التمييز (2.65) لوجيت بانحراف معياري (0.88)، ومتوسط قيم معلمة التخمين (0.45) لوجيت بانحراف معياري (0.03)، ومتوسط دوال معلومات الفقرات (0.789). ويبين الشكل (2) مقطعًا من هذه النتائج.

رقم الفقرة	INTERCEPT S.E.	SLOPE S.E.	THRESHOLD S.E.	LOADING S.E.	ASYMPTOTE S.E.	CHISQ (PROB)
10	−2.661	2.248	1.184	0.914	0.468	6.5
	1.060*	0.632*	0.210*	0.257*	0.036*	(0.3694)
11	−2.570	2.412	1.066	0.924	0.493	5.7
	1.043*	0.669*	0.208*	0.256*	0.036*	(0.4549)
12	−4.817	3.841	1.254	0.968	0.428	10.2
	1.781*	1.120*	0.152*	0.282*	0.023*	(0.1177)
13	−3.574	3.721	0.960	0.966	0.446	2.3
	1.124*	0.814*	0.156*	0.211*	0.027*	(0.8924)
14	−1.222	1.815	0.673	0.876	0.401	7.5
	0.583*	0.411*	0.222*	0.198*	0.058*	(0.2803)

الشكل (2): مقطع من تقديرات معالم الفقرات ومطابقتها وفقًا للأنموذج اللوجستي ثلاثي المعلمة

كما تم استخراج نتائج تقدير معلمة المقدرة والخطأ المعياري في التقدير لأفراد الدراسة المطابقين للأنموذج اللوجستي ثلاثي المعلمة إذ بلغ المتوسط الحسابي لمعلمة المقدرة لاختبار الفيزياء العامة (−0.0742) بانحراف معياري (0.9639)، وكانت قيمة الانحراف المعياري للجذر التربيعي لمتوسط التوزيعات البعدية (Root Mean Square Posterior Standard Deviation) تساوي (0.5322)، وتعبر هذه القيمة عن دقة تقدير قيم مقدرات الأفراد، وتزداد دقة التقدير كلما كانت هذه القيم قريبة من الصفر وفق ما أشار إليه (Natesan & Kieftenbeld, 2012).

عمومًا، يلاحظ أنه قد تم التحقق من افتراضات نظرية الاستجابة للفقرة على الاختبار محكي المرجع، وتم استخلاص خصائص الفقرات (معالم الصعوبة، والتمييز والتخمين) في ضوء الأنموذج اللوجستي ثلاثي المعلمة، كما وتم استخلاص نتائج مطابقة البيانات من خلال الاختبار الإحصائي مربع كاي (Chi Square) لفحص مؤشرات المطابقة الخاصة بالفقرات. وبصورة عامة خلصت نتائج الدراسة إلى صورة مكونة من 40 فقرة تتطابق مع الأنموذج اللوجستي ثلاثي المعلمة، ويمكن استخدامها في تقييم تحصيل الطلبة وتتمتع بدلالات مناسبة من الصدق والثبات وفاعلية الفقرات. وتتفق نتائج هذا السؤال مع نتائج دراسات أخرى مشابهة، مثل دراسة راكابو وآخرون (Rakkapao, Prasitpong & Arayathanitkul, 2016)، ودراسة البلوي (Al-Blewey, 2015) ودراسة حمادنة (Hamadneh, 2009)، والتي استخدمت الأنموذج اللوجستي ثلاثي المعلمة.

للإجابة عن السؤال الثاني: ما الخصائص السيكومترية للاختبار محكي المرجع في الفيزياء العامة الذي تم بناؤه وفق المنهاج البريطاني للصف العاشر في الأردن باستخدام الأنموذج اللوجستي ثلاثي المَعْلَمَة؟

تم بداية تحديد درجات القطع للاختبارات محكية المرجع في الفيزياء للصف العاشر وفق المنهاج البريطاني باستخدام طريقة بوكمارك (Bookmark) والتي تعتمد على قيم صعوبة الفقرات التي استخرجت لفقرات الاختبار النهائي المكون من 40 فقرة، ثم استخدام طريقة بوكمارك فمن خلالها تم تحديد درجة القطع وفق الخطوات الآتية:

- ترتيب جميع فقرات الاختبار محكي المرجع حسب صعوبة الفقرات ومن الأسهل إلى الأصعب، وتم وضعها في كتيب خاص يسمى كتيب الفقرات المرتبة Ordered Items Book (OIB)، الأمر الذي يساعد على تركيز انتباه المحكمين على أشكال الفقرات الذي يمكن أن يجيب عنها من يمتلك الحد الأدنى من الكفاية.
- توضيح المقصود بالأفراد ذوي الحد الأدنى من الكفاية الذين يقعون على الحد الفاصل بين المتقنين وغير المتقنين.
- تحديد نسبة (0.67) كمستوى مناسب لاحتمال أن يجيب المفحوص ذو الحد الأدنى من الكفاية على الفقرة بشكل صحيح، وفقًا لما أشار إليه هاينه (Huynh, 1998).
- يتم في هذه الطريقة وضع إشارة بين فقرتين ضمن كتيب الفقرات OIB إذ يتوقع من الطالب ذي الحد الأدنى من الكفاية أن يجتاز كل الفقرات قبل تلك الإشارة، ويفشل في الإجابة عن الفقرات بعدها.
- يتم تحديد درجة القطع للمحكم الواحد بجمع درجات الفقرات فوق الإشارة التي حددها كل منهم، ومن ثم يحسب المتوسط أو الوسيط لدرجات القطع لجميع المحكمين ليشير إلى درجة القطع للاختبار

وقد تم في هذه الدراسة اختيار (18) محكمًا من مدرسي مادة الفيزياء من حملة الدراسات العليا من ذوي الاختصاص والخبرة في المدارس الدولية المعتمدة للمنهاج البريطاني، وتم تزويدهم بكتيب الفقرات لكل اختبار مرتبة تصاعديًا وفقًا لقيم صعوبتها، وشرح لهم الباحث تعليمات العمل التي وردت سابقًا، وقد بلغت درجة القطع النهائية لاختبار الفيزياء العامة لطلبة الصف العاشر وفق المنهاج البريطاني (57) درجة من (100) وتعادل العلامة 22.8 من 40 للاختبار ككل.

وبناء على ذلك تم سحب عينتين عشوائيتين من الطلبة وتم مقارنة نسبة المتقنين في كل منهما للاختبار بناء على درجة القطع المعتمدة. وبلغ حجم كل عينة 50 طالبًا وطالبة لتقدير دلالات الصدق والثبات للاختبار النهائي المكون من 40 فقرة، إذ تم حساب دلالة الثبات من خلال معامل كارفر Carver ومعامل كابا Kappa وتم حساب معامل الثبات عن طريق معامل ليفنجستون Livingoston، كما تم حساب الصدق بدلالة محك من خلال معامل الارتباط بين الأداء على الاختبارات وعلامات الطلبة في مادة الفيزياء للدورة السابقة، وذلك على النحو المبين أدناه:

مؤشر ثبات كارفر (Carver Reliability Index):

تم إيجاد نسبة الطلبة المصنفين كمتقنين وغير متقنين في عينتي الطلبة، وجرى إيجاد معامل كارفر بإيجاد نسبة مجموع عدد الطلبة الذين تم تصنيفهم بشكل متسق في كلتا المجموعتين (متقن، متقن وغير متقن، غير متقن).

الجدول (1): معاملات الثبات بدلالة اتساق القرار بطريقة كارفر لاختبار الفيزياء العامة محكي المرجع للصف العاشر وفق المنهاج البريطاني

المجموعة الأولى	المجموعة الثانية		المجموع	معامل ثبات كارفر
	متقن	غير متقن		
متقن	35	5	40	0.90
غير متقن	0	10	10	
المجموع	35	15	50	

يبين الجدول (1) أن قيم الثبات بدلالة اتساق القرار بطريقة كارفر للاختبار محكي المرجع في الفيزياء العامة قد بلغ (0.90) مما يدل على أن الاختبار تتمتع بمستوى ملائم من ثبات تصنيف الطلبة إلى متقنين وغير متقنين في مرتي التطبيق وهي مقبولة بالنسبة لأغراض الدراسة.

ثبات التصنيف كابا (Kappa Coefficient):

تم استخدام طريقة كابا لحساب معامل ثبات تصنيف الأفراد (معامل الثبات بدلالة اتساق القرار) ضمن العينتين إلى متقنين وغير متقنين، مع الأخذ بعين الاعتبار الأخطاء الهامشية الناتجة عن التخمين والصدفة، ويوضح الجدول (2) معاملات الثبات بدلالة اتساق القرار بطريقة كابا للاختبار محكي المرجع في الفيزياء العامة للصف العاشر.

الجدول (2): معامل الثبات بدلالة اتساق القرار بطريقة كابا للاختبار محكي المرجع في الفيزياء العامة للصف العاشر

المجموعة الأولى	المجموعة الثانية		الاختبار	معامل كابا
	متقن	غير متقن		
متقن	35	5	40	0.74
غير متقن	0	10	10	
المجموع	35	15	50	

يبين الجدول (2) أن قيم الثبات بدلالة اتساق القرار بطريقة كابا لاختبار الفيزياء محكي المرجع قد بلغ (0.74)، وهو بذلك يتمتع بدرجة مقبولة من ثبات تصنيف الطلبة إلى متقنين وغير متقنين في مرتي التطبيق.

الثبات في ضوء نظرية الاستجابة للفقرة:

تم استخراج قيم الثبات للاختبار بصورته النهائية من خلال استخدام الأنموذج اللوجستي ثلاثي المعلمة باستخدام برنامج Bilog-MG، باستخدام مؤشر فصل الفقرات (Item Separation Index)، ويشير هذا المعامل إلى مدى ثبات المسافة على متصل السمة، وقد بلغت قيمته (0.89).

كما تم التحقق من صدق الاختبار بدلالة محك/ التلازمي باستخراج معامل ارتباط بيرسون بين درجات الطلبة على الاختبار ودرجاتهم في مادة الفيزياء في المدرسة، وقد بلغت قيمة معامل الصدق بدلالة محك/ التلازمي (0.82).

يلاحظ من النتائج السابقة أن الاختبار محكي المرجع في الفيزياء العامة يتمتع بخصائص سيكومترية ملائمة من حيث الصدق والثبات، وربما يعزى ذلك إلى الخطوات والإجراءات المنهجية المتبعة في بنائه استنادًا إلى الأطر النظرية ذات الصلة ببناء الاختبارات محكية المرجع، وهذا يجعل من المناسب استخدامه وتوظيفه في قياس تحصيل الطلبة، ولعل الأطر النظرية والعلمية المتعارف عليها في بناء مثل هذه الاختبارات توفر أرضية مناسبة لتطوير الاختبارات محكية المرجع التي تستخدم للحكم على إتقان التعلم لدى الطلبة، مما يعزز التعلم الإتقاني لديهم، وتتفق نتيجة هذا السؤال مع نتائج دراسات سابقة مثل دراسة آم وهادي (Am & Hadi, 2020) التي أظهرت أن الفقرات تتمتع بخصائص إحصائية جيدة كدراسة حمادنة (Hamadneh, 2009).

التوصيات:

في ضوء النتائج التي تم التوصل إليها في هذه الدراسة يمكن اقتراح مجموعة من

التوصيات، وهي:

- استخدام نظرية الاستجابة للفقرة في بناء مقاييس تحصيلية أخرى وتطويرها.
- توظيف الاختبار محكي المرجع المعدّ للحكم على مدى إتقان الطلبة للنتاجات التعلمية في وحدة الفيزياء العامة للصف العاشر الأساسي.
- إعداد مزيد من الدراسات التي تتحرى فاعلية هذا الاختبار، وتأثير عدد من المتغيرات ذات الصلة بتحصيل الطلبة.

References:

Ababneh, I. (2009). *Criterion referenced tests, their philosophy and the foundations of their development*, 1st ed., Amman: Dar Al-Masirah for Publishing, Distribution and Printing.

Airasian, Peter W. and Madaus, George F. (1972) *Criterion- referenced testing in the classroom. NCME 3 (4). Reprinted by permission of the National Council on Measurement in Education*, Inc., East Lansing, Mich.

Al-Blewey, A. (2015). *Constructing criterion referenced diagnostic tests in science for grades four to six and investigating their psychometric properties according to item response theory*. Unpublished Doctoral Dissertation, University of Jordan, Amman: Jordan.

Al-Hamouri, H. & Al-Kahlot, A. (2209). A proposed strategy for constructing criterion referenced diagnostic tests. *Journal of Educational and Psychological Sciences*: 10(2).

Allam, S. (2006). *Educational and psychological measurement and evaluation: Its basics, applications, and contemporary trends*. Cairo: Dar Al-Fikr Al-Arabi.

Al-Massarweh, S., & Albtoosh, K. (2021), Building a spoken-reference achievement test using the item response Theory in the psychometric course according to three-parameters logistic model, *Multicultural Education* 7(6).

Al-Mutairi, T. (2018) Constructing a criterion-reference test in research methods for students in the school of education by using item response theory: Rasch model. *Journal of the College of Education in Psychological Sciences*, 2018(42-2), 213-250.

Am, M. & Hadi, S. (2020), Implementation of Item Response Theory at final exam test in physics learning: Rasch model study, advances in social science, *Education and Humanities Research*, 541, 336-342.

Brain D. (2009). *Criterion referenced test is and a mastery test. www.moutainmeasurement.com*

Cambridge Assessment International Education, Facts and Figures (2021). *https://www.cambridgeinternational.org/about-us/what-we-do/facts-and-figures*

Cambridge IGCSE (9–1) Physics 0625/0972 (2023). Learner Guide.

Cambridge International Examinations (2017). *www.cie.org.uk*.

Cambridge International Examinations (2021). Cambridge IGCSE™ Physics (0625).

Conrad, Z. (2021). The validity of using diagnostic classification models for criterion-referenced interpretations. A dissertation for the degree of Doctor of Philosophy, University of Kansas, Lawrence, Kansas, U.S.A.

Ebel, R. L. (1972). *Essentials of educational measurement. Englewood Cliffs*. New Jersey: Prentice-Hall, INC.

Hadi, S. (2020), Implementation of item response theory at final exam test in physics learning: Rasch model study, advances in social science, *Education and Humanities Research*, 541(3) pages.

Hamadneh, I. (2009). Using item response theory in constructing a criterion referenced test in mathematics according to the three-teacher logistic model, *Journal of Educational and Psychological Sciences*, 10(2): 215-238.

Hambleton, R. & Jones, R. (1993). Comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and Practice, 12*(3), 38–47. https://doi.org/10.1111/j.1745-3992.1993.tb00543.x

Hambleton, R. & Swaminathan, H. (1985). Item response theory: Principles and applications. Boston: Kluwer Academic Publishers.

Hu, X. (2018*). Implication of item and test differential functioning in criterion referenced tests*. A dissertation for the degree of Doctor of Philosophy, Western Michigan University, Kalamazoo, Ml.

Huynh, H. (1998). On score locations of binary and partial credit items and their applications to item mapping and criterion-referenced interpretation. *Journal of Educational and Behavioral Statistics*, 23(19), 35-56.

Istiyono, E. Widihastuti; S., and Hamdi, S. (2020). Measuring creative thinking skills of senior high school male and female students in

physics (CTSP) Using the IRT-based physTCreTS. *Journal of Turkish Science Education*, 17 (4), 578-590.

Kaiser, H. (1974) An index of factorial simplicity. *Psychometrical*, 39, 31-36.

Lee, Y., Palazzo, D., Warnakulasooriya, R. and Pritchard, D. (2008). Measuring student learning with item response theory. *Phys. Rev. ST Phys. Educ. Res*. 4(1), 010102.

Murphy, K.& Davidshofer, c. (2001). *Psychological testing: Principles and applications*. New Jersev: Prentice Hall.

Natesan P. Kieftenbeld, V. (2012). Recovery of graded response model parameters: A Comparison of marginal maximum likelihood and Markov Chain Monte Carlo Estimation. *Applied Psychological Measurement*. 36(5),399-419.

Rakkapao, S., Prasitpong, S. & Arayathanitkul, K. (2016). Analysis test of understanding of vectors with the three-parameter logistic model of item response theory and item response curves technique. *Phys. Rev. Phys. Educ. Res*. 12, 020135.

Reckase, M.D. (1997). The past and future of multidimensional item response theory. *Applied Psychological Measurement*, 1, 25-36.

Sax, G. & Newton, J. (1997). *Principles of educational and psychological measurement and evaluation.* (4th ed.), U.S.A: Wadsworth Publishing company.

Soeharto S, & Rosmaiyadi R. (2018). The analysis of students' higher order thinking skills (HOTS) in wave and optics using IRT with winstep software. *J Educ Sci Technol., 1(1):145–50.*

Verbic, S. & Tomic, B. (2009). *Test item response time and the response likelihood. Physics Education, (21),* 39-52.

Walsh, W.& Betz, N. (1995): *Tests and assessment*. Englewood Cliffs, New Jersy: Printice-Hall.