قياس صعوبة بنود اختبار تحصيلي في مادة الرياضيات لتلاميذ الصف الخامس الأساسي (دراسة مقارنة بين النظرية التقليدية ونظرية الاستجابة للمفردة)

رجاء أسعد* أ.م.د. منذر بوبو**

(الإيداع: 25 تشرين الثاني 2024، القبول: 12 كانون الثاني 2025)

الملخص:

تهدف الدراسة الحالية إلى الكشف عن مدى الاختلاف في المعنى الذي يمكن إعطاؤه للسمة الكامنة من خلال الدرجة التي يحصل عليها الفرد باستخدام كل من نظرية القياس التقليدية ونظرية الاستجابة للمفردة وفق النموذج الأحادي المعلم (نموذج راش)، وذلك من خلال مقارنة درجات صعوبة البنود لاختبار تحصيلي مصمم لقياس تحصيل تلاميذ الصف الخامس الأساسي في مادة الرياضيات. ولتحقيق غرض البحث تم تصميم اختبار تحصيلي مؤلف من (25) فقرة اختبارية في الوحدة الثانية من كتاب الرياضيات لتلاميذ الصف الخامس الابتدائي ثم تم تطبيقه على عينة مؤلفة من (100) تلميذاً وتلميذة وذلك بعد التحقق من خصائص مفرداته وفق النظرية التقليدية ثم وفق النظرية الحديثة في القياس. تم مقارنة ترتيب معاملات الصعوبة في ضوء كل من النظريتين. وقد بينت النتائج تطابق ممتاز في ترتيب معاملات صعوبة البنود الاختبارية فق كلا النظريتين مع تفوق لصالح النظرية الحديثة في القياس والتقويم من خلال احتفاظها بعدد أكبر من البنود من النظرية التقليدية. وفي ضوء نتائج الدراسة تم اقتراح الاعتماد بشكل أكبر في تحديد خصائص الاختبارات على نظرية الاستجابة للمفردة، القيام بالمزيد من الدراسات على اختبارات في مقررات أخرى للتأكد من فاعلية نظرية الاستجابة للمفردة في تحديد معاملات الصعوبة للبنود الاختبارية ومقارنتها مع معاملات الصعوبة في النظرية التقليدية، استخدام نماذج نظرية الاستجابة الأخرى (معلمتين، ثلاثة معالم) في تحديد خصائص الاختبارات التحصيلية.

الكلمات المفتاحية: نظرية القياس التقليدية، نظرية الاستجابة للمفردة، معلم صعوبة المفردة، الرياضيات، نموذج راش.

*طالبة دكتوراه– قسم تربية الطفل – كلية التربية – جامعة اللاذقية.

**أستاذ مساعد –قسم القياس والتقويم –كلية التربية – جامعة اللاذقية.

Measuring the Difficulty of Achievement Test Items in Mathematics for Fifth Grade Students
(A Comparative Study between the Classical Theory and the Item Response Theory)

Raja Asaad* **Mounzer Boubou****

(Received: 25 November 2024,Accepted: 12 January 2025)

Abstract:

The current study aims to reveal the extent of the difference in the meaning that can be given to the latent trait through the degree that the individual obtains using both the traditional measurement theory and the item response theory according to the single-teacher model (Rasch model), by comparing the difficulty scores of the items for an achievement test designed to measure the achievement of fifth-grade students in mathematics. To achieve the purpose of the research, an achievement test consisting of (25) test items was designed in the second unit of the mathematics book for fifth-grade primary students, then it was applied to a sample consisting of (100) male and female students after verifying the characteristics of its vocabulary according to the traditional theory and then according to the modern theory of measurement. The order of the difficulty coefficients was compared in light of each of the two theories. The results showed excellent consistency in the order of the difficulty coefficients of the test items according to both theories, with superiority in favor of the modern theory of measurement and evaluation by retaining a larger number of items than the traditional theory. In light of the results of the study, it was suggested to rely more heavily on the item response theory in determining the characteristics of tests, conduct more studies on tests in other courses to verify the effectiveness of the item response theory in determining the difficulty coefficients of test items and compare them with the difficulty coefficients in the traditional theory, and use other response theory models (two parameters, three parameters) in determining the characteristics of achievement tests

Keywords: Traditional measurement theory, item response theory, item difficulty parameter, mathematics, Rasch model.

*PhD student - Child Development Department - Faculty of Education - University of Latakia.

**Assistant Professor - Measurement and Evaluation Department - Faculty of Education - University of Latakia.

المقدّمة الدراسة:

يعدّ التقويم من العمليات الأساسية في النظام التعليمي حيث يستخدم بهدف الحكم على جودة مدخلات هذا النظام كالبرامج والاستراتيجيات والأهداف وغيرها واتخاذ القرارات بشأنها من حيث بقائها أو تعديلها أو حذفها واستبدالها بأخرى. وتعدّ الاختبارات التحصيلية من أهم أدوات التقويم التي تستخدم على نطاق واسع في المؤسسات التعليمية من أجل معرفة مدى تحقيق التلاميذ لأهداف النظام التعليمي. ومما لا شك فيه أن الاهتمام بالاختبارات التحصيلية يترك أثره المباشر في المكونات الأخرى كالأهداف والمحتوى والأنشطة وغيرها حيث أن اتخاذ القرارات بشأن عناصر العملية التربوية والنظام التعليمي بشكل عام يعتمد على دقة المعلومات والبيانات المتوفرة حول مدى صلاحية هذه العناصر أو عدم جدوى الاستمرار فيها وهذه المعلومات نحصل عليها من خلال تطبيق أدوات التقويم المختلفة من اختبارات ومقاييس نفسية وغيرها هذه تساعدنا في اتخاذ القرارات الصائبة. وهناك علاقة طردية بين عملية القياس وفهم الظواهر المقاسة فكلما ازدادت دقة وموضوعية القياس تحسنت درجة فهم الظاهرة المدروسة وتوضحت للدارسين جميع جوانبها المبهمة وبالتالي التوصل إلى دقة التنبؤ بمستقبل هذه الظواهر وضبطها والتحكم بها وقد انصب اهتمام علماء النفس من بدء حركة القياس النفسي بتحقيق صدق الاختبارات والمقاييس النفسية وثباتها سعياً منهم لتحقيق أعلى درجة من الموضوعية في هذه الأدوات عند استخدامها في عملية القياس (حجازي والخطيب،2014،ص224) وتميّز أدبيات القياس التربوية والنفسية بين مدخلين رئيسين في بناء وتطوير الاختبارات والمقاييس وتحليل البيانات المستمدة منها وهذين المدخلين هما المدخل الكلاسيكي والمدخل المعاصر، الأول يعتمد على النظرية التقليدية للاختبارات Classical Test Theory وما تنطوي عليه من مفاهيم ومبادئ بعضها يتعلق بخصائص مفردات الاختبار والبعض الآخر يتعلق بخصائص الاختبار ككل أما المدخل الثاني فيعتمد على نظرية الاستجابة للمفردة Item Response Theory التي تهتم بالربط بين استجابة الفرد على مفردة اختبار وبين خصائص هذه المفردة.

تعدّ النظرية التقليدية من أقدم النظريات التي استخدمت لعقود في بناء الاختبارات التحصيلية وتحليل خصائصها وظلت النظرية تشكل الأساس النظري والعملي للقياس في العلوم السلوكية لسنوات عديدة مضت وقد استندت هذه النظرية على نموذج بسيط للقياس والذي ينص على أن درجة الفرد الملاحظة في الاختبار تساوي مجموع الدرجة الحقيقية للفرد في السمة المقاسة والدرجة الخطأ في الأداء على الاختبار. (اسماعيل،2007،ص3) ولكن القصور الواضح في النظرية التقليدية عن تحقيق موضوعية القياس جعلها عرضة لمجموعة من الانتقادات حيث أنه "وفق النظرية التقليدية إن فقرات الاختبار تتغير خصائصها بتغير خصائص الأفراد كما أن خصائص الأفراد تتغير بتغير خصائص الاختبار من حيث سهولة وصعوبة الفقرة". (الخياط،2012،ص88) من ناحية أخرى تعتمد النظرية التقليدية على الدرجة التي يحصل عليها المتعلم في تفسير درجة الاختبارات وهذا لا يعبّر بموضوعية وصدق عن المستوى الحقيقي لقدرته فهي هنا مرتبطة بمستوى صعوبة مفردات الاختبار فاختلاف مستوى صعوبة مفردات اختبارين يقيسان نفس القدرة عند المتعلم يؤدي إلى نتيجتين مختلفتين. (بركات،2010،ص10) إضافة إلى ذلك فإن النظرية التقليدية تركز على الاختبار ككل وتنظر إليه على أنه وحدة واحدة وبالتالي لا يسمح هذا النظام بحذف أو إضافة مفردات للاختبار لأن هذا يؤثر على صدقه وثباته. (الشواورة،2013،ص3) وقد تعدّدت المحاولات لتجاوز الثغرات في القياس التقليدي بهدف تحقيق دقة وموضوعية القياس وتقويم الظواهر التربوية والنفسية كما هو الحال في قياس الظواهر الفيزيائية حيث برزت الحاجة لإيجاد علاقات رياضية تمكّن من تحويل الدرجات الخام للطلاب المراد تفسيرها إلى مقياس مشترك بصفر مشترك ووحدة قياس مشتركة ونتيجة لتلك المحاولات انبثقت النظرية الحديثة في التقويم والقياس التربوي والنفسي والتي أطلق عليها نظرية الاستجابة للمفردة وهي عبارة عن نماذج رياضية تحدّد العلاقة بين استجابات الفرد على اختبار ما وبين السمة أو القدرة الكامنة وراء تلك

الاستجابات ومن ثم تحديد مقدار تلك السمة والاستفادة منها في التنبؤ بسلوك هذا الفرد في مواقف اختبارية مماثلة بهدف اتخاذ القرارات الخاصة بهذا الفرد.

تفترض النظرية الحديثة وجود سمات أو قدرات مشتركة بين الأفراد مع وجود اختلاف في مقدار تلك السمات بينهم والذي يستدل عليه عند الفرد من استجابته على الاختبار الذي يقيس تلك السمة. ربما الميزة الأكثر أهمية والتي تؤدي إلى هيمنة نظرية الاستجابة للمفردة في البرامج الجاهزة للاستخدام هي خاصية تقدير مواقع المفردات الفردية (الصعوبات) ومواقع الممتحنين (القدرات) بشكل منفصل ولكن على نفس المقياس. هذه الخاصية ليست متاحة في نماذج النظرية التقليدية. وقد انبثقت عن هذه النظرية عدة نماذج احتمالية تحدّد العلاقة بين أداء الفرد على فقرة اختبار أو مقياس وقدرته التي تكمن وراء هذا الأداء ثم تفسرها حسب موقعه على متصل القدرة وذلك لأن هذه النماذج تفترض أنه يمكن التنبؤ بأداء الأفراد أو يمكن تفسيره في اختبار نفسي أو تربوي معين بناء على خاصية أو خصائص معينة مميزة لهذا الأداء تسمى السمات وتحاول هذه النماذج تقدير درجات الأفراد في هذه السمات من خلال استجاباتهم الملاحظة على مفردات الاختبار الذي يقيس السمة أو القدرة المطلوبة (عيد،2004،ص246). ويعدّ نموذج راش الأحادي المعلم من أهم نماذج نظرية الاستجابة للمفردة وأكثرها استخداماً في بناء الاختبارات التحصيلية والعقلية حيث يمكّن هذا النموذج من تقدير احتمال الاستجابة الصحيحة من خلال معلمين فقط هما صعوبة الفقرة وقدرة الفرد بغض النظر عن حجم العينة وعدد فقرات الاختبار ويتمثل الافتراض الأساسي فيه أنه كلما زادت قدرة الفرد كانت الفرصة أكبر للإجابة الصحيحة على أي فقرة من فقرات الاختبار (الخياط،2012،ص94). وقد اهتمت الدراسة الحالية باستخدام نموذج راش للمقارنة بين النظرية التقليدية ونظرية الاستجابة للمفردة في استخراج معاملات صعوبة البنود لاختبار تحصيلي في مادة الرياضيات للصف الخامس الأساسي.

مشكلة الدراسة:

تعدّ مرحلة التعليم الأساسي المدخل الرئيسي لمراحل التعليم اللاحقة وحجر الأساس في عملية بناء شخصية المتعلم وقدراته العقلية والنفسية والاجتماعية وتتضمن هذه المرحلة عدداً من الصفوف تتماشى مع مراحل نمو المتعلم وتتعدد الموضوعات والمواد الدراسية في كل صف دراسي. وتعدّ الرياضيات من المواد الدراسية ذات البناء المتسلسل لموضوعاتها فكل جزء من المحتوى مستند على ما سبقه ويمهد لما يليه، فالضرب يستند على الجمع والطرح. ويمهد لتعلّم القسمة. لذلك يجب الاهتمام بجميع تفاصيل المحتوى وعدم إهمال أي جزء حتى لا يفقد التلميذ حلقة مهمة من سلسلة التعلّم الرياضي. وانطلاقاً من ذلك كان تحصيل التلاميذ في مادة الرياضيات أمراً لا بد من إيلائه جانب من الاهتمام من خلال الاهتمام ببناء اختبارات التحصيل المناسبة التي تتمتع بالخصائص السيكومترية المطلوبة والتي تحقق أهداف التعليم في المرحلة الأساسية. ومما هو معروف أن هناك مدخلين أساسيين لبناء الاختبارات والمقاييس النفسية والتربوية وطرق تحليل وتفسير البيانات المستمدة منها وهما المدخل التقليدي والمدخل المعاصر. وبشكل عام فإن لكل من المدخلين طرائقه الخاصة في تقدير الخصائص السيكومترية لمفردات الاختبار وهذه الطرائق تختلف في المدخل التقليدي عن نظرية الاستجابة للمفردة ولكل من النوعين مزايا وعيوب نتج عنها وجود جدل حول استخدامات هذين المدخلين وأيهما أفضل وقد أشارت نتائج الدراسات التي أجريت للمقارنة بينهما من حيث معالجتهما للخصائص السيكومترية للاختبارات إلى تناقض واختلاف في النتائج، ومن هنا تنبعث مشكلة الدراسة الحالية في محاولة تعرّف مدى تشابه أو اختلاف كل من النظريتين في تقدير صعوبات مفردات اختبار تحصيلي. وبناءً على ما سبق تتمثل مشكلة الدراسة في محاولة الإجابة على السؤال الرئيس التالي:

ما مستوى صعوبة بنود اختبار تحصيلي في مادة الرياضيات لتلاميذ الصف الخامس الأساسي وفق كل من النظرية التقليدية ونظرية الاستجابة للمفردة؟

يتفرع عنه الأسئلة التالية:

1) ما قيمة معاملات صعوبة المفردات وفق النظرية التقليدية؟
2) ما تقديرات معاملات صعوبة المفردات وفق نظرية الاستجابة للمفردة؟
3) هل هناك تطابق في تقديرات صعوبة المفردات وفق النظريتين؟

أهمية الدراسة:

تستمد هذه الدراسة أهميتها من:

- أهمية مادة الرياضيات وأهمية التحصيل الجيد فيها والذي يرتبط بالتقويم الجيد.
- أهمية الموضوع وهو المقارنة بين صعوبات مفردات اختبار تحصيلي المقدّرة بنموذج راش وبالنظرية التقليدية وهو مجال اهتمام كثير من الدراسات الحديثة.
- الحصول على نتائج موضوعية للمقارنة بين مدخلي القياس ولاسيما عند استخدام نموذج راش اللوغاريتمي أحادي المعلم.
- توفير اختبار تحصيلي يتضمن مفردات تقيس قدرات التلاميذ مقاسة على نفس متصل السمة.

أهداف الدراسة:

- تعرّف معاملات الصعوبة لمفردات اختبار تحصيلي والمقدّرة بنموذج راش وتلك المقدّرة بالنظرية التقليدية.
- تعرّف مدى مطابقة صعوبات فقرات اختبار تحصيلي في الرياضيات بين كل من النظرية التقليدية في القياس ونظرية الاستجابة للمفردة.

حدود الدراسة:

يقتصر تعميم نتائج الدراسة على الحدود الآتية:

- **الحدود الموضوعية:** اقتصرت الدراسة الحالية على استخدام النظرية التقليدية ونموذج راش أحادي المعلم في مقارنة تقدير خصائص مفردات اختبار تحصيلي في الرياضيات للصف الخامس الأساسي.
- **الحدود البشرية:** تلاميذ الصف الخامس الأساسي من مدارس الحلقة الأولى في مدينة بانياس
- **الحدود الزمانية:** الفصل الأول والفصل الثاني من العام الدراسي 2024/2023
- **الحدود المكانية:** تم تطبيق الاختبار التحصيلي في مدارس الحلقة الأولى مدينة بانياس للتعليم الأساسي.

مصطلحات الدراسة:

التحصيل الدراسي: يعرفه جابلن بأنه: "مستوى محدد من الإنجاز، أو براعة في العمل المدرسي يقاس من قبل المعلمين، أو بالاختبارات المقررة" (العيسوي وآخرون، 2006، ص 13). ويعرف **اجرائياً** "بانه الدرجة التي يحصل عليها التلميذ في الاختبار التحصيلي الذي أعدته الباحثة في الوحدة الثانية من كتاب الرياضيات للصف الخامس الأساسي".

- **النظرية التقليدية في القياس Classical Test Theory:** "هي نموذج بسيط يحتوي تعريف لمفهوم العلامة الظاهرية والحقيقية والخطأ العشوائي في القياس وافتراضات تتعلق بينها. وهي واحدة من نظريات القياس التي تستخدم بغرض تحديد العوامل التي تؤثر على الدرجة التي يحصل الفرد عليها في الاختبار وترتكز هذه النظرية على مفهوم الدرجة الحقيقية والدرجة الخطأ والذي يفترض أنه لو أمكن أن نجري الاختبار عدة مرات على الفرد بعناصر جديدة وتحت ظروف مختلفة فإننا نحصل على درجات ملحوظة مختلفة متوسطها هو أقرب تقدير غير متحيز لقدرة الفرد أو درجته الحقيقية". (أبو هاشم،2006،ص6)

- **نظرية الاستجابة للمفردة Item Response Theory:** "هي نظرية حديثة في القياس النفسي والتربوي يتم فيها تحديد العلاقة بين أداء المفحوص والسمة الكامنة موضع القياس وفق دالة رياضية محددة". (دبوس،2016،ص1461) "تهدف لنمذجة العلاقة بين مستوى سمة معينة لدى الفرد التي يقيسها اختبار معين واستجابته لمفردة من مفردات الاختبار". (أحمد، 2010، ص565)
- **النموذج أحادي المعلم (نموذج راش) (Rasch Model) One-Parameter Logistic Model):** "هو أحد نماذج نظرية الاستجابة للمفردة ويرتبط باسم جورج راش G. Rasch عالم الرياضيات الدنماركي بجامعة كوبنهاغن. ويعد النموذج أحادي المعلم أبسط نماذج الاستجابة للمفردة وأكثرها استخداماً في بناء الاختبارات وبنوك الأسئلة وتحليل مفرداتها ويفترض نموذج راش انعدام التخمين وتساوي معاملات التمييز ويقوم بتقدير صعوبة المفردات. ويستند نموذج راش إلى افتراض أنه كلما ازدادت قدرة الفرد عن صعوبة المفردة ازداد احتمال حدوث الاستجابة الصحيحة للمفردة والعكس صحيح وإذا تساوت قدرة الفرد مع صعوبة المفردة فإن احتمال حدوث الاستجابة الصحيحة 50% أي يتساوى مع احتمال حدوث الاستجابة الخطأ." (مسعود،2014، ص22)
- **صعوبة الاختبار:** يشير مفهوم صعوبة المفردة الاختبارية إلى عدد التلاميذ الذين فشلوا في الإجابة على المفردة على عدد الطلاب الكلي.

الدراسات السابقة:

دراسة Marie Wiberg (2004):

عنوان الدراسة: CLASSICAL TEST THEORY vs. ITEM RESPONSE THEORY An evaluation of the theory test in the Swedish driving-license test

هدف الدراسة: تفحص أي من نماذج نظرية الاستجابة للمفردة أكثر مناسبة لاستخدامه عند تطبيق الاختبار النظري في رخصة القيادة السويدية إضافة إلى مقارنة النتائج بين نظرية الاستجابة للمفردة والنظرية التقليدية في القياس.

عينة الدراسة: تكونت العينة من 5404 من المفحوصين التي طبق عليها نسخ الاختبار النظري لرخصة القيادة السويدية في كانون الثاني 2004 تكون الاختبار من 65 فقرة.

نتائج الدراسة: عند استخدام النظرية التقليدية في تفسير نتائج الاختبار تبين أن الثبات الداخلي كان عال 0.82 إضافة إلى أن قيمة p للمفردات تراوحت بين 0.13 إلى 0.97 إن تمييز المفردات لم يكن متشابهاً بين المفردات نتيجة أخرى مهمة يجب أن نضمن بارامتر التخمين طالما أن بعض الممتحنين ذوي القدرة المنخفضة يميلون إلى تخمين الجواب الصحيح على المفردات الأكثر صعوبة وعند استخدام نظرية الاستجابة للمفردة بيّن التحليل العاملي أن هناك عامل واحد مهيمن أكثر من بقية العوامل الأخرى مما يؤكد فرضية أحادية البعد وتمييز المفردة لم يكن متساوياً بين المفردات والتخمين يجب أن يضمن في النموذج طالما أن الممتحنين ذوي القدرة المنخفضة ما زالوا يستطيعون الإجابة على المفردات الصعبة بطريقة صحيحة والنتيجة أن النموذج الثلاثي البارامتر هو المناسب لهذه الدراسة ومن مقارنة النظريتين التقليدية والحديثة تبين النتائج أن النظرية الحديثة تعطي معلومات ثمينة حول المعرفة الحقيقية لدى المفحوص كما أن تقدير بارامترات المفردات مستقل عن العينة المستخدمة هذه النتيجة مهمة خاصة عندما يطبق الاختبار مرات عديدة. كل من النظريتين إذا استخدمتا في تقييم المفردات تضيف معلومات ثمينة حول الاختبار.

دراسة غادة خالد عيد (2004):

عنوان الدراسة: الدرجة الحقيقية المقدّرة باستخدام نظرية السمات الكامنة والنظرية التقليدية (دراسة سيكومترية)

هدف الدراسة: الكشف عن الفروق في الدرجات الحقيقية لليقظة العقلية المقدّرة بنموذج راش وتلك المقدّرة بالنظرية التقليدية.

عينة الدراسة: تكونت العينة من 250 طالب وطالبة من طلاب كلية التربية في جامعة الكويت.

نتائج الدراسة: بينت النتائج عدم وجود فروق ذات دلالة إحصائية بين متوسطي الدرجات الحقيقية المقدرة بنموذج راش وتلك المقدرة بالمدخل الكلاسيكي لمقياس اليقظة العقلية.

دراسة السيد محمد أبو هاشم (2006):

عنوان الدراسة: دراسة مقارنة بين النظرية التقليدية ونموذج راش في اختيار فقرات قائمة مداخل الدراسة لدى طلاب الجامعة.

هدف الدراسة: هدفت الدراسة إلى تعرّف الخصائص السيكومترية لقائمة مداخل الدراسة باستخدام النظرية التقليدية ونموذج راش والتحقق من إمكانية التنبؤ من درجات الطلاب على قائمة مداخل الدراسة المشتقة من نموذج راش بالتحصيل الدراسي لديهم.

عينة الدراسة: تكونت عينة الدراسة من 244 طالباً وطالبة بالفرقة الرابعة بكلية التربية بجامعة الزقازيق موزعة على خمسة تخصصات.

نتائج الدراسة: بينت النتائج تمتع قائمة مداخل الدراسة بمفرداتها ومكوناتها الفرعية والرئيسية بدرجة مرتفعة من الاتساق الداخلي والصدق والثبات في ضوء النظرية التقليدية للقياس. كما بينت النتائج دقة القائمة وثباتها عند تجريبها وفق أسس النظرية الحديثة وتحقق خصائصها السيكومترية طبقاً لنموذج راش. لم تتأثر تقديرات القدرة للأفراد باختلاف مستوى عينة التحليل كما أن تدريج الاختبار لا يتغير بتغير مستوى قدرة عينة الأفراد المستخدمة في الحصول على هذا التدريج. وجود أثر دال لأداء الأفراد في قائمة مداخل الدراسة بعد تطويره باستخدام نموذج راش في التنبؤ بالتحصيل الدراسي لدى طلاب الجامعة ويرجع ذلك لارتفاع موضوعية القائمة بعد تطويرها باستخدام نموذج راش.

دراسة حمدي يونس أبو جراد (2014)

عنوان الدراسة: دقة التنبؤ بحالة الغضب من سمة الغضب دراسة سيكومترية مقارنة بين النظرية التقليدية والحديثة في القياس.

هدف الدراسة: هدفت الدراسة إلى المقارنة بين نموذج سلم التقدير المنبثق عن نموذج راش و النظرية التقليدية في القياس من حيث دقة التنبؤ بحالة الغضب من سمة الغضب لدى طلبة الجامعة من خلال تقديراتهم على مقياسي سمة، وحالة الغضب.

عينة الدراسة: تم تطبيق مقياسي سمة وحالة الغضب على عينة مكونة من (125) طالبا وطالبة من طلبة جامعة القدس المفتوحة لمعايرة مفردات الاختبارين، و لغرض المقارنة فقد تم تطبيق مقياسي سمة وحالة الغضب بعد معايرة مفرداتهما على عينة مكونة من (80) طالبا وطالبة من طلبة جامعة القدس المفتوحة منهم (45) طالبا و(35) طالبة تم اختيارهم من خارج عينة التدريج.

نتائج الدراسة: بينت نتائج الدراسة أن هناك علاقة ارتباطية دالة إحصائياً بين سمة وحالة الغضب في كلا الأسلوبين وأن دقة التنبؤ باستخدام نموذج سلم التقدير كأحد نماذج النظرية الحديثة أعلى منها في النظرية التقليدية.

دراسة حجازي والخطيب (2014):

عنوان الدراسة: التوافق بين النظرية التقليدية والنموذج ثنائي المعلمة في مطابقة فقرات اختبار محكي المرجع في أحكام التلاوة والتجويد.

هدف الدراسة: الكشف عن التوافق بين النظرية التقليدية والنموذج ثنائي المعلمة في مطابقة فقرات اختبار محكي المرجع في المادة النظرية لأحكام التلاوة والتجويد.

عينة الدراسة: تم تطبيق الاختبار على عينة مكونة من 404 طالباً وطالبة موزعين على 16 مركزاً من مراكز جمعية المحافظة على القرآن الكريم.

نتائج الدراسة: أشارت نتائج الدراسة إلى مطابقة 40 فقرة للنظرية التقليدية و39 فقرة للنموذج ثنائي المعلمة كما تطابقت 39 فقرة مع كل من النظرية التقليدية والنموذج ثنائي المعلمة وقد بلغت قيمة معامل ثبات الاختبار وفق النظرية التقليدية 0.92 بينما كانت وفق النموذج ثنائي المعلمة 0.94 أما معامل الصدق فقد بلغت قيمته وفق النظرية التقليدية 0.73 بينما كانت وفق النظرية الحديثة 0.65.

تعليق على الدراسات السابقة: نلاحظ أن معظم الدراسات كانت تهدف إلى بناء اختبارات تحصيلية ونفسية باستخدام نظرية الاستجابة للمفردة ومقارنة النتائج مع نتائج نظرية القياس التقليدية. وتأتي هذه الدراسة بشكل مميز لمقارنة معامل الصعوبة وفق النظريتين باختبار في مادة الرياضيات.

الجانب النظري للدراسة:

يعدّ الوصول إلى أعلى مستويات الدقة والموضوعية في القياس النفسي والتربوي الهدف الأهم الذي تسعى الاتجاهات المعاصرة في القياس والتقويم للوصول إليه أسوة بالعلوم الطبيعية والفيزيائية التي تتجلى فيها العلاقة الدقيقة بين أداة القياس والسمة المراد قياسها وفيها تكون نتائج القياس مستقلة عن الأداة المستخدمة في التوصل إلى هذه النتائج؛ فطول غرفة الصف هو نفسه إذا استخدمنا في قياسه مسطرة خشبية أو متراً معدنياً أو متراً قماشياً أو غيرها من أدوات القياس بينما يكون ذكاء المتعلم أو قدرته التحصيلية في الرياضيات أو غيرها من الموضوعات الخاضعة للقياس والتي تختلف درجتها ومقدارها لدى نفس المتعلم عند استخدام اختبارين للذكاء أو اختبارين تحصيليين في الرياضيات والسبب أن الظواهر الفيزيائية مختلفة تماماً عن الظواهر السلوكية والنفسية والقدرات العقلية التي تخضع للقياس والتقويم حيث أن الأولى قابلة للملاحظة والقياس المباشر بينما الأخرى فهي في معظمها غير ظاهرة بل يستدل على وجودها من استجابات الفرد على مقياس أو اختبار أو غيرها "تبرز الفروق بين القياس الطبيعي والقياس النفسي من خلال الفروق الأساسية في طبيعة السمات الطبيعية والنفسية حيث أن السمات الطبيعية كالطول والوزن والمساحة هي سمات يمكن قياسها بشكل مباشر وتتحقق الموضوعية في قياسها أكثر منها في القياس النفسي وذلك لأن الأداة صادقة وثابتة بذاتها أما السمات النفسية فهي مفاهيم مجردة لا يمكن إدراكها مباشرة بالحواس ولكننا نستدل على مقدار ونوع ما يمتلكه الفرد منها" (العنزي،2010،ص6)

وبشكل عام هناك مشكلات تتعلق بطبيعة القياس النفسي والتربوي منها:

- نسبية القياس التربوي نظراً لعدم وجود وحدات متفق عليها في القياس فاتساق وحدات القياس يظل نسبياً بدرجة كبيرة وليس مطلقاً.
- القياس التربوي غير مباشر حيث أن معظم السمات أو الخصائص المراد قياسها للأغراض التربوية لا يمكن قياسها قياس مباشر وإنما يمكن فقط الاستدلال عليها.
- مشوب بالأخطاء حيث أن أدوات القياس التربوي تتباين درجة دقتها. (المرجع السابق، ص7)

نظريات القياس:

تعددت أساليب بناء الاختبارات والمقاييس النفسية والتربوية وطرائق تحليل وتفسير البيانات المستمدة منها بناء على الأسس النظرية والنماذج المستخدمة وفي الآتي توضيح لمدخلين من المداخل الأساسية في القياس التربوي:

النظرية التقليدية:

وهي أقدم النظريات التي استخدمت في وبناء الاختبارات وتطويرها وقد استمر استخدامها في بناء الاختبارات التحصيلية وتحليل خصائصها لفترة طويلة من الزمن وذلك لأكثر من 80 عاماً وقد كانت النظرية التقليدية أصلاً هي الإطار المرجعي

لتحليل وتطوير الاختبارات وتستند على أن آخذ الاختبار عنده نتيجة ملاحظة ونتيجة حقيقية، النتيجة الملاحظة عادة تبدو كتقدير للنتيجة الحقيقية زائد ناقص بعض أخطاء القياس. إحدى مزايا هذه النظرية أنها تعتمد على فرضيات سهلة التفسير، ولكنها تُنتقد لأن الدرجة الحقيقية ليست خاصية مطلقة للمتعلم لأنها تعتمد على محتوى الاختبار؛ فإذا كان هناك ممتحنون بمستويات مختلفة من القدرة في اختبار بسيط أو أكثر صعوبة سيؤدي إلى نتائج ودرجات مختلفة، نقد آخر بأن صعوبة المفردة تعتمد بشكل كبير على عينة الممتحنين الخاضعين لاختبار محدّد لذلك إنه من الصعب المقارنة بين نتائج المختبرين من خلال الاختبارات المختلفة. (Wiberg,2004,p2) كما أن نظرية القياس التقليدية بدأت من فكرة أن درجة الفرد على الاختبار تنتج من تأثير مجموعتين من العوامل عوامل مرغوبة تتكون من صفات مستقرة لدى الفرد يتم قياسها في الاختبار وعوامل غير مرغوبة تؤثر في درجة الفرد وتسمى عوامل الخطأ) (أبو هاشم، 2006، ص8)

افتراضات النظرية التقليدية:

تقوم النظرية التقليدية على مجموعة من المسلمات هي:

- أداء الفرد يمكن قياسه وتقديره.
- أداء الفرد دال لخصائصه.
- الخاصية والأداء والعلاقة بينهما تختلف من فرد إلى آخر (الفروق الفردية).
- القياس الظاهري الكلي يتكون من قياس حقيقي وآخر يرجع إلى الخطأ. (اسماعيل،2007،ص3)
- الدرجة الحقيقية للفرد يفترض أن تكون درجة أو قيمة ثابتة ذلك أنها تمثل قدرة الفرد المقاسة.
- الدرجة التي يحصل عليها الفرد ليس من الضروري أن تمثل درجته الحقيقية ولذا فالدرجة التي يحصل عليها قابلة للتغير حسب الظروف الاختبارية.
- الدرجة التي يحصل الفرد هي نتاج نوعين من الدرجات درجة حقيقية ودرجة خاطئة.

مشكلات النظرية التقليدية:

يمكن تلخيص أهم المشكلات التي تواجه القياس التقليدي في الآتي:

- حذف عناصر ووحدات الاختبار غير المميزة بين أفراد العينة مهما كانت تلك الوحدات أو المفردات ذات قيمة أو كانت ممثلة للأهداف المراد قياسها.
- الاقتصار على اتخاذ الجماعة كمعيار أوحد لتفسير درجة الاختبار مما يفقدنا البصيرة بموقع الفرد وجماعته التي ينتمي إليها على السواء بالنسبة لأهداف التعلم.
- سطوة المنحنى الاعتدالي المعياري فالابتعاد عنه يعتبر انحرافاً سواء في مدى صعوبة أو سهولة المفردات إذ يؤدي ذلك إلى التواء توزيع اعتدالها.
- تقتصر الموازنة بين الأفراد في السمة أو القدرة التي يقيسها الاختبار على تطبيق نفس مفردات الاختبار أو مجموعة مفردات مكافئة لها على كل فرد على تطبيق نفس مفردات الاختبار أو مجموعة مفردات مكافئة لها على كل فرد من الأفراد وبالتالي لا نستطيع الموازنة بين مستويات القدرة إذا أجاب الأفراد على مفردات مختلفة ومتباينة في صعوبتها.
- الدرجة الكلية للأفراد محكومة بمفردات الاختبار فإذا أخذنا عينات المفردات المتباينة في صعوبتها من مجموعة كبيرة من المفردات بحيث تقيس نفس القدرة فإن الدرجات المتوقعة للأفراد في الاختبارات التي تشتمل كل منها على إحدى هذه العينات سوف تختلف باختلاف صعوبة المفردات. وبالتالي فإن الدرجات الخام المستمدة من هذه الاختبارات ليس لها أي معنى أو دلالة في حد ذاتها حيث يختلف معناها ودلالتها باختلاف مدى صعوبة أو سهولة مفردات الاختبار.

- انعدام خطية القياس: وخطية القياس تعني وجود معدل ثابت لتدرج القياس على متصل المتغير موضوع القياس على المدى الواسع من هذا المتغير أي أن الدرجات التي يحصل عليها الفرد في مفردات اختبار يمكن جمعها كما لو كانت تمثل ميزاناً خطياً وهذا يعني عدم وجود مقياس قدرة خاص مثل قياس الطول أو الوزن يمكن تطبيقه على المدى الواسع من القدرة.

نظرية الاستجابة للمفردة:

مع الاتجاه المتزايد نحو تفريد التعليم والموضوعية في نتائج القياس والتقويم التربوي والبحث عن الأسلوب المناسب للتوصل إلى تقدير المستوى الفعلي والحقيقي لقدرات المتعلم من أجل توجيهه وإرشاده مهنياً وتعليمياً فقد تزايد الاهتمام من قبل علماء القياس ولا زال حول البحث عن أدوات أكثر مرونة وملاءمة لعملية القياس ولمواجهة الكثير من نواحي الضعف وأوجه القصور التي عجزت النظرية التقليدية عن حلها حيث ظهرت نظريات حديثة في القياس ومنها نظرية الاستجابة للمفردة والتي تعدّ من أهم نظريات القياس الموضوعي التي ظهرت في النصف الثاني من القرن العشرين والتي تطورت وانتشرت في السنوات الأخيرة بشكل مذهل حيث أصبحت البديل الوحيد في جميع الاختبارات الدولية المقننة حول العالم (مسعود،2010،ص88) ولقد طُوّرت نظرية الاستجابة للمفردة للتغلب على مشكلات النظرية التقليدية. إحدى الفرضيات الأساسية لنظرية الاستجابة للمفردة أن القدرة الكامنة لدى الممتحن مستقلة عن محتوى الاختبار، والمفحوص ذو مستوى القدرة المرتفعة سيكون عنده احتمال عال للاستجابة على المفردة بشكل صحيح، كما أنه لا تهم أي المفردات سيتم استخدامها لتقدير قدرة المفحوص، هذه الفرضية تجعل من الممكن مقارنة نتيجة المفحوص بالرغم من أنه تعرّض لنسخ مختلفة من الاختبار. (Wiberg,2004,p2)

افتراضات نظرية الاستجابة للمفردة:

1. أحادية البعد Uni–dimensionality: تفترض نظرية الاستجابة للمفردة أن أداء الممتحن على الاختبار يمكن تفسيره بواسطة قدرة واحدة أو سمة كامنة واحدة فمعظم الاختبارات تبنى لقياس سمة واحدة. وهذا الافتراض يصعب تحقيقه بشكل مطلق لأنها تتأثر بمواصفات وخصائص عينة المفحوصين المستخدمة في التحليل وبالتالي فإن التأكد منها يكون نسبياً. (مسعود،2010، 104) (كما هو الحال في أية نظرية فإن نظرية الاستجابة للمفردة تقوم على بعض الفرضيات حول ما يحدث أثناء تأدية الممتحن لاختبار ما، فهناك فرضيتان تبرزان أكثر أهمية لتصميم الاختبارات؛ أول هذه الفرضيات أن المفردات في الاختبار تقيس فقط قدرة واحدة أو مهارة وهذه هي فرضية أحادية البعد حيث أن البعد هو تعبير آخر عن القدرة) (Whitehouse, 2012, p5)
2. استقلالية القياس (الاستقلال الموضعي) Local Independence: ويقصد به أن تكون استجابات الفرد للمفردات في الاختبار مستقلة استقلالاً إحصائياً؛ وهذا يعني ألا تؤثر استجابة الفرد لإحدى مفردات الاختبار على استجابته للمفردات الأخرى أي أن تقدير صعوبة المفردة لا يعتمد على تقديرات صعوبة المفردات الأخرى المكونة للاختبار ولا على تقديرات قدرة الأفراد الذين يجيبون عليها وكذلك تقدير قدرة الفرد لا يعتمد على تقديرات قدرة أي مجموعة من الأفراد الذين يؤدون الاختبار أو على تقديرات صعوبة البنود التي يؤدونها.
3. المنحنيات المميزة للمفردات Items Characteristic Curves: وهي دوال رياضية تصف شكل العلاقة بين احتمال نجاح الفرد في الإجابة على مفردة ما والقدرة التي تقيسها مجموعة المفردات التي يشتمل عليها الاختبار ويتم ذلك من خلال دالة متزايدة مطردة تسمى أحياناً دالة الاستجابة للمفردة وتعتبر تمثيلاً جيداً للعلاقة بين احتمال الإجابة الصحيحة ومستويات القدرة موضوع القياس. (محمود،2012،ص23)
4. سرعة الأداء في الاختبار Speediness: يقصد بها مدى اعتماد درجة الفرد في الاختبار على أدائه فيه وتقاس عادة بنسبة الأفراد الذين لم يصلوا ولم يجيبوا عن آخر مفردة في الاختبار. وتفترض نظرية الاستجابة للمفردة أن

عامل السرعة لا يؤدي دوراً في الإجابة عن مفردات الاختبار بمعنى أن إخفاق المفحوصين في إجابة مفردات الاختبار يرجع إلى انخفاض قدراتهم وليس إلى تأثير عامل السرعة على إجاباتهم ويمكن تقدير ما إذا كان عامل السرعة قد لعب دوراً في الإجابة عن طريق معرفة عدد الأفراد الذين لم يتمكنوا من الإجابة عن جميع مفردات الاختبار الذي أجري عليهم. (اسماعيل،2007،ص31)

نماذج نظرية الاستجابة للمفردة:

تحتوي نظرية الاستجابة للمفردة على عدة نماذج رياضية لوغاريتمية حيث أن (نماذج نظرية الاستجابة في معظم أشكالها هي بلا شك النماذج الأكثر استعمالاً في برامج التقييم ليس فقط في الولايات المتحدة لكن في العديد من الدول الأخرى في أحدث برامج البحث في مجال القياس التربوي والنفسي) (Carlson,2013,p1) (تستخدم نظرية الاستجابة للمفردة نماذج احتمالية لتربط استجابة الفرد بالمفردة عبر دالة القدرة حيث أن المعلومات الكمية يمكن الحصول عليها من خلال العلاقة النوعية المحددة بين معالم المفردة ومستوى القدرة للممتحن وهذه العلاقة النوعية هي دالة معلومات المفردة) (Ariel,2005,p11) وتقوم نماذج نظرية الاستجابة للمفردة على افتراض (أن احتمال الاستجابة الصحيحة على مفردة ربما توصف بمعادلة رياضية تدمج مصطلحات صعوبة المفردة، تمييز المفردة، وقدرة المرشح. وبتوصيل درجات المرشحين لكل مفردة والتي تُجمع من استجاباتهم على المفردات إلى المعادلة تكون قيم صعوبات المفردات مستقلة عن قدرات المرشحين) (Whitehouse, 2012, p4) (كل نموذج من نظرية الاستجابة للمفردة يتنبأ باحتمال أن شخص معين سيعطي استجابة معينة على مفردة معينة الأفراد يمكن أن يكون لهم مستويات مختلفة للقدرة والمفردات يمكن أن تختلف في أوجه أكثر أهمية بعضها أسهل وبعضها أكثر صعوبة) (Partcher,2004,p7) وهذه النماذج هي:

- نموذج راش اللوغاريتمي أحادي المعلم One–Parameter Logistic Model: يعتبر هذا النموذج من أبسط نماذج نظرية الاستجابة للمفردة ويفترض أن لكل فرد استجابتين إما (1أو صفر) على مفردة واحدة فقط وهذا يعني أن النموذج يفترض انعدام التخمين وتساوي معاملات التمييز ويقوم بتقدير صعوبة المفردات فقط. (يتنبأ باحتمال الاستجابة الصحيحة من التفاعل بين قدرة الفرد ومعلم صعوبة المفردة) (Partcher,2004,p7)
- نموذج لورد اللوغاريتمي ثنائي المعلم Tow–Parameters Logistic Model: تم في هذا النموذج إضافة معلم جديد إلى نموذج راش وهو معلم التمييز أي أن هذا النموذج يفترض انعدام التخمين ويقوم بتقدير صعوبة المفردات ومعلم التمييز.
- نموذج بيرنبوم اللوغاريتمي ثلاثي المعلم Three–Parameter Logistic Model: يعد هذا النموذج امتداداً للنموذج اللوغاريتمي أحادي المعلم والنموذج اللوغاريتمي ثنائي المعلم وقد أضاف بيرنبوم معلماً ثالثاً أطلق عليه معلم الخط التقاربي الأدنى أو معلم التخمين. (اسماعيل،2007،ص22)

نموذج راش اللوغاريتمي أحادي المعلم:

وينسب لـ جورج راش Georg Rasch ويعد أبسط نماذج نظرية الاستجابة للمفردة وأكثرها استخداماً في بناء الاختبارات ويفترض هذا النموذج انعدام التخمين c وتساوي معاملات التمييز a ويهتم بتقدير **صعوبة** المفردات فقط b. (يعد نموذج راش حالة خاصة من نموذج بيرنبوم ثنائي البارامتر فهو يقوم على افتراض أن جميع مفردات الاختبار تميز بنفس القدر بين الأفراد ولكنها تختلف عن بعضها في الصعوبة فقط ويهدف هذا النموذج إلى تحقيق الموضوعية Objectivity في القياس بمعنى أن درجة الفرد في الاختبار لا يجب أن تكون دالة لعينة الأفراد الذين استخدموا في التدريج الأصلي للمفردات التي يشتمل عليها الاختبار كما أنه يجب أن يحصل الفرد على نفس الدرجة في كل من اختبارين يقيسان نفس السمة أو القدرة). (عيد،2004،ص250) (أي مفردة في النموذج أحادي المعلم تكون غنية بالمعلومات المفيدة للممتحنين

الذين قدرتهم مساوية لصعوبة المفردة عندما تصبح القدرة إما أقل أو أكبر من صعوبة المفردة معلومات المفردة تتناقص إذاً نحن بحاجة مفردات من صعوبات مختلفة إذا أردنا أن ننجز قياس جيد للأشخاص الذين يملكون كل الأنواع من القدرات المختلفة) (Partcher,2004,p14)

ويتميز نموذج راش عن غيره من نماذج نظرية الاستجابة للمفردة بمميزات أساسية تتعلق بمفهوم الموضوعية في القياس وهذه المميزات تجعله أكثر تطبيقاً واستخداماً في تحليل مفردات الاختبارات وهذه المميزات هي:

- إذا افترضنا توفر عينة كبيرة من المفردات الاختبارية بحيث تقيس جميعها نفس القدرة أو السمة فإنه يمكن الحصول على قيم تقديرية لقدرة الفرد وتكون هذه القيمة مستقلة إحصائياً عن عينة المفردات التي اختبر بها الفرد.
- إذا افترضنا توفر عينة كبيرة من الأفراد فإنه يمكن الحصول على قيم تقديرية لمعاملات الصعوبة ومعاملات التمييز للمفردات التي اختبروا بها وتكون هذه القيم مستقلة إحصائياً عن عينة الأفراد التي أجري عليها الاختبار.
- يمكن استخدام النموذج للحصول على معامل إحصائي يدل على مدى دقة تقدير قدرة كل فرد وربما تختلف قيمة هذا المعامل من فرد إلى آخر بحسب مستوى قدرة كل منهم.
- يعتبر أقل النماذج في عدد الافتراضات اللازم تحققها في البيانات المستمدة من الاختبار لكي نحصل باستخدامه على تقديرات دقيقة لكل من صعوبة المفردة وقدرة الفرد.
- تمكن علماء القياس من إيجاد حلول إحصائية مناسبة لمشكلة تقدير معالم النموذج في حين أنهم يواجهون مشكلات سيكومترية وإحصائية في تقدير معاملات النماذج الأخرى.
- تتساوى القيمة التقديرية لقدرات الأفراد الذين حصلوا على نفس الدرجة الكلية في الاختبار بغض النظر عن عينة المفردات التي اختبر بها كل منهم في حين أنه ليس من الضروري حدوث ذلك في حالة النماذج الأخرى. (عيد،2004،ص249)

منهج الدراسة الحالية:

تم الاعتماد في هذه الدراسة على المنهج الوصفي الذي يوفر للباحث البيانات والحقائق التي تصف ما هو موجود والحصول على تفسيرات للنتائج التي يتوصل إليها وذلك تماشياً مع كون هذه الدراسة تهدف إلى بناء اختبار تحصيلي في الرياضيات.

مجتمع وعينة الدراسة:

تكوّن مجتمع الدراسة من تلاميذ الصف الخامس الأساسي في مدارس مدينة بانياس للعام الدراسي 2024/2023 والبالغ عددهم /1318/ وفق احصائيات المجمع التربوي في مدينة بانياس وقد تكونت عينة الدراسة من (100) تلميذاً وتلميذة وتم اختيار العينة بالطريقة العشوائية البسيطة بنسبة 7.58%.

أداة الدراسة:

لتحقيق أهداف الدراسة تم بناء تجمع من مفردات اختبار تحصيلي في الرياضيات من نوع أسئلة الاختيار من متعدد مؤلف من (25) مفردة لكل مفردة أربعة بدائل واحد منها فقط هو الصحيح. تم بناء الاختبار بعد ان تم تحليل محتوى الوحدة الثانية من كتاب الرياضيات للصف الخامس والوقوف على كافة المفاهيم والمصطلحات والمخرجات المعرفية والمهارية. تم بعد ذلك بناء جدول المواصفات للاختبار بما يغطي موضوعات الوحدة الثانية من الكتاب بالإضافة إلى الأهداف التعليمية. تم بعد ذلك دراسة الخصائص السيكومترية للاختبار.

الخصائص السيكومترية للاختبار:

صدق الاختبار:

الصدق المحكمين: تم عرض الاختبار على عدد من الأساتذة في كلية التربية والعلوم في جامعة تشرين وبناء على ملاحظاتهم تم تعديل بعض الأسئلة من حيث الصياغة وتم تعديل بعض العبارات وظل الاختبار مكوناً من (25) مفردة.

وبعد ذلك تم تطبيق الاختبار على عينة استطلاعية من تلامذة الصف الخامس خارجة عن عينة البحث الأساسية بلغ عدد أفرادها /20/ تلاميذاً وذلك للتأكد من أن كافة البنود مفهومة من قبلهم؛ وقد بين التطبيق أن كافة البنود مفهومة وواضحة وأبدى أفراد العينة رغبة في الإجابة.

الاتساق الداخلي: تم التأكد من الاتساق الداخلي للبنود من خلال حساب معامل ارتباط بيرسون بين الدرجة الكلية ودرجة كل بند. لقد كانت قيم معاملات الارتباط كلها دالة إحصائيا عند مستوى دلالة اقل من /0.05/. وكانت أقل قيمة لمعامل الارتباط /0.71 / وأعلى قيمة هي / 0.92/.

الصدق التمييزي: تم ترتيب أفراد العينة الاستطلاعية وفق الدرجة الكلية للاختبار ثم تحديد مجموعتين مجموعة عليا ومجموعة دنيا. كل من المجموعتين تحوي (5) من التلاميذ. تم بعد ذلك تطبيق اختبار مان وتني لمقارنة بين متوسطي المجموعتين وكانت النتيجة كما يلي:

الجدول رقم (1): اختبار مان وتني للتحقق من الصدق التمييزي

المجموعة	العدد	المتوسط	الانحراف المعياري	قيمة Z	مستوى الدلالة	القرار
العليا	5	22.22	1.28	0.842	0.000	يوجد فرق
الدنيا	5	14.22	1.98			

بلغت قيمة مستوى دلالة الاختبار 0.000 وهي اصغر من 0.05 أي توجد فروق دالة إحصائياً بين الفئة العليا والفئة الدنيا لصالح الفئة العليا مما يشير إلى أن الاختبار قادر على التمييز بين التلاميذ الجيدين والمتوسطين والضعيفين في التحصيل.

ثبات الاختبار:

الثبات باستخدام معامل كودر–ريتشارد سون–20: تم استخدام معامل كودر ريتشاردسون (20) لحساب الثبات لان البيانات من النوع الثنائي (0–1) وهو معامل ثبات قريب من معامل الفاكرونباخ والذي أعطى قيمة /0.87/ وهي قيمة ثبات عالية. تبين أن الاختبار يتمتع ببنية داخلية ثابته.

الثبات باستخدام التجزئة النصفية: تم حسابه من خلال حساب معامل ارتباط سبيرمان براون المعدل لمعامل ارتباط بيرسون بين جزئي الاختبار بعد تقسيم الاختبار إلى قسمين يحوي القسم الأول البنود الزوجية ويحوي الجزء الثاني البنود الفردية وكانت قيمة معامل ارتباط سبيرمان براون /0.79/ وقد جاءت قمية معامل جتمان المحرر من حجم العينة /0.72/. وهي قيم ثبات جيدة.

الثبات بالإعادة: أما الثبات بالإعادة والذي يسمى ثبات استقرار الدرجات مع مرور الزمن، فقد تم التحقق منه من خلال إعادة تطبيق الاختبار على نفس العينة السابقة مرة ثانية بعد مرور أسبوعين على التطبيق الأول وقد بلغ معامل ارتباط بيرسون بين التطبيقين (0.92) وهو معامل ثبات عالٍ وقيمته جيدة، مما يدل على أن الاختبار ثابت. وبذلك يكون الاختبار صالحاً للتطبيق.

مما سبق نستنتج أن الخصائص السيكومترية للاختبار جيدة وبالتالي يمكن تطبيق الاختبار على عينة البحث النهائية والتي بغلت /100/ تلميذاً وتلميذة.

عرض نتائج الدراسة ومناقشتها:

السؤال الأول: ما تقديرات معاملات صعوبة المفردات وفق النظرية التقليدية؟

تم حساب معاملات الصعوبة لكافة البنود وفق النظرية التقليدية للقياس وقد جاءت البنود قيم معاملات الصعوبة كما في الجدول (2)

الجدول رقم (2) : قيم معاملات الصعوبة للاختبار

البند	معامل الصعوبة	البند	معامل الصعوبة	البند	معامل الصعوبة	البند	معامل الصعوبة	البند	معامل الصعوبة
A1	0.56	A6	0.69	A11	0.71	A16	0.55	A21	0.66
A2	0.86	A7	0.67	A12	0.87	A17	0.64	A22	0.66
A3	0.56	A8	0.77	A13	0.80	A18	0.74	A23	0.70
A4	0.86	A9	0.75	A14	0.60	A19	0.73	A24	0.86
A5	0.93	A10	0.75	A15	0.85	A20	0.71	A25	0.77

كانت اصغر قيمة هي /0.55/ وأكبر قيمة هي /0.90/ علماً إن قيمة معاملات الصعوبة يجب أن تكون ضمن المجال [0.15–0.85]. ومن هذه الجدول نجد أن المفردات /2–4–5–12–24/ كانت قيمتها أكبر من /0.85/ وبالتالي فهي غير ملائمة للاختبار ويجب حذفها من الاختبار قبل تطبيقه في المرة القادمة. وبذلك يصبح عدد بنود الاختبار في المرة القادمة /20/ بنداً.

السؤال الثاني: ما تقديرات معاملات صعوبة المفردات وفق نظرية الاستجابة للمفردة؟

حساب معاملات الصعوبة وفق نظرية الاستجابة للمفردة باستخدام نموذج راش:

تم بعد تصحيح الاختبار إدخال كافة البيانات إلى الحاسوب من خلال برنامج الحزمة الإحصائية SPSS ومن ثم تحوليها إلى الترميز الثنائي (0 إجابة خاطئة، 1 إجابة صحيحة) وتجهيزها لتكون جاهزة للاستخدام والتحليل.

إجراءات التحليل وفق نظرية الاستجابة للمفردة: تم استخدام برنامج Winsteps Ver. 3.75 وفق الخطوات التالية:

كتابة ملف التحكم الخاص بتعليمات البرنامج والذي يحوي الأوامر المطلوب من البرنامج تنفيذها بالإضافة للإعدادات الخاصة بالتحليل.

```
&INST
  TITLE = test2
 PERSON = Person ; persons are ...
   ITEM = Item ; items are ...
  ITEM1 = 4 ; column of response to first item in data record
     NI = 25 ; number of items
  NAME1 = 1 ; column of first character of person identifying label
NAMELEN = 3 ; length of person label
  XWIDE = 1 ; number of columns per item response
  CODES = "01 " ; valid codes in data file
 UIMEAN = 0 ; item mean for local origin
 USCALE = 1 ; user scaling for logits
 UDECIM = 2 ; reported decimal places for user scaling
        DISCRIM=YES
        ASYMPTOTE=YES
&END

END LABELS
  101111111111111111111111
  211111111110111111110011011
  301111111111111111111111
```

الشكل رقم (1): ملف الأوامر لبرنامج Winsteps لإجراء التحليل المطلوب

تم تنفيذ التحليل والتأكد من أن ملف التحكم لا يحوي أي أخطاء.

التحليل الأول بهدف تحديد الأفراد غير الصادقين (خارج حدود الملائمة التقاربية).

التحليل الثاني (بعد حذف الافراد غير الصادقين) بهدف تحديد المفردات غير الملائمة.

التحليل الثالث (النهائي) (بعد حذف كل من الأفراد غير الصادقين والمفردات غير الملائمة) والذي يهدف استخراج النتائج النهائية للتدريج.

نتيجة التحليل الأول: يتم الحصول على **الأفراد** غير الصادقين من خلال إحصاءات الملائمة (IN.ZSTD) والذي يشير إلى إحصاءات الملائمة التقاربية للأفراد (Persons Infit) حيث يتم حذف **الأفراد** اللذين تتجاوز قيم الملائمة التقاربية لهم المدى من (−2 حتى +2). **لم يتم حذف أي فرد.**

نتيجة التحليل الثاني: يتم الحصول على **المفردات** غير الصادقة من خلال إحصاءات الملائمة (IN.ZSTD) والتي تشير إلى إحصاءات الملائمة التقاربية للمفردات (Items Infit) حيث يتم حذف **المفردات** التي تتجاوز قيم الملائمة التقاربية لها المدى من (−2 حتى +2). **تم حذف المفردة رقم /1/** لان حدود الملائمة التقاربية لها كانت تساوي /5.48/ وبذلك أصبح عدد المفردات للاختبار /24/ بنداً.

نتيجة التحليل الثالث: يتم تنفيذ التحليل الثالث من أجل الحصول على **قدرات الأفراد وصعوبات المفردات والتدريج النهائي للاختبار.**

أولاً: استخراج (صعوبات) المفردات: يمثل الجدول التالي درجات صعوبة المفردات

الجدول رقم (3): درجة صعوبة المفردات مقدرة بوحدة اللوجيت لبنود الاختبار.

البند	صعوبة البند	IN.ZSTD	البند	صعوبة البند	IN.ZSTD
2	-0.78	0.34	14	0.79	0.04
3	0.99	-1.99	15	-0.70	0.22
4	-0.78	0.12	16	1.04	-1.57
5	-1.60	0.15	17	0.60	1.11
6	0.34	-1.83	18	0.06	-0.52
7	0.44	-0.87	19	0.12	0.14
8	-0.12	-0.65	20	0.23	0.65
9	0.00	-0.80	21	0.50	1.22
10	0.00	-0.30	22	0.50	0.03
11	0.23	0.67	23	0.29	0.33
12	-0.88	0.32	24	-0.78	0.59
13	-0.32	1.07	25	-0.16	-0.11

ويمثل الشكل التالي ملخصاً يحوي الحد الأدنى والحد الأعلى لتقديرات صعوبة المفردات (MEASURE) مقاسة بوحدة اللوجيت وقد تراوحت مستويات الصعوبة بين (1.6−) و (1.04) وبلغ متوسط القدرة (0.00) وفق اعدادات البرنامج التي وضعناها في ملف التحكم وكذلك الخطأ المعياري الذي يساوي 0.64 وهو أقل من الواحد حسب اعدادات البرنامج.

كما يبين الجدول أن قيمة معامل الارتباط بين الدرجة ومستوى صعوبة البند قد بلغت (0.96) وهي قيمة ارتباط مرتفعة وبلغ معامل ثبات الاختبار وفق معامل كودر – ريتشاردسون 20 (KR−20) 0.85 وهي أيضا قيمة ثبات مرتفعة.

```
     SUMMARY OF 24 MEASURED (NON-EXTREME) Items
-------------------------------------------------------------------------------
|           RAW                          MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE   ERROR      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      73.7     100.0         .00     .26      1.00    -.1   1.02     .0 |
| S.D.       9.9        .2         .64     .04       .08     .9    .16    1.0 |
| MAX.      93.0     100.0        1.04     .40      1.15    1.2   1.41    2.3 |
| MIN.      55.0      99.0       -1.60     .22       .85   -1.99    .77   -1.9 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .26  ADJ.SD     .58  SEPARATION  2.22  Item   RELIABILITY  .83 |
|MODEL RMSE    .26  ADJ.SD     .59  SEPARATION  2.27  Item   RELIABILITY  .84 |
| S.E. OF Item MEAN = .13                                                     |
-------------------------------------------------------------------------------
                DELETED:      1 Items
UMEAN=.000 USCALE=1.000
Item RAW SCORE-TO-MEASURE CORRELATION = .96
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .85
```

الشكل رقم (2): ملخص قدرات الافراد على 24 بنداً اختبارياً وفق نظرية الاستجابة للمفردة

السؤال الثالث: هل يوجد تطابق بين تقديرات صعوبة المفردات بين كل من النظرية التقليدية ونظرية الاستجابة للمفردة؟
من أجل دراسة التطابق تم ترتيب بنود الاختبار وفق درجة الصعوبة في كل من النظرية التقليدية في القياس ووفق نظرية الاستجابة للمفردة، ويبني الجدول التالي نتيجة الترتيب.

الجدول رقم (4): مقارنة بين ترتيب معاملات الصعوبة في كل من النظرية التقليدية ونظرية الاستجابة للمفردة

النظرية التقليدية		نظرية الاستجابة للمفردة	
رقم البند	معامل الصعوبة	رقم البند	معامل الصعوبة
A2	0.86	2	-0.78
A16	0.55	16	1.04
A3	0.56	3	0.99
A14	0.60	14	0.79
A17	0.64	17	0.6
A21	0.66	21	0.5
A22	0.66	22	0.5
A7	0.67	7	0.44
A6	0.69	6	0.34
A23	0.70	23	0.29
A11	0.71	11	0.23
A20	0.71	20	0.23
A19	0.73	19	0.12
A18	0.74	18	0.06
A9	0.75	9	0
A10	0.75	10	0
A8	0.77	8	-0.12
A25	0.77	25	-0.16
A13	0.80	13	-0.32
A15	0.85	15	-0.7
A4	0.86	4	-0.78
A24	0.86	24	-0.78
A12	0.87	12	-0.88
A5	0.93	5	-1.6

لقد قمنا بحذف البند الأول من الاختبار بناء على نتائج نظرية الاستجابة للمفردة وتم إبقاء البنود المحذوفة وفق نظرية الاستجابة للمفردة وتمت الإشارة إليها باللون الغامق وهي (2–4–5–12–24) وذلك فقط لضرورة المطابقة في القيم.

نلاحظ التطابق التام في ترتيب البنود وفق معاملات الصعوبة في كل من النظرية التقليدية ونظرية الاستجابة للمفردة. مما يعني ان تطابق النظريتين في ترتيب بنود الاختبار وفق معاملات الصعوبة. إلا ان النظرية الحديثة في القياس والتقويم قد تفوقت على النظرية التقليدية حيث انها نجحت في الإبقاء على /24/ بنداً من بنود الاختبار في حيث أن النظرية التقليدية أبقت على /20/ بنداً فقط.

مقترحات البحث:

- الاعتماد بشكل أكبر في تحديد خصائص الاختبارات على النظرية الحديثة في القياس والتقويم.
- اجراء المزيد من الدراسات على اختبارات في مقررات أخرى للتأكد من فاعلية النظرية الحديثة في القياس والتقويم من تحديد معاملات الصعوبة للبنود الاختبارية ومقارنتها مع معاملات الصعوبة في النظرية التقليدية.
- دراسة بنية البنود الاختبارية التي تم حذفها في كل من النظريتين للتأكد فيما إذا كان هناك خلل ما في بنية السؤال والبحث عن الأسباب التي أدت إلى حذف البنود والعمل على تحسين البنود الاختبارية التي معاملات بعيدة عن المتوسط بأكثر من ثلاثة انحرافات معيارية حتى تبقى البنود ضمن التوزع الطبيعي للبنود.
- بناء بنوك أسئلة وتدريجها وفق نظرية الاستجابة للمفردة والاستفادة منها في بناء نماذج اختبارية مختلفة ولكن متكافئة في مستوى صعوبتها.
- استخدام نماذج نظرية الاستجابة الأخرى (معلمتين، ثلاثة معالم) في تحديد خصائص الاختبارات التحصيلية.
- **المراجع:**

- أبو جراد، حمدي يونس. (2014) دقة التنبؤ بحالة الغضب من سمة الغضب دراسة سيكومترية مقارنة بين النظرية التقليدية والحديثة في القياس. *مجلة الجامعة الإسلامية*، 22(2)، الصفحات101–120.
- أبو هاشم، السيد محمد. (2006). دراسة مقارنة بين النظرية التقليدية ونموذج راش في اختيار فقرات قائمة مداخل الدراسة لدى طلاب الجامعة. *مجلة كلية التربية*، جامعة الزقازيق، العدد 52، الصفحات 1– 52.
- أحمد، محمد فتح الله سيد. (2010) بناء وتدريج بطارية لقياس الاستعداد الأكاديمي للقبول بالجامعات باستخدام الاختبارات مرجعية المحك CRT ونظرية الاستجابة للمفردة IRT . *مجلة كلية جامعة الأزهر*، العدد 144(7)، الصفحات 559– 641.
- اسماعيل، ميمي السيد أحمد. (2007) *الخصائص السيكومترية لاختبار القدرة العقلية باستخدام نموذج راش لدى طلبة المرحلة الثانوية العامة*. رسالة ماجستير، جامعة الزقازيق.
- بركات، مايا ابراهيم. (2010) *أثر تصميمات المعادلة ومتوسط صعوبة الاختبارات وتوزيع القدرة على معادلة درجات الاختبارات متعددة الأبعاد باستخدام نظرية الاستجابة لمفردة*. رسالة دكتوراه، جامعة القاهرة.
- حجازي، تغريد والخطيب، عبدالله. (2014) التوافق بين النظرية التقليدية والنموذج ثنائي المعلمة في مطابقة فقرات اختبار محكي المرجع في أحكام التلاوة والتجويد. *مجلة جامعة النجاح للأبحاث*، المجلد 28، الصفحات 2240–2270.

- الخياط، ماجد محمد. (2012) درجة مطابقة اختبار تحصيلي وفق نموذج راش أحادي المعلمة في الكشف عن مستوى المعرفة العلمية في المهارات الرياضية لدى طلبة الصف الثامن الأساسي. *مجلة جامعة الأقصى*، 16(1)، الصفحات 87– 111.
- دبوس، محمد. (2016) استخدام نظرية الاستجابة للفقرة في بناء اختبار محكي المرجع في الرياضيات بفقرات ثنائية التدريج ومتعددة التدريج وفق النموذج اللوجستي ثنائي المعلم. *مجلة جامعة النجاح للأبحاث*، المجلد 30، 2016، الصفحات 1454– 1480.
- الشواورة، ياسين سالم حماد. (2013) *دراسة مقارنة بين نموذجي التقدير الجزئي والاستجابة المتدرجة في معادلة درجات الاختبارات*، رسالة دكتوراه، جامعة القاهرة.
- العنزي، محمد طالب. (2010) *أثر شكل الفقرة على معالم الفقرة وثبات الاختبار وفقاً لنظرية استجابة الفقرة (IRT)* . رسالة ماجستير، جامعة مؤته.
- عيد، غادة خالد. (2004) الدرجة الحقيقية المقدرة باستخدام نظرية السمات الكامنة والنظرية التقليدية (دراسة سيكومترية). *مجلة جامعة أم القرى للعلوم التربوية*، 16(2)، الصفحات 228– 287.
- محمود، سومية شكري محمد. (2012) *استخدام نموذج راش في بناء بنك أسئلة لقياس التحصيل في مقرر سيكولوجية التعلم لدى طلاب كلية التربية بالمنيا* . رسالة دكتوراه، جامعة المنيا.
- مسعود، وليد. (2010) *دراسة أثر القياس الدينامي على فاعلية برنامج لتنمية الفهم القرائي باستخدام نظرية الاستجابة للمفردة*. رسالة دكتوراه، جامعة عين شمس.
- مسعود، وليد. (2014) محاضرات في القياس *(نظرية الاستجابة للمفردة مع تطبيقات عملية)*. كلية التربية، جامعة أم القرى.
- Ariel, Adelaide.(2005). *Contributions to Test Item Bank Design and Management*. Print partners I pskamp , B.V, Enschede.
- Aye, Su Mon. (2012) Item Banking for Grade 9 Mathematics : The Application Item Response Theory. *Yangon Institute of Education Research Journal*, 4(1), p1– 12
- Baker, Frank. (2001). The Basics of Item Response Theory. University of Wisconsin, Educational Resources Center, second edition, p1– 176
- Carlson, James and Davier, Matthias. (2013) Item Response Theory. ETS Research Report no. RR–13– 28,New Jersey, p1–70
- Rivera, Jennifer. (2007) Test Item Construction and Validation : Developing a Statewide Assessment for Agricultural Science Education. Cornel University, A Dissertation Presented to the Faculty of the graduate School.
- Whitehouse, Claire. (2012) The Assessment and Qualifications Alliance. Center of education research and policy.
- Wiberg, Marie.(2004) Classical Test Theory Vs. Item Response Theory An evaluation of the theory test in the Swedish driving–license test. EM No 50, p1– 26