بناء إختبار تحصيلي مقالي في مادة الكيمياء للصف الثاني الثانوي وفق نموذج التقدير الجزئي لنظرية إستجابة الفقرة

إعـــداد:

أ. إيمان عودة غنام الجهني

طالبة ماجستير تخصص القياس والتقويم قسم التربية وعلم النفس
كلية التربية والأداب جامعة تبوك المملكة العربية السعودية

إشراف : أ.د. شاهر خالد سليمان

أستاذ القياس والتقويم كلية التربية والآداب جامعة تبوك

بناء اختبار تحصيلي مقالي في مادة الكيمياء للصف الثاني الثانوي وفق نموذج التقدير الجزئي لنظرية استجابة الفقرة

أ. إيمان عودة غنام الجهني
طالبة ماجستير تخصص القياس والتقويم قسم التربية وعلم النفس
كلية التربية والأداب جامعة تبوك المملكة العربية السعودية
إشراف : أ.د. شاهر خالد سليمان
أستاذ القياس والتقويم كلية التربية والآداب جامعة تبوك

• **المستخلص:**
هدفت الدراسة إلى بناء اختبار تحصيلي مقالي في مادة الكيمياء للصف الثاني الثانوي العلمي؛ وفق نموذج التقدير الجزئي لنظرية استجابة الفقرة؛ معتمدة على منهجية البحث الوصفي التحليلي، وتمثلت أداة الدراسة في اختبار تحصيلي مقالي لمقرر الكيمياء للفصل الدراسي الأول للصف الثاني ثانوي، تألف في صورته النهائية من (١٢) سؤالا، و(٥٠) خطوة - من تصميم الباحثة- تمّ تطبيقه على عينة عشوائية مقدارها (٤٥١) طالبا وطالبة من الصف الثاني ثانوي بالمدارس الحكومية السعودية بمنطقة المدينة المنورة، للعام الدراسي (١٤٤٤هـ)، واستخدمت الأساليب الإحصائية التالية: التحليل العاملي الاستكشافي للكشف عن أحادية البعد، واستخدام برنامج (Easy Estimation) للإجابة على أسئلة الدراسة، وأشارت النتائج إلى أن الاختبار يقيس سمة واحدة، وعدم وجود ارتباط دال إحصائيًا بين فقرات الاختبار، وأن الاختبار متحرر من عامل السرعة، كما أظهرت النتائج أن هناك مطابقة بين فقرات الاختبار مع نموذج التقدير الجزئي، وتم التحقق من صدق الاختبار وثباته حيث بلغ معامل الثبات للأفراد ومعامل الثبات الامبريقي للفقرات وفق نموذج التقدير الجزئي على التوالي (٠.٩١، ٠.٩٤) مما يدلُّ على قدرة الاختبار في الفصل بين قدرات الأفراد، والتمييز بين قدراتهم على متصل السمة، وفي ضوء النتائج السابقة، قدمت الباحثة عدداً من التوصيات من أهمها: حث المختصين في مجال تطوير وتصميم المقاييس النفسية والتربوية باستخدام نماذج الاستجابة للمفردة في إعدادها وتطويرها.
الكلمات المفتاحية: الاختبار التحصيلي، الكيمياء، الصف الثاني الثانوي، نموذج التقدير الجزئي، نظرية استجابة الفقرة.

Constructing an essay achievement test in chemistry for the second year of secondary school according to the partial estimation model of the item response theory

Al-Juhani, Eman Odeh

Abstract

The study aimed to construct an essay achievement test in chemistry for the second year of secondary science. According to the partial estimation model of item response theory; It is based on the analytical descriptive research methodology, and the study tool was an essay achievement test for the chemistry course for the first semester For the second secondary grade, it consisted in its final form of (12) questions and (50) steps - designed by the researcher - it was applied to a random sample of (451) male and female students from

the second secondary grade in Saudi government schools in Al-Madinah Al-Munawwarah region, for the academic year (1444 AH).), and the following statistical methods were used: exploratory factor analysis to detect one-dimensionality, and the use of (Easy Estimation) program to answer the study questions Speed, as the results showed that there is a match between the test items with the partial estimation model, and the validity and reliability of the test were verified, as the stability coefficient for individuals and the empirical stability coefficient for the items according to the partial estimation model, respectively, reached (0.91, 0.94), which indicates the ability of The test is in separating the abilities of individuals, and distinguishing between their abilities on the continuum of the trait, and in the light of the previous results, the researcher presented a number of recommendations, the most important of which are: urging specialists in the field of developing and designing psychological and educational measures using individual response models in preparing and developing them.

Keywords: achievement test, chemistry, second secondary grade, partial estimation model, item response theory.

• المقدمة.

تمثل الاختبارات التحصيلية آلية لصيقة بأية عملية تعليمية؛ فهي العملية التي يُسند إليها إصدار الأحكام على مخرجات التعليم، وتقييم مستويات الطلاب المختلفة، فلا يستطيع المعلم الاستغناء عنها في تقييم طلابه وأيضا في تقييم أساليبه التدريسية، كما لا يستطيع المتعلمون استشعار إنجازاتهم دون الخضوع لها.

ولقد أضحت الاختبارات والمقاييس النفسية والتربوية بشكل عام، والاختبارات التحصيلية بشكل خاص في العصر الحديث، من أكثر محاور العملية التربوية التعليمية عنايةً واهتماماً من قبل مسؤولي نظم التربية والتعليم؛ نظراً لحجم الدور الجديد الذي غدت تلعبه الاختبارات والمقاييس التحصيلية كأدوات أساسية وفاعلة في تقويم أداء المتعلمين وتطوير العملية التعليمية بجميع أبعادها التربوية والعلمية، حيث يعتمد عليها مسؤولو التربية والتعليم الحكومية، في تحديد مدى تقدم الطالب في التحصيل، ومن ثم اتخاذ القرارات التربوية المناسبة الخاصة بهذا الطالب وبعناصر المنظومة التعليمية، بحسب إسهام كلّ عنصر في هذه العملية؛ الأمر الذي دفع الباحثين التربويين وخبراء القياس والتقويم النفسي والتربوي في السنوات الأخيرة نحو تكثيف وتوسيع جهودهم البحثية في مجال تطوير الامتحانات وأساليب التقويم وبناء الاختبارات والمقاييس النفسية والتربوية؛ وفق معايير عصرية يحددها المختصون في مجال القياس التربوي(الخولي،٢٠١٣).

ورغم الاهتمام الكبير الذي منيت به الاختبارات التحصيلية في سياق اتجاهات إصلاح وتطوير التعليم؛ إلا أنّ الاختبارات التحصيلية واجهت ولا تزال مشكلات عدة أفقدتها في كثير من الأحيان الدقة والموضوعية، سيما وتفسير درجات مثل هذه الاختبارات يعتمد على الدرجة الكلية التي لا تعبر بصدق عن المستوى الحقيقي لقدرة المتعلم، إذ ترتبط بمعامل صعوبة المقررات، حيث لوحظ أن المقررات المختلفة في مستوى الصعوبة - وبالرغم من أنها تقيس ذات المحتوى- تؤدي إلى مستويات مختلفة من القدرة عند ذات المتعلم، هذا من جهة، ومن جهة أخرى فإن تقدير معامل صعوبة المفردات في هذه الاختبارات يعتمد على مستوى أداء المتعلمين في العينة، ومعاملات صعوبة المفردة الواحدة مختلفة باختلاف مستوى أداء أفراد العينة (الخولي،٢٠١٣).

ولهذا فقد سعت المؤسسات التعليمية الحكومية، مثل مديريات التعليم خلال السنوات الأخيرة نحو الاهتمام بمعايرة (Calibration) الفقرات وبنوك الأسئلة (Item Bank) لتطوير فقرات بنوك الأسئلة، وخاصة بعد التقدم الكبير في استخدام برامج الحاسوب؛ وبنك الأسئلة عبارة عن تجمع من الفقرات ذات خصائص سيكومترية معلومة، تودع الفقرات في البنك بطريقة تسهل على مطور الاختبار سحب مجموعة منها تستخدم في بناء اختبارات مناسبة لأغراض مختلفة في القياس والتقويم التربوي، وتحقق الهدف من الاختبار بدقة وبأقل جهد.

وفي سياق مراجعة الأدب المعرفي بمجال القياس والتقويم التربوي، سنجد أن هناك أسلوبان لتحليل الفقرات وتكوين بنوك الأسئلة، يتمثل الأسلوب الأول في النظرية الكلاسيكية التقليدية في القياس (Classical Test Theory) وتشير المراجع في النظرية الكلاسيكية إلى مؤشرات تستخدم في انتقاء فقرات الاختبارات والمقاييس منها معاملات التمييز والصعوبة للفقرات، فتقدير معامل التمييز للفقرة من وجهة النظر الكلاسيكية هو معامل الارتباط بين الأداء على الفقرة والأداء على الاختبار بجميع فقراته، أما معامل صعوبة الأسئلة فيتم تقديره حسب النظرية الكلاسيكية في القياس بحساب نسبة النجاح على الفقرة (الزعبي، ٢٠٠٥). وأما الأسلوب الثاني فيعتمد على نظرية الاستجابة للفقرة (Item Response Theory)

ويعتمد إعداد معظم الاختبارات على نظرية القياس الكلاسيكية (CTT) (Test Theory Classical) والتي تستند على أنّ درجة الاختبار عبارة عن مجموعة الدرجات الحقيقة بالإضافة إلى الخطأ العشوائي، لذلك قد يحصل فرد على نسبة مختلفة في اختبارين للذكاء بسبب اختلاف مقدار الخطأ، كما تفترض أن دقة القياس ثابتة عبر مدى السمة (Domino,2006)

حيث إنّ هذه المقاييس قد تكون أكثر ملاءمة للتمييز بين الأفراد الذين يقعون في المدى المرتفع للسمة فقط، وبالتالي تتوزع هذه المقاييس بشكل غير متساوٍ في الدقة عبر المدى الطبيعي، ويمكن أن تكون دالة المعلومات للمقاييس غير متكافئة لتجمع قيم الصعوبة والتمييز للمفردات في منطقة ضيقة من مدى السمة (Fraley et al. 2000).

وعلى الرغم من انتقادات الباحثين المستمرة للنظرية الكلاسيكية والإشارة الى العديد من أوجه وجوانب القصور فيها؛ إلا أن تطبيقها هو السائد في عملية بناء الاختبارات وتطبيقها وتحليل نتائجها وتفسيرها؛ دون تمكنها من معالجة بعض القضايا المهمة في القياس والتقويم مثل بناء بنوك الأسئلة بمفهومها الجديد؛ كما أنّ إحصائيات تحليل الفقرة بالطرق الكلاسيكية لا تقدم معلومات حول كيفية استجابة الفقرة من قبل الأفراد المختلفين في مستوى القدرة المقاسة بالفقرة، وإحدى الطرق لتحليل فقرات الاختبار التي تعطي صورة أكثر شمولية حول استجابة الفقرة من قبل الأفراد المختلفين في مستوي القدرة تعرف بنظرية الاستجابة للفقرة (Item Response Theory) وهي نظرية تتضمن افتراضات ومعالجات يمكن بها التغلب على كثير من جوانب القصور في النظرية الكلاسيكية (الزعبي، ٢٠٠٥). (الخولي، ٢٠٠٦).

حيث يشير بعض الباحثين المتخصصين إلى أنّ أبرز عيوب نظرية القياس الكلاسيكية، تتمثل في: اعتمادها على الخصائص السيكومترية للاختبار وصعوبة وتمييز المفردات على عينة الأفراد التي يطبق عليها الاختبار، وتركيزها بشكل أساسي على تقييم درجات الاختبار الكلية، وتفترض أنه كلما زادت درجةً الفرد دل ذلك على زيادة السمة أو القدرة المقاسة، ولكن الواقع يشير إلى أن بعض الأفراد ذوي القدرة العالية يحصلون أحيانا على درجات منخفضة في الاختبارات والعكس صحيح، وتفترض أن الاختبارات تقيس متغيرا أحادي البعد، ولكن يصعب تقدير أحادية البعد في كثير من الاختبارات النفسية، كما أنها تقيد الدرجة الكلية للاختبار بمفرداته، وتغير تكوين الاختبار ومعنى مفرداته بمرور الزمن، وصعوبة تحقيق مفهوم الاختبارات المتكافئة عمليا لحساب ثبات الاختبار، كما أنّ الأفراد لا يكونون في نفس الظروف الاختبارية عند استخدام طريقة إعادة التطبيق، كما أنها تعطي قيمة تقديرية كلية لمصادر أخطاء القياس المتعددة (سليمان، ٢٠٠٩)، و(Yang and kao, Luo et al,2014)

ونتيجة لهذا القصور الكبير الذي يغلب على نظرية القياس الكلاسيكية، فقد اتجه علماء القياس والتقويم نحو البحث عن طرائق جديدة في القياس تحقق الموضوعية وتملك القدرة على العامل مع متطلبات القياس الجديدة في

القياس تحقّق الموضوعية، وتملك القدرة على التعامل مع متطلبات القياس الجديدة؛ فأثمر عن هذا البحث ظهور نظرية الاستجابة للمفردة (IRT) (Item Response Theory) التي انبثق عنها مجموعة من النماذج الرياضية، تمّ تصنيفها إلى نماذج أحادية البعد (Unidimensional Models) تفترض وجود سمة أو قدرة واحدة تكمن وراء أداء المتعلم في الاختيار، ونماذج متعددة الأبعاد (Multidimensional Models) تفترض وجود سمات متعددة. (الخولي، ٢٠١٣). بحيث ينبغي التأكيد على عدم الاكتفاء باستخدام الأساليب السيكومترية الكلاسيكية في التحقق من الخصائص السيكومترية للمقاييس النفسية، وإنما يجب أن يمتدّ التحليل للاعتماد على نماذج نظرية الاستجابة للمفردة وبخاصة المقاييس التي تستخدم على نطاق واسع في البحوث النظرية والجوانب التطبيقية (هيبة، ٢٠١٦).

ويمكن استخدام نظرية الاستجابة للمفردة لبناء وتقييم أدوات القياس الجديدة، وعلى الرغم من استخدامها بشكل أساسي في الاختبارات المعرفية واسعة النطاق، إلا أنه يزداد اهتمام الباحثين باستخدامها لتطوير المقاييس النفسية (Gnamba & Batinic,2011).

كما يؤكد بعض الباحثين على أنه ينبغي استخدام نظرية الاستجابة للمفردة في القياس من أجل تدرج بعض الاختبارات والمقاييس المستخدمة في العلوم الإنسانية التي يمكن من خلالها تقدير بارامترات الأفراد والمفردات بأقل قدر من الخطأ (اللهيبي، ٢٠١٨).

إذ إنه ينبغي الاهتمام بنظرية الاستجابة للمفردة بصفة خاصة، وبأساليب القياس الحديثة بصفة عامة، وتوفير متطلبات التعامل مع هذه الأساليب، مثل: الحاسبات الآلية وبرامجها، وتدريب الباحثين على فنيات تلك البرامج، نظراً لما تتمتع به نظرية الاستجابة للمفردة من مميزات بالمقارنة بنظرية القياس الكلاسيكية (عكاشة وآخرون، ٢٠٢٠).

وقد توصلت نتائج دراسة (الحاج، وأحمد، ٢٠٢١) إلى أنّ أهمية استخدام نماذج نظرية الاستجابة للمفردة يكمنُ أيضاً وبدرجة أساسية في إعداد وتصميم الاختبارات والمقاييس النفسية المختلفة، لما تتميزُ به من خصائص يمكن الاستفادة منها في تحقيق موضوعية القياس ودقته، وتعميم النتائج، وإمكانية المقارنة بين مجموعات متباينة من الأفراد

وتمّ تطوير نماذج لنظرية الاستجابة للمفردة أحادية البعد متعددة الاستجابة، لأنّ المفردات ثنائية الاستجابة تضع قيوداً غير مقبولة في الواقع التربوي والنفسي الميداني، وتحدّ من تنوع صيغ التقويم التي ينادي بها خبراء القياس في الوقت الحاضر، فتقويم مهارات مثل: حلّ المشكلات، وكتابة مقال،

وتجميع جهاز، وكذلك تصميم مفردات استبانات الاتجاهات والشخصية يتطلب مفردات اختبارية متعددة الاستجابات (حبيب، وعزيز، ٢٠١٨).

ومن نماذج الاستجابة للمفردة أحادية البعد متعددة الاستجابة نموذج التقدير الجزئي (partial Credit Model) الذي يصلح لتدريج المفردات التي تختلف في بارامترات الصعوبة وعتبات الصعوبة داخل المفردة الواحد، ونموذج الاستجابة المتدرجة (Response Credit Model) الذي يصلح لتدرج مفردات المقاييس التي تختلف في بارامترات الصعوبة والتمييز وعتبات الصعوبة داخل المفردة الواحدة (Robitzsch&Steinfeld,2018).

ويشير الخولي (٢٠١٣) إلى تزايد الاهتمام بنموذج التقدير الجزئي لاستخدامه في تطبيقات متعددة كتصميم بنوك الأسئلة، ومعادلة درجات مفردات الصور الاختبارية، وتصميم الاختبارات الموائمة المحوسبة، وتحليل استجابات الأفراد على مقاييس الاتجاهات والشخصية والاختبارات التحصيلية المصممة على المستوى القومي.

لذا ومن خلال - ما سبق - تحاول الدراسة الحالية بناء اختبار تحصيلي في مادة الكيمياء للصف الثاني ثانوي وفق نموذج التقدير الجزئي لنظرية استجابة الفقرة.

• مشكلة الدراسة:

على الرغم من انتشار استخدام نظرية القياس التقليدية في القرن العشرين لبناء الاختبارات النفسية والتربوية، وتحليل البيانات المستمدة منها، إلا أنّ هذه النظرية لم تستطع التغلب على كثير من التطبيقات العملية المستحدثة، ولذلك أصبحت نظرية الاستجابة للمفردة تستخدم في تطوير الاختبارات التحصيلية المقننة، واختبارات الذكاء والاستعدادات، ومقاييس الشخصية والاتجاهات والسمات الوجدانية الأخرى (علام، ٢٠٠٥).

حيث أشارت الدراسات الحديثة ذات العلاقة، إلى أهمية وضرورة استخدام نماذج نظرية الاستجابة للمفردة في إعداد الاختبارات والمقاييس والتحقق من خصائصها السيكومترية، وقد أوصت دراسة كلا من (الشافعي؛ ونورالدين، ٢٠٠٧) باستخدام نماذج الاستجابة للمفردة في بناء الاختبارات النفسية والتربوية، وذلك لما تتمتع به هذه النماذج من خصائص يمكن من خلالها رفع كفاءة أدوات القياس بما يوفر موضوعية ودقة وجودة هذا القياس.

كما أوصت أيضاً دراسة (Yang and Kao,2014) باستخدام نظرية الاستجابة للمفردة للتحقق من صدق المقاييس، لأنها توفر طريقة لحل العديد من تحديات القياس التي يجب معالجتها عند إنشاء اختبار أو مقياس.

كما أنّ الباحثة من خلال احتكاكها العملي بالميدان التربوي، قد رصدت بالملاحظة المباشرة، حدوث الكثير من المشكلات التعليمية المتعلقة بإعداد

الاختبارات والمقاييس التربوية التحصيلية، التي تتولد نتيجة امعان المعلمين في بناء اختباراتهم التحصيلية على ضوء نظرية القياس التربوي التقليدية، وعزوفهم شبه التام عن استخدام نظرية الاستجابة للفردة وكذا تطبيق نماذجها الملائمة، ولا سيما نموذج التقدير الجزئي.

وفي ضوء -ما سبق- ونظراً لقلة الدراسات والبحوث العربية التي تناولت نموذج التقدير الجزئي في البيئة العربية بشكل عام والمحلية بشكل خاص، جاءت هذه الدراسة لتحاول بناء اختبار تحصيلي مقالي في مادة الكيمياء للصف الثاني الثانوي العلمي وفق نموذج التقدير الجزئي لنظرية استجابة الفقرة

ينبثق من مشكلة الدراسة مجموعة من الأسئلة التي تتمثل في الآتي:

- ما درجة مطابقة نموذج التقدير الجزئي لفقرات لاختبار تحصيلي مقالي في مادة الكيمياء للصف الثاني الثانوي؟
- ما قيم معالم الفقرة (معلمة الصعوبة) لاختبار تحصيلي مقالي في مادة الكيمياء للصف الثاني الثانوي الذي تم تطويره وفق نموذج التقدير الجزئي؟
- ما قيم معلم قدرة الأفراد للإجابة عن الاختبار التحصيلي المقالي في مادة الكيمياء للصف الثاني الثانوي الذي تم تطويره وفق نموذج التقدير الجزئي؟
- ما دلالات ثبات اختبار تحصيلي مقالي في مادة الكيمياء للصف الثاني الثانوي الذي تم تطويره وفق نموذج التقدير الجزئي؟
- ما دلالات صدق اختبار تحصيلي مقالي في مادة الكيمياء للصف الثاني الثانوي الذي تم تطويره وفق نموذج التقدير الجزئي؟

• أهداف الدراسة:

هدفت الدراسة الحالية بصورة رئيسية إلى بناء اختبار تحصيلي مقالي في مادة الكيمياء للصف الثاني الثانوي العلمي وفق نموذج التقدير الجزئي لنظرية استجابة الفقرة؛ وذلك ممن خلال تحقيق الأهداف الفرعية التالية:

- التحقق من مدى مطابقة نموذج التقدير الجزئي لفقرات لاختبار تحصيلي مقالي في مادة الكيمياء للصف الثاني الثانوي.
- الكشف عن قيم معالم الفقرة (معلمة الصعوبة) لاختبار تحصيلي مقالي في مادة الكيمياء للصف الثاني الثانوي الذي تم تطويره وفق نموذج التقدير الجزئي.
- الكشف عن قيم معلم قدرة الأفراد للإجابة عن الاختبار التحصيلي المقالي في مادة الكيمياء للصف الثاني الثانوي الذي تم تطويره وفق نموذج التقدير الجزئي.

- الكشف عن دلالات ثبات اختبار تحصيلي مقالي في مادة الكيمياء للصف الثاني الثانوي الذي تم تطويره وفق نموذج التقدير الجزئي.
- الكشف عن دلالات صدق اختبار تحصيلي مقالي في مادة الكيمياء للصف الثاني الثانوي الذي تم تطويره وفق نموذج التقدير الجزئي.

• أهمية الدراسة:

• الأهمية النظرية:

تتمثل أهمية الدراسة النظرية في الآتي:

- تسليط الضوء على أهمية الاختبارات في مجالات التقييم في مدارس وزارة التعليم، وكون البحث منصبا عليها لأهميتها في القياسات النفسية والتربوية، وخاصة تلك التي يتم بناؤها في ضوء نماذج رياضية مختلفة ترتبط بالنظرية الحديثة في القياس (IRT) التي يعول عليها كثيرا في بناء الاختبارات وبنوك الأسئلة الحديثة.
- توضيح الاهتمام الكبير الذي أصبحت توليه المؤسسات التربوية المختلفة في بناء فقرات الاختبار والتي تتمتع بخصائص سيكومترية - صعوبة الفقرة والتمييز- لا تتأثر بخصائص عينة المفحوصين (القدرة) ويمكن تطبيقها على عينات مختلفة من المفحوصين دون تحيز إلى مجموعة دون غيرها، أو تطبيق عينات مختلفة من الفقرات على نفس عينة المفحوصين دون أن تتأثر خصائص الفقرة بخصائص قدرة المفحوصين – القدرة-.
- إنّ تصميم بنوك الأسئلة وفق نماذج نظرية الاستجابة للفقرة يمكننا من توفير المال والجهد والوقت المبذول في تصميم الاختبارات المختلفة كلّ عام، حيث إن هذه البنوك تسمح بحذف مفردات قديمة وإضافة مفردات جديدة فرضتها ظروف تغير المنهاج الخاص بمادة ما مع مرور الزمن، وذلك وفق آلية محددة تسمح بتدرج المفردات الجديدة على مقياس بنك الأسئلة نفسه، ودون اللجوء إلى تصميم اختبار جديد يتوافق مع المتغيرات الحاصلة في هذا المنهاج، كما هو الحال بالنسبة إلى الاختيارات المصممة وفق نظرية القياس التقليدية.
- استثمار التقنيات والبرمجيات المناسبة في تطوير نظام التقويم بما يواكب التطور الحاصل في تقنيات المعلومات والاتصالات.

• الأهمية التطبيقية:

تتمثل الأهمية التطبيقية للدراسة في الآتي:

- تعريف القارئ العربي بنموذج التقدير الجزئي الذي يعتبر من أبسط النماذج متعددة التدريج في نظرية الاستجابة للمفردة.
- لفت انتباه المسئولين عن العملية التعليمية في وزارة التعليم إلى إمكانية تصميم بنوك أسئلة ذات المفردات ثنائية ومتعددة الاستجابة، وبالتالي حل

مشكلة وضع أسئلة ثنائية ومتعددة الاستجابة في الاختبار الواحد، مما يسمح بتقدير المعلومات الجزئية عند المتعلم وبالتالي تشخيص أداء المتعلم بصورة أكثر دقة وموضوعية تسمح بوضع الخطط التعليمية لمعالجة أوجه قصور تحصيل المتعلم.

- نشر ثقافة نظرية الاستجابة للمفردة وميزاتها التربوية وإجراءات بنائها واستخدامها في عملية القياس والتقويم لدى المعلمين والموجهين والقيادات التربوية والمهتمين.
- دعم وزارة التعليم بمجموعة من الاختبارات لمادة الكيمياء وفق نموذج التقدير الجزئي لنظرية استجابة الفقرة تتصف بالشمولية والموضوعية والدقة اللازمة، والتي يمكن بواسطتها تكوين نماذج اختباريه عند الحاجة لقياس قدرات الطلاب في أبعاد مختلفة.
- توفير وقت وجهد معلم الكيمياء بتوفير اختبارات تحصيلية جاهزة للمادة طبقا لمواصفات وأهداف المقرر الدراسي بدلا من إعدادها كل مرة للمقرر نفسه.

• مصطلحات الدراسة:

• الاختبار التحصيلي [Attainment Test]

يعرفه "براون" بأنه الاختبار يقيس التعلم الذي يحدث نتيجة الخبرة في موقف تعلم يتم التحكم فيه مثل صف مدرسي أو برنامج تدريبي، وهو الإطار المرجعي للاهتمام بما تعلمه الطلاب (صلاح الدين، ٢٠١٢)، ويعرف الاختبار التحصيلي بأنه طريقة منظمة لتحديد مستوى تحصيل الطلبة لمعلومات ومهارات في مادة دراسية درسوها مسبقا بصبغة رسمية من خلال إجابتهم عن عينة من الأسئلة (عودة، ٢٠٠٥).

وتعرف الباحثة الاختبارات التحصيلية إجرائياً: بأنها إجراء منظم لقياس ما اكتسبه المتعلمون من حقائق ومفاهيم وتعليمات ومهارات نتيجة لدراسة موضوع ما، أو وحدة تعليمية معينة، وتقع الاختبارات التحصيلية في فئتين: الأولى تمثل الاختبارات المقننة، والأخرى تمثل الاختبارات التي يعدها المعلم التي تعرف بالامتحانات.

• المفردة ثنائية الاستجابة [Binary Response Item]:

وهي المفردات التي تنقسم فيها الإجابة إلى طريقة يمنح فيها الفرد عند استجابته درجتين مختلفتين وهي التي تستخدم في تحليل المفردات التي يجيب المفحوص عنها بصح أو خطأ، أو بالاختيار بين (نعم أو لا) وتكون الدرجة التي يحصل عليها المفحوص إما (١) أو (صفرا) (علام، ٢٠٠٥).

• المفردة متعددة الاستجابة [Mullti-Response Item]:

وهي فقرات اختباريه تكون الاستجابة عليها متعددة (١، ٢، ٣، ٤، ٥) حيث يكلف المفحوص بتحديد الاستجابة التي يراها مناسبة(علام، ٢٠٠٥).

- **نماذج الاستجابة للمفردة (Item Response Models)**

هي مجموعة دوال رياضية احتمالية تهدف إلى تحديد العلاقة بين أداء الفرد في اختبار محدد وبين السمة الكامنة وراء هذا الأداء (الخولي، ٢٠١٣).

- **نموذج التقدير الجزئي متعدد التدريج (Polytomous Partial Credit Model)**

و يرمز له (PCM) أحد نماذج نظرية الاستجابة للمفردة الذي يعد امتدادا لنموذج (Rasch) أو النموذج اللوغاريتمي أحادي البارامتر، الذي يفترض أن معامل الصعوبة هو البارامتر الوحيد المؤثر على أداء المفحوصين، وتتوقف عليه إجابة الفرد بصورة صحيحة أو خطأ، بينما بارامتر التمييز يكون ثابتا لجميع المفردات، وقيمة بارامتر التخمين تساوي صفر (فرحات، ٢٠٢٢).

- **نظرية الاستجابة للمفردة [Item Response Theory]**

وهي مجموعة الطرق الإحصائية التي تستخدم في حساب معالم الفقرات والأفراد الخاصة بالاختبار –اداة الدراسة الحالية- من صعوبة وتخمين ودالة معلومات(زكري، ٢٠٠٨)، والمكونة من نماذج أحادية البعد ويقصد بها: نموذج راش الأحادي البارامتر، نموذج لورد ثنائي البارامتر، نموذج الثلاثي البارامتري (الحاج محمود، ٢٠١٨).

• محدّدات الدراسة:

- الحد الموضوعي: اقتصرت الدراسة على بناء اختبار تحصيلي في مادة الكيمياء للصف الثاني ثانوي وفق نموذج التقدير الجزئي لنظرية استجابة الفقرة.
- الحد الديموغرافي(البشري): اقتصرت الدراسة على عينة من طلاب وطالبات الصف الثاني ثانوي المنتظمين في مدارس (المدينة المنورة) في العام الدراسي ١٤٤٤هـ.
- الحدّ المكاني: اقتصر تطبيق الدراسة على مدارس المرحلة الثانوية بمنطقة المدينة المنورة في المملكة العربية السعودية.
- الحدّ الزماني: طبقت الدراسة في الفصل الدراسي الثاني للعام ١٤٤٤هـ-2023م.

• الإطار النظري:

• المبحث الأول: الاختبارات التحصيلية Academic Achievement

• المطلب الأول: تعريف الاختبارات التحصيلية:

تهدف الاختبارات التحصيلية إلى قياس مدى تحقق الأهداف التعليمية، وهي تعدّ بواسطة المعلمين في نهاية كل فصل دراسي، وفي نهاية كل مرحلة دراسية. والاختبارات التحصيلية كما يعرفها البعض هي: "عبارة عن أدوات

تستخدم لقياس مدى فهم، ومعرفة، وتحصيل الطلبة للمعلومات والمهارات في مادة دراسية بهدف قياس نواتج التعلم، وتكون في صورة مختلفة تراعي فيها قواعد القياس والتيقن من صدقٍ وثباتٍ وموضوعية وقدرة على التمييز"(خنيش، ٢٠٢١، ص١٤٦).

وتعرف بأنها: "قياس مدى ما يكتسبه الطالب من معلومات ومعارف وخبرات ومهارات نظرية أو تطبيقية، نتيجة ادائه في الاختبارات خلال مرحلة دراسية معينة منظمة، ويمكن قياس التحصيل من خلال التعرف على متوسط الدرجة التي يحصل عليها الطالب في الاختبار" (الزبيدي، ٢٠١٨م: ص٤٦٤).

وعُرفت بأنها:" الاختبارات الذي يحقق المعايير التي يسعى إليها نظام التعليم، ويلتزم بها المعلمين عند إعداد الاختبارات التحصيلية"(البادي، ٢٠٢١م: ص٧٩).

وتعرفُ الاختبارات التحصيلية بأنها: "أداة تستخدم لتحديد مستوي اكتساب المتعلم المعلومات والمهارات في مادة دراسية تعلمها مسبقا، من خلال إجاباته على عينة من الأسئلة التي تمثل محتوي المادة الدراسية"(أيوب، ٢٠٢٢م: ص٤٢٨).

- **المطلب الثاني: أهمية الاختبارات التحصيلية.**

وتحدد أهمية الاختبارات التحصيلية لكل من الطلبة والمعلمين وصناع القرار، على النحو الآتي:

- **أولاً: أهمية الاختبارات التحصيلية للطلبة [Abu Foudeh, 2020) حيث إنها:**
  - تزيد من دافعية الطلبة لتعلم المواد الدراسية والمراجعة المستمرة للمعلومات في المواد الدراسية التي سبق تعلمها.
  - تؤثّرُ الاختبارات في العمليات المعرفية التي ستخدمها الطلبة أثناء تعلمهم.
  - توفرُ الاختبارات خبرات تعلم ذاتها وبذاتها للطلبة.
  - توفرُ الاختبارات تغذية راجعة عن مدى تعلم الطلبة.
- **ثانياً: أهمية الاختبارات التحصيلية للمعلم [Abu Foudeh, 2020):**

فهي تساعد المعلم في التعرف على مستوى التحصيل الدراسي الذي وصل إليه الطلبة، ومراقبة التغيرات التي تطرأ على عملية تعلم الطلبة، بالإضافة في أنها تسهم في رفع مستوى طموح المعلم وزيادة الجهد الذي يبذله لرفع جودة العملية التعليمية

- **ثالثاً: أهمية الاختبارات التحصيلية للمؤسسة التعليمية [الزبيدي، ٢٠١٨]:**

وتأتي أهمية الاختبارات من كونها وسيلة للحكم على فاعلية العملية التعليمية والتربوية، من خلال ما تقوم به من دور فعال فيها، حيث تساعد

على تشخيص العقبات والمشكلات التعليمية، وتقديم الحلول المناسبة وأوجه حلها، وهي مهمة للمؤسسة التعليمية بحيث تدفعها لمراجعة أهدافها، ومدى ملاءمة المنهج لتحقيق هذه الأهداف، كما تساعد على معرفة مدى تحقيق الخطة التعليمية للأهداف الخاصة بها

ولهذا يتبين أن الاختبارات التحصيلية تلعب دوراً مهما نحو جميع أطراف ومحاور العملية التربوية والتعليمية بما في ذلك الإدارة، حيث تزود صناع القرار التربوي بمعلومات تستخدم في إصدار القرارات التربوية الإدارية، والتي قد تتمثل في رفع أو نقل الطلبة من مستوى دراسي إلى مستوى دراسي أخر.

- **المطلب الثالث: أهداف الاختبارات التحصيلية**

أن الاختبارات التحصيلية تسمح للمعلم بوضع محصلة جرد للكفاءات المكتسبة أو تركيب للتعليمات، حيث تتخذ قرارات المتعلقة بتوجيه التلاميذ حسب مكتسباتهم والمقصود بالتوجيه هنا توجيه تعلم التلاميذ. وتهدف الاختبارات التحصيلية كما يرى (فاتح، وآخرون، ٢٠١٩)، (مزارة، ٢٠٢١) إلى ما يلي:

- قياس مدى استيعاب الطالب للمعرفة، والفهم والمهارات المتعقلة بالمادة الدراسية في وقت معين، ونقصد بالمعرفة ما يملكه الطالب من معلومات، والفهم يتضمن القدرة على التعبير بطرق عديدة والمهارات هي معرفة عمل شيء معين.
- التوصلُ بالتلميذ إلى أنسبِ الطرق للوصول إلى أفضل أداء مُمكن في التحصيل.
- اكتساب الاستعدادات العقلية والمزاجية المختلفة المتوفرة لدى التلميذ وتتبع عملية نموه بجوانبها المختلفة العقلية والتحصيلية.
- توجيهٍ الطالب دراسيا لنوع الدراسة والتخصص الملائم لقدارته العقلية ومهنيا بالنسبة للمهن والأعمال.
- معرفة مدى ملائمة المناهج الدراسية لمستويات العقلية من عدمه.
- الكشف عن حاجات الطلبة ومشكلاتهم، وقدراتهم، وميولهم، والطلبة المتأخرين، وأسباب هذا التأخير.
- مساعدة المعلم على تطوير أساليبه وطرائقه في التدريس من خلال معرفته لمدى تقدمه نحو بلوغ أهدافه.
- معرفة الفروق بين التلاميذ أنفسهم وأوجه القصور في تخلف التلاميذ في بعض المواد الدراسية.
- تعد مؤشرات جيدة عن قدرة وإمكانية الطلبة ومستوي تحصيلهم ونشاطهم من خلال التمكن من تصنيفهم ومعرفة مستواهم من أجل وضع خطط وأساليب تدريسية واستثارة دافعيتهم للتعلم، وجعلهم قادرين على التفاعل الصفي وتحسين عملية التعلم وتسهيل مهمتها.

وعلى ضوء وظائف الاختبارات التحصيلية يرى (الزبيدي، ٢٠١٨) أن أهدافها الرئيسية تتمثل في الآتي:

- التشخيص: أي محاولة تعرف جوانب القوة والضعف لدى الطالب في جانب من جوانب التحصيل للاستفادة من النتائج في تدعيم جوانب القوة ومعالجة جوانب الضعف مع ما يستدعيه ذلك من تقويم لأسلوب التدريس أو المناهج أو المرافق التعليمية المختلفة ومصادر التعلم، مع الإشارة إلى أنّ هناك اختبارات خاصة بالتشخيص، ولكن ذلك لا ينفي عن الاختبارات التحصيلية العادية وظيفة التشخيص.
- التصنيف: أي تصنيف الطلاب إلى تخصصات مختلفة: أكاديمي، تجاري، صناعي، زراعي، وما إلى ذلك، أو تصنيفهم إلى مجموعات اعتمادا على قدرتهم العقلية أو ميولهم، وبطبيعة الحال لا يكون هذا التصنيف ممكنا إلا بالاعتماد على نتائج الطلاب في اختبارات تحصيلية، أو اختبارات خاصة، أو وسائل قياس أخرى في مقابلات شخصية أو استبانات وقوائم وما إلى ذلك.
- قياس مستوي التحصيل: مدى تحقق الأهداف التعليمية لدى المتعلم في مادة دراسية بعينها، وفي المواد الدراسية جميعها، والاختبارات التحصيلية في معظمها إنما تنصبّ على تحقيق هذا الهدف؛ بقصد الأخذ بنتائجه في تحسين مستوي التعليم وترفيع الطلاب إلى صفوف أعلى، وما إلى ذلك.
- تنشيط دافعية التعلم: والنقل من صف إلى آخر، ومنح الدراجات والشهادات.
- التعرف على مجالات التطوير للمناهج والبرامج والمقررات الدراسية.

• المطلب الرابع: تصنيف الاختبارات التحصيلية:

يمكن تصنيف الاختبارات التحصيلية بطرق مختلفة أو جوانب متعددة ولكل منها أبعاده الهامة أو الرئيسية التي تختلف باختلاف مستوي الأداء التعليمي المراد تحقيقه، ووفقا لما تضمنته دراسات كلا من (فاتح، وآخرون، ٢٠١٩)، (مزارة، ٢٠٢١) (خنيش، ٢٠٢١) (الزبيدي،٢٠١٨) فإن الاختبارات التحصيلية تتصنف على النحو الآتي:

• أولاً: تصنيف الاختبارات على وفق الأغراض التي نستخدم فيها:

يتمثل هذا النوع في استخدام الاختبارات الآتية:

- الاختبارات التي تستخدم في بداية الفصل أو المرحلة الدراسية.
- الاختبارات التي تستخدم في عملية قياس نمو التلميذ التحصيلي أثناء تطبيق وتنفيذ المقرر الدراسي، وتلك تُسمّى الاختبارات البنائية.
- الاختبارات المتعلقة بتشخيص الجوانب الإيجابية والسلبية لتحصيل الطلبة والصعوبات التي تواجههم ويطلق عليها أحيانا بالاختبارات التشخيصية.

- **ثانياً: تصنيف الاختبارات وفق استجابات التلاميذ المرجوة:**

وفي هذا المجال تصنف الاختبارات التحصيلية إلى ثلاثة أنواع هي:

- الاختبارات التحريرية (القلم والورقة): وهي إحدى وسائل التقويم الأكثر انتشارا بين أوساط المعلمين بهدف قياس وتقويم المستوي التحصيلي لتلاميذهم، وتكون على نوعين هما:
  - ✓ اختبارات التعرف: هي تلك الأداة التقويمية التي يتم من خلالها قيام الطالب باختبار الإجابة الصحيحة من بين اجابات أو بدائل لقياس نواتج التعلم من البسيط إلى المعقد ويطلق عليها بالاختبارات الموضوعية، لأنّ إجابات التلاميذ عليها محددة لا يختلف اثنان على تصحيحها، وتشمل:
    - ✓ اختبارات المطابقة (المزاوجة).
    - ✓ اختبارات الصواب والخطأ.
    - ✓ اختبارات الاختيار من متعدد.
  - ✓ اختبارات الاستدعاء (الاسترجاع): في هذا النمط الاختباري يستخدم المعلم بعض الأسئلة التي يعدها مسبقا لتلاميذه التي قد تتطلب إجاباتهم عليها إجابة قصيرة أو تكملة عبارة أو جملة تلك الفراغات فيها أو أسئلة مقاليه (تقليدية) ذات إجابات غير محددة أي مستفيضة بحيث تتيح للطالب حرية التعبير في تنظيم المعلومات أو الأفكار التي يختارها للإجابة على هذا النمط من الأسئلة لقياس نواتج التعلم.
- الاختبارات العملية (المهاراتية): يهدف هذا النوع من الاختبارات إلى تقييم أداء أو مهارة المتدرب وفعاليته ودقته في العمل ؛ أي أنها ترتبط بالجوانب ذات العلاقة بالمهارات الحركية لتحقيق الأهداف المعرفية والمقررات التي تتعلق بالجانب الأدائي ذات الصفة العملية أكثر من الاهتمام بالجوانب الأخرى ويتجلى استخدام هذا النوع من الاختبارات في الميادين المختبرية والصناعية والموسيقية وآلات التدريب على الطباعة والكتابة بحيث تكون جزءا متمما للعمل الدراسي لما تنعكس فوائدها على ترجمة الجوانب النظرية إلى واقع عملي مدروس في موضوعات الفيزياء والكيمياء وعلوم الحياة وتطبيقاتها المختبرية وبرامج التدريب المهني والموسيقي.
- الاختبارات الشفوية (التعبيرية): هي تلك الوسيلة الاختبارية التي هي تلك الوسيلة الاختبارية التي تستخدم في عملية التقويم التحصيلي للطلاب بهدف قياس قدرتهم على النطق والتعبير اللفظي ومدى فهمهم أو استيعابهم لموضوع الدرس عن طريق توجيه بعض الأسئلة الشفوية لهؤلاء التلاميذ والطلب منهم الإجابة عنها بصورة شفوية أيضا، ولقد استخدمت هذه الوسيلة على نطاق واسع بين أوساط المعلمين منذ القدم، وهي

الوسيلة التقويمية المفضلة لديهم كونها تعد المناسبة لتقويم بعض الأهداف التربوية جنبا إلى جنب مع الاختبارات التحريرية.

- **المطلب الخامس: خطوات بناء الاختبار التحصيلي:**

تتمثل خطوات بناء الاختبار التحصيلي كما وردت لدى (فاتح، وآخرون، ٢٠١٩)، (أحمد، ٢٠٢٠)، (عتوم، ٢٠١٤) في الخطوات الاجرائية الرئيسية التالية:

- **أولاً: تحديد الغرض من الاختبار**
- **ثانياً: تحديد أهداف الاختبار**
- **ثالثاً: تحديد مادة الاختبار**
- **رابعاً: إعداد جدول المواصفات:**

ومن أهم فوائد جدول المواصفات (أحمد، ٢٠٢٠):

- يضمن صدق الاختبار؛ لأنه يجبر المدرس على توزيع أسئلته على مختلف أجزاء المادة وعلى جميع الأهداف.
- يضمن تمثيل جميع مستويات الأهداف عند بناء الاختبار.
- يعطي كلّ جزءٍ من المادة وزنه الحقيقي، وذلك بالنسبة للزمن الذي أنفق في تدريسه وكذلك حسب أهميته.
- عرض تحليل المحتوى للمادة الدراسية والوزن النسبي لكلٍّ من المحتوى والأهداف التعليمية.

- **خامساً: وضع الأسئلة**

يُراعى في وضع الأسئلة اعتبارات كثيرة منها (فاتح، وآخرون، ٢٠١٩):

- الابتعاد عن نقل عبارات معينةٍ من الكتب في صيغة أسئلة.
- أن يكون كلّ سؤال مستقلا عن الآخر، بحيث لا تبني الإجابة عن سؤالٍ معين على الإجابة على سؤال سابق.
- أن تكون اللغة صحيحة وواضحة.

ويراعي في تصميم الأسئلة محاولة صياغتها في أصلح الصور التي تضمن الوصول إلى الغرض من القياس، كأن يكون بعضها من النوع الذي تطلق فيه الحرية للإجابة وبعضها مع تكميل الجمل إلى غير ذلك.

- **سادساً: صياغة تعليمات الاختبار:**
  - لابُدّ أن تكون تفصيلية وسهلة وواضحة وموجزة.
  - تعدّ قبل كتابة مفردات الاختبار.
  - تعدّ قبل تجربة الاختبار.
  - يجبُ أن تشمل الهدف من القياس وتتضمن البيانات الخاصة بالأفراد وطريقة الإجابة على الأسئلة (أحمد، ٢٠٢٠).

• **سابعاً: التجربة الاستطلاعية للاختبار:**

يجبُ تطبيق الاختبار في صورته المبدئية على عينة غير منتقاة من التلاميذ أي تمثلُ جميع مستوياتهم. فالتجربة الاستطلاعية للاختبار تحدد مدي صلاحية الاختبار للاستخدام، وذلك لحصول على الآتي (أحمد، ٢٠٢٠):

- حساب ثبات الاختبار.
- حساب صدق الاختبار.
- تحديد زمن الاختبار.

وبشكل عام يجب عند تصميم الاختبار التحصيلي أن يناسب الغرض منه، حيث تتعدد أغراض الاختبارات التحصيلية ويتم تفسيرها بحرص وأن تستخدم هذه النتائج لتحسين عملية التعلم، فينبغي للمتعلم أن يعرف أن الاختبار ليس مجرد وسيلة لمعرفة مستوي التحصيل، بل إن هدفه الرئيسي تحسين تعلمه وتقويم نواحي القصور.

ومن خلال ما سبق يتضح أن خطوات بناء الاختبار تتمثل في الآتي:

- تحديد الغرض من الاختبار: فلابد من تحديد الغرض من الاختبار، لاسيما وأن هذه الخطوة تسهم في التوجيه للخطوات التالية في عملية إعداد الاختبار.
- تحديد محتوى الاختبار: وذلك بتحديد الموضوعات التي ستتضمن في الاختبار.
- تحديد الأهداف التعليمية: حيث يتمّ تحديد الأهداف السلوكية (نواتج التعليم) المراد تحقيقها من تدريس المادة الدراسية.
- تحليل المحتوى: وذلك لضمان تحقيق الشمولية والتوازن في تغطية جميع موضوعات المادة الدراسية.
- إعداد جدول المواصفات: والذي هو عبارة عن مخطط تفصيلي، يحدد محتوى الاختبار ويربط محتوى المادة الدراسية بالأهداف التعليمية، ويحدد الأوزان النسبية لكل موضوع من موضوعات المحتوى التعليمي، والأوزان النسبية للأهداف التعليمية في مستوياتها المختلفة.
- كتابة الاختبار: حيث تكتب فقرات الاختبار بصورة أولية، ثم تراجع لحين صياغتها بصورة نهائية.
- تحليل الاختبار: وهي الخطوة التي تتم بعد الانتهاء من تطبيق الاختبار؛ لحساب قيم مؤشرات النزعة المركزية والتشتُّت، للكشف عن مدى جودة الفقرات لاستبعادها أو التعديل عليها.

• **المطلب السادس: صفات الاختبارات التحصيلية السليمة**

تعتبر الاختبارات واحدة من الطرق الأربع المهمة في جمع البيانات الكمية عن السمات أو الخصائص التي يتمتع بها الطلبة، كما أنها تعد من أهم

الوسائل التي يعتمد عليها في قياس وتقويم قدرات الطلبة والكشف عن نقاط القوة لتعزيزها ونقاط الضعف لوضع الخطط لمعالجتها، وبالتالي الارتقاء بالعملية التربوية ومخرجاتها، كما تلعب دورا هاما ومميزا في الأبحاث التربوية الوصفية والارتباطية والتجريبية.

ولكي تكون الاختبارات التحصيلية ناجحة وفاعلة في تحقيق أهدافها ووظائفها والغايات التربوية والتعليمية منها، فلا أن تتحقق فيها مجموعة من المعايير والمواصفات العلمية والمنهجية. وتلخص الباحثة أهمها كما وردت عند (عتوم، ٢٠١٤)، (ساعد، وآخرون، ٢٠١٧) في المواصفات التالية:

- الصدق: وهو أن يقيس الاختبار ما أُعدّ لقياسه
- الثبات: وهو أن يعطي النتائج نفسها على نفس المجموعة إذا طبّق في نفس الظروف؛ بشرط أن لا يحدث تعلم أو تدريب في الفترات بين مرات إجراء الاختبار.
- الموضوعية: أن تكون علامات الاختبار معبرة عن الأداء الفعلي وأن لا تختلف علامته باختلاف طريقة التصحيح أو المصحح أي لا تتدخل ذاتية المصحح أو ظروف التصحيح بالنتائج.
- الوضوح: أن تكون فقرات الاختبار واضحة وغير مخادعة وغير مموهة.
- العدالة: أن تكون فقرات الاختبار مناسبة لمستوى الطلبة العقلي والزمني.
- الواقعية: أن يكون الاختبار في حدود الظروف والإمكانات المتاحة وأن يتناسب طوله مع الوقت المسموح به للتطبيق.
- سهولة التطبيق: أن تكون تعليمات الاختبار واضحة ومفهومة لأن عدم وضوح التعليمات يؤدي إلى انخفاض في علامات الطالب.
- سهولة التصحيح: أن يكون مفتاح الإجابة معدّ مسبقا حتى لا تتدخل ذاتية المصحح في النتائج وخاصة في الأسئلة المقالية.

وعليه من خلال -ما سبق- يتضح أنّ من أهم شروط الاختبار الجيد مايلي:

- أن يتمّ التخطيط لها بشكل جيد، ويتم بناؤها وفق أسس علمية دقيقة.
- أن تحددَ مخرجات التعليم التي سيقيسها الاختبار بشكل دقيق.
- أن تتنوع الأسئلة في الاختبار بما يتناسب لقياس نواتجً ومخرجات التعلم المطلوب قياسها.
- أن تكون أسئلة الاختبار ممّا درسه الطالب في المحتوى الدراسي.
- أن يسعى الاختبار لقياس وتوظيف للمعارف والحقائق والمبادئ المختلفة التي درسها الطلبة.
- أن يكون الاختبار مناسبًا لمستوى الطلبة الذين نسعى لتطبيق الاختبار عليهم.

- يجبُ أن يحتوي الاختبار على أسئلة موضوعية وأخرى مقاليه، لتستطيع قياس قدرة الطلبة على التعبير، وتنظيم الأفكار، والتحليل وتطبيق القوانين والقواعد على مواقف جديدة.

• المبحث الثاني: نظرية إستجابة الفقرة

• المطلب الأول: مفهوم نظرية إستجابة الفقرة.

تمثل نظرية استجابة الفقرة أساس نظري جديد في القياس النفسي والتربوي يقوم على افتراض وجود علاقة بين قدرة الفرد واحتمالية الإجابة الصحيحة على الفقرة الاختبارية؛ وفق دالة رياضية محددة، مستندة في ذلك على نظرية الاحتمالات، ولهذه النظرية مجموعة من الفروض لابدّ من تحققها حتى يمكن استخدامها، كما نستطيع من خلالها معادلة الفقرات الاختبارية على نفس التدرج عند وجود أكثر من نموذج اختباري (عالم، وآخرون، ٢٠٢٠). وتقوم نظرية الاستجابة للفقرة على مبدأ وجود سمة واحدة تفسر أداء الطالب، وتوضح العلاقة بين اجابته على اختبار معين والسمة المتمثلة وراء تلك الاستجابة (Casabianca& Lewis.2015).

ولهذا تعرّف نظرية الاستجابة للفقرة بأنها إحدى أهم الثورات الحديثة في مجال القياس النفسي والتربوي، التي ارتبطت بنظرية السمات الكامنة، وعدت أسلوبا مبتكرا لنمذجة البيانات، من خلال إنشاء نماذج تؤدي دورا مهمًّا في القياس والاختبار التفاعلي في الوقت الحالي ؛ وقد اشار (Hambleton and Jones,1993) إلى أن نظرية الاستجابة للفقرة تعتبر من التطورات الحديثة والمهمة جدا، لما تمتاز به من دقة وكفاءة وموضوعية في قياس السلوك، لأنها تغلبت على العديد من العيوب والمشكلات وجوانب القصور والضعف في النظرية التقليدية في القياس، التي كانت تعيق تطوير وبناء الاختبارات التربوية والنفسية لفترة طويلة، ولكن ثبت أن لديها قصور في العديد من الجوانب والنواحي المختلفة (العطيان، ٢٠١٧).

وتعد نظرية الاستجابة للفقرة (Item Response Theory) من التطورات الحديثة في القياس النفسي والتربوي، بسبب ما قدمته من طرق سيكومترية ذات فاعلية كبيرة في بناء المقاييس النفسية والتربوية، وفي طريق تفسير الدرجات على هذه المقاييس بالطريقة التقليدية في القياس (بني عواد، ٢٠١٨).

وقيل بأنها: "نظرية حديثة في القياس النفسي والتربوي، ويتم فيها تحديد العلاقة بين أداء المفحوص والسمة الكامنة موضوع القياس وفق دالة رياضية محددة"(الحاج محمود، ٢٠١٨:ص ١٢٧).

وقيل بأنها: "أساس نظرية جديد في القياس النفسي والتربوي يقوم على افتراض علاقة بين قدرة الفرد واحتمالية الإجابة الصحيحة على الفقرة

الاختبارية وفق دالة رياضية محددة، مستندة في ذلك على نظرية الاحتمالات، ولهذه النظرية مجموعة من الفروض لابد من تحققها حتى يمكن استخدامها" (عالم، ٢٠١٩م: ص١٦٦).

وقيل كذلك بأنها: "نظرية حديثة في مجال القياس تعتمد على نماذج رياضية، توضح العلاقة بين القدرة واحتمال استجابة الطالب الذي يمتلك قدرة معينة استجابة صحيحة على الفقرة" (المرواني، وآخرون، ٢٠١٩م: ص٧٠).

وتعرف أيضاً هذه النظرية بأنها: "نظرية إحصائية عامة بشأن الفقرة الاختبارية وأداء الفرد المفحوص على الاختبار وكيفية ارتباط الأداء بالقدرات التي تقيسها فقرات الاختبار، ويمكن أن تكون استجابات الفقرة منفصلة أو مستمرة، ويمكن أن تكون الدرجات ثنائية أو متعددة، ويمكن أن تكون درجة أقسام الفقرة مرتبة أو غير مرتبة، ويمكن أن تكون هناك قدرة واحدة أو العديد من القدرات الكامنة خلف الأداء على الاختبار" (علي، ٢٠٢١م: ص١٤٢).

ولهذا فإنّ هذه النظرية - كما يعرفها البعض- قد هدفت هذه النظرية إلى التغلب على مكامن الضعف في النظرية الكلاسيكية للقياس، ووفرت طرقاً واليات دقيقة لمعالجة القضايا الرئيسية في القياس التربوي والنفسي، نحو بنوك الأسئلة، والقياس التكيفي، وغيرها من التطبيقات (Casabianca& Lewis.2015).

- **المطلب الثاني: افتراضات نظرية الاستجابة للفقرة [Item Response Theory]:**

تقوم نظرية الاستجابة للفقرة على مجموعة من الافتراضات، حيث تفترض هذه النظرية أن أداء المفحوصين في الاختبار يمكن تفسيره عن طريق السمة أو السمات الكامنة (Latent Traits) المراد قياسها، والتي لا يمكن قياسها بصورة مباشرة، إذ يتم استخدام الدرجات التي تم تقديرها للمفحوص في تلك السمة في التنبؤ بأدائه في اختبار ما أو في فقرة من الاختبار، لأنّ العلاقة الحقيقية بين درجات المشاهدة (الخام) للمفحوص، والسمة المراد قياسها لا يمكن الحصول عليها بطريقة مباشرة، ومن هنا، فإن نظرية الاستجابة للفقرة تصف هذه العلاقة بواسطة دالة تعتمد على مجموعة من الافتراضات.

حيث بنيت نظرية الاستجابة للفقرة (IRT) على مجموعة من الافتراضات ذكرها هامبلتون وسواميثان (Hambleton & Swaminathan, 1985). وهي افتراض أحادية البعد، والاستقلال الموضعي، ومنحنى خصائص الفقرة، وافتراض السرعة (Ha:2017). Fontanella, L. (Villano, & Donato .2015)، وتفصيل ذلك على النحو الآتي:

- الافتراض الأول: افتراض أحادية البعد (Unidimentionlity):
- الافتراض الثاني: فيتمثل بالاستقلال الموضعي (Local Indpendence):
- الافتراض الثالث: فيتمثل بمنحنى خصائص الفقرة (Item Characteristic Curve):
- الافتراض الرابع: فيتمثل بافتراض السرعة (Speediness):

وفي ضوء- ما سبق- تستند نظرية استجابة الفقرة على مسلمتين رئيسيتين:

- الأولى: إمكانية التنبؤ بأداء الفرد على فقرة من فقرات الاختبار عن طريق عامل أو عدّة عوامل؛ تسمّى السمات أو القدرات الكامنة.
- الثانية: يمكن وصف العلاقة بين أداء الفرد على أحد فقرات الاختبار بواسطة دالة تسمّى الدالة المميزة للمفردة (Item Characteristic function) أو المنحنى المميز للمفردة (Item Characteristic) (الشمراني، ٢٠١٦).

• المطلب الثالث: مميزات نظرية استجابة الفقرة.

وجهت الأنظار إلى نماذج نظرية الاستجابة للفقرة لما تتميزُ به من وجود مجموعة كبيرة من فقرات الاختبار تقيس نفس السمة، ويكون تقدير قدرة الفرد مستقلا عن عينة الفقرات (**Item Free**) التي تطبق عليه، ووجود مجتمع كبير من الأفراد، بحيث تكون الخصائص السيكومترية للفقرات مستقلة عن عينة الأفراد (**Person Free**) التي استخدمت في تقدير هذه الخصائص، كما أنّ تفسير درجات الأفراد يتمّ في ضوء الفقرات، وليس في ضوء الجماعة المرجعية كما في النظرية الكلاسيكية، كما أنه يتم تقدير الخطأ المعياري لكلّ مفحوص على حدة وليس باعتباره خطأً معياريا واحدا لكل المفحوصين، مع إمكانية المقارنة بين أداء الأفراد الذين اختبروا باختبارات مختلفة تقيس نفس السمة أو القدرة دون أن تكون هذه الاختبارات متوازنة (أبو وردة، ٢٠٢٠).

تمتاز نظرية استجابة الفقرة بعدة خصائص تجعلها أكثر ملائمة لتلبية حاجات الباحثين ومطوري الاختبارات، فقد أشار البعض إلى هذه المزايا التي تتمثل في الآتي (الشمراني، ٢٠١٦) (الهواوشة، ٢٠١٥).

- إمكانية تقدير الخصائص السيكومترية للفقرات، مثل الصعوبة والتمييز والتخمين بشكل مستقل عن عينات الأفراد المستخدمة في تقدير هذه الخصائص.
- تقدير قدرة الفرد مستقل عن عينة الفقرات التي طبق عليها الاختبار.
- توفر النظرية مؤشرا إحصائيا لتقدير درجة قياس قدرة كل فرد.
- خصائص الفقرة لا تعتمد على عينة المفحوصين.

- خصائص المفحوصين لا تعتمد على عينة الفقرات.
- قدمت النظرية الحديثة نماذج رياضية تعبر عن مستوى الأداء على الفقرة بدلا من التعبير عن مستوى أداء الفرد عن الاختبار.
- قدمت النظرية الحديثة نماذج رياضية تعطينا دقة القياس عند كل مستوى للقدرة، بعكس النظرية الكلاسيكية التي تعطي دقة القياس للاختيار ككل.

• المطلب الرابع: نماذج نظرية الاستجابة للفقرة:

لقد انبثقت عن هذه النظرية مجموعة من النماذج تعرف بنماذج السمة الكامنة (**Latent Traits Model**) وهي نماذج رياضية تستخدمُ لتقدير المعالم التي تصف الفقرة والقدرات الكامنة التي تصف الأفراد، نظرا لأنّ مستويات السمة غير ملاحظة بطبيعتها وتؤثر في المتغيرات الملاحظة، ومن خلال النماذج يمكن اشتقاق تقديرات للسمة تنطوي عليها لمجموعة من الفقرات، وهذه التقديرات لها خصائص تفسيرية (تفسير العلاقة بين الاستجابة للفقرة والسمة الكامنة التي تقيسها) ومن خلال الصيغ الرياضية تحدد احتمالية الاستجابة الصحيحة لفقرات مختلفة.

ومن أهم نماذج نظرية الاستجابة للفقرة الشائعة (عودة، ٢٠١٦):

- أولاً: النموذج اللوغاريتمي أحادي المعلم (نموذج راش)
- ثالثا: النموذج اللوغاريتمي ثلاثي المعلم (نموذج لورد):

• المطلب الخامس: نماذج نظرية الاستجابة للفقرة المتعددة.

- نموذج الاستجابة المتدرجة (GRM):
- نموذج التقدير الجزئي (partial credit model):
- نموذج سلم التقدير (Rating Scale Model):

• المطلب السادس: المعالم في نظرية الاستجابة للفقرة:

توجد ثلاثة معالم رئيسية مستخدمة في نظرية الاستجابة للفقرة، وهي:

- أولاً: معلمة الصعوبة (موقع الفقرة): (Item location or item difficulty parameter):
- ثانياً: معلمة التمييز (discrimination parameter):
- ثالثاً: معلمة التخمين (guessing parameter):
- رابعا: معلمة اللامبالاة (Carelessness parameter):

• المطلب السابع: تقدير المعالم للفقرات والأفراد وفق نظرية الاستجابة للفقرة

• أولاً: طريقة الأرجحية العظمى:

ويوجد ثلاث طرائق تعتمد عليها الأرجحية العظمى تستخدم في التقدير منها:

- طريقة الأرجحية العظمى المشتركة (Estimation Joint Maxmum Likelihood):
- طريقة الأرجحية العظمى الشرطية (Estimation Conditional Maxmum Likelihood):
- طريقة الأرجحية العظمى الهامشية (Estimation Maxmum Likelihood):

• ثانياً: الطريقة البيزية [BME: Bayesian Method Estimation]:

تعتمد هذه الطريقة على تقديرات بيزن(**Bayesian Estimation**) وتستخدم معلومات عن العينة، وتطبق عندما لا يستطيع المفحوص أن يجيب عن فقرات الاختبار إجابة خاطئة أو صحيحة، أي عندما لا يكون بالإمكان تطبيق طريقة الأرجحية العظمى، وتعطي تقديرات جيدة للمعالم بحال توافر معلومات عن توزيع القدرة عند الأفراد المفحوصين، وسميت بهذا الاسم لأنها تستخدم نظرية بيز في حساب الاحتمال المشروط(العطيان، ٢٠١٧)، (أبو السل، ٢٠١٦).

وتستخدم هذه الطريقة عندما لا نستطيع تطبيق طريقة الأرجحية العظمى، وتتميز هذه الطريقة بالإضافة إلى استخدامها معلومات عن العينة، بأنها تستخدم معلومات أولية متوافرة من خبرات سابقة.

ومن أجل اختيار النموذج الأمثل لمطابقة بيانات البحوث المختلفة، يقترح (**1985,Hambleton & Swaminathan**) و(**et al Hambleton, 1991**) معايير لتقييم نماذج نظرية الاستجابة للمفردة وهي:

- التحقق من افتراضات النموذج.
- التحقق من الخصائص المتوقعة للنموذج.
- التحقق من قدرة النموذج على التنبؤ بالدرجات الفعلية

• الدراسات السابقة.

أ] الدراسات العربية:

دراسة العلي(٢٠٠٤م) بعنوان: " فاعلية نموذج التقدير الجزئي في بناء أسئلة من فقرات متعددة الخطوات في مادة الكيمياء في الصف الثاني الثانوي العلمي حيث تلخصت أهم نتائج الدراسة في: إن الخطأ المعياري لتقدير درجة صعوبة الأسئلة يزيد بزيادة عدد خطوات الاجابة. تبيّن وجود فروق ذات دلالة إحصائية بين متوسطات الفروق المعايرة لتقديرات معالم صعوبة الأسئلة المحسوبة في المجموعات الأربع المختلفة، وهي: المجموعة الكلية والمجموعة العشوائية، والمجموعة العليا، والمجموعة المتدنية. كما وجدت فروق ايضا في متوسطات الفروق المعايرة بين تقديرات قدرات أفراد المجموعة العشوائية والعليا والمتدنية؛ فيما أظهرت نتائج الدراسة عدم وجود فروق دال إحصائيا

في وسط الفرق المعاير بين تقديرات قدرات المجموعة الكلية المحسوبة في الاختبار الصعب والاختبار السهل. كذلك توصلت الدراسة إلى أن الكفاية النسبية للاختبار بشكل عام تزيد بزيادة خطوات الإجابة، حيث تتساوى الكفاية النسبية للاختبارً الصعب مع الاختبار الكلي عند أصحاب القدرات العالية لكنها تزداد للاختبار السهل عند أصحاب القدرة المتدنية جداً.

دراسة الخولي(٢٠٠٦م) بعنوان "أثر طرق معادلة درجات الاختبار في تدرج مفردات بنك الأسئلة باستخدام نموذج التقدير الجزئي". حبث تلخصت أهم نتائج الدراسة في أن استخدام طرق مختلفة لمعادلة درجات مفردات بنك الأسئلة الموزعة على اختبارين أحدهما سهل (X) والآخر صعب (Y) والمدرجة باستخدام نموذج التقدير الجزئي يؤدي إلى إنتاج مقاييس مشتركة لمفردات الاختبارين (Y،X) ولكنها مختلفة فيما بينها من جهة وفيما بينها وبين المقياس المرجعي من جهة أخرى.

دراسة العبدالله(٢٠٠٩م) بعنوان: "أثر حجم العينة وعدد المفردات المشتركة على تدرج مفردات بنك الأسئلة باستخدام نموذج التقدير الجزئي". وقد توصلت الدراسة إلى نتائج، من أهمها: تتناسب دقة تدرج مفردات بنك الأسئلة طردا مع حجم العينة، حيث بلغت نسبة المفردات غير الملائمة (٥,١٪) من العدد الكلي لمفردات البنك، وذلك عند انخفاضٍ حجم العينة من (١٠٠٠) إلى (٨٠٠)، في حين أنّ ارتفاع هذه النسبة كان واضحاً عند انخفاض حجم العينة إلى (٤٠٠)، حيث بلغ (٨,٥٪)، وهذه النسبة تعتبر غير مقبولة بناءً على المحكات التي حدّدتها الدراسة، وتتناسب دقة تدرج مفردات بنك الأسئلة طردا مع عدد المفردات المشتركة والمستخدمة في معادلة الصور الاختبارية.

دراسة سليمان؛ والصالح(٢٠١٦م) بعنوان: "بناء اختبار تحصيلي محكي المرجع في القياس والتقويم باستخدام النموذج اللوغاريتمي ثلاثي المعلم لنظرية الاستجابة للمفردة الاختبار". وتوصلت نتائج الدراسة إلى أن افتراضات نظرية الاستجابة للمفردة الاختبارية قد تحققت في بيانات الدراسة، ومطابقة الاستجابات لعدد (٣٣) مفردة من مفردات الاختبار لافتراضات النموذج اللوغاريتمي ثلاثي المعلم ، حيث كانت الاستجابات لتلك المفردات بصورتها النهائية ضمن حدود المطابقة، وتم التأكد من تمتع الاختبار بمظاهر متعددة من الصدق، وحذف (٧) مفردات اختبارية لم تطابق النموذج، كما أظهرت النتائج المتعلقة بتقديرات معالم المفردة (الصعوبة، التمييز، التخمين) أنها مقبولة ضمن المعايير التي أوردتها أدبيات القياس والتقويم في هذا المجال.

دراسة القطاونة(٢٠١٦م) بعنوان: " الثبات في الاختبارات محكية المرجع في الرياضيات للصف العاشر حسب النظرية الكلاسيكية ونظرية استجابة

الفقرة حسب النموذج ثنائي المعلم، دراسة مقارنة". حيث تلخصت أهم نتائج الدراسةٍ في: أن معاملات الثبات المقدرة وفق النظرية الكلاسيكية غير دالة إحصائيا وأن معامل الثبات الخاص بالفقرات ومعامل الثبات الإمبريقي، وعند مقارنة معاملات الثبات المحسوبة وفق النظرية الكلاسيكية مع معاملات الثبات المحسوبة وفق نظرية استجابة الفقرة (النموذج ثنائي المعلم في هذه الدراسة) فقد أعطي معامل ثبات لفنجستون ثبات أعلى من كل من: معاملات الثبات الخاص بالفقرات، ومعامل الثبات الأمبريقي، ولكنه كان أقلّ من معامل ثبات الأفراد.

دراسة الزغبي، وسوالمه(٢٠١٧م) بعنوان: " بناء اختبار لقياس مهارات التفكير الناقد باستخدام نظرية استجابة الفقرة". وقد تمّ التحقق من افتراض أحادية البعد، والصدق والثبات للاختبار. وتم استخدام نموذج ثلاثي المعلمة، وقد خلصت الدراسة إلى حصول المقياس على نتائج دالة ومعنوية قوية تؤكد ملاءمة محتوى الاختبار وصدق وثباته وقدراته كمقياس مناسب لما صمّم لأجله.

دراسة يوسف(٢٠١٨م) بعنوان: " تصميم تعليمي – تعلمي على وفق نماذج النظرية البنائية و أثره في تحصيل مادة الكيمياء والمهارات الحياتية والتعلم المنظم ذاتيا عند طالبات الصف الثاني المتوسط"، حيث أظهرت النتائج وجود فرق دال إحصائيا في متغيري التحصيل الدراسي في مادة الكيمياء والمهارات الحياتية بين مجموعتي البحث التجريبية والضابطة لمصلحة المجموعة التجريبية التي درست بالتصميم التعليمي- التعلمي وفق نماذج النظرية البنائية وبحجم أثر كبيرين ، في حين لم يظهر فرق دال إحصائيا بين مجموعتي البحث في متغير التعلم المنظم ذاتيا ، وأوصى البحث بضرورة اعتماد التصميم التعليمي –التعلمي وفق نماذج النظرية البنائية عند تدريس مادة الكيمياء للصف الثاني المتوسط.

دراسة عبدالوهاب (٢٠١٩م) بعنوان: " أثر طريقة تقدير الدرجات ونموذج الاستجابة للمفردة في دقة تقدير بارامترات الأفراد والمفردات لاختبار صواب وخطأ متعدد". حبث أسفرت النتائج عن وجود أثر لكل من طريقة تقدير الدرجات ونوع النموذج المستخدم والتفاعل بينهما في دقة تقدير بارامترات صعوبة المفردات وبارامترات القدرة للطلاب، كما وجد أثر لطريقة تقدير الدرجات في دقة تقدير بارامترات تمييز المفردات وفق نموذج الاستجابات المتدرجة، وقد أوصى البحث بأفضلية الاعتماد على الطريقة التي تراعي المعرفة الجزئية بأكبر قدر ممكن في تقدير الدرجات على المفردات، وكذلك أفضلية استخدام نموذج الاستجابات المتدرجة في تدريج هذه المفردات.

دراسة حسين؛ والزهرة(٢٠٢١م) بعنوان: "_ كفايات بناء الاختبارات التحصيلية الجيدة لدى أساتذة التعليم الابتدائي دراسة ميدانية على عينة

من أساتذة التعليم الابتدائي بولاية الوادي _". وتلخصت أهم نتائج الدراسة في أن مستوى كفايات بناء الاختبارات التحصيلية الجيدة لدى أساتذة التعليم الابتدائي منخفض ولم يصل إلى حدّ الاتقان والمحدّدة بـ٧٥٪. كما تبين عدم وجود فروق ذات دلالة إحصائية بين أساتذة التعليم الابتدائي في كفايات بناء الاختبارات التحصيلية الجيدة تبعا لمتغير الجنس. ولا توجد فروق ذات دلالة إحصائية بين أساتذة التعليم الابتدائي في كفايات بناء الاختبارات التحصيلية الجيدة تبعا لمتغير الأقدمية في التدريس.

دراسة خليل؛ وإسماعيل(٢٠٢١م) بعنوان: " _الأداء التفاضلي لاختبار تحصيلي الكتروني في مقرر القياس والتقويم لدى طلبة الفرقة الثالثة بكلية التربية بالجامعة باستخدام نموذج التقدير الجزئي _". حيث أظهرت نتائج البحث ما يلي: امتدت تقديرات صعوبة مفردات الاختبار التحصيلي الالكتروني بين (٢.٥٤- ١.٦٨) باستخدام الطريقة التامة لتقدير الدرجات وفق نموذج التقدير الجزئي وجميعها ضمن المدى المقبول وهو بين (٢.٩٥- ٣.٩٥). وتبين أن جميع قيم مربع كاي لمفردات الاختبار التحصيلي غير دالة إحصائيا مما يدل على مطابقة البيانات لافتراضات نموذج التقدير الجزئي باستخدام الطريقة التامة لتقدير الدرجات، وجاءت تقديرات القدرة باستخدام الطريقة الجزئية في تقدير الدرجات أعلى من الطريقة التامة في تقدير الدرجات، وجاء الخطأ المعياري لتقديرات القدرة باستخدام الطريقة التامة لتقدير الدرجات أعلى من الخطأ المعياري لتقديرات القدرة باستخدام الطريقة الجزئية لتقدير الدرجات، كما أنه لا توجد فروق ذات دلالة إحصائية بين الذكور والإناث في تقدير الأداء التفاضلي المنتظم لمفردات الاختبار التحصيلي الإلكتروني باستخدام طريقة الانحدار اللوجستي ، ما عدا المفردة رقم (١) ، وأن قيمة محك دالة المعلومات باستخدام (AIC) (Akaike's) تدل على أن أفضل نموذج للحكم على كفاءة المعلومات هو النموذج الذي يعتمد على الطريقة الجزئية في تقدير الدرجات.

دراسة محمد(٢٠٢١م) بعنوان: " الاختبارات التحصيلية كأسلوب فعّال للتقويم التربوي". حيث خلصت الدراسة بتحديد أهم صفات الاختبار التحصيلي الجيد (الصدق، الثبات، الموضوعية، الشمولية، التمييز) والدور الذي تلعبه في نجاح عملية التقويم، كما أكدت الدراسة على احترام الأستاذ لأهم خطوات بناء الاختبار التحصيلي حتى يكون صادقا. كما ختمنا مداخلتنا بمجموعة من المقترحات والتوصيات.

دراسة عليمات(٢٠٢٢م) بعنوان: " _بناء اختبار محكي المرجع في الرياضيات وفق النموذج اللوجستي ثلاثي المعلمة_".حيث أظهرت نتائج الدراسة أن تقديرات معالم الفقرات (الصعوبة، والتمييز، والتخمين) كانت

مقبولة.وحددت درجة القطع بطريقة أنجوف وكانت تساوي (٢٢). وأوصى الباحث باستخدام الاختبار أو أي اختبار جزئي منه من قبل معلمي الرياضيات، ومن قبل الباحثين، وأوصى بإفادة من المنهجية المستخدمة في بناء الاختبار لبناء اختبارات لمقررات أخرى.

دراسة فرحات(٢٠٢٢م) بعنوان: "بناء وتدريج مقياس التأمّل الناقد في ضوء نموذجي التقدير الجزئي والاستجابة المتدرجة لدى طلاب جامعة الأزهر". حيث تلخصت أهم نتائج الدراسة في: عدم تطابق (٢٠) مفردة من مفردات المقياس مع نموذج التقدير الجزئي، و(٦) مفردات، مع نموذج الاستجابة المتدرجة، وأنّ نموذج الاستجابة المتدرجة أكثر ملاءمة للبيانات من نموذج التقدير الجزئي في ضوء محك عدد المفردات المتطابقة وغير المتطابقة، ومؤشر ٢-Log Likelihood، وأعطى معاملات ثبات هامشي أعلى، ودالة معلومات أفضل عند المستوى المنخفض والمتوسط من القدرة، بينما أعطى نموذج التقدير الجزئي دالة معلومات أفضل عند المستوى المنخفض من القدرة.

ب] الدراسات الأجنبية

دراسة ونجي وهارفي (Wagner & Harvey، 2003) بعنوان: "تطوير اختبار للتفكير الناقد في ضوء نظرية استجابة الفقرة وفق اختبار واتسون جليسر". وهدفت الدراسة إلى تطوير أداة جديدة لاختبار التفكير الناقد باستخدام نظرية استجابة الفقرة، وذلك باستخدام أوجه التفكير الناقد لاختبار واتسون – جليسر. وتكونت عينة الدراسة من (٤٠٧) طالباً وطالبة من الطلبة الجامعيين، من بينهم (٢٧٩) طالبة، تتراوح أعمارهم بين ٢٧-١٨سنة. وتم تطبيق اختبار واتسون – جليسر النموذج أ، بالاضافة إلى الاختبار المطور WAT، حيث توصلت الدراسة إلى نتائج ايجابية، اشارت إلى وجود تحسن ودقة القياس خصوصاً بين المفحوصين متدني القدرة.

دراسة فاجات وباليا (Bhagat & Baliya، 2015) بعنوان: "بناء وصدق اختبار تحصيلي في العلوم " وقد تحدد الهدف العام للدراسة في: بناء اختبار تحصيل علمي موحد في مادة العلوم"، حيث أجريت الدراسة الحالية لبناء وتوحيد اختبار تحصيل في العلوم لطلاب الصف السابع لقياس إنجازاتهم. يتكون الاختبار من ١٤٦ عنصراً تغطي جميع جوانب كتاب العلوم للمعيار السابع الذي حدده C.B.S.E. بعد إجراء تحليل العناصر، تم الاحتفاظ بـ ٨٠ عنصراً في المسودة النهائية للاختبار. تم حساب الثبات بطريقة Test-retest ووجدت قيمة معامل الارتباط ٠.٨٤. تم تحديد صلاحية اختبار التحصيل بطريقة صدق المحتوى.

دراسة شافات (Shafaat، 2015) بعنوان: " بناء الاختبار وتقييمه: مراجعة موجزة " حيث توصلت الدراسة إلى حقيقة جوهرية مفادها أنّ التدريس

والاختبار لا ينفصلان، ولكي تكون مهنيًّا سليمينٍ في الحكم على أداء الطالب، فإن من الأهمية بمكان أن يكون المعلم ملما وعارفا بقواعد ومعايير إعداد وبناء الاختبارات التحصيلية الموضوعية، ولديه التزام عال بمبادئ واخلاقيات التقييم التربوي.

دراسة هيمنت وساريتا (Hemant & Sarita، 2018) بعنوان: " بناء وتقنين اختبار تحصيلي في العلوم ". حيث تبين من خلال التطبيق والتحليل الاحصائي لنتائج الاختبار أن جميع عناصر الاختبار كانت موثوقية الاختبار-إعادة الاختبار المحسوبة وقيمة معامل الارتباط ٠.٨٢ أي ان الاختبار التحصيلي الذي تم بنائه موثوق بنسبة ٨٢٪، وقد تم تحديد صدق اختبار التحصيل بواسطة طريقة تحليل المحتوى (طريقة الكفاية).

دراسة الخوالدة والعلاونة والزبون (Al Khawaldeh & Alawneh & Alzboun ، 2022) بعنوان: " توافر معايير الجودة لبناء اختبارات التحصيل العلمي من وجهة نظر لجان الامتحانات" حيث أظهرت نتائج التحليل الاحصائي لاستجابات الخبراء عدم وجود فروق إحصائية معنوية بين استجابات الخبراء على أداة الاستبانة، بحيث تعزى لمتغير الخبرة أو درجة المؤهل العلمي؛ وفي ضوء نتائج الدراسة تم تقديم عدد من التوصيات والمقترحات، أهمها: الاستمرار في توفير الوقت المناسب للرد على اختبار الأسئلة، وضرورة الاستمرار في تقديم إجابة نموذجية للأسئلة بشكل واضح.

دراسة جرمت (Gurmit، 2022) بعنوان: " بناء وتقنين اختبار تحصيلي في العلوم" وقد تحدد الهدف العام للدراسة في: بناء وتقنين نموذج اختبار موضوعي في مادة العلوم، حيث تم تطوير اختبار التحصيل في العلوم وتوحيده للصف التاسع، من قبل الباحث تحت إشراف مجلس تعليم مدرسة البنجاب. وقد تكون الاختبار في ضوء اجراءات التطوير المتعلقة ببناء اختبار الإنجاز في العلوم، من التخطيط والإعداد واختيار العناصر والتجربةٍ والتسجيل وتحليل العناصر. ويتألف الاختبار في البداية من ١٨٠ فقرة (عنصرا) في صورته الأولى؛ وبعد المناقشة مع الخبراء والإدارة وتطبيقه على الطلاب، تم تخفيض العناصر إلى ٩٠ عنصراً. وفي ضوء الاجراءات النهائية للتحليل الشامل لمحتوى الاختبار تم تجهيز الصورة النهائية للاختبار والتي تكونت من ٦٠ فقرة.

• المنهجية والإجراءات

• منهج الدراسة:

تماشيا مع طبيعة مشكلة الدراسة وتساؤلاتها؛ اتبعت الدراسة الحالية منهج البحث الوصفي بأساليبه التحليلية النوعية والاحصائية الوصفية والاستدلالية المختلفة؛ كونه المنهج المناسب لتحقيق أهداف الدراسة والإجابة على تساؤلاتها.

حيث استخدمت الباحثة الأساليب البحثية الرئيسية التالية:

- أسلوب تحليل المضمون في ضوء نظرية عناصر/مكونات المنهج الحديث (الأهداف، المفاهيم، القواعد والقوانين والحقائق)؛ وذلك لتحليل المحتوى المنهجي المستهدف ببناء اختبار تحصيلي فيه.
- أسلوب جدول المواصفات لتحديد معايير ومؤشرات بنية الاختبار الذي سيتمّ بنائه ليحقق معايير الشمولية للمحتوى المنهجي.
- نموذج التقدير الجزئي لنظرية الاستجابة للفقرة في تحديد نوع وطريقة بناء فقرات الاختبار وترتيبها ووضع العلامات المعيارية لدرجاتها.
- أساليبُ التحليل الاحصائي الاستدلال لاختبار فرضيات الدراسة.

• مجتمع الدراسة.

يتكون مجتمع الدراسة الكلي، من جميع طلاب وطالبات الصف الثاني ثانوي في المدارس الحكومية لوزارة التعليم في منطقةِ المدينة المنورة للعام الدراسي (١٤٤٤هـ)، وقد بلغ عددهم (٢٥٠٧٠) طالبًا وطالبة ، منهم (١٢٩٥٨) طالبًا و (١٢١١٢) طالبة، (مصدر البيانات: وزارة التعليم، ٢٠٢٣، احصائيات التعليم الحكومي التعليم في منطقة المدينة المنورة للعام الدراسي١٤٤٤هـ)؛ حيث تمّ اختيار هذه المرحلةٍ لتحقيق الغرض من الاختبار في قياس قدرة الطلاب في مادة الكيمياء وفقا لمنهج الكيمياء المقررة من قبل وزارة التعليم في المملكة العربية السعودية .

• عينة الدراسة:

نظرا لكبر مجتمع الدراسة بحيث يصعب مسح جميع أفراده، فقد لجأت الباحثة إلى اختيار عينة بحثية جزئية من مجتمع البحث الكلي؛ بأسلوب العينة العشوائية العنقودية، وفق اعتبارات فرضتها طبيعة الدراسة من حيث أهمية الحصول على تباين في استجابات الطلاب عند كلّ خطوة من خطوات الاستجابة. حيث جرت عملية اختيار أفراد العينة على مرحلتين:

- المرحلة الأولى: تمّ اختيار عينة الدراسة من المدارس الثانوية الحكومية التابعة لمدينة (المدينة المنورة)، مع مراعاة التكافؤ قدر الامكان بحسب نوع جنس المدرسة، وقد انتهت هذه المرحلة باختيار (٩) مدارس ثانوية منها (٥) مدارس بنسبة (٥٥.٥٦٪) مدارس ذكور، و (٤) مدارس بنسبة (٤٤.٤٤٪) مدارس إناث.
- المرحلة الثانية: النزول الميداني إلى مدارس العينة المختارة، واختيار عينة عشوائية من طلبة الصف الثاني الثانوي بكلّ مدرسة. وتمت هذه الخطوة بالتنسيق مع ادارة المدرسة ومعلم الكيمياء للصف الثاني الثانوي فيها، حيث استعانة الباحثة في اختيار الطلبة بناءً على كشوفات الطلاب الرسمية المنتظمين خلال فترة تطبيق الدراسة، واختيار شعبتين من كل مدرسة.

في ضوء ذلك تم الاختيار بالطريقة العشوائية العنقودية، حيث بلغ حجم العينة الكلي (٤٥١) طالبا وطالبة، منهم (٢١٥) طالبا في (٥) مدارس، و(٢٣٦) طالبة في (٤) مدارس اناث، ملحق (١). كما في الجدول التالي:

جدول (٢) توزيع عينة الدراسة الأساسية وفق لمتغير النوع

النوع	التصنيف	التكرار	النسبة المئوية	المجموع
	ذكر	215	٪47.67	451
	انثى	236	٪52.33	

حيث تبين أن توزيع عينة الدراسة تبعاً لمتغير النوع الاجتماعي (ذكور-إناث) يشير احصائيا إلى أن الطالبات الاناث يمثلن (52.33٪) من اجمالي العينة الكلية، وهي أعلى من نسبة تمثيل الطلاب الذكور، حيث بلغت نسبة الطلاب الذكور (47.67٪) من إجمالي العينة الكلي

- **أداة الدراسة.**
- **التعريف بأداة الدراسة:**

نظرا لأنّ الهدف من هذه الدراسة بناء اختبار تحصيلي مقالي في مادة الكيمياء وفق نموذج التقدير الجزئي لنظرية استجابة الفقرة، قامت الباحثة بتصميم اختبار تحصيلي مكون من (١٢) سؤالاً مقاليًا في مقرر الكيمياء للفصل الدراسي الأول للصف الثاني ثانوي للوحدتين الدراسيتين الأولى والثانية من الكتاب: الوحدة الأولى (الحسابات الكيميائية)، والوحدة الثانية (الإلكترونات في الذرة)، ملحق (٢).

- **خطوات بناء وإعداد أداة الدراسة:**

تمّ بناء اختبار تحصيلي مقالي في مادة الكيمياء للصف الثاني ثانوي في ضوء نموذج التقدير الجزئي لنظرية استجابة الفقرة، وفق الخطوات الآتية:

- تحديد الغرض من الاختبار وهو قياس تحصيل الطلاب في مادة الكيمياء وفق نموذج التقدير الجزئي لنظرية استجابة الفقرة.
- التحليل النوعي لمنهج الكيمياء المقرر من وزارة التعليم للصف (الثاني ثانوي مسارات) للفصل الدراسي الأول للعام الدراسي (١٤٤٤هـ) لتحديد الأهداف السلوكية، وقد شمل كتابة عناصر البناء المفاهيمي لمادة الكيمياء المفاهيم والحقائق والنظريات والقوانين والتطبيقات في الوحدتين (١،٢)، ملحق (٣).
- بناء جدول المواصفات لمحتوى المنهج كمياً لاعتماده كمرجعية منهجية لتحديد حجم وتوزع الاختبار وتحقيق التوازن في الاختبار، والتأكد من أنه يقيس عينة ممثلة لأهداف الاختبار ومحتوى المادة الدراسية، وتحديد عدد البنود الخاصة بكل هدف وإعطاء وزن نسبي لكل هدف، بالاعتماد على تصنيف هرم بلوم المعرفي، ملحق رقم (٤).
- كتبت البنود في صورة أسئلة مقالية (١٢) سؤال، حيثُ إنّ الاستجابة عليها تتمّ بخطوات متسلسلة ومضبوطة بعدد محدد من المستويات من الأداء

بحيث لا يمكن للطالب الوصول إلى المستويات العليا من الأداء دون الوصول إلى المستويات الدُّنيا، وتمّ مناقشة صيغ الأسئلة، وعدد خطوات الاستجابة، والإجابات المتوقعة من الطلبة، والفترة الزمنية اللازمة لتطبيق الامتحان من خلال ورشة عمل تُخصّصُ لهذا الغرض، ويختلف عدد خطوات الاستجابة في أسئلة الاختبار، وتتراوح خطوات الاستجابة في السؤال ما بين ٣ إلى ٦ خطوات.

ويبين الجدول التالي الموضوعات وعدد الأسئلة على كل موضوع، وعدد خطوات الاستجابة لكل سؤال.

جدول (٣) تكرارات الأسئلة والفقرات (خطوات الاستجابة) لكل وحدة والاجمالي

الوحدة الدراسية	عدد الأسئلة	توزيع الأسئلة حسب عدد الخطوات				الكلي
		٦خطوات	٥خطوات	٤خطوات	٣خطوات	
الحسابات الكيميائية	٧	٢	–	٤	١	٣١
الإلكترونات في الذرة	٥	–	١	٢	٢	١٩
المجموع	١٢	٢	١	٦	٣	٥٠

ويلاحظ من جدول (٣) تفاوت في عدد الأسئلة في الموضوعات المختلفة، وذلك لأن الأسئلة كان يجب أن يتم اعدادها بشكل مضبوط مسبقا، بحيث تتدرج الاستجابة عليها بمستويات محددة. وقد تم إجراء تقييم مدى مطابقة وتحقيق الاختبار لمعايير جدول المواصفات بمقارنة نتائج تحليل مكونات الاختبار وتوزع فقراته على الدروس والمستويات المعرفية بحسب مخرجات نتائج تحليل المحتوى التي تضمنها الملحق الرابع، حيث تم الوصول إلى أن الاختبار المصمم يطابق معايير جدول المواصفات بنسبة ١٠٠٪، ملحق رقم (٥). ويبين الجدول (٤) مستويات الاستجابة لكل سؤال.

جدول (٤):رقم السؤال ومستويات الاستجابة

رقم السؤال	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢
مستويات الاستجابة	٤+١	٤+١	٣+١	٤+١	٦+١	٤+١	٦+١	٤+١	٤+١	٣+١	٣+١	٥+١

حيث يبين الجدول أنّ لكلّ سؤال عدد من الخطوات المحددة للحل، وبناءً على عدد هذه الخطوات يكون عدد فئات أو أقسام الاستجابة فيه، عند زيادة المستوى الصفري على خطوات الحل لكل سؤال نحصل على عدد فئات أو أقسام الاستجابة (١+عدد الخطوات =عدد أقسام الاستجابة) ففي السؤال (س١) على سبيل المثال، يصل الطالب في الحل إلى أحد المستويات التالية:

- المستوى الصفري: وفيه لا ينجز الطالب أيّ عمل صحيح، ويمنح الطالب العلامة (صفر).
- المستوى الأول: وفيه ينجز الطالب الخطوة الأولى من خطوات الحل، ويمنح الطالب العلامة (١).
- المستوى الثاني: وفيه ينجز الفرد الخطوة الثانية من خطوات الحل، ويمنح الطالب العلامة (٢).
- المستوى الثالث: ينجز الطالب الخطوة التالية الثالثة من خطوات الحل، ويمنح الطالب العلامة (٣).

- المستوى الرابع: ينجز الطالب الخطوة الأخيرة من خطوات الحل، ويمنح الطالب العلامة (٤).

ولذلك يكون عدد المستويات خمسة ٤،٣،٢،١،٠ وعدد الخطوات أربعة.

• نتائج الدراسة ومناقشتها:

• أولاً: النتائج المتعلقة بالتحقق افتراضات نماذج الاستجابة للمفردة.

١- افتراض أحادية البعد [Unidimensional]

تمّ التحقق من افتراض أحادية البعد، لما له من أثر على دقة التقديرات فكلما زادت احتمالية تحقيق افتراض أحادية البعد في الاختبار، زادت دقة وصدق تفسير نتائج الدراسة، من خلال التحليل العاملي الاستكشافي Factor (Analysis) عن طريق برنامج SPSS باستخدام طريقة المكونات الرئيسية (Principle Component Analysis) والتدوير المتعامد (Varimax). ويظهر جدول(٥) قيمة الجذر الكامن (Eigen Values) ونسبة التباين المفسر (Explained Variance) للعامل الأول والثاني وناتج قسمة جذر العامل الأول على جذر العامل الثاني والنتائج كما في الجدول الآتي:

جدول (٥): قيم الجذور الكامنة ونسب التباين المفسر

النموذج	العامل الأول	العامل الثاني	
الجذر الكامن	6.21	0.99	6.27
التباين المفسر	٪51.75	٪8.27	

ويتبين أن العامل الأول يفسر 51.75٪ من التباين الكلي، ويفسر العامل الثاني 8.27٪ من التباين، ومجموع مايفسره العامل الأول والثاني معا 60.02٪ من التباين الكلي، وأن ناتج قسمة قيمة الجذر الكامن للعامل الثاني أكبر من (2) مما يشير إلى وجود سمة سائدة ؛ لذا يمكن القول أن افتراض أحادية البعد تتحقق في الاختبار (Hattie, 1985).

٢- افتراض الاستقلال الموضعي [Local Independendence]

أشار هامبلتون وسوامينثان (Hambelton & Swaminathan, 1985) إلى أن افتراض الاستقلال الموضعي يكافئ افتراض أحادية البعد، وهذا يعني أنه إذا تحقق افتراض أحادية البعد في الاختبار، فإن الاختبار يتحقق فيه افتراض الاستقلال الموضعي، مما يدعم من اعتبار كون أن فقرات الاختبار تكون غير مرتبطة فلا توجد علاقة ارتباطية بين الاستجابات على الفقرات وبالتالي فإن إستجابة الفرد على فقرة لا تؤثر على أدائه على فقرة أخرى سلبا أو إيجابا.

٣- افتراض التحرر من السرعة:

تم إعطاء الوقت الكافي للإجابة عن فقرات الاختبار، حيث لم يتم تحديد وقت محدد للانتهاء من الإجابة عن مفردات الاختبار، وبالتالي لا يؤثر عامل السرعة في أداء المستجيب/ة على الاختبار.

- **ثانياً: النتائج المتعلقة بالإجابة عن أسئلة الدراسة**
- **النتائج المتعلقة بالإجابة عن السؤال الأول: ينص السؤال الأول على: " ما درجة مطابقة نموذج التقدير الجزئي لفقرات لاختبار تحصيلي مقالي في مادة الكيمياء للصف الثاني الثانوي ؟"**

وللإجابة على هذا السؤال قامت الباحثة بالتحقق من درجة مطابقة الفقرات لنموذج التقدير الجزئي، باستخدام برنامج (EasyEstimation) لتحليل البيانات، وبعد حذف استجابات الطلاب والطالبات الذين يقعون خارج حدود المطابقة (0.50 -1.5) ، وبعد ذلك تم إعادة التحليل مرة أخرى، وأظهرت النتائج مطابقة جميع الفقرات وعددها(12) للنموذج حيث إن قيمة إحصائي المطابقة الخارجية أو الداخلية المقابلة لتقديرات معالم الصعوبة لهذه الفقرات تقع داخل حدود المطابقة الملائمة (0.50 -1.5) كما في الجدول (٦) .

جدول (٦): مطابقة النموذج للفقرات

م	المطابقة الخارجية (OutFit)	الانحراف المعياري للمطابقة الخارجية (StdOutFit)	المطابقة الداخلية (InFit)	الانحراف المعياري للمطابقة الداخلية StdInFit)
Item001	0.91	−1.46	0.92	−1.22
Item002	1.23	3.25	1.10	1.39
Item003	0.87	−2.03	0.89	−1.68
Item004	0.97	−0.46	0.97	−0.40
Item005	0.87	−2.05	0.88	−1.70
Item006	0.89	−1.63	0.90	−1.43
Item007	0.73	−4.39	0.77	−3.59
Item008	0.74	−4.19	0.81	−2.97
Item009	0.89	−1.75	0.88	−1.71
Item010	1.20	2.89	1.10	1.42
Item011	0.98	−0.32	1.03	0.45
Item012	0.93	−1.04	0.90	−1.41

ويتضح من الجدول رقم(٦) أن جميع الفقرات وعددها(12) فقرة مطابقة لنموذج التقدير الجزئي، لأن قيمة إحصائي المطابقة الخارجية والداخلية المقابلة لتقديرات معالم الصعوبة لهذه الفقرات تقع داخل حدود المطابقة الملائمة(0.50–1.5). حيث بلغت أعلى قيمة للمطابقة الخارجية (1.23) وكانت أقل قيمة (0.73) بمتوسط (0.93) بإنحراف معياري مقدارها (0.15)، بينما بلغت أعلى قيمة للمطابقة الداخلية (1.10) وكانت أقل قيمة (0.77) بمتوسط (0.93) وبإنحراف معياري (0.09) كما يوضحها الجدول (٧).

جدول (٧): قيم الوسط الحسابي، والانحراف المعياري لإحصائي المطابقة الداخلية والخارجية لقيم صعوبة الأسئلة

القيم	المطابقة الخارجية Outfit	المطابقة الداخلية Infit
أعلى قيمة	1.23	1.10
أقل قيمة	0.73	0.77
المتوسط الحسابي	0.93	0.93
الإنحراف المعياري	0.15	0.09

وقد اتفقت هذه النتيجة مع دراسة العلي (٢٠٠٤)، ودراسة عبدالوهاب (٢٠١٩) التي أشارت إلى تطابق جميع الفقرات مع نموذج التقدير الجزئي.

- **النتائج المتعلقة بالإجابة عن السؤال الثاني : ينص السؤال الثاني على: "ما قيم معالم الفقرة [معلمة الصعوبة] لاختبار تحصيلي مقالي في مادة الكيمياء للصف الثاني الثانوي الذي تم تطويره وفق نموذج التقدير الجزئي ؟"**

باستخدام برنامج (Easy Estimation) تم حساب معلمة الصعوبة للاختبار وفق نموذج التقدير الجزئي ونلاحظ وجود 3 معلمات صعوبة نسبية (3عتبات فارقة) للفقرات رقم (11.10.3)، و4 معلمات صعوبة نسبية(4عتبات فارقة) للفقرات رقم(9.8.6.4.2.1)،و5 معلمات صعوبة نسبية(5عتبات فارقة) للفقرة رقم (12)، و6 معلمات صعوبة نسبية(6عتبات فارقة) للفقرتين رقم (7.5)، كما يوضح ذلك الجدول (٨).

جدول (٨): أرقام الأسئلة، وتقدير درجة صعوبة كل خطوة من خطوات الاستجابة (الصعوبة النسبية)

رقم السؤال	الصعوبة النسبية الأولى (1-0)	الصعوبة النسبية الثانية (2-1)	الصعوبة النسبية الثالثة (3-2)	الصعوبة النسبية الرابعة (4-3)	الصعوبة النسبية الخامسة (5-4)	الصعوبة النسبية السادسة (6-5)
1	-1.65	-0.99	0.17	0.88	——	——
2	-0.78	-0.27	0.68	1.32	——	——
3	-1.45	-0.42	0.71	—	——	——
4	-1.90	-0.85	0.12	1.05	——	——
5	-0.96	-0.51	0.12	0.69	1.34	2.14
6	-1.80	-0.93	0.15	1.16	——	——
7	-0.90	-0.29	0.21	0.61	1.05	2.09
8	-1.18	-0.67	-0.16	1.28	——	——
9	-1.01	-0.17	0.71	1.54	——	——
10	-1.48	-0.65	0.18	—	——	——
11	-0.88	0.13	0.91	—	——	——
12	-1.67	-0.88	-0.17	0.38	1.15	——

ويتضح من نتائج الجدول (٨) أن أصعب خطوة من خطوات استجابة أسئلة الاختبار كانت الخطوة السادسة من السؤال رقم 5 حيث بلغت (2.14) لوجيت، وأن أسهل خطوة من خطوات استجابة أسئلة الاختبار كانت الخطوة الأولى من السؤال رقم 4 حيث بلغت (1.90-) لوجيت، إي أن مدى الصعوبة النسبية في خطوات الاستجابة في أسئلة الاختبار بلغت (4.04) لوجيت والتي بالإمكان أن نقيس بها قدرات الأفراد، إي أن هذا المدى من الصعوبة النسبية في خطوات الاستجابة في أسئلة الاختبار يسمح لنا من استخدام أسئلة الاختبار في تقدير قدرات مدى مقبول من القدرة، حيث أن الأسئلة متباينة في صعوبتها وبالتالي يمكننا تقديمها إلى الأفراد كلا على حسب قدرته.

ويلاحظ من جدول (٨) أن قيم معلمة الصعوبة النسبية الأولى (1-0) تراوحت من (1.90-) لوجيت إلى (0.78-) لوجيت ، بمتوسط حسابي (1.31-)

وانحراف معياري (0.40)، فيما تراوحت قيم معلمة الصعوبة النسبية الثانية (2-1) من (0.99-) لوجيت إلى (0.13) لوجيت، بمتوسط حسابي (0.54-) وانحراف معياري (0.35)، فيما تراوحت قيم معلمة الصعوبة النسبية الثالثة(3-2) من (0.17-) لوجيت إلى (0.91) لوجيت، بمتوسط حسابي (0.30) وانحراف معياري (0.36)، بينما تراوحت قيم معلمة الصعوبة النسبية الرابعة (4-3) من (0.38) لوجيت إلى (1.54) لوجيت ، بمتوسط حسابي (0.99) وانحراف معياري (0.38)، و تراوحت قيم معلمة الصعوبة النسبية الخامسة (5-4) من (1.05) لوجيت إلى (1.34) لوجيت، بمتوسط حسابي(1.18) وانحراف معياري (0.15)، بينما تراوحت قيم معلمة الصعوبة النسبية السادسة(6-5) من (2.09) لوجيت إلى(2.14) لوجيت، بمتوسط حسابي (2.12) وانحراف معياري (0.04)، كما يوضحها جدول (٩)

جدول (٩): أقل قيمة وأعلى قيمة للصعوبات النسبية والوسط الحسابي والانحراف المعياري لهم

القيم	الصعوبة النسبية الأولى	الصعوبة النسبية الثانية	الصعوبة النسبية الثالثة	الصعوبة النسبية الرابعة	الصعوبة النسبية الخامسة	الصعوبة النسبية السادسة
أقل قيمة	-1.90	-0.99	-0.17	0.38	1.05	2.09
أكثر قيمة	-0.78	0.13	0.91	1.54	1.34	2.14
المتوسط	-1.31	-0.54	0.30	0.99	1.18	2.12
الانحراف	0.40	0.35	0.36	0.38	0.15	0.04

- **النتائج المتعلقة بالإجابة على السؤال الثالث : ينص السؤال الثالث على: " ما قيم معلم قدرة الأفراد للإجابة عن الاختبار التحصيلي المقالي في مادة الكيمياء للصف الثاني الثانوي الذي تم تطويره وفق نموذج التقدير الجزئي ؟"**

بهدف التحقق من درجة مطابقة الأفراد لنموذج التقدير الجزئي، قامت الباحثة باستخدام برنامج (EasyEstimation)لتحليل البيانات،وأظهرت نتائج التحليل عدم مطابقة (60) مستجيب/ة من أفراد عينة الدراسة وأرقامهم:

3	12	22	25	32	36	51	62	64	68	76	77
78	79	91	94	95	97	98	103	106	107	108	114
120	121	140	189	231	245	273	285	287	291	294	303
310	312	314	320	323	324	325	333	341	345	346	360
372	374	375	378	395	415	434	435	436	439	442	446

بمعنى أنّ قيمة إحصائي المطابقة الخارجية أو الداخلية المقابلة لتقديرات هؤلاء الأفراد تقع خارج حدود المطابقة (0.5 -1.5)، أي أن الأفراد قد تجاوزوا الحد المقبول إحصائيا بإجاباتهم صوابا عن الأسئلة التي تزيد في مستوى صعوبتها عن مستوى قدراتهم، أو يفشلوا في الإجابة صوابا عن فقرات تتطلب مستوى قدرة أقل من قدراتهم، أو عدم جديتهم في الاستجابة، وتم إعادة التحليل بعد حذف استجابات الأفراد الغير مطابقين للنموذج باستخدام برنامج Easy Estimation، وتم بعد ذلك حساب معلم القدرة والخطأ

المعياري لتقدير قدراتهم وإحصاءات المطابقة الداخلية والخارجية. والجدول (١٠) يوضح ذلك:

جدول (١٠): المتوسط الحسابي لقدرة الطلبة ومتوسط الخطأ المعياري لتقدير القدرة وإحصاءات المطابقة الداخلية والخارجية

إحصاءات القدرة			إحصاءات المطابقة الخارجية Outfit		إحصاءات المطابقة الداخلية Infit	
متوسط الخطأ المعياري	القدرة Theta	العدد	متوسط الانحراف المعياري	متوسط Outfit	متوسط الانحراف المعياري	متوسط Infit
0.32	0.003	391	-0.38	0.93	-0.28	0.90

ويلاحظ من الجدول (١٠) أن متوسط المطابقة الداخلية بلغت (0.90) بانحراف معياري قدره (0.28-)، كما بلغت قيمة متوسط المطابقة الخارجية (0.93) بانحراف معياري قدره(0.38 -)، وأن قدرات الأفراد المستخرجة من البرنامج بلغت (0.003)، ومتوسط الخطأ المعياري(0.32)، وتشير هذه القيمة إلى أن قدرات الأفراد تقع معظمها في نطاق القدرة المتوسطة.

- **النتائج المتعلقة بالاجابة على السؤال الرابع: ينص السؤال الرابع على : "ما دلالات ثبات اختبار تحصيلي مقالي في مادة الكيمياء للصف الثاني الثانوي الذي تم تطويره وفق نموذج التقدير الجزئي ؟"**

من خلال التحليل الإحصائي باستخدام نموذج التقدير الجزئي تم الحصول على معامل الثبات المتعلق بالأفراد (Person – Reliability)، ومعامل الثبات الامبريقي المتعلق بفقرات الاختبار (Test - Reliability)، حيث يشير الثبات إلى مدى الدقة في تحديد مواقع كل من قدرة الأفراد ، ودرجة صعوبة الأسئلة على متصل السمة، ويتم تحديد ذلك بمؤشر الفصل بين الأفراد(Separation Index – Person Gp)، وهو عبارة عن ناتج قسمة الانحراف المعياري لقيم معامل الصعوبة للفقرات على الوسط الحسابي للخطأ المعياري في تقدير معلمة الصعوبة.

ومن الملاحظ أن معامل الفصل بين الأفراد يزداد كلما قلت قيمة الأخطاء المعيارية في تقدير معلمة الصعوبة. ومؤشر الفصل بين الأسئلة (Separation Index – Item Gi)، وهو عبارة عن ناتج قسمة الانحراف المعياري لقيم معامل تقديرات القدرة على الوسط الحسابي للخطأ المعياري في تقديرات القدرة.

ومن الملاحظ أن معامل الفصل بين الفقرات يزداد كلما قلت قيمة الأخطاء المعيارية في تقديرات القدرة؛ والذي يجب أن تكون قيمة أكبر من (2) فإذا لم تزد قيمته عن ذلك يصعب عندئذ تقدير قدرة الأفراد بواسطة هذه الأسئلة أو تقدير درجة صعوبة الأسئلة من خلال عينة الأفراد (Wright & Linacre,1993).

وقد بلغت قيمة Gp مؤشر الفصل للأفراد (3.26) وهي أكبر من (٢) ولذا فإن عينة كافية لتقدير درجة صعوبة أسئلة الاختبار.

وبلغت قيمة Gi مؤشر الفصل بين الأسئلة (3.96) وهي أيضاً تزيد عن (2)، مما يدل على قدرة أسئلة الاختبار لتقدير قدرة الأفراد .

وبتطبيق العلاقة (Wright & Linacre,1993)

$$R_P = \frac{G_p^2}{1 + G_P^2}$$

بلغ معامل الثبات للأفراد Rp الذي هو عبارة عن نسبة التباين المشاهد الذي لا يعزى إلى أخطاء القياس(0.91) وهذه القيمة كبيرة وأعلى من المتوسط مما يدل على قدرة الأفراد الذين تم تطبيق الاختبار عليهم في الفصل بين الأسئلة وتعريف متصل السمة التي تقيسه الأسئلة.

وبتطبيق العلاقة (Wright & Linacre,1993)

$$G_i = \frac{G_i^2}{1 + G_i^2}$$

بلغ معامل الثبات للفقرات Ri (0.94) وهي كبيرة وأعلى من المتوسط، مما يشير إلى قدرة الفقرات في الفصل بين قدرات الأفراد، وقدرتها في التمييز بين قدرات الأفراد على متصل السمة، و تتفق هذه النتيجة مع دراسة العلي (٢٠٠٤)، حيث بلغ معامل الثبات للأفراد ومعامل الثبات للفقرات على التوالي (0.91-0.99) وهي قيم مرتفعة.

- **النتائج المتعلقة بالاجابة على السؤال الخامس: ينص السؤال الخامس على: " ما دلالات صدق اختبار تحصيلي مقالي في مادة الكيمياء للصف الثاني الثانوي الذي تم تطويره وفق نموذج التقدير الجزئي ؟"**

يتحقق صدق الاختبار في نظرية استجابة الفقرة إذا تحققت افتراضات النموذج المستخدم، وإذا تم مطابقة جميع الأفراد والفقرات لنموذج التقدير الجزئي لبيانات الاختبار بصورته النهائية بعد حذف الأفراد غير المطابقين والفقرات غير المطابقة لافتراضات نموذج التقدير الجزئي، وقد تحققت افتراضات نموذج التقدير الجزئي لنظرية استجابة الفقرة.

- **صدق الاختبار الظاهري [صدق المحكمين]:**

تمّ عرض تحليل المحتوى والأهداف وجدول المواصفات وأسئلة الاختبار في صورتها الأولية ملحق (٦) ، وتوزيع النسخ التحكيمية على عددٍ من المحكمين المختصين في التربية وعلم النفس، وعلى عددٍ من المعلمين في مادة الكيمياء للصف الثاني ثانوي ممن يحملون درجة البكالوريوس في مادة الكيمياء

وخبرة لا تقلّ عن خمس سنوات في تدريس مادة الكيمياء في مدارس (المملكة العربية السعودية) ملحق رقم (٧)، حيث طلب منهم تحكيم الأداة علميا من خلال ابداء آرائهم وملاحظاتهم المنهجية والشكلية، والهدف من الاحتكام إلى المحكمين إعطاء حكم على فقرات و نتائج الاختيار ومدى تغطيتها للمجال السلوكي المحدد في المنهاج والكتاب المقرر من خلال الاستبانة المعدة للتحكيم، وقد بلغت نسبة الاتفاق بين المحكمين (93.65٪) على مناسبة الفقرات وأبعادها ووضوح التعليمات وطريقة الإجابة عليها، ونسبة من عدم الاتفاق تساوي(6.35٪)، وبهذا أصبح جاهزا لتطبيقه على العينة، ملحق (٨).

- **صدق المحتوى:**

من الضروري أن يكون هناك ارتباط بشكل كبير بين صدق المحتوى وجدول المواصفات، فكلما كانت فقرات الاختبار تقيس الأهداف التعلمية المحددة في جدول المواصفات، والمرتبطة بالمحتوى التعليمي المراد قياسه، يكون صدق المحتوى عاليا، وتم التحقق من صدق المحتوى من خلال مطابقة جدول المواصفات للمحتوى والأهداف وإجماع المحكمين بنسبة (93.65٪)، ملحق (٥).

• التوصيات.

في ضوء النتائج التي تمّ التوصّل إليها، توصي الباحثة بالآتي:

- حث المختصين في مجال تطوير وتصميم المقاييس النفسية والتربوية باستخدام نماذج الاستجابة للمفردة في إعدادها وتطويرها.
- استخدام نماذج التقدير الجزئي في تطوير وبناء المقاييس والاختبارات، والتحقق من الخصائص السيكومترية لها للتغلب على أوجه النقد والقصور التي وجهت لها.
- ضرورة الاهتمام بتوفير البرامج الإحصائية الحديثة المناسبة لتدريج المقاييس والاختبارات في ضوء نظرية الاستجابة للفقرة.
- ضرورة الاهتمام بتدريب الباحثين والمشاركين في عملية التقويم على كيفية التعامل مع البرامج الإحصائية الحديثة، وإدخال البيانات وتفسير النتائج.
- إجراء دورات تدريبية للباحثين في كيفيّة استخدام نموذج التقدير الجزئي في تحليل نماذج الاختبارات، وتوفير الإمكانات اللازمة لذلك.

• ثالثاً: المقترحات.

تقترح الباحثة القيام بالدراسات الآتية:

- بناءُ اختبار تحصيلي مقالي في (الرياضيات، العلوم) لمختلف الصفوف الدراسية؛ وفق نموذج التقدير الجزئي لنظرية استجابة الفقرة.
- تدريج ومعايرة مقاييس نفسية تربوية؛ وفق نموذج التقدير الجزئي لنظرية الاستجابة للمفردة.

• المراجع العربية.

- أبو السل، محمد شحادة، (٢٠١٦م) بناء مقياس دافع الإنجاز لطلبة المرحلة الثانوية في دمشق وفق نظرية الاستجابة للفقرة (IRT) مجلة اتحاد الجامعات العربية للتربية وعلم النفس، كلية التربية، جامعة دمشق، المجلد (١٤)، العدد (٤)، ١٩٤-٢٢٨.
- أبو شندي، يوسف عبدالقادر على (٢٠٢١م) استخدام نموذج استجابة الفقرة المتدرجة في تقدير الخصائص السيكومترية لمقياس قلق المستقبل المهني لدى طلبة الجامعات في سلطنة عمان، مجلة جامعة النجاح للأبحاث، العلوم الإنسانية، جامعة النجاح الوطنية، المجلد (٣٥)، العدد (٦)، ٩٠١-٩٢٦.
- أبو فودة، باسل خميس سالم (٢٠١٦م) التوافق بين النظرية التقليدية في القياس ونظرية استجابة الفقرة في مطابقة فقرات اختبار محكي المرجع في وحدة الهندسة التحليلية، دراسة عربية في التربية وعلم النفس، رابطة التربويين العرب، العدد (٧٣)، ١٨٩-٢١٥.
- أبو وردة، إيمان عبدالله حسين (٢٠٢٠م) بناء اختبار لقياس القدرة على التفكير المجرد باستخدام نظرية الاستجابة للفقرة، رسالة دكتوراه، كلية التربية، جامعة اليرموك، الأردن.
- أحمد، دعاء عوض سيد (٢٠٢٠م) قضايا تشخيص صعوبات التعلم ما بين الاختبارات التشخيصية والاختبارات التحصيلية، المجلة العربية لعلوم الإعاقة والموهبة، المؤسسة العربية للتربية والعلوم والآداب، العدد (١٤)، ٦٣١- ٦٤٨.
- الإدارة العامة للاختبارات والقبول (٢٠١٥). دليل نظم وإجراءات الاختبارات في التعليم العام. وزارة التعليم، المملكة العربية السعودية.
- امطانيوس، نايف مخائيل (٢٠١٦م) القياس والتقويم في التربية وعلم النفس، دار الاعصار العلمي للنشر والتوزيع، عمان، الأردن، الطبعة الأولى.
- أيوب، حامدي (٢٠٢٢م) كفايات بناء الاختبارات التحصيلية لدي أساتذة التعليم المتوسط خرجي المدارس العليا، مجلة دراسات نفسية وتربوية، جامعة قاصدي مرباح، المجلد (١٥)، العدد (١)، ٤٢٥- ٤٣٩.
- البادي، بدر بن حمد بن حميد، الحدابي، داوود، عبد الملك يحيي (٢٠٢١م) مدى كفاية معلمي المعاهد الإسلامية في بناء الاختبارات التحصيلية وفق معايير الاختبار الجيد من وجهة نظرهم، المركز القومي للبحوث، غزة، المجلد (٥)، العدد (٢١)، يونيو، ٧٦- ٩٦.
- براسنة، محمود إبراهيم عوفان (٢٠١٥م) أثر طريقة التعويض عن الاستجابات المفقودة على دقة معادلة نموذج اختبار باستخدام نظرية الاستجابة للفقرة، رسالة دكتوراه، كلية التربية، جامعة اليرموك، الأردن.
- بشير، حسين (٢٠٢٢م) نظرية الاستجابة للمفردة الاختبارية، افتراضاتها ونماذجها، مجلة العلوم الإنسانية والاجتماعية، جامعة عبد الحميد مهري، قسنطينة (٢)، المجلد (٨)، العدد (١)، ١٩٨- ٢١١.
- بني عواد، على (٢٠١٨م) سلوك دالة المعلومات بتغير نسب فقد الاستجابة وحجم العينية في ضوء استجابة الفقرة، مجلة الزرقاء للبحوث والدراسات الإنسانية، جامعة الزرقاء، عمادة البحث العلمي، المجلد (١٨)، العدد (٣)، ٤٢٩- ٤٤٠.
- الثبيتي، عمر عواض عوض (٢٠٢١م) استخدام نموذج الاستجابة المتدرجة في نظرية الاستجابة للفقرة في بناء مقياس الاستعداد الأكاديمي للطالب الجامعي، مجلة العلوم التربوية والنفسية، جامعة البحرين، مركز النشر العلمي، المجلد (٢٢)، العدد (٤)، ٢٠٩- ٢٤١.
- جري، خضير (٢٠١٦م) التقنيات التربوية تطورها، تصنيفاتها، وأنواعها، اتجاهاتها، مؤسسة ثائر العصامي، بغداد، الطبعة الثانية.
- الحاج محمود، فريال محمود محمد (٢٠١٨م) الأداء التفاضلي لفقرات وممومات اختبار محكي المرجع في العلوم للصف السادس الابتدائي مصمم وفق نظرية استجابة الفقرة، مجلة جامعة القدس المفتوحة للبحوث الإنسانية والاجتماعية، جامعة القدس المفتوحة، العدد (٤٤)، آذار، ١٢٣- ١٣٥.

- الحاج، فريال، وأحمد، سمية (٢٠٢١م) بناء مقياس للاتجاه نحو الاختبارات الإلكترونية وفق نموذج الاستجابة المتدرجة، المجلة السعودية للعلوم النفسية، جامعة الملك سعود، الجمعية السعودية للعلوم التربوية والنفسية، العدد(٦٧).
- حبيب، صفاء، وعزيز، صادق (٢٠١٨م) بناء المقاييس النفسية وفقاً لنظرية الاستجابة للفقرة باستعمال الأنموذج الكشفي المعمم، الدار المنهجية للنشر والتوزيع.
- حسين، ضيف؛ والزهرة، الأسود (٢٠٢١). "كفايات بناء الاختبارات التحصيلية الجيدة لدى أساتذة التعليم الابتدائي دراسة ميدانية على عينة من أساتذة التعليم الابتدائي بولاية الوادي"، مجلة العلوم النفسية والتربوية، مجلة العلوم النفسية والتربوية، عدد ٧ (٢)، الجزائر، جامعة الوادي، الجزائر، ١٢٧ — ١٣٩.
- خليل، محمد المري محمد إسماعيل، وإسماعيل، ميمي السيد أحمد (٢٠٢١م) الأداء التفاضلي لاختبار تحصيلي إلكتروني في مقرر القياس والتقويم لدى طلبة الفرقة الثالثة بكلية التربية بالجامعة باستخدام نموذج التقدير الجزئي، المجلة التربوية، كلية التربية، جامعة سوهاج، المجلد (٨٧)، ٦٥٩- ٧٠٠.
- خنيش، يوسف (٢٠٢١م) مدي امتلاك الأستاذ الجامعي للكفايات المعرفية لإعداد الاختبارات التحصيلية، مجلة البحوث التربوية والتعليمية، المدرسة العليا للأساتذة بوزريعة، المجلد (١٠)، العدد (٢)، ١٤١- ١٥٨.
- الخوالدة، عبدالإله ناجي عقيل (٢٠١٦م) تطوير صورة أردنية لاختبار الترابطات المتباعدة باستخدام نظرية الاستجابة للفقرة، رسالة ماجستير، كلية التربية، جامعة اليرموك، الأردن.
- الخوالدة، نايف عارف، النصراوين، معين سلمان سليم (٢٠٢٠م) دقة تقدير معالم الفقرات عند استخدام حالتين من النموذج اللوجستي الثلاثي المعلمة، مجلة جامعة عمان العربية للبحوث، سلسلة البحوث التربوية والنفسية، جامعة عمان العربية، عمادة البحث العلمي والدراسات العليا، المجلد (٤)، العدد (١)، ٣٩٩- ٤٢٨.
- الخولي، زياد (٢٠٠٦م) أثر معادلة درجات الاختبار في تدريج مفردات بنك الأسئلة باستخدام نموذج التقدير الجزئي، رسالة دكتوراه، معهد الدراسات والبحوث التربوية، مصر.
- الخولي، زياد (٢٠١٣م) اتجاهات معاصرة في القياس الموضوعي (نموذج التقدير الجزئي متعدد التدريج) مجلة الآداب، كلية الآداب، جامعة بغداد، العدد (١٠٦) ،٦٥٣- ٦٨٠.
- دبوس، محمد طالب (٢٠١٦م) استخدام نظرية الاستجابة للفقرة في بناء فقرات اختبار محكي المرجع في الرياضيات بفقرات ثنائية التدرج ومتعددة التدرج وفق النموذج اللوجستي ثنائي المعلم، مجلة جامعة النجاح للأبحاث، العلوم الإنسانية، جامعة النجاح الوطنية، المجلد (٣٠)، العدد (٧)، ١٤٥٣- ١٤٨٠.
- زارع، نسرين محمد سعيد، الحربي، أمل بنت معيبد بن محمد (٢٠٢٢م) أثر اختلاف فئات الاستجابة وحجم العينة على الخصائص السيكومترية لقائمة التفكير البنائي باستخدام نموذج سلم التقدير، المجلة السعودية للعلوم النفسية، جامعة الملك سعود، الجمعية السعودية للعلوم التربوية والنفسية، العدد (٧) ، ١٧٧- ١٩٥.
- الزبيدي، محمد بن حسين يحيى (٢٠١٨). " بناء أداة لتقويم الاختبارات التحصيلية في الجامعات السعودية من وجهة نظر الطلاب"، مجلة كلية التربية جامعة أسيوط المجلد٣٤ العدد(٢)، ٤٥٧-٤٩٧.
- الزعبي، امال احمد؛ وسواله، يوسف محمد (٢٠١٧). " بناء اختبار لقياس مهارات التفكير الناقد باستخدام نظرية استجابة الفقرة"، مجلة اتحاد الجامعات العربية للتربية وعلم النفس، المجلد١٥، العدد(٣).
- الزعبي، محمد محمود البشير (٢٠٠٥م) فاعلية نموذج التقدير الجزئي في معايرة فقرات متعددة الخطوات، جرش للبحوث والدراسات، جامعة جرش، المجلد (١٠)، العدد (١).

- زكري، علي بن محمد عبدالله (٢٠٠٨). " الخصائص السيكومترية لاختبار أوتيس لينون للقدرة العقلية مقدرة وفق القياس الكلاسيكي ونموذج راش لدى طلبة المرحلة المتوسطة بمحافظة صبيا التعليمية"، رسالة ماجيستير غير منشورة، جامعة أم القرى: السعودية.
- ساعد، صباح، وعامر، وسيلة. (٢٠١٢). تقويم الكفاية المعرفية لدى معلمي المرحلة الابتدائية في مجال بناء الاختبارات التحصيلية الموضوعية. دفاتر المخبر: جامعة محمد خيضر بسكرة – كلية العلوم الإنسانية والاجتماعية – مخبر المسألة التربوية في ظل التحديات الراهنة، العدد(٩)، ١٥١ – ١٦٤.
- ساعد، صباح؛ وبن عمر، وسيلة (٢٠١٧). "تقييم كفاءة بناء الاختبارات التحصيلية لدى أساتذة التعليم الجامعي وفق معايير الاختبار الجيد"، دراسة تحليلية للاختبارات التحصيلية للسداسيين الأول والثاني للمستويات الجامعية من ٢٠١٢ -٢٠١٣ م) مجلة العلوم الانسانية والاجتماعية، جامعة قاصدي مرباح ورقة، العدد(٢٨).
- سليمان، أمين (٢٠٠٩م) القياس والتقويم في العلوم الإنسانية أسسه وأدواته وتطبيقاته، دار الكتاب الحديث.
- سليمان، شاهر خالد؛ الصالح، علي محمد خالد (٢٠١٦). بناء اختبار تحصيلي محكي المرجع في القياس والتقويم باستخدام النموذج اللوغاريتمي ثلاثي المعلم لنظرية الاستجابة للمفردة الاختبارية"، مجلة التربية، جامعة الأزهر مصر، المجلد٣٥، ١٦٨ جزء١.
- السيد، سعيد يوسف سعيد (٢٠٢٢م) تقنين مقياس العوامل الستة المنقح للشخصية (R- pi- HEXACO) للبيئة الأردنية باستخدام نظرية الاستجابة للفقرة والتحقيق من بنيته العاملية، رسالة دكتوراه، كلية التربية، جامعة اليرموك، الأردن.
- الشافعي، محمد، ونور الدين، أمين (٢٠٠٧م) استخدام نموذج التقدير الجزئي الرياضي اللوغاريتمي في تطوير اختبارات القدرات العقلية الأولية على عينة من البيئة السعودية، مجلة كلية التربية، جامعة الزقازيق، العدد (٥٦) ٢٤٥- ٣٤٦.
- الشريفين، نضال كمال (٢٠٠٣). "مدى تحقق معايير الفاعلية في معادلة اختبارين أحدهما ثنائي التدريج والآخر متعدد التدريج وفق نماذج النظرية الكلاسيكية والنظرية الحديثة في القياس، اطروحة دكتوراه غير منشورة، جامعة عمان العربية: الأردن.
- الشمراني، محمد موسي (٢٠١٦م) أثر حجم العينة وطول الاختبار والتفاعل بينهما على تقدير معلمة الصعوبة والتمييز والتخمين والخطأ المعياري باستخدام نظرية استجابة الفقرة، مجلة كلية التربية، كلية التربية، جامعة طنطا، المجلد (٦٤)، العدد (٤)، أكتوبر، ٢٦٤- ٣٠٦.
- الضمور، سالم عبدالحميد؛ وعوض، رضا سمير؛ وأبو علام، رجاء محمود (٢٠١٥). " مدى توافر معايير التقويم التربوي العالمية في أساليب القياس والتقويم المستخدمة لدى معلمي الرياضات بالمملكة الأردنية الهامشية من وجهة نظرهم"، عالم التربية، المؤسسة العربية للاستشارات العلمية وتنمية الموارد البشرية، المجلد١٦. العدد(٥٠).
- ضيف، حسين والأسود، الزهرة (٢٠٢١). "كفايات بناء الاختبارات التحصيلية لدى أساتذة التعليم الابتدائي: دراسة ميدانية على عينة من أساتذة التعليم الابتدائي بولاية الوادي. مجلة العلوم النفسية والتربوية. ٧ / ٢ ، جامعة الوادي، الجزائر.
- الطراونة، صبري حسن خليل (٢٠١٦م) بناء مقياس للاستقواء لطلبة الصفوف الأساسية العليا في الأردن باستخدام نموذج سلم التقدير، مجلة العلوم التربوية والنفسية، جامعة البحرين، مركز النشر العلمي، المجلد (١٧)، العدد (٣)، ٥٤٧- ٥٦٨.
- الطراونة، صبري حسن خليل (٢٠١٨م) تطوير مقياس التدين لدى طلبة الجامعة باستخدام نموذج سلم التقدير في نظرية الاستجابة للفقرة، مجلة العلوم التربوية والنفسية، مركز النشر العلمي، جامعة البحرين، المجلد (١٩)، العدد (٣).
- الطراونة، صبري حسن خليل، المناسية، جميلة على محمد (٢٠١٨م) أثر عدد مموهات فقرات الاختيار من متعدد على تقدير الخصائص السيكومترية للفقرات وفق نظرية الاستجابة

للفقرة، مجلة جامعة الحسين بن طلال للبحوث، جامعة الحسين بن طلال، عمادة البحث العلمي والدراسات العليا، المجلد (٤)، العدد (١)، يونيو.
- عالم، توفيق على (٢٠١٩م) توظيف التعليم الإلكتروني في العملية التعليمية، سلسلة التعليم والتعلم والتدريب، مركز الحزمي للتصوير، صنعاء، الطبعة الرابعة.
- عالم، توفيق على، وأبو هادي، سوسن حسن (٢٠٢٠م) فاعلية برنامج تصميم حاسوبي لإدارة بنوك أسئلة في مادة الرياضيات مبنية وفق نظرية استجابة الفقرة، المجلة العلمية، كلية التربية، جامعة ذمار، العدد (١٦)، ١٧٢- ٢٠٨.
- عبابنة، عماد غصاب (٢٠١٩م) بناء مقياس اتجاهات معلمي العلوم نحو التعلم النشط المتمركز حول الطالب باستخدام نموذج سلم التقدير لراش، مجلة جامعة النجاح للأبحاث، العلوم الإنسانية، جامعة النجاح الوطنية، المجلد (٣٣)، العدد (٩)، ١٤١٣- ١٤٤٨.
- عبد الوهاب، محمد محمود محمد (٢٠٢٠م) أثر فاعلية بدائل مفردات الصواب والخطأ المتعدد على معالم المفردات والأفراد في ضوء نظرية القياس التقليدية ونموذج الاستجابات المتدرجة، المجلة السعودية للعلوم النفسية، جامعة الملك سعود، الجمعية السعودية للعلوم التربوية والنفسية، جامعة الملك سعود، الجمعية السعودية للعلوم التربوية والنفسية، العدد (٤).
- العبدالله، زياد أحمد (٢٠٠٩م) أثر حجم العينة وعدد المفردات المشتركة على تدرج مفردات بنك الأسئلة باستخدام نموذج التقدير الجزئي، رسالة ماجستير، معهد الدراسات والبحوث التربوية، جامعة القاهرة.
- عبدالوهاب، محمد (٢٠١٩). "أثر طريقة تقدير الدرجات ونموذج الاستجابة للمفردة في دقة تقدير بارامترات الأفراد والمفردات لاختبار صواب وخطأ متعدد"، رسالة دكتوراه غير منشورة، جامعة ام القرى، السعودية.
- عتوم، محمد صالح (٢٠١٤م) تقييم الاختبارات التحصيلية من إعداد معلمي العلوم في محافظة جرش، دراسة تحليلية لنتائج الطلبة للفصل الدراسي الثاني للعام (٢٠١٢/٢٠١٣م) في مبحث العلوم، جامعة جرش، المجلد (١٥)، العدد (٢)، ٣٦١- ٣٧٥.
- عثمان، ـ (٢٠١٤) "أساليب التقويم التربوي"، دار أسامة، عمان.
- العطيان، إخلاص (٢٠١٧م) أثر التعامل مع القِيم الشاذة في بيانات اختبار تيمس (TIMSS) على دقة تقدير معالم الفقرات والأفراد وفقاً لنظرية الاستجابة للفقرة، رسالة دكتوراه، كلية التربية، جامعة اليرموك، الأردن.
- عكاشة، محمود، والبنا، عادل، وأبو حلاوة، محمد، ومجلي، محمد (٢٠٢٠م) تدرج مقياس الذكاء الإنفعالي الجمعي باستخدام نظرية الاستجابة للمفردة على عينة من طلاب الجامعة، مجلة الدراسات التربوية والإنسانية، كلية التربية، جامعة دمنهور، المجلد (٤)، العدد (١٢)، ١٧٩- ٢٤٨.
- علام، صلاح الدين محمود (٢٠٠٥م) نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي، دار الفكر العربي، القاهرة، الطبعة الأولى، ٢٠٠٥م.
- علام، صلاح الدين محمود (٢٠١٢). "الاختبارات والمقاييس التربوية والنفسية، دار الفكر، عمان الأردن، ط٢
- علي، سعدي أحمد (٢٠٢١م) تأثير اختلاف مستوي القدرة في الأداء التفاضلي وفقاً للأنموذج ثنائي المعلم لنظرية الاستجابة للفقرة، مجلة الآداب، كلية الآداب، جامعة بغداد ١٣٧- ١٦٢.
- العلي، محمد محمود البشير (٢٠٠٤). "فاعلية نموذج التقدير الجزئي في بناء أسئلة من فقرات متعددة الخطوات في مادة الكيمياء في الصف الثاني الثانوي العلمي"، رسالة دكتوراه غير منشورة، جامعة عمان العربية، الاردن.
- عليمات م. م. (٢٠٢٢). بناء اختبار محكي المرجع في الرياضيات وفق النموذج اللوجستي ثلاثي المعلمة. مجلة جامعة القدس المفتوحة للأبحاث والدراسات التربوية والنفسية، ١٣(٣٨).
https://doi.org/10.33977/1182-013-038-004

- عودة، أحمد (٢٠٠٥) القياس والتقويم في العملية التدريسية. ط ٣، دار الأمل: اربد -الأردن.
- عودة، موسي عزت (٢٠١٦م) تقدير معالم الفقرة والقدرة بالطرق التي تستند إلى نظرية الاستجابة للفقرة تبعاً لمتغيري طول الاختبار وعد البدائل، رسالة دكتوراه، كلية الدراسات العليا، الجامعة الأردنية، الأردن.
- عيلبوني، سمير فؤاد إيليا، والأشقر، وفاء محمد على خضر (٢٠١٥م) فحص تحقق فرض الاستقلال الموضعي بتحقق فرض أحادية البعد باختلاف حجم العينة وطول الاختبار في نظرية استجابة الفقرة، مجلة كلية التربية، كلية التربية، جامعة عين شمس، المجلد (٣)، العدد (٣٩)، ٤٧٧- ٥٣٢.
- فاتح، الورعادي، بن نابي، نصيرة (٢٠١٩م) معايير وشروط بناء الاختبارات التحصيلية، مجلة الحكمة للدراسات التربوية والنفسية، مؤسسة كنوز الحكمة للنشر والتوزيع، العدد (١٧)، مارس.
- فرحات، رمضان السيد (٢٠٢٢م) بناء وتدرج مقياس التأمل الناقد في ضوء نموذجي التقدير الجزئي والاستجابة المتدرجة لدى طلاب جامعة الأزهر، المجلة المصرية للدراسات النفسية، الجمعية المصرية للدراسات النفسية، المجلد (٣٢)، العدد (١١٤).
- قارة، مريم؛ وساعد، صباح (٢٠٢١). "الاختبارات التحصيلية مرجعية المحك خطوات اجرائية لبناء اختبار تحصيلي مرجعي المحك في مادة الرياضيات لتلاميذ السنة الخامسة ابتدائي، مجلة علوم الإنسان والمجتمع، كلية العلوم الإنسانية والاجتماعية، مجلد٢٢، جامعة محمد خضير بسكرة: الجزائر.
- القرشي، خديجة ضيف الله (٢٠١٦م) استخدام نموذج التقدير الجزئي في انتقاء فقرات مقياس اتجاهات طلبة الدراسات العليا نحو برنامج (spss) في منطقة مكة المكرمة، المجلة التربوية الدولية المتخصصة، دار سمات للدراسات والأبحاث، المجلد (٥)، العدد (٩)، ٥٢٩- ٥٥٢.
- القطاونة، آلاء خالد (٢٠١٦م) بدراسة الثبات في الاختبارات محكية المرجع في الرياضيات للصف العاشر حسب النظرية الكلاسيكية ونظرية استجابة الفقرة حسب النموذج ثنائي المعلم، دراسة مقارنة، رسالة ماجستير، كلية العلوم التربوية، جامعة مؤتة، الأردن.
- الكرامنة، محمد صالح محمد (٢٠١٥م) بناء اختبار تكييفي محوسب للذكاء المنطقي الرياضي للطلبة الموهوبين باستخدام نظرية الاستجابة للفقرة، رسالة دكتوراه، كلية الدراسات العليا، الجامعة الأردنية، الأردن.
- اللهيبي، عبد الله (٢٠١٨م) استخدام نموذج الاستجابة للمفردة الاختبارية في تدريج مفردات مقياس الاتجاه نحو الرياضيات، مجلة جامعة شقراء، العدد (٩).
- محمد، بلعالية (٢٠٢١م). " الاختبارات التحصيلية كأسلوب فعّال للتقويم التربوي"، مجلة الخلدونية، العدد١٣، المجلد١.
- مرشود، محمد فايق سالم (٢٠٢٠م) بناء مقياس للرضا الوظيفي من وجهة نظر الأكاديميين في جامعة عمان الأهلية وفق نظرية الاستجابة للفقرة، مجلة البلقاء للبحوث والدراسات، جامعة عمان الأهلية، عمادة الدراسات العليا والبحث العلمي، المجلد (٢٣)، عدد خاص.
- المرواني، أشواق ضيف الله سليم، وسليمان، شاهر خالد (٢٠١٩م) أثر موقع البديل الصحيح في اختبار اختيار من متعدد على دقة تقدير معالم الفقرة وفق النموذج ثنائي المعلمة لنظرية الاستجابة للفقرة، المجلة الدولية للدراسات التربوية والنفسية، مركز رفاد للدراسات والأبحاث، المجلد (٦)، العدد (١)، آب.
- مزارة، نعيمة (٢٠٢١م) التقويم التربوي وفنيات الاختبارات التحصيلية، عرض نموذج كيفية تصميم اختبار في وحدة مدخل علوم التربية، المجلة العلمية للعلوم التربوية والصحية النفسية، المؤسسة العلمية للعلوم التربوية والتكنولوجية والتربية الخاصة، المجلد (٣)، العدد (٢)، يونيو ٢٥٧- ٢٨٠.
- النعيمي، عز الدين عبدالله عواد (٢٠٢١م) الخصائص السيكومترية لمقياس بك للاكتئاب"الصورة الثانية" باستخدام نموذج الاستجابة المتدرجة، مجلة الشمال للعلوم

الإنسانية، جامعة الحدود الشمالية، مركز النشر العلمي والتأليف والترجمة، المجلد (٦)، العدد (٢)، ٣٣٣- ٣٦٩.
- الهواوشة، هليل خليل (٢٠١٥) "بناء مقياس للكشف عن مرض الفصام في الأردن وفق نظرية استجابة الفقرة؛ رسالة ماجستير غير منشورة، جامعة مؤتة: الأردن.
- هيبة، محمد (٢٠١٦م) الخصائص السيكومترية لمقياس ضبط الذات في ضوء نظرية الاستجابة للمفردة، مجلة كلية التربية، جامعة الأزهر، المجلد (٤) العدد (١٦٨)، ٤٥٠- ٤٩٠.
- يوسف، جيهان فارس (٢٠١٨). "تصميم تعليمي – تعلمي على وفق نماذج النظرية البنائية وأثره في تحصيل مادة الكيمياء والمهارات الحياتية والتعلم المنظم ذاتيا عند طالبات الصف الثاني المتوسط"، رسالة دكتوراه غير منشورة، جامعة بغداد، العراق.

• المراجع الأجنبية.

- Abu Foudeh, Basel Khamis Salem (2020). A Training Program Based on the quality Standards of Achievement Tests to Improve the Competence of Teaching Staff in Constructing Tests at Middle East University in Jordan, Universal Journal of Educational Research 8(4): 1212-1217, 2020. DOI: 10.13189/ujer.2020.080411, http://www.hrpub.org/
- Boopathiraj, C. & Chellamani, K. (2013). Analysis of Test Items on Difficulty Level and Discrimination Index in the Test for Research in Education. International Journal of Social Science and Interdisciplinary Research, 2 (2), 189-193.
- Casabianca, J. & Lewis, C. (2015). IRT Item Parameter Recovery with Marginal Maximum Likelihood Estimation Using Loglinear Smoothing Models. Journal of Educational and Behavioral Statistics, 40 (6), 547-578.
- Domino, G. & Domino, M.(2006). Psychological testing: An introduction (2 nd ed) . Cambridge University Press.
- Fontanella, L. Villano, P. & Donato, M. (2015). Attitudes owardsRomapeopleandmigrants:acomparisonthroughaBayesianmultidimensionalIRTmodel.Quality&Quantity,49(1), 1-20
- Fraley, R.C. and Shaver, P.R. (2000) Adult Romantic Attachment: Theoretical Developments Emerging Controversies and Unanswered Questions. Review of General Psychology, 4, 132-154. http://dx.doi.org/10.1037/1089-2680.4.2.132
- Gnambs, T. & Batinic, B.(2011). Evaluation of measurement precision with Rasch-type models: The case of the short generalized opinion leadership scale. Personality and Individual Differences, 50, 53- 58.
- Gurmit Singh (2022) " Original Research Paper Education Construction and Standardization of Achievement Test in Science " Indian Journal of Research Volume-8 | Issue-2
- Ha, D. (2017). Applying Multidimensional Three-Parameter Logistic Model (M3PL) in Validating a Multiple-Choice Test.

International Journal of Scientific and Research Publications, 7 (2), 2250-3153.
- *Hambleton, R. (1989). Principles and Selected applications of item response theory, In Linn, Robert, L. (Ed.), Educational Measurement (3rd ed., PP. 147-201). New York: American Council on Education, Macmillan Publishing Company.*
- Hambleton, R., Swaminthan, H. & Rogers, H.(1991). *Fundamentals of Item Response Theory, International Educational and Professional*. Publisher Newbury Park.
- Hemant Lata Sharma & Sarita (2018), Construction and standardization of an achievement test in Science, International Journal of Research and Analytical Reviews 1037 VOLUME5 I ISSUE 3 I JULY – SEPT 2018] e http://ijrar.com .
- Pooja Bhagat , J. N. Baliya, (2015), " Construction and Validation of Achievement Test in Science " International Journal of Science and Research, Volume5 , Issue6, www.ijsr.net
- Robitzsch, A. & Steinfeld, J.(2018). Item response models for human ratings: overview, estimation methods, and implementation in R. Psychological Test and Assessment Modeling, 60(1), 101-139.
- Saleh Ayed Al Khawaldeh & Dr. Yousef Jabber Alawneh & Dr. Mamon Saleem Alzboun (2022) " The Availability of Quality Standards for the Construction of Science Achivement Tests from the Point of View of the Examination Committees ", Journal of Hunan University（Natural Sciences）Vol. 49. No. 09
- Shafaat Hussain (2015) " Test Construction and Evaluation: A Brief Review", Indian Journal of Applied Research 5(6):725-729.
- Tutz, Berger(2016) -Item focused tress for the identification of item in differential item functioning psychometrika.81.(3).727- 750
- Wagner, Teresa & Harvey, R. (2003). Developing A new critical Thinking Test Using item Response Theory. Paper presented at the 2003 Annual conference of the Society for Industrial Organizational Psychology. Orlando. Retrieved www. Critical thinking.org
- Yang, F. & Kao, S.(2014). Item response theory for measurement validity.
- Yu, C. (2011). A Simple Guide to the Item Response Theory (IRT) and Rasch Modeling. Retrieved30/04/2018from www. Creative wisdom.com/computer/sas/irt.pf.