مجلة جامعة الأقصى (سلسلة العلوم الإنسانية) المجلد السادس عشر، العدد الأول، ص 87- 111 ،يناير 2012
ISSN 2070 - 3147

درجة مطابقة اختبار تحصيلي وفق نموذج راش أحادي المعلمة في الكشف عن مستوى المعرفة العلمية في المهارات الرياضية لدى طلبة الصف الثامن الأساسي

د. ماجد محمد الخياط *

الملخص

هدفت الدراسة إلى التحقق من فاعلية النموذج اللوغارتمي ذي المعلمة الواحدة" نموذج راش" ومطابقة البيانات للنموذج لاختبار(TIMSS)والذي يقيس المهارات الرياضية لمستوى الصف الثامن الأساسي، ولتحقيق هدف الدراسة تم تطبيق فقرات الاختبار على عينة من طلبة الصف الثامن الأساسي بلغ عددها (599) طالباً وطالبة؛ بعد إيجاد الخصائص السيكومترية (الصدق والثبات)لاختبار(TIMSS)، وللإجابة على أسئلة الدراسة تم استخدام البرنامج الإحصائي (Rumm, 2010)؛ وأشارت نتائج الدراسة إلى تمتع الاختبار المقدم بالصدق والثبات؛ حيث بلغت قيمة معامل الثبات للاختبار باستخدام معامل كرونباخ الفا(0,90)، أما معامل الصدق بدلالة المحك بين درجات الطلبة على الاختبار ودرجاتهم في مادة الرياضيات فقد بلغت قيمته(0,87). وأشارت النتائج أيضاً إلى مناسبة (25) فقرة من فقرات الاختبار والبالغ عددها (35) فقرة للبيئة المحلية بحيث تمتعت بخصائص إحصائية مناسبة، من حيث معامل صعوبة الفقرة.

ABSTRACT

The purpose of this study was to investigate the effectiveness of the one parameter model "rasch model" and fitness of data for the model to (TIMSS) test, which measures mathematical skills to the level of eighth grade. The sample was chooses for (599) students from eighth grade at the Directorate of Education in Balqa students Governorate / Jordan after the psychometric properties (validity and reliability) from the (TIMSS) test were created.

To answer the questions of this study the researchers used statistical program (Rumm 2010), The test showed that the results provided were valid and reliable, as the value of reliability coefficient for the test using the coefficient of Cronbach Alpha was (0, 90), while the coefficient of criteria-

* جامعة البلقاء التطبيقية – الأردن.

related validity between the degrees of Students (TIMSS Test) and Students marks in mathematics has reached (0, 87).

The results indicated to the suitability of (25) items from the test's original (35) items to the local environment, with appropriate statistical properties, in terms of the item difficulty coefficient.
Keywords: Fitness, Rasch Model, TIMSS Test, Mathematical Skills.

مقدمة:

أهتم علماء النفس منذ أن بدأت حركة القياس النفسي بتحقيـق صـدق وثبـات الاختبـارات والمقاييس النفسية؛ سعياً منهم لتحقيق أعلى درجة مـن الموضـوعية لهـذه الاختبـارات عنـد استخدامها في عملية القياس، فوفق النظرية الكلاسيكية(Classical Theory)يمكن التعبير عـن قدرة الفرد من خلال الدرجة الحقيقية؛ والتي تنتج عن أدائه على الاختبار مع وجود خطأ؛ وبنـاء على ذلك ستتغير درجة الطالب بتغير مستوى الاختبـار؛ بحيـث أن فقـرات الاختبـار تتغيـر خصائصها بتغير خصائص الأفراد، كما أن خصائص الأفراد تتغير بتغير خصائص الاختبار؛ من حيث سهولة وصعوبة الفقرات، وقد أسفرت جهود العلماء في مجال القياس عن ظهـور بعـض الاتجاهات الحديثة في مجال القياس؛ ومن بـين هـذه الاتجاهـات نظريـة الاسـتجابة للفقـرة Item Response Theory (IRT) أو نظريـة السـمات الكامنـة Latent Trait Theory (LTT)، وحظي هذا المدخل الجديد باهتمام الباحثين؛ حيث يتغلب على كثير من مشكلات القياس التقليدية.

وتُعد النظرية الحديثة في القياس التي يطلق عليها نظرية الاستجابة على الفقرة (Item Response Theory "IRT") نقطة تحول في هذا التطور (Warm, 1978) ؛ وذلك بسبب ما قدمته من طرق ذات فعالية كبيرة فيما يتعلق ببناء الاختبارات وتصحيحها وتحليل نتائجها مقارنة بما قدمته النظرية الكلاسيكية في القياس (Mislevy & Bock, 1990).

وقد هدفت هذه النظرية على التغلب على مشكلات النظرية الكلاسيكية؛ إذ أن الاختبارات التحصيلية التي تُصمم في إطار النظرية الحديثة تكون على درجة عالية من الموضوعية، ويمكن تعميم نتائجها، كما أنها تتيح الفرصة للمقارنات بين مجموعات متباينة من الأفراد؛ بحيث يمكن تقدير قدرة الفرد الذي يطبق عليه الاختبار بدقة، ومقارنته بغيره من الأفراد.

إن بناءَ وتطوير المقاييس باستخدام النظرية الكلاسيكية يتضمن انتقاء الفقرات وفقاً لمحتواها وخصائصها المتعلقة بالصعوبة، والتمييز، وفاعلية بدائلها؛ حيث يتم انتقاء الفقرات ذات الدلالة التمييزية العالية، ومستوى الصعوبة المناسب؛ والذي يناسب الغرض من المقياس المقدم، وتوزيع القدرة في المجموعة التي يطبق عليها المقياس أو الاختبار(Wolfe & Smith,2007).

وبما أن المؤشرات الكلاسيكية للفقرات غير ثابتة عبر المجتمعات التي تختلف في مستوى القدرة (Bond & Fox,2007)؛ لذلك فإن الإجراءات الكلاسيكية في انتقاء الفقرات يعتمد على المجموعة المستخدمة في تحديد مؤشرات الفقرة، ودرجة تمثيل المجموعة للمجتمع الذي سيطبق عليه الاختبار(Mislevy & Bock, 1990)، (أي أن اختلاف المجموعة المطبق عليها الاختبار يؤثر في الخصائص الإحصائية للفقرات)، بعكس النظرية الحديثة والتي يتم فيها انتقاء الفقرات اعتماداً على مجموعة كبيرة من الأفراد (أي أن الخصائص الإحصائية للفقرات غالباً ما تبقى ثابتة حتى لو تغيرت مجموعة الأفراد). ولقد انبثق عن نظرية الاستجابة للفقرة مجموعة من النماذج الرياضية، يعتمد كل نموذج منها على معادلة رياضية تحدد علاقة أداء الفرد على فقرة المقياس بقدرته التي تكمن وراء هذا الأداء، وتفسير قدرة الفرد حسب موقعه على سلم القدرة؛ إلا أن اختيار النموذج المناسب يعتمد على مطابقة البيانات(المستمدة من العينة)على النموذج نفسه.

ويعد نموذج راش الأحادي المعلمة من أفضل النماذج إذا ما تم تحقيق فروضه؛ حيث إن انتقاء الفقرات التي تتميز بخصائص إحصائية مناسبة من المشكلات التي تواجه مطوري الاختبارات، وتبرز هذه المشكلة بشكل واضح عند الحاجة لتوفير عدد كبير من الفقرات (Item Pool)، أو بنك من الفقرات(Item Banking) لتشكيل هذه الاختبارات، فمطور الاختبار لا يهتم فقط بتحديد نوع أو صيغة الفقرات؛ بل أيضاً بتحديد خصائص الفقرات التي سيتضمنها الاختبار مثل الصعوبة، والتمييز، والتخمين، ودالة معلومات الفقرة، ودالة معلومات الاختبار، وغيرها من الخصائص الإحصائية الخاصة بالفقرات.

مشكلة الدراسة:

يعتمد معظم الباحثين المختصين في مجال القياس والتقويم عند انتقاء فقرات لاختباراتهم على قياس مستوى صعوبة فقرات الاختبار، والدلالة التمييزية بناء على النظرية الكلاسيكية في

القياس، ولقلة وجود اختبارات تم غربلة فقراتها وفق النظرية الحديثة في القياس؛ والتي تعتمد على قدرة المفحوص عند تقديم فقرات الاختبار كمحك أساسي عند تقديم فقرات الاختبار؛ حيث أظهرت نتائج العديد من الدراسات دقة وموضوعية نتائج نماذج النظرية الحديثة في القياس في إعداد الاختبارات، وقياس مستوى أداء الطلاب بدقة؛ بالرغم من اختلاف نوعية وأهداف الاختبارات المستخدمة، وأن الفقرات التي يتم انتقاؤها وفقاً لنوع النموذج تختلف إذا كان مدى تمييز فقرات الاختبار متسعاً، فالنماذج الكلاسيكية تستبعد فقط الفقرات ذات التمييز المنخفض، أما نماذج النظرية الحديثة في القياس فتستبعد الفقرات ذات التمييز المتطرف (المنخفض أو المرتفع)، وهذا التفاوت في انتقاء الفقرات ربما يرجع إلى اختلاف الهدف من تحليل الفقرات في كل من الحالتين، فالهدف في حالة النماذج الكلاسيكية يكون موجهاً نحو الكشف عن فقرات الاختبار ذات مستوى صعوبة ودلالة تمييزية غير مناسبة، وتقدير البارامترات المتعلقة بكل فقرة، واستخدامها في تقدير الخصائص السيكومترية للاختبار ككل، أما في حالة نماذج النظرية الحديثة في القياس فيكون الهدف موجهاً نحو الحصول على تدريج لفقرات الاختبار (Items Calibration)، أي تقدير صعوبة الفقرات (Items Difficulty) وتقييم حسن مطابقة درجاتها للنموذج المستخدم، والإفادة من ذلك في تقدير سمات الأفراد (صلاح الدين علام، 1987 :29)،ونظراً لقلة الدراسات التي عملت على تحليل فقرات لاختبارات غير تحصيلية عن طريق إخضاع بيانات الاختبارات لنموذج راش، ولشيوع استخدام هذا النموذج في البيئة الأجنبية؛ شجع ذلك للباحث دافعاً قوياً لإجراء هذه الدراسة، إضافة إلى الحاجة لوجود صورة أردنية لاختبار (TIMSS) يتمتع بخصائص سيكومترية مناسبة لتطبيقه على البيئة الأردنية؛ لذا جاءت هذه الدراسة لقياس فاعلية اختبار(TIMSS)وفق نموذج راش أحادي المعلمة في الكشف عن مستوى المعرفة العلمية في المهارات الرياضية لدى طلبة الصف الثامن الأساسي؛ ولمعرفة مدى امتلاك الطلبة في مستوى الصف الثامن الأساسي للمهارات الرياضية.

أسئلة الدراسة: تسعى الدراسة الحالية إلى الإجابة عن الأسئلة الآتية:

1- ما درجة تحقيق بيانات اختبار Timss لافتراضات النموذج أحادي المعلم(راش) لدى طلبة الصف الثامن الأساسي؟

2- ما مؤشرات المطابقة الإحصائية لفقرات الاختبار المقدم TIMSS لافتراضات النموذج

أحادي المعلم (راش) لدى طلبة الصف الثامن الأساسي؟

3- ما مستوى القدرة المتوقعة لدى طلبة الصف الثامن الأساسي وفق نموذج راش أحادي المعلمة؟

أهمية الدراسة:

1- **الأهمية العلمية**: إن استخدام نموذج راش(أللوغاريتمي أحادي المعلم) يضيف أهمية لهذه الدراسة في تحقيق الموضوعية في القياس؛ وذلك من خلال تدريج معالم الفقرات التي تعبر عن أداء الطلبة على مقياس السمة الكامنة بشكل مستقل عن خصائص العينة؛ وذلك لأن تحليل الفقرات لمقياس ما وفق النظرية التقليدية في القياس يعطي معالم غير ثابتة لقيم تمييز، وصعوبة الفقرات باختلاف عينة الأفراد التي يطبق عليها الاختبار، إضافة إلى تقديرات غير ثابتة لأداء الأفراد، وباختلاف مجموعة الفقرات التي تتم استجاباتهم عليها، كما أنه من الناحية النظرية فإن استقصاء فاعلية نموذج راش؛ والذي يستخدم في تحليل فقرات ثنائية الاستجابة؛ يعتبر ذو فائدة في إضافة دليل على أهمية هذا الأسلوب في تحليل الفقرات، هذا إلى جانب توفير اختبار يتضمن فقرات تعكس قدرات الطلبة مقاسه على نفس المتصل؛ بحيث يمكن إجراء مقارنة مباشرة بين قدرات الطلبة على اختلاف قيمها في متصل القدرة.

2- **الأهمية التطبيقية:** حيث ستسهم هذه الدراسة في تحقيق الجوانب التالية:

أ- فحص افتراضات نظرية الاستجابة للفقرة وتحديداً نموذج راش فيما يتعلق بتوضيح أثر خصائص المفحوصين والفقرات، على دقة التقدير لمعالم المفحوصين والفقرات.

ب- زيادة قدرة معدي(معلمي الرياضيات) اختبارات مادة الرياضيات على تحديد الطريقة التي بواسطتها يمكن اختيار الفقرات الملائمة؛ وذلك للحصول على دقة أكبر في تقدير قدرات عينة من المفحوصين(في حالة توافر بنك للأسئلة).

ج-توفير اختبار يقيس المهارات الرياضية بعد تدريجه وفق نموذج راش يصلح للاستخدام في البيئة الأردنية.

أهداف الدراسة:

تسعى هذه الدراسة لتحقيق الأهداف التالية:

1- الكشف عن درجة مطابقة بيانات اختبار TIMSS لافتراضات نموذج راش.

2- انتقاء فقرات من اختبار TIMSS للرياضيات مناسبة لمرحلة الصف الثامن الأساسي.

حدود الدراسة:

تتحدد نتائج الدراسة الحالية في الأمور التالية:

1-طبيعة عينة الدراسة اقتصرت على طلبة الصف الثامن الأساسي في مديرية التربية والتعلـيم/ مديرية تربية السلط.

2- الأبعاد المشمولة بالأدوات والتي تم تطويرها من قبل الباحث.

3-اقتصرت نتائج الدراسة على مقياس اختبار TIMSS والذي طوره الباحث.

التعريفات المفاهيمية والإجرائية:

تم تعريف مصطلحات الدراسة بشكل إجرائي وهي نمـوذج راش، المطابقـة، اختبـار TIMSS، على النحو الآتي:

1- **اختبار TIMSS:**

وهو اختبار تحصيلي يقيس مستوى الطلبة ممن أنهوا الصف الثامن الأساسـي فـي المهـارات الرياضية.

2- **نموذج راش Rasch Model:**

يطلق عليه النموذج أحادي المعلمة One Parameter Model ، وقد اقترحه جورج راش، ويهتم بتحديد موقع الفقرة الاختبارية على سلم صعوبة لجميع الفقرات التي تشكل الاختبار، كمـا يهتم بتدريج مستويات قدرة الفرد باختبار معين على نفس متصل الفقرات (علام، 2000).

3- طلبة الصف الثامن:

تبلغ أعمارهم (14) عاماً من الذكور والإناث المسجلين في المرحلة الأساسية بالمدارس الحكومية في مديرية تربية السلط بالأردن.

الإطار النظري والدراسات السابقة

لقد ظهرت نظرية الاستجابة للفقرة كنظرية بديلة للنظرية الكلاسيكية والتي يفترض فيها تلافي عيوب النظرية الكلاسيكية، كما تم توضيحه سابقاً، وقد بين هامبلتون (Hambleton, et .al, 1985) ثلاث مزايا رئيسية لنظرية الاستجابة للفقرة هي:

1- بافتراض وجود عدد كبير من فقرات الاختبار الذي يقيس سمة معينة؛ فإن تقدير قدرة الفرد يتم مستقلاً عن عينة الفقرات التي تطبق عليه؛ أي أن تقديرات القدرة للأفراد تحرر إحصائياً من خصائص الفقرات المستخدمة في تقدير القدرة " Item – Freed".

2- بافتراض وجود عدد كبير من الأفراد يكون تقدير معالم الفقرات مستقلاً إحصائياً عن عينة الأفراد التي استخدمت في تقدير هذه المعالم " Sample – Freed ".

3- يمكننا الحصول على إحصاء (Statistic) لدرجة الدقة في قياس قدرة كل فرد، وتختلف قيمته من فرد إلى آخر.

ومن الجدير بالذكر أن الاستفادة من المزايا الثلاث تحصل عندما يكون هناك مطابقة بين البيانات الاختبارية وأحد نماذج السمات الكامنة المستخدمة في القياس؛ وذلك لأن بنية هذه النماذج في حالة المطابقة تتضمن تحقق المزايا الثلاث السابقة.

ويعد نموذج راش الأحادي المعلمة من أكثر النماذج شيوعاً في تصميم وبناء الاختبارات والمقاييس التربوية والنفسية، ويهتم هذا النموذج بتحديد موقع الفقرة على تدريج صعوبة جميع فقرات الاختبار، بالإضافة إلى اهتمامه بتدريج مستويات قدرة الفرد في اختبار معين على نفس مستويات القدرة المقاسة ، كما يفترض النموذج أن الفرد لا يلجأ إلى التخمين العشوائي في إجابته على فقرات الاختبار(علام ، 2000)، ونموذج راش "Rasch Model" يوفر متطلبات القياس الموضوعي عندما تتحقق فروضه، وهي:

- أحادية البعد: حيث يعرف المتغير" السمة أو القدرة " بواسطة مجموعة من فقرات الاختبار أحادية البعد؛ أي أن فقرات الاختبار لا تختلف فيما بينها إلا من حيث مستوى الصعوبة فقط، كما يتحدد الأفراد ذوو قدرة معينة من مستوى أدائهم على الاختبار.

- استقلال الإجابة على فقرة معينة عن باقي فقرات الاختبار، والأفراد الذين يجيبون عليه؛ بمعنى أن إجابة الفرد على أي فقرة من فقرات الاختبار لا تؤثر على إجابته على الفقرات

الأخرى من الاختبار نفسه.

- توازى المنحنيات المميزة لخصائص الفقرات؛ بمعنى أن جميع فقرات الاختبار يكون لها قوة تمييزية متساوية (كاظم،1996).

ويعد نموذج راش من أهم نماذج نظرية الاستجابة على الفقرة؛ والـذي يتميـز بسـهولة استخدامه وموضوعيته، وهو من أكثر نماذج هذه النظرية استخداماً، ولا يتطلب استخدامه إجراء عمليات رياضية معقدة، فضلاً عن أنه يمكن تقدير احتمال الاستجابة الصحيحة من خلال معلمين فقط هما: صعوبة الفقرة، وقدرة الفرد؛ بغض النظر عن حجم العينة، وعـدد فقـرات الاختبـار (Hambelton & Traub, 1973)، ويستخدم النموذج في تحليل الفقرات التي درجتها إما واحد صحيح أو صفر، ويتمثل الافتراض الأساس للنموذج بأنه كلما زادت قدرة الفرد كانت الفرصـة أكبر للإجابة الصحيحة على أي فقرة من فقرات الاختبار، وأن احتمال الإجابة الصـحيحة علـى الفقرة الأسهل أكبر من احتمال الإجابة الصـحيحة علـى الفقـرة الأصـعب (Chang, et. al, 2010)؛ وهذا يعني أن الفرد الذي يجيب بصورة صحيحة على الفقرات الصعبة فـي الاختبـار سيجيب إجابة صحيحة على الفقرات الأسهل ، كذلك فإن الفرد ذا القدرة الأعلى سـوف يجيـب بصورة صحيحة عن عدد أكبر من الفقرات من الفرد ذي القدرة الأدنى(Lord, 1971).

وإذا ما تحقق مطابقة بيانات الاختبار لنموذج راش فإنه يعد من أفضل النماذج التي يمكن استخدامها في إيجاد معالم الفقرات والأفراد على اختبارات التحصيل المختلفة؛ حيـث إن انتقـاء فقرات الاختبار استناداً إلى النظرية الكلاسيكية في القياس يسبب أخطاء قياس؛ وتعود مصادر تلك الأخطاء تعود إما إلى الاختبار نفسه، أو إلى الأفراد، أو إلى الموقف الاختباري، أو إلى واضـع الاختبار، ومن هنا حاولت النظرية الحديثة في القياس(IRT) التغلب على مصادر الخطأ السابقة، وغيرها من المشكلات؛ بحيث يمكننا من انتقاء فقرات الاختبـار بمـا يتناسـب وقـدرة الفـرد المستجيب، وتتميز هذه الفقرات عندئذ بخصائص إحصائية مناسبة.

وقد كانت مشكلة انتقاء فقرات الاختبارات استناداً إلى نموذج راش محور بحث العديد من الباحثين، حيث قام دعنا (2005) بدراسة هدفت لبناء اختبار مكيف هرمي في الرياضيات للصف الثامن الأساسي وفقاً لنموذج راش، وتم بناء اختبارين من نوع الاختيار من متعدد وبأربعة بدائل لكل فقرة تقيس فقراتها المستويات الخمسة الأولى من تصنيف بلوم في المجال المعرفي، الأول

يتعلق بمحتوى الفصل الدراسي الأول، والثاني يتعلق بمحتوى الفصل الدراسي الثاني، وطبق الاختباران على عينة الدراسة التي اشتملت على (376) طالباً اختيروا بطريقة عشوائية ذات مرحلتين، وحللت الفقرات باستعمال برنامج Rascal ، وتم إجراء تكافؤ لصعوبة الفقرات لتكون متدرجة تدريجاً موحداً مشتركاً، بلغ عدد فقراته (62) اختير منها (55) فقرة استخدمت في بناء الاختبار المكيف؛ الذي تألف من (10) مستويات اختبارية مرتبة ترتيباً هرمياً وفق صعوبته، وبعدد فقرات مساوٍ لرقم المستوى الاختباري.

وفي دراسة (Koh, et.al 2006) والتي هدفت إلى التحقق من أن فقرات اختبار (ARAT) أحادية البعد باستخدام التحليل الإحصائي لنموذج راش، تم تطبيق الاختبار على (351) مريضاً في تايوان، وأظهرت النتائج فاعلية (19) فقرة من الاختبار لقياس كل المواضيع الخاصة بالاختبار؛ إضافة إلى دعم النتائج لافتراض أحادي البعد للفقرات السابقة، تم استخدام نموذج (Mokken) الذي اشتمل على التحليل؛ والذي أظهر عدم مطابقة (4) فقرات من الاختبار لنموذج راش.

وفي دراسة (Leonidas , et. al 2006) والتي هدفت إلى دراسة الملل الحاصل في المدارس، تم استخدام استبانة استخدمها كثير من الدراسات السابقة وتميزت بالصدق؛ وكذلك مقياس يتعلق بسلوكات الأفراد. وكان الهدف أيضاً من الدراسة فحص تصميم وصدق البناء والثبات للمقياسين، تكونت عينة الدراسة من (335) فرداً، (160) ذكراً متوسط أعمارهم(11.9)، و(175) أنثى متوسط أعمارهن(12.8)، الذكور، في (21) صف للمرحلة الابتدائية، تم تحليل النتائج باستخدام نموذج راش، وأظهرت النتائج أن كلا الاختبارين يتطابقان مع نموذج راش فيما يتعلق بالثبات والصدق، كما تم تحليل نتائج الاختبارين وبشكل منفصل وللعينتين (الذكور والإناث)، وأظهرت النتائج تمتع الاختبارين بصدق البناء والثبات، وهذا يدل على فاعلية نموذج راش لمطابقة البيانات مع عينة الدراسة.

وفي دراسة (Barbara, et. al, 2006) والتي هدفت لفحص الصدق والثبات لاختبار القدرة الذاتية (SEE) واختبار التوقعات للمخرجات التعليمية (OEE)، تكونت عينة الدراسة من (166) فرداً من الإناث، وتم تحليل النتائج باستخدام نموذج راش، والتحليل العاملي، وأظهرت نتائج التحليل أن بعض الفقرات تتميز بثبات مرتفع.

وهدفت دراسة (Worth , et. al, 2007) لتطوير طريقة جديدة لقياس مستوى القدرة لدى الأفراد (TBI) واستخدام نموذج راش للتوصل إلى القياس الخطي، أحادي البعد باستخدام برنامج (Win steps, 3.45) ، وأيضاً توقعات القدرة، وتم تطبيق اختبار (TBI) على (160) فرداً (52 أنثى، 111 ذكر)، أظهرت النتائج عودة (31%) من أفراد العينة للمدرسة أو العمل تطابق بيانات الاختبار مع نموذج راش.

وهدفت دراسة (Pallant, et. al, 2007) إلى إثبات وبرهنة استخدام التحليل الإحصائي لنموذج راش عن طريق تطبيق اختبار (HADS) لقياس القلق والإحباط، وطبق الاختبار على (296) مريض، وتم تحليل نتائج الاختبار باستخدام نموذج راش عن طريق برنامج (Rumm, 2020) لمطابقة النموذج مع البيانات، أظهرت النتائج دعماً واضحاً لاستخدام اختبار (HADS – 14) لقياس الضغوط النفسية، وأظهرت أيضاً مطابقة البيانات مع النموذج باستثناء بعض الفقرات والتي كانت غير مطابقة، وأظهرت البيانات زوجين من الاختبارات الفرعية للاختبار الكلي هما اختبار للقلق وآخر للإحباط.

وهدفت دراسة (Gill, et. al, 2007) لتطوير مقياس (Wis) في مجال التمريض، وتكونت عينة الدراسة من (48) فرداً؛ تم إجراء مقابلات فردية معهم للقيام بتطوير المقياس الخاص بعملية التمريض، تم استخدام نموذج راش لتحليل بيانات الاختبار، وأظهرت النتائج تمتع (30) فقرة من فقرات المقياس تتميز بالخصائص الإحصائية المناسبة المتمثلة في صعوبة الفقرات، وتمييزها، ومطابقتها نموذج راش .

وفي دراسة (Wilson, et. al, 2007) والتي هدفت لقياس المهارات الرياضية باستخدام اختبار موضوعي (اختيار من متعدد) تكونت عينة الدراسة من (566) طالب، معظم الطلبة تم إعطائهم مهارات في الجبر والحساب تعتمد على المهارات الرياضية، تم استخدام نموذج راش عند تحليل البيانات لاعتماد الاختبار إضافة إلى وصف مستويات الطلبة في المهارات الرياضية أظهرت النتائج أن درجات الطلبة على المقياس تعتمد على الأكبر سنا ومعرفة بالمهارات الرياضية إضافة إلى مستوياتهم السابقة بالقدرات الرياضية، مفهوم الذات والعمر، وأظهر النموذج تطابق في البيانات المجمعة من عينة الدراسة.

وفي دراسة (Peter, et. al, 2007) والتي هدفت لتقييم فقرات اختبار (ESS) الخاص

بالمرضى، البيانات من هذه الدراسة تم أخذها من دراسات أخرى سابقة أجريت على المرضى وعددهم (180) مريضاً، والبيانات المجمعة كانت جيدة، وتتمتع بثبات عالي قدره (0،84)، وأظهر نموذج راش المستخدم في التحليل تطابق في البيانات مع النموذج، وتم تحديد درجة في ضوئها يتم الحكم على أداء الفرد الخاضع للاختبار سواء أكانت درجة عليا أو دنيا باستخدام نموذج راش.

وفي دراسة (Chang, et. al, 2010) والتي هدفت إلى تحليل فقرات اختبار قرائي في تايوان خاص بالقبول في كلية الطب، تكونت عينة الدراسة من (36) شخص، وتكون الاختبار من (100) فقرة متخصصة، تم تحليل فقرات الاختبار باستخدام نموذج راش؛ وذلك بالطلب من كل فرد من عينة الدراسة تقييم مدى فاعلية وصعوبة كل فقرة من فقرات المقياس وتدريجها على مقياس لوغاريتمي، دلت نتائج الدراسة أن متوسط قدرة المفحوصين كانت أعلى من متوسط صعوبة فقرات المقياس، حيث دلت النتائج على وجود (60) فقرة من أصل (100) فقرة دون مستوى قدرة المفحوصين، وعدم فاعلية نموذج راش المستخدم في تحليل نتائج المقياس المستخدم.

وفي دراسة ادواردز والكوك (Edwards& alcock, 2010) والتي هدفت إلى تحليل نتائج اختبار للرياضيات تكون من (11) فقرة في بريطانيا، تكونت عينة الدراسة من (164) طالب وطالبة من طلبة السنة الأولى في المرحلة الجامعية، طلب منهم ترتيب فقرات الاختبار وفق مستوى صعوبة كل فقرة، تم تحليل نتائج الدراسة وفق نموذج راش لتقييم مدى مطابقة الفقرات لنموذج راش آحادي المعلمة، دلت النتائج على مطابقة فقرات الاختبار لستة أشخاص فقط من عينة الدراسة بعد إعادة ترتيب فقرات الاختبار وفق نموذج راش.

وفي دراسة فرانك جوني وآخرون(Franchignoni, et. al, 2011) والتي هدفت إلى تحليل فقرات اختبار داش (DASH) في ايطاليا؛ وفق النظرية الكلاسيكية في القياس ونموذج راش آحادي المعلمة، تكونت عينة الدراسة من (283) مريضاً، أظهرت نتائج الدراسة مطابقة عشرة فقرات لنموذج راش آحادي المعلمة، ودلت النتائج على فاعلية نموذج راش في مطابقة الفقرات بشكل أفضل من التحليل العاملي وفق النظرية الكلاسيكية في القياس.

ومن هنا يتضح أنه يمكن استخلاص بعض المعطيات من نتائج الدراسات السابقة فيما يتعلق بدقة تقدير المعالم في بعض نماذج نظرية الاستجابة للفقرة؛ حيث بينت بعض الدراسات أن

دقة تقدير المعالم في نماذج نظرية الاستجابة للفقرة تتأثر بطول الاختبار، وحجم عينة المفحوصين، وأن هناك عدم اتفاق في نتائج الدراسات المختلفة حول فعالية اختيار الفقرات وفق النظرية الكلاسيكية في القياس، أو وفق الاستجابة للفقرة، كدراسة (koh , 2006)، إضافة إلى أن دقة تقدير المعالم تتغير بشكل واضح تبعاً للطريقة المستخدمة في تقدير المعالم، والبرنامج الحاسوبي المستخدم في عملية التقدير، كما أن بعض الدراسات السابقة استندت على أسلوب المحاكاة (Simulation) لفحص فرضياتها حول دقة تقديرات المعالم، بيد أن هذه الدراسة الحالية ستسند إلى بيانات ميدانية حقيقية لفحص فرضياتها.

الطريقة والإجراءات:

مجتمع الدراسة:

تكون مجتمع الدراسة من طلبة المرحلة الأساسية في مديرية تربية السلط للعام الدراسي 2012/2011، موزعين على (28) مدرسة، بعدد (292) فصلاً دراسياً، والجدول (1) يبين توزيع أفراد مجتمع الدراسة حسب المرحلة الدراسية:

جدول 1. أعداد الطلبة حسب المرحلة الدراسية في مديرية تربية السلط

المرحلة	ذكور	إناث	المجموع
ثامن أساسي	5745	7459	13204
تاسع أساسي	5866	7104	12970
عاشر أساسي	5897	6509	12406
المجموع	17508	21072	38580

عينة الدراسة:

لاختيار عينة ممثلة لمجتمع الدراسة، قام الباحث بمراجعة قوائم المدارس للمرحلة الأساسية في مديرية تربية السلط للعام الدراسي 2012/2011، وتم اختيار مرحلة الثامن الأساسي كمرحلة دراسية لغرض الدراسة الحالية (لأن اختبار TIMSS فقراته وضعت لمستوى الصف الثامن الأساسي)، وقد تم اختيار (5) مدارس ذكور بشكل عشوائي من أصل (32) مدرسة، واختيار(5) مدارس إناث من أصل (70) مدرسة، وتم تطبيق الاختبار على الطلبة المتواجدين في تلك المدارس أثناء فترة التطبيق؛ والبالغ عددهم (599) طالباً وطالبة، وبواقع (311) طالبة،

و(288) طالباً.

إجراءات تطوير الاختبار:

تتلخص الإجراءات التي اتبعها الباحث في تطوير الاختبار وفق هذه الطريقة بالخطوات الآتية:

1- ترجمة فقرات اختبار (TIMSS) للرياضيات؛ والتأكد من ذلك عن طريق عرضها على متخصصين في الرياضيات.

2- اختيار الفقرات متعددة الإجابة واستبعاد الفقرات التي تحتاج إلى حل بأسلوب إنشائي أو الفقرات التي تحتوي على سلّم إجابة نعم أو لا، وذلك بسبب أن نموذج راش أحادي المعلمة يتعامل مع الاختيار من متعدد كنمط أسئلة يتم تحويل سلّم الإجابة إلى (صح أو خطأ) عند تصحيحها أي (واحد أو صفر).

3- عند كتابة فقرات الاختبار راعى الباحث أن تكون الفقرة (السؤال) واضحة لجميع المفحوصين.

وبعد إعداد الاختبار بصورته الأولية للتأكد من وضوح وملائمة صياغة الفقرات لغوياً، ومدى دقتها في قياس ما وضعت لقياسه، تم عرض الفقرات على (10) محكمين، حيث تكون الاختبار في صورته الأولى من (72) فقرة في صيغة اختيار من متعدد ، تحتوي كل فقرة على أربعة أو خمسة بدائل للإجابة، وفي ضوء أراء المحكمين تم تعديل فقرات الاختبار من حيث ملائمة الفقرة و صياغتها لغوياً ومدى ملائمتها للبيئة الأردنية، ثم قام الباحث بطباعة الفقرات في شكل اختبار أولي بغرض تطبيقه على العينة الاستطلاعية.

4- تجريب الفقرات: بعد أن أعد الاختبار ووضع في صورة ملائمة للتجريب، بدأت الخطوة الرابعة وهي خطوة تجريب فقرات الاختبار بهدف التعرف على خصائصها السيكومترية، أما مرحلة ما قبل التجريب فقد هدفت إلى تطبيق الاختبار بصورته الأولية على عينة صغيرة من المفحوصين بلغ عددهم (20) مفحوصاً، تم اختيارهم من نفس مجتمع الدراسة، وفضل الباحث أن تتم عملية التطبيق بشكل فردي لملاحظة كل ما يجري أثناء التطبيق، ولمعرفة الصعوبات التي يمكن أن تحدث أثناء عملية تطبيق الاختبار في الحصة الصفية؛ وذلك لوضع الحلول المناسبة،

وكان من أهداف هذه المرحلة التأكد من وضوح تعليمات الاختبار، ووضوح صياغة الفقرات، وفهمها من قبل المفحوصين، إضافة إلى جانب تحديد الوقت الكافي لإجراءات التطبيق.

أما مرحلة تجريب الفقرات والتي هدفت إلى التحقق من إحصائيات فقرات الاختبار فقد طبقت الصورة الأولى للمقياس على (60) طالب تم اختيارهم بشكل عشوائي من أربعة فصول دراسية لمرحلة الثامن الأساسي، وبعد أن تم تطبيق الاختبار على العينة المذكورة؛ جمعت البيانات وحللت باستخدام الرزمة الإحصائية للعلوم الاجتماعية SPSS، بعد ذلك تم تصحيح فقرات الاختبار، وحسبت معاملات تمييز الفقرات؛ والتي هي معاملات ارتباط درجة كل فقرة من فقرات الاختبار مع الدرجة الكلية للمقياس Item-Total Correlation، إضافة إلى حساب معامل صعوبة الفقرة؛ والذي هو عبارة عن عدد الإفراد الذين أجابوا على الفقرة بشكل صحيح مقسوماً على العدد الكلي للأفراد، وبناء عليه تم حذف وتقليل الفقرات التي قل معامل تميزها عن (30،.) تمشياً مع وجهات النظر السائدة في مجال القياس(Thorndike, 1982)، إضافة إلى مراعاة صعوبة الفقرة؛ حيث تم حذف الفقرات التي معامل صعوبتها اقل من (30.) أو أكثر من (80،.)،حيث تم استبعاد (32) فقرة من الاختبار، ليبقى الاختبار بصورته الأولية مكوناً من (35) فقرة.

تم تطبيق الاختبار بصورته المعدلة والذي يتكون من (35) فقرة على عينة تم اختيارها بشكل عشوائي من مرحلة الصف الثامن الأساسي؛ والذي بلغ عددها(599) طالباً وطالبة؛ من أجل اختبار افتراض أحادية البعد للبيانات، وقد حث الباحث أفراد الدراسة على الاستجابة بكل دقة وموضوعية على جميع فقرات الاختبار؛ مع تجنب إحداث أي تأثير من قبل الباحث على نمط إجابة الطلبة، ثم أتيحت الفرصة لهم للإجابة عن فقرات الاختبار التي استغرقت (50) دقيقة تقريباً.

تصحيح الاختبار:

بعد إجراء عملية تطبيق الاختبار على أفراد عينة الدراسة تم تصحيح فقرات الاختبار؛ وذلك بإعطاء قيم رقمية للإجابة الصحيحة (1) وقيمة (صفر) للإجابات الخطأ، وتمت إجراءات التصحيح بواسطة الحاسوب، وباستخدام برنامج SPSS وبعدها تم إدخال البيانات الخاصة بالإجابات إلى الحاسوب؛ وباستخدام برنامج (RUMM, 2010) لإيجاد المعالم للأفراد والفقرات.

الخصائص السيكومترية للمقياس:

أولا: الصدق:

تم التوصل إلى دلالات صدق الاختبار(TIMSS) بالطرق التالية:

1- تحقق للمقياس صدق المحتوى من خلال تحديد الفقرات، وماذا تقيس كل فقرة ؟ ومصادر الحصول على الفقرات؟ وطريقة صياغتها ؟ وطرق تحكيمها كما أوضحنا سابقاً.

2- الصدق بدلالة المحك الذي تم التحقق منها بحساب قيم معاملات الارتباط بين درجات المفحوصين على الاختبار (TIMSS)، ودرجات الطلبة في مادة الرياضيات لعينة الدراسة بأكملها؛ والبالغ عددها (599) طالبا وطالبة؛ حيث تم إيجاد معامل الارتباط بين درجات الطلبة على الاختبار ودرجاتهم في مادة الرياضيات؛ وبلغ قيمته(0.87).

ثانيا: الثبات:

تم إيجاد ثبات الاختبار بإيجاد معاملات الارتباط بين الأداء على الفقرة والاختبار ككل، باستخدام معامل كرونباخ ألفا؛ حيث تم تحليل فقرات الاختبار لجميع أفراد العينة، حيث بلغت قيمة معامل الثبات للاختبار باستخدام معامل كرونباخ ألفا (0.90).

نتائج الدراسة:

تمحورت أسئلة هذه الدراسة حول إيجاد معالم الفقرات للاختبار المقدم؛ إضافة إلى قياس مدى ملاءمة البيانات للنموذج آحادي المعلمة للاختبار، وقد تم تحليل البيانات باستخدام برنامج (RUMM. 2010).

وللإجابة على السؤال **الأول** والذي ينص على" ما درجة تحقيق بيانات اختبار TIMSS افتراضات النموذج أحادي المعلم (راش) لدى طلبة الصف الثامن الأساسي؟" تم إجراء التحليلات الإحصائية المناسبة من حيث مستوى صعوبة الفقرات، وقيمة الخطأ المعياري للفقرات، والجدول (2) يبين هذه التحليلات.

جدول(2) مستوى صعوبة الفقرات وقيمة الخطأ المعياري لاختبار الملائمة

الفقرة	مستوى الصعوبة	الخطأ المعياري	البواقي	درجات الحرية
1	0. 382	0.096	1.840	652.28
2	0.501	0.094	1.661	652.28
3	0.613	0.094	1.483	652.28
4	0. 700	0.096	1.551	652.28
5	0.681	0.096	0.932	652.28
6	0.553	0.094	1.801	652.28
7	0.583	0.095	1.665	652.28
8	0.581	0.094	2.508	652.28
9	0. 363	0.097	0.972	652.28
10	0. 543	0.097	1.435	652.28
11	0.411	0.096	2.410	652.28
12	0. 562	0.096	1.459	652.28
13	0. 531	0.095	2.751	652.28
14	0.311	0.095	0.869	652.28
15	0.455	0.122	1.846	652.28
16	0. 342	0.096	1.582	652.28
17	0.392	0.095	1.077	652.28
18	0. 782	0.097	1.449	652.28
19	0.441	0.097	0.030	652.28
20	0.647	0.100	1.899	652.28
21	0.706	0.094	2.544	652.28
22	0.364	0.095	1.713	652.28
23	0.319	0.097	2.649	652.28
24	0. 542	0.096	2.703	652.28
25	0.559	0.099	2.298	652.28
26	0. 611	0.096	1.911	652.28
27	0.589	0.095	1.529	652.28
28	0.621	0.095	0.982	652.28
29	0.600	0.094	0.624	652.28
30	0. 541	0.096	0.457	652.28
31	0.512	0.099	2.421	652.28
32	0.301	0.095	2.141	652.28
33	0.524	0.095	0.468	652.28
34	0.305	0.097	0.312	652.28
35	0.371	0.096	0.859	652.28
Mean	0.000	0.097		
Std Dev	0.308	0.004		

مستوى الدلالة الإحصائية ($0,05 \geq \alpha$)

يبين الجدول مستوى الصعوبة لجميع فقرات الاختبار؛ حيث تراوح مستوى صعوبة الفقرات ما بين (782 .0) للفقرة رقم(18)،إلى(0.305)للفقرة رقم(34)،مما يشير أن الاختبار بشـكل عـام يصنف على أنه متوسط الصعوبة.

وللإجابة على السؤال الثاني والذي ينص على" ما مؤشرات المطابقة الإحصائية لفقرات الاختبار المقدم TIMSS لافتراضات النموذج أحادي المعلم(راش)لدى طلبة الصف الثامن الأساسي؟ تم إيجاد قيم كآي تربيع، ومستوى الدلالة، والجدول(3) يوضح ذلك.

جدول 3. اختبار مطابقة الفقرات لنموذج راش

df درجات الحرية	Prob مستوى الدلالة	Chi Square قيم كاي تربيع	Item الفقرة
9	0.052000	378.947	1
9	0.030000	84.257	2
9	0.000000	19.529	3
9	0.011022	27.821	4
9	0.060423	30.092	5
9	0.055405	30.207	6
9	0,000000	25,762	7
9	0,057000	165,311	8
9	0,083202	31,961	9
9	0,055000	66,753	10
9	0,070001	21,734	11
9	0,678323	6,604	12
9	0,072506	17,422	13
9	0,065011	39,592	14
9	0,068221	382,679	15
9	0,072510	34,067	16
9	0,070414	19,619	17
9	0,090402	54,220	18
9	0,820222	47,300	19
9	0,070011	16,432	20
9	0,011696	26,493	21
9	0,060015	38,310	22
9	0, 048860	23,653	23
9	0,080101	58,204	24

Item الفقرة	Chi Square قيم كاي تربيع	Prob مستوى الدلالة	df درجات الحرية
25	17,662	0,080001	9
26	20,279	0,086266	9
27	18,076	0,054305	9
28	26,918	0,070000	9
29	28,484	0,000792	9
30	32,831	0,000143	9
31	50,241	0,000000	9
32	30,176	0,060409	9
33	34,090	0,064086	9
34	15,925	0,068474	9
35	13,023	0,016158	9

مستوى الدلالة الإحصائية ($\alpha \geq 0,05$)

يبين الجدول (3) قيم اختبار كآي تربيع ودلالتها الإحصائية عند مستوى دلالة ($\alpha \geq 0,05$) ؛ حيث نلاحظ أن (25) فقرة من أصل عدد فقرات الاختبار الكلي للاختبار غير دالة إحصائياً، مما يشير إلى مطابقتها لنموذج راش، وهذه الفقرات هي: 34,33,32,28,27,26,25,24,22,20,19,18,17,16,15,14,13,12,11,10,9,8,6 ،5,1)، أما بالنسبة للفقرات (2، 3، 4، 7، 21، 23، 29، 30، 31، 35) فقد دلت قيمة كآي تربيع لمطابقة البيانات للنموذج إلى وجود دلالة إحصائية عند مستوى دلالة ($\alpha \geq 0,05$) ؛ وهذا يشير إلى عدم مطابقة هذه الفقرات للنموذج أحادي المعلمة.

كما تم إيجاد مستوى الصعوبة والدلالة التمييزية وفق النظرية الكلاسيكية معبراً عنها بمعامل الارتباط الثنائي المتسلسل Point Biserial ؛ والجدول (4) يوضح مستويات الصعوبة والدلالة التمييزية لفقرات الاختبار.

جدول 4. مستوى صعوبة الفقرة حسب النظرية الكلاسيكية ومعامل ارتباط الفقرة بالدرجة الكلية (معامل الارتباط المتسلسل POINT BISERIAL)

الفقرات	الكلي	مستوى صعوبة الفقرة	معامل بايسيريال
1	318	0, 38	0,120
2	387	0,50	0,433
3	391	0,61	0,602
4	317	0, 70	0,701

الفقرات	الكلي	مستوى صعوبة الفقرة	معامل بايسيريال
5	325	0,68	0,643
6	393	0,55	0,459
7	349	0,58	0,621
8	407	0,58	0,257
9	300	0, 36	0,646
10	308	0, 54	0,556
11	323	0,41	0,670
12	332	0, 56	0,623
13	351	0, 53	0,664
14	360	0,31	0,581
15	46	0,45	0,184
16	329	0, 34	0,700
17	341	0,39	0,594
18	300	0, 78	0,760
19	298	0,44	0,570
20	259	0,64	0,650
21	399	0,70	0,641
22	365	0,36	0,599
23	292	0,31	0,662
24	320	0, 54	0,708
25	269	0,55	0,678
26	332	0, 61	0,644
27	357	0,58	0,604
28	345	0,62	0,599
29	393	0,60	0,562
30	318	0, 54	0,586
31	275	0,51	0,696
32	346	0,30	0,676
33	352	0,52	0,499
34	302	0,30	0,591
35	328	0,37	0,618

مستوى الدلالة الإحصائية ($\alpha \geq 0,05$)

نلاحظ من الجدول السابق أن مستويات صعوبة الفقرات تراوحت ما بين (0,30-0,78) ؛ مما يدلّل على أن فقرات الاختبار تميز بين الأفراد ذوي مستويات القدرة المتباينة.

وللإجابة على السؤال الثالث والذي ينص على" ما مستوى القدرة المتوقعة لدى طلبة

الصف الثامن الأساسي وفق نموذج راش أحادي المعلمة؟" تم تقدير مستوى القدرة المتوقعة لأفراد العينة وفق الجدول الآتي:

جدول 5. تقدير القدرة المتوقعة للأفراد والتكرارات حسب الدرجة الكلية على الاختبار

الدرجة الكلية	التكرار	التكرار التراكمي	النسبة التراكمية	التقدير	الخطأ المعياري
0	0	0	0,0	4,954	1,676
1	0	0	0,0	2, 28	1,014
2	4	4	0,6	2, 01	0,728
3	1	5	0,7	1,513	0,603
4	3	8	1,2	1, 50	0,530
5	13	21	3,1	1,581	0,482
6	27	48	7,2	1,551	0,447
7	23	71	10,6	1,782	0,421
8	60	131	19,6	1,581	0,399
9	53	184	27,5	2, 36	0,383
10	54	238	35,5	1, 643	0,370
11	43	281	41,9	1,611	0,360
12	38	319	47,6	1, 565	0,351
13	31	350	52,2	1, 634	0,344
14	7	357	53,3	2,311	0,338
15	9	366	54,6	1,454	0,333
16	1	367	54,8	2, 342	0,329
17	0	367	54,8	2,039	0,327
18	1	368	54,9	1, 682	0,325
19	1	369	55,1	2,441	0,324
20	1	370	55,2	1,747	0,324
21	1	371	55,4	1,706	0,324
22	1	372	55,5	2,464	0,326
23	0	372	55,5	2,419	0,328
24	0	372	55,5	1, 543	0,331
25	6	378	56,4	1,459	0,335
26	17	395	59,0	1, 313	0,341
27	14	409	61,0	1,389	0,347
28	39	448	66,9	1,020	0,355
29	36	484	72,2	1,230	0,364
30	35	519	77,5	1, 892	0,376
31	29	548	81,8	1,512	0,389
32	29	577	86,1	2,311	0,407
33	7	584	87,2	1,724	0,428
34	9	593	88,5	2,405	0,456
35	6	599	89,4	2,447	0,491

مستوى الدلالة الإحصائية ($\alpha \leq 0,05$)

مناقشة النتائج:

هدفت هذه الدراسة لتدريج اختبار تحصيلي (TIMSS) يقيس مستويات المهارات الرياضية الأساسية للصف الثامن الأساسي؛ بعد تدريج فقراته باستخدام نموذج راش، وتّم تحليل فقرات الاختبار بالطريقة الكلاسيكية في القياس، ثم تّم تحليل الفقرات بالصورة النهائية بالنظرية الحديثة في القياس (نموذج راش) عن طريق البرنامج الإحصائي(Rumm, 2010).

ويتضح من نتائج هذه الدراسة أن البيانات المستمدة من الاختبار تطابق نموذج راش كما في جدول(3)، كما أشارت النتائج على أن فقرات الاختبار صالحة لقياس القدرة الكامنة التي يهدف إليها الاختبار وهي القدرة الرياضية، وذلك من خلال مطابقة فقرات الاختبار لنموذج راش، رغم وجود عدد من الفقرات لم تطابق النموذج؛ وإن الاعتماد عليها يؤدي إلى تقديرات غير دقيقة لهذه القدرة؛ بمعنى زيادة الأخطاء المعيارية لهذه التقديرات.

والمتفحص لنتائج هذه الدراسة يلاحظ أن النتائج الخاصة بالإجابة على السؤال **الأول** والمتعلق بدرجة مطابقة الاختبار لنموذج راش، دلت على مستوى الصعوبة لجميع فقرات الاختبار، حيث تراوح مستوى صعوبة الفقرات ما بين (0,782) للفقرة (18)، إلى(30,5) للفقرة (34)، مما يشير أن الاختبار بشكل عام يصنف على أنه متوسط الصعوبة، هذه النتيجة تنطبق مع دراسة دعنا (2005)، وتدلّ على إمكانية التوصل إلى اختبارات متوسطة الصعوبة بغض النظر عن حجم العينة المستخدمة، كما نلاحظ أن(18) فقرة من أصل عدد فقرات الاختبار الكلي للاختبار غير دالة إحصائياً؛ لكن السبب في عدم مطابقة عدد أكبر من الفقرات لنموذج راش قد يعود إلى أكثر من سبب، أهمها وجود عامل التخمين في الإجابة على الفقرات التي هي من نوع الاختيار من متعدد، إضافة إلى أنّه قد تكون الفقرات غير ملائمة للنموذج الأحادي وملائمة أكثر للنموذج الثنائي أو الثلاثي نظراً لوجود التخمين، كما دلت عليه دراسة (Eyrec, et.al 2005) ، كما أن عدد الفقرات في الاختبار قد تكون سبباً في عدم مطابقة كثير من البيانات للنموذج؛ حيث دلّت عدد من الدراسات إلى أن استخدام عدد قليل من الفقرات يساهم في مطابقة البيانات لنموذج راش، كما أنه قد تختلف مطابقة البيانات للنموذج حسب البرنامج الإحصائي المستخدم؛ رغم مطابقة (18) فقرة من أصل (40) فقرة للإحصائي المستخدم لبيانات

الاختبار بصورة ممتازة، ولكن بصورة عامة فإن مطابقة بيانات الاختبار لنموذج راش في هـ ـذه الدراسة كانت جيدة، وهذا ما أشار إليه البرنامج الإحصائي المستخدم في عملية التحليل.

أما بالنسبة للسؤال **الثاني** والمتعلق بمؤشرات المطابقة الإحصائية لفقرات الاختبار من حيث صعوبة الفقرات، والدلالة التمييزية لها، فقد دلّت نتائج الدراسة تمتع فقرات الاختبار بمستوى صعوبة مناسبة تراوحت بين(0,38-0,61)، وهذا يدل على ملائمة فقرات الاختبار من حيث معامل الصعوبة لعينة الدراسة المطبق عليها الاختبار، وهذا ما دلت عليه دراسة جيل (Gill , 2007)، كما يلاحظ من خلال تفحص إحصائيات فقرات الاختبار تمتع هذه الفقرات بمستويات صعوبة جيدة، (انظر جدول (5,4)، وأن متوسط الفقرات(أي صعوبة الفقرة) وتباينها كان ينخفض كلما قل عدد فقرات الاختبار، وأن معامل ألفا المحسوب (95,0) لا يتأثر إلا بشكل طفيف جداً في حالة حذف أي فقرة من فقرات الاختبار؛ وكذلك بالمثل فأنه في حالة حذف أي فقرة من فقرات الاختبار؛ فإن متوسط الاختبار أو قيمة تباينه لا يتأثر إلا بدرجة طفيفة، أما بالنسبة لتقدير مستويات القدرة لأفراد الدراسة فقد دلّت النتائج على أن تكرارات الأفراد حسب الدرجة الكلية على الاختبار، وتقدير مستويات القدرة (Estimate)، والخطأ المعياري المرتبط بتقدير القدرة الحقيقية كانت مناسبة؛ حيث أن مستويات القدرة تتراوح ما بين(-3,33) للدرجة الكلية (2)،و(3,83) للدرجة الكلية (39) على الاختبار، وهذا مؤشر واضح يدل على توزيع أفراد الدراسة على سلّم مستويات القدرة ولجميع الأفراد، مما يدلل على مناسبة فقرات الاختبار المختارة لعينة الدراسة؛ إضافة إلى معرفة الفقرات المناسبة للأفراد الذين سيتم تطبيق فقرات الاختبار عليهم مستقبلاً، كوننا قد عرفنا مستوى القدرة للأفراد؛ بمعنى فصل مستوى القدرة لدى الأفراد عن الفقرات والعكس صحيح.

إن فقرات الاختبار المستخدمة والمنتقاة للصورة النهائية وفقاً لتحليل الفقرات عن طريق النظرية الكلاسيكية والنظرية الحديثة(نموذج راش) قد جاء منسجماً ومتناغماً مع المعايير الإحصائية المقبولة للنظريتين، وهذا مؤشراً على مناسبة فقرات الاختبار، والبرنامج الإحصائي المستخدم في عملية التحليل.

التوصيات:

يوصي الباحث القيام بالدراسات التالية:

- بناء مقاييس تقيس المهارات الرياضية، منسجماً مع المنهاج الدراسي، ولجميع المراحل الدراسية ولكلا الجنسين، وتحليل النتائج وفق مطابقة البيانات للنموذج المستخدم.
- إعادة تطبيق اختبار(TIMSS) على عينة مسحية تشمل جميع مناطق المملكة، وتحليل النتائج وفق نموذج راش .
- إعادة تطبيق اختبار(TIMSS)على عينات أخرى، وتحليل النتائج وفق النموذج الثنائي؛ أو الثلاثي المعلمة.
- إجراء دراسة مسحية بالاعتماد على اختبار(TIMSS)،ومقارنة النتائج وفق النظرية الكلاسيكية والنظرية الحديثة في القياس.

قائمة المراجع:

المراجع العربية:

1- دعنا، زينات (2005). بناء اختبار محبوك في الرياضيات للصف الثامن الأساسي وفق نموذج راش في نظرية السمة الكامنة، مجلة دراسات العلوم التربوية، 32، (1) الأردن.

2- علام، صلاح الدين محمود (2000). القياس والتقويم التربوي والنفسي: أساسيته وتطبيقاته وتوجهاته المعاصرة، دار الفكر العربي، القاهرة.

3- علام، صلاح الدين (1987)0 دراسة موازنة ناقدة لنماذج السمات الكامنة والنماذج الكلاسيكية في القياس النفسي والتربوي، جامعة الكويت، المجلة العربية للعلوم الإنسانية، 27(1): 18-44.

4- كاظم، أمينة (1996). دراسة نظرية نقدية حول القياس الموضوعي للسلوك، (أنور الشرقاوي وزملائه، اتجاهات معاصرة في القياس والتقويم النفسي والتربوي، القاهرة: مكتبة الانجلو المصرية.

المراجع الأجنبية:

- Barbara, K. Orwig, Denise, Zimmerman, Sheryl, Hawkes, William. (2006). Retrieved 23/4/2011 from: Brain Injury.

- Bond TG, Fox CM. (2007). Applying the Rasch model - Fundamental measurement in the human sciences. 2nd ed. Mahwah, NY: Lawrence Erlbaum Associates.
- Chang, K-Y. Tsou M-Y. Chan, K-H. Chang, S-H.Tai, J. J. Chen H-H. (2010). Item analysis for the written test of Taiwanese board certification examination in anaesthesiology using the Rasch model, British Journal of Anaesthesia, 104 (6): 717–2
- Crocker, L & Algina, J. (1986). Introduction to Classical & Modern Test Theory. New York: Holt, Pine hart & Winston.
- Edwards. H & alcock. L. (2010). Using Rasch analysis to identify uncharacteristic responses to undergraduate assessment, Teaching Mathematics and Its Applications, 29, 165- 175.
-Eyres, Sophie, (2005), Construct Validity and Reliability of the River Mead Post-Concussion Symptoms Questionnaire, Retrieved 20/6/2011 from: ebsco.html.
- Franchignoni, Franco, Ferriero, Giorgio, Giordano, Andrea, Sartorio, Francesco, Vercelli, Stefano, Brigatti, Elisa. (2011). Psychometric properties of QuickDASH - A classical test theory and Rasch analysis study, Manual Therapy, 16 (2011) 177-182.
- Hambelton, R, K. & Traub, R.E. (1973). Analysis of Empirical Data, N. Y.
- Hambleton R.K & Swaminathan. H. (1985). Item Response Theory Principles and Applications, Boston, Kluwer. Nijh off Publishing.
- Hambleton R.K, & Jones, R.W. (1993). Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development, NCME Items: The Instructional Topics in Educational Topics in Educational Measurement Series, Model 16.
- http: //nces. Ed. gov /timss/.
- Koh, Chia-Lin,Hsueh, I-Ping, Wang, Wen-Chung,Sheu, Ching-Fan,Yu, Tzu-Ying , Wang, Chun-Hou & Hsieh, Ching-Lin. (2006). Validation of the Action Research Arm Test Using Item Response Theory in Patients after Stroke, Retrieved 22/5/2011 from: Journal of Rehabilitation Medicine**.**
- Leonidas & Kyriakides, Kaloyirou, Chrystalla & Lindsay, Geoff. (2006). An Analysis of The Revised Olweus Bully/Victim Questionnaire Using The Rasch Measurement Model, Retrieved 30/4/2011 From: British Journal of Educational Psychology**.**
-Lord, F .M. (1971). Self Scoring of Flexi level Test, Journal of Educational

Measurement, 8:147-151.
- Mislevy, R, j, &Bock, R, D, (1990). PC-Bilog: Item Analysis and Scoring With Binary, Logistic Model Mooresville, IN, Scientific Software.
- Pallant, Julie F & Tennant, Alan. (2007). An Introduction to The Rasch Measurement Model: An Example Using The Hospital Anixter and Depression Scale (HADS). Retrieved 30/8/2011 from: British Journal of Clinical Psychology.
- Peter, H & Broman, Jan-Erik. (2007), Measurement Properties and Hierarchical Item Structure of The Epworth Sleepiness Scale in Parkinson's disease, retrieved 28/8/2008 From: Journal of Sleep Research.
- Thorndike, R .L (1982). Applied Psychometrics, Boston: Houghton-Mifflin.
- Wilson, T.M & Mac Gillivray, H.L. (2007), Counting on the Basics: Mathematical Skills among Tertiary Entrants, Retrieved 15/10/2011 from International Journal of Mathematical Education in Science & Technology.
- Wolfe EW, Smith Jr EV.(2007). Instrument development tools and activities for measure validation using rasch models: part II - validation activities. Journal of Applied Measurement; 8:204- 34.
Worth.IL,Gill,Bhakta,Bipin,Eyres,Sophie,Carey,AmyChamberlainM.Anne &Tennant,Alan.(2007). Development and Psychometric Testing of the Nurse-Work Instability Scale (Nurse-WIS), Retrieved 23/4/2011 from Journal of Advanced Nursing.