The use of the Rasch model in grading a reference test in mathematics for the fifth grade

Dr. Mounzer Boubou*
Raja Assad**

(Received 4 / 12 / 2022. Accepted 25 / 5 / 2023)

□ ABSTRACT □

The current research aims to use the Itme response theory (Rasch model) in constructing a referenced achievement test for the first unit in mathematics for the fifth grade. To achieve this, the descriptive approach was adopted and a set of test items (25 items) were prepared that measure achievement in mathematics in the cognitive domain of the multiple choice type. The research sample consisted of (100) male and female students who were selected by simple random method. The psychometric properties of the achievement test were verified as a tool for the research, and the results showed that the values of the vocabulary difficulty coefficients ranged between (-1.46) and (1.44), and the individual abilities values ranged between (1.32) and (4.70). The results also showed that the paragraphs matched the one-parameter model. In light of the results, a set of recommendations were proposed.

Keywords: item response theory, the one-parameter model, the criterion references test.

Copyright **:**Tishreen University journal-Syria, The authors retain the copyright under a CC BY-NC-SA 04

* **Associate Professor, Department of Measurement and Evaluation, Faculty of Education, Tishreen University, Lattakia, Syria.**
** **PhD student, Department of Child Education, Faculty of Education, Tishreen University, Lattakia, Syria.**

استخدام نموذج راش في تدريج اختبار محكّي المرجع في الرياضيات للصف الخامس الأساسي

د. منذر بوبو*

رجاء أسعد**

(تاريخ الإيداع 4 / 12 / 2022. قبل للنشر في 25 / 5 / 2023)

ملخّص

يهدف البحث الحالي إلى استخدام نظرية الاستجابة للمفردة (نموذج راش) في بناء اختبار تحصيلي محكّي المرجع للوحدة الأولى في مادة الرياضيات للصف الخامس الأساسي. ولتحقيق ذلك تم اعتماد المنهج الوصفي وتم إعداد مجموعة من الفقرات الاختبارية (25 فقرة) تقيس التحصيل في الرياضيات في المجال المعرفي من نوع الاختيار من متعدد. وقد تكونت عينة البحث من (100) تلميذاً وتلميذة تم اختياريهم بالطريقة العشوائية البسيطة. تم التحقق من الخصائص السيكومترية للاختبار التحصيلي كأداة للبحث، وقد أظهرت النتائج أن قيم معاملات الصعوبة للمفردات تراوحت بين القيمة (−1.46) والقيمة (1.44) وتراوحت قيم قدرات الأفراد بين القيمة (1.32) والقيمة (4.70). كما بيّنت النتائج مطابقة الفقرات للنموذج أحادي المعلم. وفي ضوء النتائج تم اقتراح مجموعة من التوصيات.

الكلمات المفتاحية: نظرية الاستجابة للمفردة، النموذج أحادي المعلم، اختبار محكّي المرجع.

حقوق النشر : مجلة جامعة تشرين- سورية، يحتفظ المؤلفون بحقوق النشر بموجب الترخيص CC BY-NC-SA 04

* **أستاذ مساعد ، قسم القياس والتقويم، كلية التربية، جامعة تشرين، اللاذقية، سورية.**

** **طالبة دكتوراه ، قسم تربية الطفل، كلية التربية، جامعة تشرين، اللاذقية، سورية.**

مقدمة

يشتمل المنهاج التعليمي للصف الخامس من مرحلة التعليم الأساسي في سورية على مجموعة مهمّة من المواد التعليمية كالعلوم واللغات والرياضيات، وتشكّل الرياضيات مجموعة من الأنظمة وتطبيقات هذه الأنظمة في جميع نواحي الحياة والتخصّصات العلمية، وتعدّ الرياضيات من المواد الدراسية الصعبة مقارنة بغيرها، فهي مادة ذات طبيعة مجرّدة وبشكل خاص في عملياتها ومفاهيمها إضافة إلى التسلسل المنطقي لموضوعاتها التي تبدو وكأنها مقامة بشكل بناء محكم بحيث لا يمكن الوصول إلى مستوى محدّد قبل إتقان المستوى الأدنى منه فلا يمكن فهم عملية القسمة قبل إتقان عمليات الجمع والطرح والضرب وهذا بخلاف المواد الدراسية الأخرى، وبشكل عام لا تعطي الاختبارات التحصيلية (الجماعية/ المحكّية) المستخدمة على نطاق واسع في المدارس لتحقيق عملية قياس الأهداف التعليمية؛ لا تعطي معلومات دقيقة عن المستوى الحقيقي للمتعلم حيث لا يكون فيها لدرجة المتعلم معنى ما لم تقارن بمعيار يعتمد على مستوى الجماعة التي ينتمي إليها ويكون لهذه الدرجة تفسير فقط في ضوء خصائص هذه الجماعة أو يتم مقارنتها بمحك خارجي محددّ.

ولكن مقارنة نتائج أداء المتعلم بأداء جماعته أو مقارنة أدائه بمعايير خارجية محدّدة كمحك للأداء لا تكفي لتحقيق جميع الأهداف التي تتوخاها عملية القياس والتقويم التربوي حيث تقتصر أساليب القياس جماعية المرجع ومحكية المرجع على هدف واحد من أهداف القياس هو التمييز بين مستويات الأفراد في حالة القياس جماعي المرجع وقياس الاكتساب في حالة القياس محكي المرجع وتهمل أهدافاً أخرى كثيرة مثل مقارنة أداء الفرد بأدائه السابق أو اللاحق وتقدير معدل نمو تحصيل الفرد خلال فترة زمنية محددة. (محمود،2012،ص6)

إضافة إلا أن هذه الأساليب تجعل نتائج القياس بعيدة عن الموضوعية وذلك لأن أهداف القياس السلوكي لا تقتصر على مجرد التمييز بين أداء الأفراد (القياس جماعي المرجع) أو على مجرد قياس البعد عن مستوى معين من الأداء (القياس محكي المرجع)، وحتى يمكن التوصل إلى صورة القياس التي تمكّن من تحقيق جميع أهداف القياس المحتملة ينبغي أن يكون القياس موضوعياً ولكي يكون كذلك ينبغي أن لا يتأثر باختلاف الأداة المستخدمة كما ينبغي أن لا يتأثر أيضاً بالعناصر التي استخدمت هذه الأداة في تقديرها، وأن تتدرج هذه الأداة بوحدة قياس مطلقة ثابتة تتوافق مع تدرج مستويات المتغير موضوع القياس. (السيد ابراهيم، 2010، ص9)

وذلك طالما أن الهدف الأساسي للقياس التربوي والنفسي هو تحديد كم هو مقدار السمات الكامنة التي يمتلكها الفرد والسمات تعني القدرات مثل القراءة والقدرات الرياضية والحسابية وهذا التعبير العام يستخدم مع نظرية الاستجابة للمفردة للإشارة إلى هذه القدرات. (Baker,2001,p5)

ويشكل 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 حيث تكون نتائج القياس مستقلة عن الأداة المستخدمة في التوصل إليها؛ فطول غرفة الصف هو نفسه إذا استخدمنا في قياسه مسطرة خشبية أو متراً معدنياً أو متراً قماشياً أو غيرها من أدوات القياس، وفي المجال التربوي والنفسي إذا كان تحصيل المتعلم في مادة دراسية معينة أو قدرته على الفهم القرائي أو قدرته على التعبير لغوياً أو غيرها هي الموضوع المقاس فإن المطلوب أن تكون نتيجة قياسه ثابتة مهما استخدمنا من اختبارات مختلفة مناسبة في هذا المجال.

وإذا كان الشخص يريد قياس كم مقدار السمات الكامنة التي يمتلكها الفرد فمن الضروري أن يكون لديه ميزان للقياس؛ فكما للمسطرة نظام متري محدّد يجب تعريف ميزان للقياس له نقطة وسط هي الصفر ووحدة قياس واحدة ويمتد من اللانهاية الموجبة إلى اللانهاية السالبة. (Baker,2001,p6)

وكما هو معروف فإن هناك مدخلين أساسيين لبناء المقاييس النفسية والاختبارات التحصيلية هما النظرية الكلاسيكية ونظرية الاستجابة للمفردة كما توجد أيضاً نظريات أخرى مازال علماء القياس مهتمين بتطويرها لتناسب عملية القياس الموضوعي ولكن الأكثر انتشاراً هما هذين المدخلين.

لقد ظلت النظرية الكلاسيكية تشكل الأساس النظري والعملي للقياس في العلوم السلوكية لسنوات عديدة مضت مستندة على نموذج بسيط للقياس ينص على أن درجة الفرد الملاحظة في الاختبار تساوي مجموع الدرجة الحقيقية للفرد في السمة المقاسة والدرجة الخطأ في أدائه على الاختبار ولكن القصور الواضح في هذه النظرية عن تحقيق موضوعية القياس جعلها عرضة لمجموعة من الانتقادات.

إن المشكلة الرئيسية هي أن جميع الخصائص السيكومترية للاختبارات الجماعية والمحكية في ظل النظرية الكلاسيكية تعتمد على خصائص عينة الأفراد التي يجرى عليها الاختبار وعلى خصائص عينة المفردات التي يشتمل عليها الاختبار مما يؤدي إلى نظام قياس نفسي وتربوي غير مستقر، فمعاملات صعوبة المفردات تتذبذب وتتغير بتغير سمات أو قدرات أفراد العينة المراد اختبارها، وقياس سمات أو قدرات الأفراد تتذبذب وتتغير بتغير صعوبة مفردات الاختبار، وثبات الدرجات المستمدة من الاختبار تتذبذب وتتغير بتغير مستوى وانتشار سمات أو قدرات عينة الأفراد، والدرجات المستمدة من هذا الاختبار لا يكون لها معنى أو دلالة في ذاتها إذ يختلف معناها باختلاف سهولة أو صعوبة مفردات الاختبار وباختلاف ضيق أو اتساع مدى الاختبار. (علام،2006،ص680)

ونتيجة لنواحي القصور المتعددة التي تنتاب النظرية الكلاسيكية وأساليبها في القياس والتي لم تعد تفي بمتطلبات القياس الموضوعي فقد تركزت الجهود حول تطوير أدوات موضوعية دقيقة لقياس الخصائص المميزة للاختبارات التحصيلية التي تقيس الأبعاد المتعلقة بعمليتي التعليم والتعلم؛ ومن هذه الأدوات نظرية الاستجابة للمفردة.

إنّ أهمية هذه النظرية تكمن في تحقيقها لموضوعية القياس فدرجة الفرد في اختبار ما هي دالة رياضية مرتبطة بقدراته ومستقلة عن مجموعة المفردات الاختبارية التي يختبر بها من جهة وعن درجات أفراد المجموعة الاختبارية التي ينتمي إليها من جهة ثانية مما يجعل تفسير درجة الفرد ثابتاً لأنه مرتبط بقدرة الفرد وبمستوى أداء مطلق ومتحرر من المفردات الاختبارية ومن المجموعة الاختبارية التي ينتمي إليها.

وتتضمن هذه النظرية عدداً من النماذج تهدف إلى تحديد العلاقة بين أداء الفرد في اختبار ما وبين السمة أو القدرة التي تكمن وراء هذا الأداء في صورة دالة رياضية احتمالية تقوم على افتراض أنه كلما ازدادت قدرة الفرد عن صعوبة المفردة ازداد احتمال حدوث الاستجابة الصحيحة للمفردة والعكس صحيح. (مسعود، 2010، ص93).

إن نماذج نظرية الاستجابة للمفردة في معظم أشكالها هي بلا شك النماذج الأكثر استعمالاً في برامج التقييم وفي أحدث برامج البحث في مجال القياس التربوي والنفسي وربما الميزة الأكثر أهمية والتي تؤدي إلى هيمنة نظرية الاستجابة للمفردة هي خاصية تخمين مواقع المفردات الفردية (الصعوبات) ومواقع الممتحنين (القدرات) بشكل منفصل ولكن على نفس المقياس، هذه السمة ليست متاحة في نظرية القياس الكلاسيكية. (Carlson & Davier,2013,p5)

ويعد نموذج راشRasch أحد أهم تلك النماذج بسبب سهولته وموضوعيته ونظراً لاستخدامه على نطاق واسع في كثير من البلدان في بناء وتطوير الاختبارات. النموذج الأحادي البارامتر يتنبأ باحتمال الاستجابة الصحيحة من التفاعل بين قدرة الفرد ومعلم صعوبة المفردة. (Partchev,2004,p10)

يتميز هذا النموذج بأن معلم القدرة يمكن أن يقدّر دون تحيّز لأي مجموعة من الفقرات المختارة من النطاق الشامل للفقرات المطابقة للأنموذج وتعتمد دقة القياس على الاختيار الملائم لصعوبة الفقرة ويتميز كذلك بإمكانية فحص صدق النموذج بشكل مستقل عن قدرة الأفراد وصعوبة الفقرات لهذه الأسباب حاز هذا النموذج على اهتمام العديد من الباحثين وبشكل خاص في بناء الاختبارات التحصيلية. (السامرائي والخفاجي،2012،ص969)

ويهتم البحث الحالي باستخدام هذا النموذج في تدريج اختبار تحصيلي محكّي المرجع للوحدة الأولى في مادة الرياضيات للصف الخامس الأساسي.

مشكلة البحث:

نظراً لأهمية الاختبارات التحصيلية كأداة من أدوات التقويم التي اتجهت أهدافه في الآونة الأخيرة إلى الكشف عن القدرات الكامنة لدى المتعلم والتي تدفعه للاستجابة بطريقة محدّدة على مفردات اختبار تحصيلي محدّد، ينبغي إعادة النظر في طرق وأساليب التقويم حيث أنه ما زالت الاختبارات الصفية المستخدمة في قياس مستويات المتعلمين اختبارات يعدّها المعلم بدون الاستناد إلى تحليل محتوى أو جدول مواصفات فهي إذاً لا تمثل المحتوى تمثيلاً حقيقياً من ناحية ومن ناحية أخرى فهي بلا شك تتبع النظرية الكلاسيكية أي أنها اختبارات تشتمل على مفردات غير متدرجة الصعوبة على مقياس خطي ولا تهتم بحساب قدرات الأفراد المقابلة لمفردات الاختبار.

كما تبنى الاختبارات التحصيلية في غالبية الأحيان دون الاستناد إلى نظرية أو نموذج متطور في القياس التربوي وتفسر درجاتها دون تحديد واضح للإطار المرجعي الذي يعطي معنى ودلالة لهذه الدرجات وبذلك لا تشكل هذه الأسئلة والمفردات بمجملها اختباراً تحصيلياً بمفهومه العلمي نظراً لأنها لا تمثل النطاق المعرفي والسلوكي الشامل للمعارف والمهارات التي ينطوي عليها المحتوى الدراسي وبالتالي يصعب الإفادة من نتائج مثل هذه الاختبارات في اتخاذ قرارات صائبة تتعلق بالمتعلم أو المعلم أو استراتيجيات التعليم. (علام،2006،ص725)

ونظراً لكون العملية التعليمية جزء من المجتمع التربوي الذي تأثّر بالتطور العلمي والتكنولوجي في جميع المجالات وبدء يستخدم التكنولوجيا المتطورة لتحقيق أهدافه التي تغيرت بدورها لتناسب التغيرات في جوانب المجتمع الأخرى ففي إطار الأهداف المستقبلية للتربية وفي ضوء المنظور الرحب للتحصيل الدراسي لا تكون عملية التقويم غاية في ذاتها لتحديد نجاح المتعلمين وانتقالهم إلى الصفوف الأعلى أو المراحل التعليمية اللاحقة أو رسوبهم وإنما تصبح جزءاً متكاملاً من عملية التعليم، توجهّها وتعزّزها وتصحّح مسارها بما يحقق لها أهدافها، وهذا يتطلب التحوّل من أساليب ونظم الامتحانات التقليدية السائدة والدرجات والتقديرات الاعتبارية التي تقتصر على الموازنة بين أداء المتعلم وأداء أقرانه في اختبارات تهتم بإبراز الفروق الفردية وتؤكدها إلى أساليب ونظم تؤكد الشخصية المتكاملة والمتوازنة للمتعلم وما يمتلكه من الكفايات والمهارات الوظيفية. (المرجع السابق،ص726)

والواقع الذي يجب ألا نغفله هو أن التعليم لم يعد مقتصراً على مجرد التمييز بين الطلبة في السمة المقيسة وإنما أخذ يركز على اكتسابهم مهارات معينة وتحقيق أهداف محددة بل وإتقان تلك المعلومات والمهارات وبالتالي مقارنة أداء الفرد بمستوى أداء معين في مجال من السلوك وهذا ما تهدف إليه الاختبارات محكية المرجع. (حجازي والخطيب،2014،ص2242)

وبالنظر إلى المدخلين الرئيسين لبناء الاختبارات التحصيلية وتحليل خصائصها فإن النظرية الكلاسيكية تختلف عن النظرية الحديثة في كونها تقتصر على تحقيق هدفٍ واحد فقط من أهداف القياس وتغفل أهداف القياس الأخرى، ومن ناحية أخرى فإن القياس المعتمد على النظرية الكلاسيكية يكون معتمداً على خصائص الاختبار المستخدم وعلى مجموعة الأفراد المختبرين مما يجعل النتائج تتغير بتغير المفردات وعينة الأفراد وبالتالي عدم تحقق موضوعية القياس.

وتواجه النظرية الكلاسيكية مشكلات في تصميم الاختبار وتحديد المفردات المتحيزة وتعادل درجات الاختبارات لأنها لا تأخذ في الاعتبار تغاير استجابات الممتحنين للمفردات مختلفة الصعوبة وتفترض تساوي تباين الخطأ في القياس بالنسبة لجميع الممتحنين وفيها تعتمد الخصائص السيكومترية للاختبارات على خصائص عينة الممتحنين التي يطبق عليها الاختبار وتكون المقارنة بين الممتحنين ممكنة فقط عندما يؤدون الاختبار نفسه وتفترض أن العلاقة بين درجة الممتحن ومستوى قدرته علاقة خطية وتتعامل مع الدرجات كما لو كان الفرق الثابت بين درجتين من درجات الاختبار لا يختلف معناه بحسب موقع الدرجات على متصل القدرة. (محمود،2012،ص25)

في حين أن النظرية الحديثة تعتبر استجابة كل فرد على كل مفردة موقفاً منفرداً بذاته وتتيح وجود مقياس للصعوبة مستقل عن العينة مما يجعل من الممكن مقارنة الأفراد عند أدائهم اختبارات مختلفة وأن العلاقة بين درجة الفرد ومستوى صعوبته هي علاقة احتمالية وليست خطية كما أنها تحقق استقلالية القياس حيث لا تختلف تقديرات الأفراد باختلاف الصور الاختبارية المستخدمة كما توفّر هذه النظرية خطية القياس؛ حيث تتدرج كل من صعوبة المفردات وقدرة الأفراد بوحدة قياس معرّفة رياضياً، ومن ثم يصبح الاختبار صالحاً للمقارنة وحساب التغير على مرّ الزمن وعلى عينات غير عينة الدراسة. (مسعود،2010، ص12)

كما تركّز النظرية الكلاسيكية على الاختبار ككل وتنظر إليه على أنه وحدة واحدة، وبالتالي لا يسمح هذا النظام بحذف أو إضافة مفردات لأن هذا يؤثر على صدقه وثباته بينما في النظرية الحديثة يكون التركيز على المفردة باعتبارها الوحدة الأساسية ويسمح هذا النظام بإضافة مفردات أخرى للاختبار أو حذف بعضها دون أن يؤثر ذلك على صدقه وثباته.(الشواورة،2013، ص2)

وفي هذا السياق وجدت الباحثة أن معظم الدراسات تتجه نحو استخدام النظرية الحديثة في بناء الاختبارات التحصيلية وتحليل الخصائص السيكومترية لمفرداتها فلقد أجرت كاظم(1988) دراسة هدفت إلى استخدام نموذج (راش) في بناء اختبار موضوعي جديد في مقرر المدخل في علم النفس وتفسير أداء الأفراد على هذا الاختبار ثم التحقق من مدى موضوعية هذا التفسير كما يتمثل في فرضيات نموذج (راش) ويعني هذا تقييم لمدى صدق النموذج في تحقيقه لموضوعية القياس.

كما أجرت نداء علي (2011) دراسة هدفت إلى دراسة فعالية اختبار محكي المرجع وفق النموذج الأحادي المعلم لراش في مقرر القياس والتقويم الذي يدرسه طلاب السنة الرابعة معلم صف في كلية التربية جامعة دمشق. كما أجرت هند الحموري (2011) دراسة هدفت إلى استقصاء إمكان استخدام نماذج نظرية الاستجابة للمفردة في بناء الاختبارات التحصيلية ذات العدد القليل من الفقرات التي يعدها المعلمون بالإضافة إلى توضيح طريقة بناء الاختبار وتطويره وفق هذه النظرية.

كما أجرى محمد أنور السامرائي وأحمد شاكر الخفاجي (2012) دراسة هدفت إلى بناء اختبار تحصيلي محكي المرجع في مادة علم نفس الخواص مبني وفق نظرية السمات الكامنة ونماذج القياس المنبثقة عنها (أنموذج راش) تحديداً.

كما أجرى محمد دبوس(2016) دراسة هدفت إلى استخدام نظرية الاستجابة للفقرة في بناء تجمع فقرات اختبار محكي المرجع في الرياضيات وفق النموذج ثنائي المعلم وروعي في بناء هذه الفقرات أن تقيس تحصيل الطالب في الرياضيات في المعرفة المفاهيمية والمعرفة الإجرائية وحل المشكلات.

وبناء على ما سبق يسعى البحث الحالي إلى استخدام أحد نماذج نظرية الاستجابة للمفردة (نموذج راش أحادي المعلم) في بناء وتدريج اختبار تحصيلي محكي المرجع في الوحدة الأولى من كتاب الرياضيات للصف الخامس الأساسي ومن ثم اختبار مدى تحقّق فرضيات نظرية الاستجابة للمفردة في بيانات البحث وحساب تقديرات معالم الفقرات وقدرات عينة الأفراد وفق النموذج أحادي المعلم.

أهمية البحث وأهدافه:

تتجلى أهمية هذا البحث بالآتي:

- تأمين اختبار تحصيلي في مادة الرياضيات للصف الخامس الأساسي يحقق مستوى مقبول من الموضوعية.
- التشجيع على ضرورة بناء اختبارات محكية المرجع في المواد الدراسية المختلفة وفق نظرية الاستجابة للمفردة.

هدف البحث:

هدف البحث إلى استخدام نظرية الاستجابة للمفردة الاختبارية في بناء تجمع فقرات اختبار تحصيلي محكي المرجع في الرياضيات للصف الخامس الأساسي وفق النموذج أحادي المعلم.

أسئلة البحث:

- ما مدى تحقق افتراضات نظرية الاستجابة للمفردة في بيانات البحث وفق النموذج أحادي المعلم؟
- ما قيم تقديرات معالم فقرات الاختبار (الصعوبة) وفق النموذج أحادي المعلم لنظرية الاستجابة للمفرد؟
- ما درجة مطابقة الاستجابات عن فقرات الاختبار لتوقعات النموذج أحادي المعلم لنظرية الاستجابة للمفردة؟

مصطلحات البحث:

- نظرية الاستجابة للمفردة **Item Response Theory** :

نظرية حديثة في القياس النفسي والتربوي يتم فيها تحديد العلاقة بين أداء المفحوص والسمة الكامنة موضع القياس وفق دالة رياضية محددة. (دبوس،2016،ص1461)

وتهدف هذه النظرية 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 أبسط نماذج نظرية الاستجابة للمفردة وأكثرها استخداماً في بناء الاختبارات وبنوك الأسئلة وتحليل مفرداتها ويفترض هذا النموذج انعدام التخمين وتساوي معاملات التمييز ويقوم بتقدير صعوبة المفردات. ويستند 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،2017، ص22)

- اختبار محكّي المرجع **Criterion- Referenced Test** :

هو الاختبار الذي يتم فيه مقارنة أداء الفرد بمستوى أداء محدد مسبقاً دون مقارنته بمتوسط أداء جماعة مرجعية. (حجازي والخطيب،2014،ص2247)

حيث يتم تحديد مستوى الطالب بالنسبة إلى محك ثابت دون الرجوع إلى أداء الآخرين وهذا المستوى يرتبط عادة بالأهداف السلوكية للمادة الدراسية.(السامرائي والخفاجي،2012،ص970)

ويستخدم مثل هذا النوع من الاختبارات لتحديد مكان المتعلم على متصل الكفاية لمجال سلوكي معرف تعريفاً دقيقاً وبالتالي يميز بين المختبرين بحسب درجة إتقانهم لهذه المهارة والتي تم صياغتها في ضوء محك محدد مسبقاً ويصف بدرجة عالية من الدقة مدى ما أتقنه المتعلم. (أحمد، 2010، ص546)

ويعرّف إجرائياً بأنه مجموعة المفردات الاختبارية من نوع الاختيار من متعدد من إعداد الباحثة والذي سيستخدم في تحديد قدرات التلاميذ في الرياضيات.

حدود البحث:

الحدود الزمانية: تم إجراء هذا البحث في الفصل الدراسي الأول من العام الدراسي 2021/2020

الحدود المكانية: مدارس مدينة بانياس للتعليم الأساسي.

الحدود الموضوعية: الوحدة الأولى من كتاب التلميذ للصف الخامس الابتدائي في مادة الرياضيات.

الجانب النظري:

يرى كثير من علماء النفس المعرفي وخبراء القياس أن عمليات القياس وتصميم الاختبارات وعلاقتها بالعملية التربوية سوف تتخذ في القرن الحادي والعشرين توجهاً جديداً يختلف اختلافاً جوهرياً عما هو سائد. فالتطورات التي حدثت في علم النفس المعرفي وعلم القياس النفسي والتربوي وتقنيات المعلومات والاتصالات وما تفرضه التحديات المعاصرة، يتطلب إعادة النظر في ممارسات التعليم ونظم وأساليب القياس والتقويم لمجتمع المتعلمين، وتحقيق ذلك يتطلب توافر معلومات تختلف عما تزودنا به أساليب التقويم التي تقتصر على تقديم مؤشرات عامة تتعلق بمجموعات من الأفراد بهدف المقارنة بينهم أو التنبؤ بأدائهم الأكاديمي. فعملية التقويم سوف تصبح جزءاً لا يتجزأ من عملية التعليم، أي أنهما سيصبحان عملية واحدة متكاملة؛ فالاختبارات الحالية التي تستخدم فقط في أغراض انتقاء الطلاب للالتحاق بمراحل التعليم المتتالية ولا تقدم معلومات يمكن الاستفادة منها في التقدّم التربوي سوف تكون قليلة الجدوى، وذلك لأننا لا نستطيع باستخدامها قياس نمو وتطور الكفايات الإنسانية قياساً يتطلب التحقق من الفروق الجوهرية بين الأداء المتميز والأداء الضعيف عن طريق تقييم بنية المعرفة والعمليات المعرفية التي تكشف عن درجات كفايات الفرد في مجال دراسي معين.(علام،2006،ص716)

بعض نواحي القصور في الاختبارات مرجعية الجماعة :**Norm Referenced Tests**

يرى علام (2006) أن هناك الكثير من نواحي الضعف والعيوب التي تنتج عن استخدام الاختبارات التي تعتمد على جماعة الممتحنين كأساس لاستنتاج الخصائص السيكومترية لاختبار تحصيلي معين واتخاذ قرارات بناءً على درجات هذا الاختبار ومن نواحي الضعف هذه نذكر الآتي:

- إن أساليب القياس المستندة إلى الفلسفة التربوية التي تحثّ على التنافس الفردي من أجل التفوق على الأقران تجعل عملية تقويم التحصيل محدودة الفائدة وذات تأثير سلبي على المتعلمين والمناخ المدرسي وبذلك لا تسهم في

تحقيق عملية التقويم لوظائفها المرجوة فالاختبارات التحصيلية التي تبنى وفق هذا التوجه التنافسي تستهدف التمييز والتفاضل بين الطلاب من أجل وضع تقديرات لهم في المادة الدراسية أو اتخاذ قرارات الانتقاء.

- يفترض هذا التوجه أن تعلّم الطالب يعتمد اعتماداً أساسياً على قدرته والجهد الذي يبذله وعلى العكس من ذلك يفترض أن إخفاق الطالب يرجع إلى انخفاض قدرته أو سلبية اتجاهه نحو التعلم أو المادة الدراسية.
- إن التأكيد الأساسي فيها يكون للدرجات الكلية وبذلك يقلّ الاهتمام فيها على دقائق المحتوى بينما يزيد التركيز على التجريدات الأكثر عمومية. فمثلاً يحدّد اختبار تحصيلي الدرجة الكلية في العمليات الحسابية مثلاً ولكنه لا يوضح شيئاً عن نوع العملية الحسابية التي يستطيع الطالب أو لا يستطيع إجراءها فالدرجة الكلية يصعب استخدامها في تحديد مواطن الضعف أو الأخطاء الشائعة أو فجوات التحصيل.
- إن انتقاء مفردات الاختبارات مرجعية الجماعة يعتمد بدرجة كبيرة على الأساليب الإحصائية في تحليل مفرداتها حيث تستبقى فقط المفردات التي تميّز بدرجة أفضل بين الطلاب وتسمح بترتيبهم بالنسبة لبعضهم البعض وانتقاء المفردات على هذا الأساس يقلّل من صدق محتوى الاختبار أي مدى تمثيل مفرداته لنطاق المحتوى الذي يقيسه ويحيد بالاختبار عن الغرض منه كأداة قياس لذلك فإن عملية إضافة أو حذف المفردات بهذا الشكل من أجل التحكم في مستوى صعوبتها أو تمييزها يؤثر تأثيراً بالغاً في صدق الاختبار. (علام،2006،ص729)

مميزات الاختبارات محكية المرجع:

تتميز هذه الاختبارات بعدد من المميزات جعلتها أكثر استخداماً من الاختبارات مرجعية الجماعة في قياس تحصيل المتعلمين ومن هذه المميزات نذكر الآتي:

- إن الاختبارات محكية المرجع لا تقتصر على تقويم أداء الطالب بشكل فردي وخلال تتابع المنهج الدراسي فقط بل تستعمل لأغراض التقويم التربوي في مستويات عليا مثلاً عند تقويم أداء المدرسة ككل.
- تُعتمد الاختبارات المحكية المرجع أساساً في تحديد مجموعة من المخرجات التعليمية المرغوب فيها ومستويات إنجازها المحدّدة، لذا فإن هذا النوع من الاختبارات يكون مفيداً عندما تختلف المؤسسات التربوية في تقديم محتويات مختلفة لطلبتها لأنها ستفرض على هذه المؤسسات وجوب أن يبلغ الطلبة مستويات إتقان محدّدة في اكتساب معلومات ومهارات محدّدة.
- تفيد الاختبارات محكية المرجع في تقدير الوضع التحصيلي الراهن للطلبة كما أنها تساعد في تحديد أسباب عدم إتقان الطالب لبعض الأهداف التي تبنى في ضوئها هذه الاختبارات ووصف الطرق المناسبة لعلاج حالات الإخفاق في تحقيق الأهداف.
- إن مصطلح المحك لا يشير بالضرورة إلى الإتقان التام على المتصل بل يمكن تحديد المستويات المحكية في أي نقطة وستكون هذه النقطة عبارة عن سلوك محدّد يستعمل لوصف مهمات محدّدة يجب أن يكون الطالب قادراً على أدائها ويتم المقارنة في ضوء هذه المستويات كما يمكن تغيير هذه المستويات من فترة إلى أخرى تبعاً لنمو الطالب.
- تعبّر درجة الطالب في الاختبار محكي المرجع عن مدى تحقيقه لأهداف تعليمية مرجوة فهو يمدنا بمعلومات محددة وتفصيلية عن تحصيل كل طالب في موضوع دراسي معين أو مهارة معينة لذلك فإن هذا النوع من الاختبارات له قيمة تشخيصية في تحصيل كل طالب لمجموعة محددة من المعارف والمهارات. (السامرائي والخفاجي،2012،ص973)

– الاختبارات مرجعية المحك تهتم بقياس مدى تحقق مجموعة متكاملة محددة من المهام وتزود المتعلم بأدلة عن درجة تمكنه منها كما توضح تقدمه أثناء دراسة مطلب تعليمي محدّد وما يعترضه من صعوبات تعوق هذا التقدم. (علام،2006،ص730)

– إن الاختبارات محكية المرجع لا تهتم بمقارنة أداء الفرد بأداء أقرانه وإنما تهتم باتخاذ قرارات تتعلق بمدى إتقان المتعلم للمحتوى الدراسي المحدد بدقة والذي يقيسه الاختبار من خلال تصنيف المتعلمين إلى مجموعتين إحداهما متقنة لهذا المحتوى والأخرى غير متقنة وبتشخيص جوانب الضعف لدى المتعلمين غير المتقنين لتوجيههم إلى البرامج العلاجية المناسبة. (حجازي والخطيب،2014،ص2243)

أوجه القصور في المقاييس المعتمدة على فلسفة القياس جماعي/ محكي المرجع فقط:

بالرغم من المميزات التي ذكرت عن الاختبارات محكية المرجع إلا أنه يمكن تحديد عدداً من نواحي القصور في هذا النوع من الاختبارات وفي الاختبارات مرجعية الجماعة أيضاً وهذه السلبيات هي:

– اقتصار كل نوع من هذه المقاييس على تحقيق هدف واحد من أهداف القياس وإهمال أهداف أخرى للقياس مثل مقارنة مستوى تحصيل الأفراد بالرغم من اختلاف الاختبار المستخدم أو مستوى عينات الأفراد.

– تقل فاعلية الاختبارات جماعية المرجع في تقويم المتعلمين تقويماً تشخيصياً يحدّد جوانب القوة والضعف في تعلمهم.

– إن حذف المفردات غير المميزة من الاختبار جماعي المرجع بغض النظر عن أهميتها أو تمثيلها للأهداف يؤثر على البناء الصحيح للاختبار.

– يقوم حساب صدق الاختبار محكي المرجع على مقارنة الاكتساب الواقعي بالاكتساب المتوقع ولا توجد طريقة كافية لتحديد هذا الاكتساب المتوقع.

– تختلف قيمة معامل الثبات في القياس جماعي/ محكي المرجع من مجموعة لأخرى من الأفراد مما يحدّ من إمكانية تعميمه وذلك لأن بعض درجات الاختبار قد تتضمن أخطاء للقياس أكثر من الأخرى.

– تظهر في القياس محكّي المرجع مشكلة المحكّ ومن يحدّده وعلى أي أساس.

– عدم وجود وحدة قياس متساوية لتدريج كل من قدرات الأفراد وصعوبة المفردات واعتبار الدرجة الخام الكلية أساساً للقياس مما يؤدي إلى بعض المشكلات عدم تحقيق خطية القياس. (أحمد،2010،ص578)

نظرية الاستجابة للمفردة **Item Response Theory**:

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)

لقد بنيت هذه النظرية على افتراض أنّ هناك علاقة بين ما يمتلكه المفحوص من السمة ومعامل صعوبة فقرة ما واحتمال الحصول على الإجابة الصحيحة عند مستوى قدرة معين وأن هذه العلاقة تتخذ شكل المنحنى الرياضي حيث يتم تدريج قدرات المفحوصين وصعوبة الفقرات على متصل واحد والذي يحدد مقدار احتمال الاستجابة الصحيحة أو الخطأ على الفقرة هو الفرق بين قدرة الفرد وصعوبة الفقرة. (العنزي،2010،ص14)

ويرجع الفضل في وضع الأسس النظرية لنظرية الاستجابة للمفردة إلى عالم القياس النفسي التربوي لورد Lord الذي ظلت أعماله لا يلتفت إليها ابتداء من عام 1952 وحتى عام 1968 الذي نشر فيه كتابه "النظرية الإحصائية لدرجات

الاختبارات العقلية" الذي جمع فيه أسس النظرية الكلاسيكية والحديثة في القياس. ويعتبر عالم الرياضيات الدنماركي جورج راش G.Rasch أستاذ الرياضيات والإحصاء بجامعة كوبنهاغن بالدنمارك(1945) أول من وضع أولى نماذج هذه النظرية وهو النموذج أحادي البارامتر وساعده في تطويره وتطبيقه عالم القياس النفسي بنجامين رايت B.Wright بجامعة شيكاغو. (مسعود،2010،ص92)

أهم افتراضات نظرية الاستجابة للمفردة:

تقوم هذه النظرية على مجموعة من الافتراضات التي يجب توافرها في المفردات التي يتم تحليل خصائصها السيكومترية وهذه الافتراضات هي:

1- أحادية البعد Unidimensionality:

يشير اصطلاح الأبعاد إلى عدد السمات الكامنة المسؤولة عن أداء الأفراد في الاختبار... وتفترض معظم نماذج نظرية الاستجابة للمفردة أن هناك سمة أو قدرة واحدة فقط كافية لتفسير وتوضيح الفروق بين أداءات الأفراد على الاختبار وتسمى تلك النماذج بالنماذج أحادية البعد، أما النماذج التي تفترض وجود أكثر من قدرة واحدة تكمن خلف أداء الفرد تسمى نماذج متعددة الأبعاد. إن أحادية البعد تعني أن أي مفردة من المفردات المتدرجة الصعوبة تتطلب في حلها نفس النوع من الإجراءات والعمليات السلوكية ولكنها تختلف فيما بينها من حيث تدرج صعوبتها فقط.

2- التحرر من عامل السرعة في الإجابة Speediness:

إن الاختبارات التي تكون ملائمة لنماذج نظرية الاستجابة للمفردة لا يتم إعدادها تحت شرط السرعة أي أن إخفاق بعض الأفراد في الإجابة على بعض مفردات الاختبار يرجع أساساً إلى انخفاض قدراتهم وليس إلى تأثير عامل السرعة في إجاباتهم وهذا الافتراض لا يعلن عنه كأحد افتراضات النموذج في معظم الأحوال حيث أنه متضمن في الافتراض الخاص بأحادية البعد وذلك لأنه إذا اعتبرت السرعة عاملاً مؤثراً في أداء الأفراد على الاختبار فإنه يكون على الأقل هناك عاملان يؤثران في الأداء(سرعة الأداء والسمة المقاسة من خلال محتوى الاختبار) وهذا لا يتفق مع أحادية البعد التي تمثل افتراضاً أساسياً لنماذج نظرية الاستجابة للمفردة. (مسعود،2010، 104)

3- استقلال موضع المفردة Local Item Independence:

يعني أن لا تؤثر استجابة الفرد لإحدى مفردات الاختبار على استجابته للمفردات الأخرى. أي أن تقدير صعوبة أي مفردة من مفردات الاختبار لا تعتمد على أي من:

- صعوبة المفردات الأخرى
- قدرة الأفراد الذين يجيبون عن المفردات
- قدرة أي مجموعة أخرى من الأفراد الذين يطبق عليهم الاختبار(محمود،2012،ص23)

نماذج نظرية الاستجابة للمفردة:

تحتوي نظرية الاستجابة للمفردة على عدة نماذج رياضية لوغاريتمية تفترض وجود علاقة بين قدرة الفرد وبين استجابته الملاحظة على مفردة اختبارية محددة يتم وصف هذه العلاقة بدالة رياضية احتمالية تفترض أنه كلما زادت قدرة الفرد عن صعوبة المفردة زاد احتمال الاستجابة الصحيحة للمتعلم على المفردة الاختبارية والعكس صحيح. وهذه النماذج هي:

1- النموذج اللوغاريتمي ثلاثي المعلم Three Parameter Logistic Model:

وينسب للعالم بيرنبوم Birnbaum وبناء على هذا النموذج يوجد لكل مفردة ثلاثة معالم وهي الصعوبة *b* والتمييز *a* والتخمين *c*.

2- النموذج اللوغاريتمي ثنائي المعلم Tow Parameter Logistic Model:

وينسب لـ لورد ويعد هذا النموذج حالة خاصة من النموذج الثلاثي لأنه يفترض انعدام معامل التخمين أي أن *c* = صفراً ويهتم هذا النموذج بتقدير صعوبة المفردات *b* ومعامل التمييز *a*.

3- النموذج اللوغاريتمي أحادي المعلم one Parameter Logistic Model:

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*. اهتم راش بالربط بين نتائج الاختبارات السيكولوجية التي تطبق على مجموعة من الأفراد والتي تضمنت مجموعات مختلفة من الفقرات حيث كان راش يهدف إلى تحقيق مفهوم الموضوعية بمعنى أن درجة الفرد في الاختبار لا يجب أن تكون دالة لعينة الأفراد التي استخدمت في التدريج الأصلي للفقرات التي يشتمل عليها الاختبار كما أنه يجب أن يحصل الفرد على نفس الدرجة في كل من اختبارين يقيسان نفس السمة أو القدرة. (السامرائي والخفاجي،2012،ص977)

مميزات نموذج راش الأحادي المعلم:

يتميز هذا النموذج بمجموعة من المميزات لا تتوافر في النماذج الأخرى :

- هو النموذج الوحيد من بين النماذج الاحتمالية الذي تتوافر فيه خاصية استقلالية بارامتراته أي عدم تباين تقديرات قدرة الأفراد بتباين عينة المفردات التي يختبرون بها وعدم تباين الخصائص السيكومترية للمفردات بتباين عينة الأفراد التي أجري عليها الاختبار ويتطلب ذلك التحقق بصورة دورية من عدم حدوث تغيرات ربما تطرأ على القيم التقديرية لبارامترات النموذج نتيجة لتغير خصائص العينة أو تغير طبيعة محتوى أو معنى المفردات الاختبارية بالنسبة إليهم بمرور الزمن.

- يعد أبسط النماذج الاستاتيكية الاحتمالية حيث يشتمل على بارامتر واحد يتعلق بمفردات الاختبار وهو صعوبة المفردة فعلى الرغم من تباين صعوبة مفردات الاختبار فإن النموذج يفترض أن جميع المفردات تتساوى في قدرتها على التمييز وبالتالي فإن المفردات لا يجب أن تميز بين مستويات القدرة المقاسة تمييزاً دالاً فحسب بل إن جميع مفردات الاختبار يجب أن تكون قدرتها على التمييز متساوية تقريباً والميزة التي تنجم عن ذلك هي أن المجموع غير الموزون لدرجات مفردات الاختبار يعد إحصاءً كافياً لتقدير قدرة الفرد المختبر.

- يؤدي إلى منحنيات مميزة للمفردات تكون متوازية أي لا تتقاطع وكذلك المنحنيات المميزة للأفراد وهذا ييسر عملية التقدير الإحصائي لبارامترات النموذج والوصول إلى ما يسمى بالتقارب، فالمنحنى المميز للمفردة هو دالة رياضية تربط بين احتمال الإجابة الصحيحة لفرد ما عن مفردة اختبارية وبين القدرة التي يقيسها الاختبار والمنحنى المميز للفرد وهو تلك الدالة التي تربط بين احتمال الإجابة الصحيحة لمجموعة من الأفراد المتساوين تقريباً في القدرة المقاسة وبين صعوبات المفردات وتوازي هذه المنحنيات لا يحدث إلا إذا حققت البيانات الافتراضات التي يستند إليها النموذج وبخاصة فرضية البعد الأحادي للاختبار وتساوي تمييز مفرداته وعندئذ يمكن تفسير مستويات وأنماط القدرة المقاسة.

- سهولة التقدير الإحصائي ففي النموذج ثلاثي البارامتر تتقاطع المنحنيات المميزة للمفردات مما يؤدي إلى صعوبات في عمليات التقدير الإحصائي لبارامتراته. (الشوربجي، د.ت، ص83)

الدراسات السابقة:

1- دراسة أمينة كاظم (1984):

عنوان الدراسة: استخدام نموذج راش في بناء اختبار تحصيلي في علم النفس وتحقيق التفسير الموضوعي للنتائج

هدف الدراسة: استخدام نموذج (راش) في بناء اختبار موضوعي جديد في مقرر المدخل في علم النفس وتفسير أداء الأفراد على هذا الاختبار.

عينة الدراسة: تم إجراء الاختبار على عينة التقنين التي بلغت 418 فرداً من طلبة وطالبات جامعة الكويت. نتائج الدراسة: التحقق إجرائيا من تحقيق الاختبار لفرضيات النموذج أحادي المعلم كما تمكنت الدراسة من بناء اختبار موضوعي (باستخدام نموذج راش) في مقرر المدخل في علم النفس تتدرج بنوده بوحدة قياس مطلقة ويصل إلى التفسير الموضوعي لدرجة الفرد على هذا الاختبار أو أي اختبار فرعي آخر مأخوذ من نفس هذا الاختبار ويصبح هذا بمثابة بنك لأسئلة هذا المقرر نستطيع أن نسحب منه أي عدد من البنود التي تحقق أهداف القياس كما يمكن أن نضيف إليه بنود جديدة تشترك معه في تدريج واحد مشترك وصفر واحد مشترك.

2- دراسة **Jennifer** (2007):

عنوان الدراسة : Test Item Construction and Validation: Developing A statewide Assessment for Agricultural Science Education

هدف الدراسة: تطوير نظام بناء المفردات مستند على الاختبار محكي المرجع.

عينة الدراسة: بلغت 226 طالباً وطالبة من طلاب التعليم الزراعي في أقسام النبات والحيوان في ولاية نيويورك.

نتائج الدراسة: بعد مراجعة الأدب المتعلق ببناء الاختبارات وأساليب التقويم تبين أن الاختبارات محكية المرجع هي طرق ستستخدم بدلاً من الاختبارات جماعية المرجع حيث يتم استخدام الاختبارات المحكية المرجع CRT Criterion referenced testing بالاعتماد على حقيقة أن كل سؤال مشتق من هذه الطريقة مرتبط بعنصر معين من المحتوى وهذه الاختبارات تنوي قياس إجادة المتعلم للأهداف التعليمية وليس مقارنة المتعلمين بعضهم مع بعض على طول المنحنى المعياري. تجميع المفردات لبناء الاختبار من أجل هذه الأهداف يجعل CRT الاختيار المثالي.

3- دراسة حمدي يونس أبو جراد (2008):

عنوان الدراسة: استخدام نموذج راش في تطوير اختبار كاتل الثالث للذكاء الصورة (أ)

هدف الدراسة: إعادة تدريج اختبار الذكاء الثالث لكاتل الصورة (أ) باستخدام نموذج راش الأحادي المعلم وعمل معايير مختلفة تفسر من خلالها مستويات القدرة للأفراد.

عينة الدراسة: مجموعة من طلبة كلية التربية بجامعة القدس المفتوحة بلغ عددها 240 طالباً وطالبة.

نتائج الدراسة: توصلت النتائج إلى حذف فقرتين من فقرات اختبار التصنيف لعدم ملاءمتها لنموذج راش وعليه فقد بلغ عدد فقرات الاختبار بعد تدريجه باستخدام نموذج راش 48 فقرة كما تم التوصل إلى معايير للاختبار عن طريق إيجاد الرتب المئينية والدرجات التائية ونسب الذكاء الإنحرافية المناظرة للتقديرات المختلفة لقدرة الأفراد لكل من الذكور والإناث.

4- دراسة نداء علي (2011):

عنوان الدراسة: فاعلية استخدام نموذج راش في بناء اختبار تحصيلي محكي المرجع لمقرر القياس والتقويم في التربية.

هدف الدراسة: دراسة فعالية اختبار محكي المرجع وفق النموذج الأحادي المعلم لراش في مقرر القياس والتقويم الذي يدرسه طلاب السنة الرابعة معلم صف في كلية التربية جامعة دمشق.

عينة الدراسة: شملت (426) طالباً من طلاب كلية التربية قسم معلم الصف في جامعة دمشق يدرسون مقرر القياس والتقويم للعام الدراسي 2009/ 2010.

نتائج الدراسة: معامل الثقة للمفردات بلغ 0.98 وللأفراد 0.94 وكلاهما عال. كما أظهرت النتائج أن المقياس المشترك للاختبار محكي المرجع في مقرر القياس والتقويم مع نموذج راش يحقق فرضيات نظرية الاستجابة للمفردة وهي أحادية البعد واستقلالية القياس والتحرر من عامل السرعة. كما أظهرت النتائج أن قيمة معلومات الأداء للاختبار توفر القدر الأكبر من المعلومات في المستويات المتوسطة للقدرة والمفردات التي تعطي المعلومات الأكثر في الموضوع مقرر القياس والتقويم هي (45، 46) وكمية المعلومات هي(20، 39) بينما حجم كمية المعلومات تنخفض في مستويات القدرة المنخفضة.

5- دراسة هند الحموري (2011):

عنوان الدراسة: دراسة استكشافية لملاءمة نماذج نظرية الاستجابة للمفردة في بناء اختبار من إعداد المعلم.

هدف الدراسة: استقصاء إمكان استخدام نماذج نظرية الاستجابة للمفردة في بناء الاختبارات التحصيلية ذات العدد القليل من الفقرات التي يعدها المعلمون بالإضافة إلى توضيح طريقة بناء الاختبار وتطويره وفق هذه النظرية.

عينة الدراسة: تكونت من 284 طالباً وطالبة من طلبة الصف الخامس في مديرية تربية وتعليم الزرقاء الأولى.

نتائج الدراسة: عدم إمكانية استخدام نظرية الاستجابة للمفردة في بناء الاختبارات ذات العدد القليل من الفقرات.

6- دراسة محمد أنور السامرائي وأحمد شاكر الخفاجي (2012):

عنوان الدراسة: بناء اختبار تحصيلي محكي المرجع في مادة علم نفس الخواص لطلبة أقسام العلوم التربوية والنفسية.

هدف الدراسة: بناء اختبار تحصيلي محكي المرجع في مادة علم نفس الخواص مبني وفق نظرية السمات الكامنة ونماذج القياس المنبثقة عنها (أنموذج راش) تحديداً.

عينة الدراسة: كانت عينة طبقية عشوائية بلغت 349 طالباً وطالبة في عدد من جامعات العراق وفي عدد من الكليات.

نتائج الدراسة: وجود 13 فقرة كانت قيمة مربع كاي لعشرة فقرات منها دالة إحصائياً لذا تم استبعادها من الاختبار فيما استبعد برنامج راسكال ثلاث فقرات في الصفحة الأولى من التحليل الإحصائي لعدم مطابقتها لأنموذج راش.

7- دراسة محمد دبوس(2016):

عنوان الدراسة: استخدام نظرية الاستجابة للفقرة في بناء فقرات اختبار محكي المرجع في الرياضيات بفقرات ثنائية التدريج ومتعددة التدريج وفق النموذج اللوجستي ثنائي المعلم.

هدف الدراسة: استخدام نظرية الاستجابة للفقرة في بناء تجمع فقرات اختبار محكي المرجع في الرياضيات وفق النموذج ثنائي المعلم.

عينة الدراسة: تكونت من (502) طالباً وطالبة من طلبة الصف الثامن الأساسي في مديرية تربية وتعليم نابلس للعام الدراسي 2014/2013 .

نتائج الدراسة: أظهرت النتائج تحقق افتراضات نظرية الاستجابة للفقرة كما بينت نتائج التحليل فيما يتعلق بمطابقة الفقرات الثنائية التدريج للنموذج الثنائي المعلم أن جميع الفقرات كانت مطابقة للنموذج ثنائي المعلم وأن الفقرات المتعددة التدريج كانت مطابقة لنماذج الاستجابة المتدرجة.

منهج البحث:

انطلاقاً من كون هذا البحث يهدف لبناء اختبار محكي المرجع في الرياضيات فإن المنهج الملائم هو المنهج الوصفي لأنه يمد الباحث بالبيانات والحقائق لوصف ما هو كائن أثناء البحث والحصول على تفسير لتلك البيانات.

مجتمع وعينة البحث:

تكوّن مجتمع البحث من تلاميذ الصف الخامس الأساسي في مدارس مدينة بانياس للعام الدراسي 2019/2018. وتكونت عينة البحث من (100) تلميذاً وتلميذة وتم اختيار العينة بالطريقة العشوائية.

تم اختيار مدرستين من مدارس مدينة بانياس هما مدرسة القصور الجنوبي وتشتمل على ثلاثة شعب للصف الخامس والبالغ عددهم 72 تلميذاً وتلميذة ومدرسة مساكن المصفاة وتشتمل على ثلاثة شعب وعدد تلاميذ 80 تلميذاً وتلميذة في الصف الخامس. تم تطبيق البحث على شعبتين من مدرسة القصور الجنوبي وشعبتين من مدرسة المساكن والبالغ عددهم 100 تلميذ وتلميذة.

أداة البحث:

لتحقيق هدف البحث تم بناء تجمع من مفردات اختبار محكي المرجع في الرياضيات من نوع أسئلة الاختيار من متعدد مؤلف من (25) مفردة لكل مفردة أربعة بدائل واحد منها فقط هو الصحيح.

صدق الاختبار:

صدق المحتوى: تم عرض الاختبار على عدد من مدرسي مادة الرياضيات وبناء على ملاحظاتهم تم تعديل بعض الأسئلة من حيث الصياغة وتم تعديل بعض البدائل بسبب التشابه وظل الاختبار مكوناً من (25) مفردة.

وبعد ذلك تم تطبيق الاختبار على عينة استطلاعية من تلامذة الصف الخامس خارج عينة البحث الأساسية بلغ عدد أفرادها /10/ تلاميذ في مدرسة القصور الجنوبي وذلك للتأكد من أن كافة البنود مفهومة من قبلهم؛ وقد بين التطبيق أن كافة البنود مفهومة وواضحة وأبدى أفراد العينة رغبة في الإجابة.

وبعد الانتهاء من التطبيق على العينة الاستطلاعية تم الحصول على تسهيل مهمة من مديرية التربية في طرطوس لتطبيق الاختبار في مدارس المحافظة، بعد التطبيق تم إدخال البيانات باستخدام برنامج ((spss

الاتساق الداخلي: تم التأكد من الاتساق الداخلي للبنود من خلال حساب معامل ارتباط بيرسون بين الدرجة الكلية وعلامة الفرد على كل بند. لقد كانت قيم معاملات الارتباط كلها دالة إحصائيا عند مستوى دلالة اقل من /0.05/. وكانت اقل قيمة لمعامل الارتباط / 0.67/ وأعلى قيمة هي / 0.89/.

ثبات الاختبار:

الثبات باستخدام معامل كودر- ريتشارد سون-21: تم استخدام معامل كودر ريتشاردسون (21) 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 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.

النتائج والمناقشة

ما درجة ملاءمة افتراضات نموذج راش اللوغاريتمي لبيانات الاختبار المستمدة من استجابات العينة؟

التحقق من افتراض أحادية البعد

تم اختبار فرضية أحادية البعد من خلال طريقة تحليل نموذج راش للمكونات الأساسية المعتمدة على البواقي كما ورد في دليل استخدام برنامج Winsteps ، وبما أن نموذج راش يفترض أحادية البعد إلا أن أحادية البعد ليس مطلقة، ويجب أن ينظر إلى مفهوم أحادية البعد كما يعنيها التحليل العاملي حيث يختلف الهدف بينهما. فالتحليل العاملي يهدف إلى تحديد العوامل التي يتكون منها الاختبار ولكن نظرية الاستجابة للمفردة تهدف تعرّف الانحرافات عن السمة المقاسة فيما اذا كانت ترقى إلى أن تكون عامل مستقل أو لا. يوضح الجدول التالي التحليل العاملي للمكونات الأساسية للبواقي باستخدام نموذج راش وقد تم استخراجها باستخدام برنامج (Winsteps).

```
    Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)
                                             -- Empirical --     Modeled
Total raw variance in observations     =      36.6 100.0%          100.0%
  Raw variance explained by measures   =       4.6  12.7%           12.7%
    Raw variance explained by persons  =       4.3  11.7%           11.7%
    Raw Variance explained by items    =        .3   0.9%              .9%
  Raw unexplained variance (total)     =      32.0  87.3% 100.0%    87.3%
    Unexplned variance in 1st contrast =       1.7   4.7%   5.4%
    Unexplned variance in 2nd contrast =       1.5   4.2%   4.8%
    Unexplned variance in 3rd contrast =       1.5   4.0%   4.5%
    Unexplned variance in 4th contrast =       1.4   3.8%   4.4%
    Unexplned variance in 5th contrast =       1.3   3.6%   4.2%
```

جدول (1) التحليل العاملي للبواقي وفق نموذج راش

حجم تباين البواقي المعيارية	الجذر الكامن	الملاحظ	المتوقع
التباين الكلي في الاستجابات	36.6	100.0%	100.0%
التباين الذي فسره العامل الرئيسي	4.6	12.7%	12.7%
التباين المفسر بواسطة الأفراد	4.3	11.7%	11.7%
التباين المفسر بواسطة المفردات	0.3	0.9%	0.9%
مجموع التباين غير المفسر	32.0	87.3%	87.3%
التباين الذي فسره العامل الأول	1.7	4.7%	5.4%
التباين الذي فسره العامل الثاني	1.5	4.2%	4.8%

تفسير الجدول السابق:

إن قيمة التباين الكلي غير المفسر بواسطة Raw unexplained variance أكبر أو تساوي (60%) فإن ذلك يعد مؤشراً قوياً لأحادية البعد. ونجد هنا أن قيمته قد بلغت (87.3%) ويسمى ببعد راش الأول.

البعد الثاني الأكبر يفسر (4.7%) من التباين والذي يشار إليه من القيمة Unexplned variance in 1^{st} contrast وبما أن قيمته أقل من (5%) فهذا مؤشر أخر على أحادية البعد.

نعتمد على محك ثالث هو قيمة الجذر الكامن لنسبة التباين الذي يفسره العامل الثاني إذ يجب ألا تتعدى أو تكون أقل من (3) وهذا المحك هو Unexplned variance in 2^{nd} contrast وقد بلغت قيمته هنا (1.5)

التحقق من استقلالية الموضع:

إن استجابة المفحوص على مفردات الاختبار تكون مستقلة إحصائياً عندما يؤخذ مستوى المفحوص بعين الاعتبار، أي أن استجابة المفحوص على مفردة ما يجب ألا يؤثر على مفردة أخرى، أي أن تقدير صعوبة أي مفردة لا يعتمد على تقديرات صعوبة المفردات الأخرى ولا يعتمد على قدرة الأفراد الذين يجيبون عليها، وكذلك لا يعتمد على تقدير قدرة أي مجموعة أخرى من الأفراد الذين طبق عليه الاختبار، ولا يعتمد على قيم صعوبة المفردات التي يجيبون عليها، وهذا الافتراض يوضح أن قدرة المفحوص وخصائص المفردة هما اللتان تؤثران في الأداء أو الاستجابة، وبتحقق هذا الافتراض فإن احتمال الحصول على أي تسلسل من العلامات لمجموعة من المفردات هو ببساطة حاصل ضرب الاحتمالات لكل هذه المفردات، ومن خلال نتيجة التحليل التباين تبين أن بيانات الاختبار المستمدة من استجابات العينة جاءت ملائمة لافتراضات نموذج راش حيث أن قيم التباين مرتفعة وذلك مؤشر قوي لأحادية البعد واستقلالية الموضع.

الحكم على ملائمة المفردات للنموذج:

يعتمد الحكم على المفردة على نوعين من الإحصاءات

للحكم على ملائمة المفردات من خلال مؤشر الملائمة التقاربية **Infit** ومؤشر الملائمة التباعدية**Outfit**

مؤشر الملائمة التقاربية Infit: يهتم هذا المؤشر بالعلاقة بين أداء المفردة الفعلي وأداءها المتوقع مع الأفراد ذوي المستوى القريب من صعوبتها. وتعتبر قيمة الملائمة التقاربية Infit Mean-Square مثالية إذا كانت تساوي (1). ويتم الحكم على دلالة قيمة الملائمة التقاربية من خلال الدرجة المعيارية لها Standardized Infit Mean-Square والتي تتراوح بين (-2.5) و (+2.5) وذلك عند مستوى ثقة (95%). وعندما تكون القيمة أكبر من (+2.5) فهذا يشير إلى عيب في صياغتها أو عدم صدقها في قياس ما تقيسه باقي المفردات. أما إذا كانت سالبة أقل من (-2.5) فذلك يعني أنها شديدة التشابه أو الاعتماد على بعض المفردات في المقياس. وفي أغلب الأحيان يعتمد الباحثين على أحد المؤشرين في الحكم على ملائمة المفردات للنموذج المستخدم، وللحكم على ملاءمة المفردات للنموذج جرى التحليل على مرحلتين:

التحليل الأول: تم 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 (-2.5 حتى +2.5). نتيجة التحليل تبين إن الأفراد غير الملائمين والذين تم حذفهم من البيانات قد بلغ عددهم (16) فرداً وبذلك يصبح عدد الأفراد (84).

التحليل الثاني: للتعرف على المفردات غير الملائمة بنفس الطريقة السابقة. وقد تم حذف المفردة رقم (18) لأن حدود الملائمة لها تجاوزت حدود الملائمة المطلوبة. وعند القيام بالتحليل للمرة الثانية لم يتم استبعاد أي مفردة انما تم استبعاد فرد واحد فقط هو الفرد رقم (46). وبذلك يصبح لدينا (83) فرداً و (24) مفردة في الاختبار النهائي.

السؤال الثالث:

ما تدريج صعوبة مفردات الاختبار باستخدام نموذج راش أحادي المعلم؟

للإجابة على هذا السؤال تم تحليل استجابات الأفراد باستخدام برنامج Winsteps المصمم لتحليل النموذج أحادي المعلم وكما ورد سابقا بأنه لم توجد حالات متطرفة في الأفراد أو المفردات. وتم تحليل (24) مفردة و(83) فرد،

والجدول التالي يبين معاملات الصعوبة مقدرة بوحدة اللوجيت.

يوضح الجدول صعوبة المفردات وأخطاؤها المعيارية مرتبة تنازليا وفق صعوبة المفردات

جدول (2) صعوبة المفردات وأخطاؤها المعيارية مرتبة تنازليا

معامل التمييز	ارتباط بوينت بيسريال	الملائمة التقاربية		الخطأ	الدرجة الخام	الصعوبة	المفردة
DISCRIM	PTME	IN.ZSTD	IN.MSQ	ERROR	SCORE	MEASURE	ENTRY
0.94	0.18	0.27	1.05	0.41	76	−1.46	1
0.69	0.28	1.18	1.13	0.26	57	0.38	2
1.17	0.46	−0.78	0.92	0.25	54	0.58	3
1.19	0.48	−0.86	0.92	0.25	45	1.14	4
0.82	0.37	0.71	1.07	0.25	51	0.77	5
0.86	0.17	0.47	1.09	0.34	72	−0.9	6
1.06	0.36	−0.19	0.95	0.32	70	−0.68	7
0.96	0.34	0.45	1.06	0.28	63	−0.05	8
1.4	0.54	−1.34	0.88	0.25	40	1.44	9
1.25	0.51	−1.25	0.75	0.33	71	−0.78	10
0.35	0.27	2.2	1.21	0.25	47	1.01	11
0.99	0.26	−0.02	0.97	0.39	75	−1.3	12
0.91	0.27	0.13	1.01	0.31	68	−0.48	13
1.42	0.56	−1.91	0.76	0.28	62	0.02	14
1.12	0.44	−0.29	0.97	0.25	53	0.64	15
1.19	0.46	−0.6	0.93	0.26	56	0.45	16
0.88	0.14	0.39	1.09	0.39	75	−1.3	17
0.74	0.33	1.09	1.11	0.25	53	0.64	19
1.07	0.38	−0.07	0.98	0.31	68	−0.48	20
0.73	0.32	0.92	1.09	0.25	54	0.58	21
1.15	0.43	−0.65	0.87	0.31	69	−0.57	22
1	0.32	0.15	1.02	0.31	69	−0.57	23
1.34	0.53	−1.4	0.82	0.28	62	0.02	24
0.4	0.27	1.84	1.18	0.25	49	0.89	25

ويمثل الجدول التالي نتائج تحليل الأفراد والمفردات للقيم المتحررة من قدرات الأفراد ومن صعوبة المفردات.

جدول (3) نتائج تحليل الأفراد والمفردات للقيم المتحررة من قدرات الأفراد ومن صعوبة المفردات.

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	17.3	24.0	1.26	.55	1.00	.1	1.06	.1
S.D.	3.8	.0	.98	.13	.14	.7	.56	.8
MAX.	23.0	24.0	3.42	1.03	1.49	1.8	4.71	2.0
MIN.	6.0	24.0	-1.26	.44	.73	-1.8	.51	-1.7

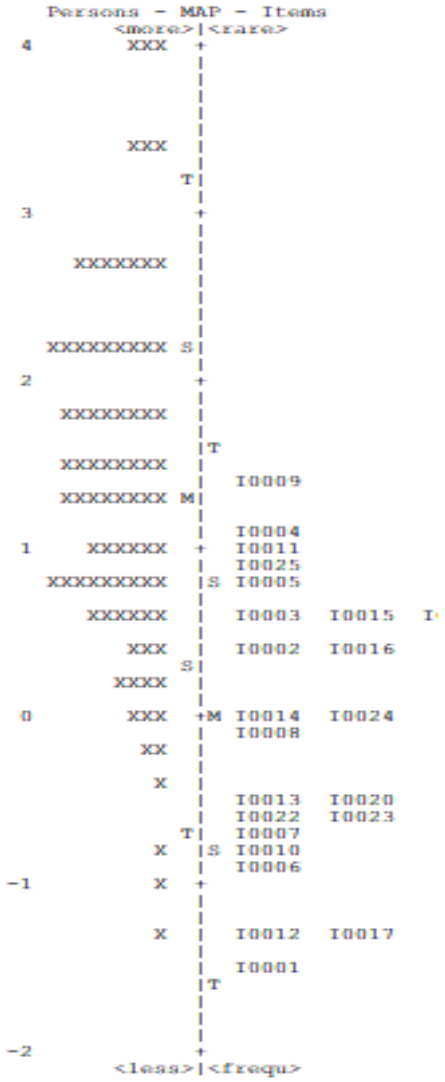
REAL RMSE	.58	ADJ.SD	.80	SEPARATION	1.38	Person RELIABILITY	.66	
MODEL RMSE	.56	ADJ.SD	.81	SEPARATION	1.43	Person RELIABILITY	.67	
S.E. OF Person MEAN = .11								

يلاحظ من الجدول السابق أن متوسط توزيع القدرة بلغ (1.26) بانحراف معياري قدره (0.98) والخطأ المعياري هو (0.13) وهو قريب مما يفترضه النموذج. مما يشير إلى دقة تحديد مواقع الأفراد على متصل السمة كما يبين الشكل أن قيم مؤشرات الملائمة الإحصائية لمتوسط المربعات التقاربية والتباعدية لمفردات الاختبار إلى أنها قريبة من الوضع المثالي الذي يفترضه النموذج وهو القيمة (1). حيث كان قيمة المتوسط حسابي (1.00) وانحراف معياري مقداره (0.14). والتي هي ضمن القيم التي حددها دليل عمل البرنامج والتي تقع بين (1.30) والقيمة (0.70) وتدل هذه المؤشرات على اقتراب المنحنى الملاحظ من المنحنى المتوقع (أفضل منحني له يطابق النموذج) وبذلك يتوافر لهذا المنحني ما يتوافر للمتوقع من النموذج من استقلالية معلمة صعوبة المفردات عن خصائص توزيع القدرة للأفراد.

الإجابة على السؤال الرابع

ما مؤشرات صدق الاختبار المحكي المرجع في مادة الرياضيات وفق نظرية الاستجابة للبند؟

وللتحقق من صدق الاختبار تم استخدام خريطة رايت (Wright) بعد التدريج النهائي للاختبار ويوضح الشكل التالي الخريطة



تظهر خريطة "رايت" بأن مفردات الاختبار تميل إلى أن تكون متسلسلة، فالزيادة في متوسط عتبة صعوبة المفردات تشير إلى أنها ليست منفصلة. ونجد أن كل مفردات الاختبار قد وقعت بين انحرافين معياريين عن المتوسط. وعدد المفردات التي تقع في المجال أقل من المتوسط مساو لعدد المفردات التي تقع أعلى من المتوسط.

ما مؤشرات ثبات الاختبار المحكي الرجع في مادة الرياضيات وفق نظرية الاستجابة للمفردة؟

مؤشرات ثبات كل من المفردات والأفراد:

ثبات الأفراد: تشير قيم ثبات الأفراد إلى مدى اتساق الترتيب النسبي لتقدير قدرات الأفراد في حالة إعادة التدريج. وتشير القيم الضعيفة إلى ضيق مدى قدرات الأفراد أو صغر حجم الاختبار. ونلاحظ إن قيمة الثبات قد بلغت (0.66)

REAL RMSE	.58	ADJ.SD	.80	SEPARATION	1.38	Person RELIABILITY	.66
MODEL RMSE	.56	ADJ.SD	.81	SEPARATION	1.43	Person RELIABILITY	.67
S.E. OF Person MEAN = .11							

ثبات المفردات:

تشير قيم ثبات المفردات إلى مدى اتساق الترتيب النسبي لتقدير صعوبات المفردات في حالة إعادة التدريج، ولا يوجد مؤشر يقابله في النظرية الكلاسيكية. لذا تنفرد نظرية الاستجابة للبند بحسابه. وتشير القيم الضعيفة إلى ضيق مدى صعوبات المفردات أو صغر حجم العينة.

ونلاحظ من الجدول أن قيمة الثبات قد بلغت (0.86)

جدول (4)

SUMMARY OF 24 MEASURED (NON-EXTREME) Items

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	60.8	83.0	.00	.29	.99	.0	1.06	.1
S.D.	10.3	.0	.82	.05	.12	1.0	.36	1.2
MAX.	76.0	83.0	1.44	.41	1.21	2.2	2.25	2.3
MIN.	40.0	83.0	-1.46	.25	.75	-1.9	.54	-1.9
REAL RMSE	.30	ADJ.SD	.76	SEPARATION	1.51	Item RELIABILITY	.86	
MODEL RMSE	.30	ADJ.SD	.76	SEPARATION	1.57	Item RELIABILITY	.87	
S.E. OF Item MEAN = .17								

DELETED: 1 Items

UMEAN=.000 USCALE=1.000

Item RAW SCORE-TO-MEASURE CORRELATION = -.99

1920 DATA POINTS. LOG-LIKELIHOOD CHI-SQUARE: 1834.39 with 1817 d.f. p=.3826

معامل الفصل:

يتيح برنامج Winsetps مؤشر أخر هو معامل الفصل Separation Index كمؤشر للثبات. ويشير معامل الفصل إلى نسبة التمييز لفيشر حيث يعبر عن النسبة بين التباين الحقيقي وتباين الخطأ. ويجب أن لا تتجاوز قيمته (2).

ونلاحظ من الجدول السابق أن قيمته لم تتجاوز (2) في كلتا الحالتين. فقد بلغت بالنسبة للافراد /1.38/ كما بلغت بالنسبة للمفردات /1.51/.

ما مؤشرات القدرة التمييزية للاختبار المحكي المرجع في مادة الرياضيات وفق نظرية الاستجابة للمفردة؟

استخراج إحصاءات تمييز وتخمين المفردات:

يوفر برنامج Winsteps تقدير كل من التمييز والتخمين. ونظراً إلى أن نموذج راش يفترض أن تميز المفردات ثابت لذا يجب أن يراعى أن تكون قيم معاملات التمييز متسقة ويفضل حذف القيم الشاذة كما يفضل بشكل عام أن تكون القيم أكبر من (0.5) أما إذا كانت القيم غير متسقة فيجب عند ذلك استخدام النموذج الثنائي أو الثلاثي المعلم

جدول (5) قيم معاملات التمييز للبنود المتبقية

ENTRY NUMBER	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	ESTIM DISCR	ASYMPTOTE LOWER	ASYMPTOTE UPPER	Item
1	76	83	-1.46	.41	1.05	.3	1.27	.7	.18	.25	91.3	91.2	.94	.00	1.00	I0001
2	57	83	.38	.26	1.13	1.2	1.22	1.2	.28	.39	66.3	72.0	.69	.10	.96	I0002
3	54	83	.58	.25	.92	-.8	.96	-.2	.46	.40	68.8	70.0	1.17	.00	1.00	I0003
4	45	83	1.14	.25	.92	-.9	1.00	.0	.48	.43	76.3	67.4	1.19	.00	1.00	I0004
5	51	83	.77	.25	1.07	.7	1.07	.5	.37	.41	67.5	68.5	.82	.08	1.00	I0005
6	72	83	-.90	.34	1.09	.5	1.51	1.3	.17	.29	86.3	86.5	.86	.00	.98	I0006
7	70	83	-.68	.32	.95	-.2	.83	-.4	.36	.31	85.0	84.4	1.06	.00	1.00	I0007
8	63	83	-.05	.28	1.06	.5	.93	-.2	.34	.36	75.0	77.0	.96	.07	1.00	I0008
9	40	83	1.44	.25	.88	-1.3	.83	-1.4	.54	.44	67.5	67.7	1.40	.00	1.00	I0009
10	71	83	-.78	.33	.75	-1.2	.54	-1.4	.51	.30	88.8	85.4	1.25	.00	1.00	I0010
11	47	83	1.01	.25	1.21	2.2	1.23	1.7	.27	.42	58.8	67.5	.53	.10	.97	I0011
12	75	83	-1.30	.39	.97	.0	1.14	.4	.26	.26	90.0	90.0	.99	.00	1.00	I0012
13	68	83	-.48	.31	1.01	.1	1.52	1.6	.27	.33	85.0	82.2	.91	.00	.99	I0013
14	62	83	.02	.28	.76	-1.9	.62	-1.9	.56	.37	85.0	76.1	1.42	.00	1.00	I0014
15	53	83	.64	.25	.97	-.3	.90	-.6	.44	.41	67.5	69.5	1.12	.00	1.00	I0015
16	56	83	.45	.26	.93	-.6	.84	-1.0	.46	.39	71.3	71.4	1.19	.00	1.00	I0016
17	75	83	-1.30	.39	1.09	.4	2.25	2.1	.14	.26	90.0	90.0	.88	.00	.99	I0017
18	DELETED															I0018
19	53	83	.64	.25	1.11	1.1	1.08	.6	.33	.41	60.0	69.5	.74	.09	1.00	I0019
20	68	83	-.48	.31	.98	-.1	.76	-.7	.38	.33	82.5	82.2	1.07	.00	1.00	I0020
21	54	83	.58	.25	1.09	.9	1.16	1.1	.32	.40	73.8	70.0	.73	.00	.96	I0021
22	69	83	-.57	.31	.87	-.6	.76	-.7	.43	.32	86.3	83.3	1.15	.00	1.00	I0022
23	69	83	-.57	.31	1.02	.2	.93	-.1	.32	.32	81.3	83.3	1.00	.08	1.00	I0023
24	62	83	.02	.28	.82	-1.4	.65	-1.7	.53	.37	77.5	76.1	1.34	.00	1.00	I0024
25	49	83	.89	.25	1.18	1.8	1.33	2.3	.27	.42	63.8	68.0	.54	.08	.97	I0025
MEAN	60.8	83.0	.00	.29	.99	.0	1.06	.1			76.9	77.1				
S.D.	10.3	.0	.82	.05	.12	1.0	.36	1.2			10.0	8.1				

ونلاحظ من الجدول السابق أن قيم معاملات التمييز للبنود المتبقية متسقة فيما بينها ومحصورة بين (1.420.53-) ونلاحظ أن القيم جميعها أكبر من (0.5)

استخراج معاملات التخمين:

يفترض نموذج راش أن معاملات تخمين المفردات تساوي الصفر. لذا يجب أن تراعى قيم تخمين المفردات بحيث لا تتعدى القيمة (0.10) وفي هذه الحالة يجب حذف المفردة.

ويبين الشكل السابق قيم معاملات التخمين والتي جميعا ضمن المجال المقبول

إذن نلاحظ أن قيم معاملات التخمين للبنود المتبقية لم تتجاوز /0.10/ وكانت اكبر قيمة هي /0.10/.

ماموشرات صعوبة الفقرات للاختبار المحكي المرجع في مادة الرياضيات؟

استخراج نتائج (صعوبات) المفردات:

يمثل الجدول التالي العلاقة بين الدرجة الكلية المتوقعة والقدرة المقابلة لكل درجة خام كلية مع الخطأ المعياري لتقدير القدرة. وهذه القيم محسوبة بوحدة اللوجيت لأننا حددنا في ملف التحكم أن تكون كافة القيمة محسوبة بوحدة اللوجيت.

جدول (6) العلاقة بين الدرجة الكلية المتوقعة والقدرة المقابلة لكل درجة خام كلية مع الخطأ المعياري لتقدير القدرة.

SCORE	MEASURE	S.E.	SCORE	MEASURE	S.E.
0	-4.69	1.84	13	0.2	0.44
1	-3.44	1.04	14	0.4	0.44
2	-2.67	0.76	15	0.6	0.45
3	-2.19	0.64	16	0.8	0.46
4	-1.82	0.57	17	1.02	0.48
5	-1.52	0.53	18	1.26	0.5
6	-1.26	0.5	19	1.52	0.53
7	-1.02	0.48	20	1.82	0.57
8	-0.8	0.46	21	2.18	0.64
9	-0.59	0.45	22	2.66	0.75
10	-0.38	0.45	23	3.42	1.03
11	-0.19	0.44	24	4.67	1.84
12	0.01	0.44			

الاستنتاجات والتوصيات:

وفي ضوء نتائج البحث تم اقتراح الآتي:

1. تطبيق البحث على مواد دراسية أخرى غير الرياضيات وعلى صفوف دراسية أخرى وذلك لتوسيع مدى إمكانية تعميم النتائج والتأكيد على أهمية استخدام نظرية الاستجابة للمفردة كمدخل حديث في القياس التربوي.

2. نتيجة لموضوعية القياس التي حققها نموذج راش في البحث الحالي، يوصي البحث باستخدام نماذج الاستجابة المفردة الأخرى (ثنائي، ثلاثي المعلم) في تدريج المقاييس التربوية لتحسين بنائها.

المراجع:

- أبو جراد، حمدي يونس. استخدام نموذج راش في تطوير اختبار كاتل الثالث للذكاء الصورة (أ). مجلة الجامعة الإسلامية (المجلد السادس عشر)، العدد الثاني، 2008، ص533-ص583
- حجازي، تغريد والخطيب، عبدالله. التوافق بين النظرية الكلاسيكية والنموذج ثنائي المعلمة في مطابقة فقرات اختبار محكي المرجع في أحكام التلاوة والتجويد. مجلة جامعة النجاح للأبحاث، المجلد 28، 2014، ص2240- ص2270
- الحموري، هند عبد المجيد. دراسة استكشافية لملائمة نماذج نظرية الاستجابة للمفردة في بناء اختبار تحصيلي من إعداد المعلم. مجلة العلوم التربوية والنفسية، المجلد 12، العدد2، 2011، ص48- ص81
- دبوس، محمد. 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، ص1454- ص 1480
- السامرائي، محمد أنور محمود والخفاجي، أحمد محمد شاكر. بناء اختبار تحصيلي محكي المرجع في مادة علم نفس الخواص لطلبة أقسام العلوم التربوية والنفسية. مجلة الأستاذ، العدد 203، 2012، ص 964- ص1002
- السيد ابراهيم، نسرين اسماعيل. تأثير بعض المتغيرات على دقة القياس باستخدام نظرية الاستجابة للمفردة (دراسة سيكومترية). رسالة دكتوراه، جامعة عين شمس، 2010، ص5- ص446
- الشواورة، ياسين سالم حماد. دراسة مقارنة بين نموذجي التقدير الجزئي والاستجابة المتدرجة في معادلة درجات الاختبارات، رسالة دكتوراه، جامعة القاهرة، 2013، ص1-ص9
- علام، صلاح الدين محمود. القياس والتقويم التربوي والنفسي (أساسياته وتطبيقاته وتوجهاته المعاصرة). دار الفكر العربي، القاهرة، 2006، ص31- ص748
- علي، نداء بهاء الدين. فاعلية استخدام نموذج راش في بناء اختبار تحصيلي محكي المرجع لمقرر القياس والتقويم في التربية. رسالة دكتوراه، جامعة دمشق،2011
- العنزي، محمد طالب. أثر شكل الفقرة على معالم الفقرة وثبات الاختبار وفقاً لنظرية استجابة الفقرة (IRT) . رسالة ماجستير، جامعة مؤته، 2010، ص1-ص110
- كاظم، أمينة. استخدام نموذج راش في بناء اختبار تحصيلي في علم النفس وتحقيق التفسير الموضوعي للنتائج. مطبوعات جامعة الكويت، 1988، ص7- ص127

- محمود، سومية شكري محمد. استخدام نموذج راش في بناء بنك أسئلة لقياس التحصيل في مقرر سيكولوجية التعلم لدى طلاب كلية التربية بالمنيا. رسالة دكتوراه، جامعة المنيا، 2012، ص1-ص 199

- مسعود، وليد. دراسة أثر القياس الدينامي على فاعلية برنامج لتنمية الفهم القرائي باستخدام نظرية الاستجابة للمفردة. رسالة دكتوراه، جامعة عين شمس، 2010، ص1- ص496

- مسعود، وليد. محاضرات في القياس (نظرية الاستجابة للمفردة مع تطبيقات عملية). كلية التربية، جامعة أم القرى، 2014، ص1- ص76

Reference

- Abu Jarad, Hamdi Younis. The use of the Rasch model in the development of the third Cattell test for intelligence (a). Journal of the Islamic University (Volume Sixteen), No. 2, 2008, pp. 533-p. 583

-Hijazi, Taghreed and Al-Khatib, Abdullah. The agreement between the classical theory and the two-parameter model in matching the items of the referenced test in the provisions of recitation and intonation. An-Najah University Research Journal, Volume 28, 2014, pp. 2240- pp. 2270

-Al-Hamouri, Hind Abdul-Majid. An exploratory study of the suitability of individual response theory models in constructing an achievement test prepared by the teacher. Journal of Educational and Psychological Sciences, Volume 12, Issue 2, 2011, pp. 48-p. 81

-Dabbous, Muhammad. The use of paragraph response theory in constructing a reference test in mathematics with two- and multi-graded items according to the two-teacher logistic model. An-Najah University Research Journal, Volume 30, 2016, pp. 1454- pp. 1480

-Al-Samarrai, Muhammad Anwar Mahmoud, and Al-Khafaji, Ahmed Muhammad Shaker. Constructing a reference based achievement test in the subject of trait psychology for students of educational and psychological sciences departments. Al-Ustad Magazine, Issue 203, 2012, pp. 964- pp. 1002

-Mr. Ibrahim, Nasreen Ismail. The effect of some variables on the measurement accuracy using the individual response theory (a psychometric study). PhD dissertation, Ain Shams University, 2010, pp. 5- p. 446

-Al-Shawawra, Yassin Salem Hammad. A comparative study between the partial assessment and graded response models in the equation of test scores, Ph.D. dissertation, Cairo University, 2013, pp. 1-p. 9

-Allam, Salahuddin Mahmoud. Educational and psychological measurement and evaluation (its basics, applications and contemporary trends). Dar Al-Fikr Al-Arabi, Cairo, 2006, pp. 31-pg. 748

- Ali, Bahauddin's call. The effectiveness of using the Rasch model in constructing a referenced achievement test for the measurement and evaluation course in education. PhD thesis, Damascus University, 2011

-Al-Anazi, Muhammad Talib. The effect of vertebra shape on vertebra parameters and test stability according to vertebra response theory (IRT). Master Thesis, Mu'tah University, 2010, pp. 1-p. 110

-Kazem, Amina. Using the Rasch model in constructing an achievement test in psychology and achieving objective interpretation of the results. Kuwait University Press, 1988, pp. 7 - pp. 127

-Mahmoud, Soumya Shukri Muhammad. Using the Rasch model in building a question bank to measure achievement in the learning psychology course for students of the Faculty of Education in Minia. PhD thesis, Minia University, 2012, pp. 1-p. 199

-Masoud, Walid. Studying the effect of dynamic measurement on the effectiveness of a program for developing reading comprehension using the word response theory. Ph.D. Thesis, Ain Shams University, 2010, pp.1-p.496

-Masoud, Walid. Lectures on measurement (theory of response to the individual with practical applications). College of Education, Umm Al-Qura University, 2014, pp. 1-pg. 76

− Baker, Frank. The Basics of Item Response Theory. University of Wisconsin, Educational Resources Center, second edition, 2001, p1- p176

− Carlson, James and Davier, Matthias. Item Response Theory. ETS Research Report no. RR-13- 28,New Jersey, 2013, p1- p70

− Partcher, Ivalio. A visual guide to item response theory . Friedrich Schiller, University at Jena, p1- p61

− Rivera, Jennifer. 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, 2007, p1- p178

− Wiberg, Marie. Classical Test Theory Vs. Item Response Theory An evaluation of the theory test in the Swedish driving-license test. EM No 50, 2004, 2004, p1- p26