جامعة قناة السويس

كلية التربية بالسويس

بسم الله الرحمن الرحيم

الخصائص السيكومترية للاختبار وفقراته التي يتم انتقاؤها لبناء اختبار محكي المرجع لتقدير علامة المجال وفق النظرية الحديثة في القياس

إعداد

د/ عمر صالح مفضي بني ياسين

أستاذ مساعد (قياس وتقويم)

كلية عجلون الجامعية– جامعة البلقاء التطبيقية

مجلة كلية التربية بالسويس– المجلد الخامس– العدد الثالث– إبريل ٢٠١٢م

الخصائص السيكومترية للاختبار وفقراته التي يتم انتقاؤها لبناء اختبار محكي المرجع لتقدير علامة المجال وفق النظرية الحديثة في القياس

إعداد

د/ عمر صالح مفضي بني ياسين

أستاذ مساعد (قياس وتقويم)

كلية عجلون الجامعية– جامعة البلقاء التطبيقية

الملخص

هدفت هذه الدراسة إلى معرفة الخصائص السيكومترية للاختبار وفقراته التي تستخدم لبناء اختبار محكي المرجع لتقدير علامة المجال وفق النظرية الحديثة في القياس. ولتحقيق ذلك تم بناء اختبار تحصيلي في مادة الرياضيات للصف الثامن للفصل الدراسي الأول ٢٠١٠|٢٠١١، وتكون الاختبار من (٥٠) فقرة من نوع الاختبار من أربع بدائل. طبق على (٤٦٠) طالباً وطالبة موزعين على (١٢) شعبة دراسية تم اختيارها بالطريقة العشوائية العنقودية. أشارت نتائج الدراسة إلى مطابقة (41) فقرة من فقرات الاختبار لافتراضات نموذج راش، وتم حذف (٩) فقرات لأنها لم تطابق النموذج. وبلغ معامل ثبات الأفراد (0.96) ومعامل ثبات الاختبار(٠.٨٨). وتم التأكد من تمتع الاختبار بمظاهر متعددة من الصدق. كما أشارت النتائج إلى أن أفضل الفقرات التي ينبغي أن تختار لتقدير علامة المجال هي الفقرات متوسطة الصعوبة حيث تكون أكثر دقة في قياس تحصيل الطلبة، وتكون أخطاؤها المعيارية أقل ما يمكن،كما يمكن اختيار عدد قليل من الفقرات السهلة والصعبة لتغطية كامل المجال إذا لزم ذلك.

الكلمات المفتاحية: اختبار محكي المرجع، علامة المجال، الخصائص السيكومترية للاختبار وفقراته.

The psychometric characteristics of test and it's items which used to construct criterion – referenced test to estimate the domain score according to the modern theory of measurement

Dr: Omer Saleh Bani Yaseen

Division of Educational sciences

Ajloun College - Al.Balqa Applied University

Abstract

This study aimed to know the psychometric properties of the criterion – referenced test and its items which are used to estimate the domain score according to the modern theory of measurement. For this purpose, an achievement test was constructed in the mathematics of eight class of the first semester of2010-2011. the test consisted of (50) multiple choice items. The test was applied on (460) students distributed into (12) classes. The results revealed that (41) items fit rasch model assumptions, whereas (9) items where deleted because they didn't fit rasch model assumptions. The coefficients of reliability estimated for person and items were (.96) and (.88) respectively. The test validity was ensured through several approaches. The results also showed that the best items we choose to estimate the domain score are the items with intermediate difficulty because they are more precise to measure students achievement and also their standard errors are the least. also we can choose small number of easy and difficult items to represent all the domain if there is a need.

***Key words*:** Criterion- Referenced test, Domain Score, psychometric Characteristics.

مقدمة:

تعتبر الاختبارات أداة رئيسية من أدوات التقويم لهذا فإن الحديث عن الاختبارات يستلزم التركيز على نوعين أساسين منها هما: (popham, 1979, p:211).

١- الاختبارات معيارية المرجع:

وهذا النوع من الاختبارات يهتم في جوهرة في تصنيف الأفراد حسب مواقعهم النسبية بين أقرانهم في القدرات المختلفة، وتقدير الفروق للأداء، وأحياناً يطلق على هذا النوع من الاختبارات القياس جماعي المرجع لأنه يسعى إلى تقدير مستوى أداء الفرد بالنسبة لأداء مجموعة مرجعية أو معيارية في القدرة التي يقيسها الاختبار.

٢- الاختبارات محكية المرجع:

وهذا النوع من الاختبارات يهدف إلى مقارنة أداء الفرد بمستوى أداء مطلق ومحدد في مجال من السلوك، أو بمستوى كفاية معين، دون الحاجة إلى مقارنة أداء الفرد بأداء الآخرين، ولهذا يعتمد الاختبار محكي المرجع في جوهره على تقدير مستوى أداء الفرد بالنسبة لمحك أو مستوى إتقان محدد في مجال سلوكي محدد.

ولا يقتصر الاختلاف بين هذين النوعين من الاختبارات على هدف القياس فقط بل يتعدى ذلك إلى بناء الاختبار نفسه، فهما يختلفان من حيث اختيار الفقرات ومستويات صعوبتها وقدرتها على التمييز بين الأفراد، وكذلك من حيث شروط صدقها وثباتها ومعاييرها فأفضل الفقرات في الاختبارات معيارية المرجع هي الأقدر على التمييز بين الأفراد، وذات قيم صعوبة متوسطة، وأسوأ الفقرات هي التي لا تستطيع التمييز بين الأفراد أما في الاختبارات محكية المرجع فيراعى أن تمثل الفقرات المجال السلوكي خير تمثيل، وتساهم في تحديد مستوى الأداء أو الإتقان المتوقع من المفحوص (Crocker & Algina, 1986, p:193).

ويشير بوك ولورنس (Buck & Lawrence, 1978, p:312) إلى أن الاختبارات محكية المرجع تكون فعالة عند استخدامها في برامج التدريب، فهي وسليه جيدة للكشف عن نقاط القوة والضعف، وتمد واضعي المقررات والبرامج التعليمية بتغذية راجعة لتساعدهم في تعديلها وتطويرها.

ولذلك يمكن تقويم فاعلية انتقاء الاختبار محكي المرجع؛ أي فاعلية بناء هذا النوع من الاختبارات وفق معيارين هما (Hambleton & swaminathan, 1985, p:65):

١- إلى أي درجة يمكن أن تشكل العلامة على اختبار محكي المرجع تقديراً دقيقاً للعلامة التي يمكن أن تتحقق على مجال السلوكي بمجمله؟

٢- إلى أي درجة يمكن للاختبار محكي المرجع تصنيف المفحوصين عند نقطة تمثل مستوى الإتقان، كما يصنفهم اختبار المجال بمجمل فقراته عند نفس نقطة القطع؟

ويرى أصحاب النظرية الحديثة في القياس (نظرية الاستجابة للفقرة) أنها تقدم طريقة فعالة في بناء فقرات الاختبار محكي المرجع وذلك اعتماداً على كيفية المعلومات التي تشارك بها فقرة الاختبار في المعلومات الكلية للاختبار عن المفحوصين (item information function)، فعلى سبيل المثال فإنه من الممكن للفقرة أن تكون مطابقة تماماً للنموذج الثنائي أو الثلاثي المعلمة ولكن معلمة التمييز لهذه الفقرة تكون منخفضة إلى الدرجة التي تكون عديمة الفائدة في أي اختبار (Stocking, 1999, pp:322).

ومن أهم الملامح لهذه الطريقة أن مشاركة كل فقرة في دالة معلومات الاختبار يمكن أن تحدد بشكل مستقل عن باقي الفقرات في الاختبار (Lord, 1980, p:92).

تقدير علامة المجال السلوكي للاختبار محكي المرجع:

يقصد بعلامة المجال للاختبار محكي المرجع العلامة التي يحصل عليها الممتحن عند إجابته عن فقرات اختبار بحيث تمثل هذه الفقرات جميع جوانب المجال السلوكي لهذا الاختبار(Hambleton & swaminathan, 1985, p:61) ويعرف كروكروالجينا (Crocker, Algina, 1986, p:109) علامة المجال حسب النظرية الكلاسيكية في القياس بأنها العلامة الحقيقية للفرد. وبما أنه ليس ممكناً دائماً أن يمتحن المفحوص في كل الفقرات التي تمثل مجالاً معينا، خاصة إذا كان عددها كبيراً، فإنه يمكن سحب عينة من الفقرات عشوائياً من مجتمع الفقرات التي تمثل هذا المجال، ثم يمتحن المفحوص بها فتكون علامته على هذا الاختبار تقديراً لعلامته على المجال كله(Bock, thissen & Zimoski, 1997, p:186)، وفي النظرية الحديثة في القياس فإنه علامة المجال تقدر كما يأتي.

$$\hat{\pi} = \frac{1}{n}\sum_{i=1}^{n} ui$$

حيث $\hat{\pi}$: علامة المفحوص

Ui هي استجابة المفحوص عن الفقرة i ، وتأخذ العلامة صفر أو واحد.

وتعرف العلامة الحقيقة π بأنها القيمة المتوقعة ($\hat{\pi}$) E لعلامة المفحوص كما يأتي:

$$E(\hat{\pi}) = \pi = \frac{1}{n}\sum_{i=1}^{n} E(Ui)$$

وتربط نظرية الاستجابة للفقرة بين علامة المجال(العلامة الحقيقية) وبين قدرة المفحوص (θ) كما يأتي:
لمفحوص بقدرة (θ) فإنه يعبر عن علامة المجال له بالمعادلة الآتية:(Hambleton & Jons, 1993, p:38).

$$E(\pi)/\theta = \pi(\theta) = \frac{1}{n}\sum_{i=1}^{n} E(Ui/\theta)$$

فعندما يمثل المجال بعدد كبير من الفقرات يمكن أخذ مجموعة من الفقرات من تجمع الفقرات الكبير الممثل للمجال وتقديمها للمفحوص، وتكون علامة المفحوص على الاختبار تقديراً لعلامته على المجال، لأن تقدير قدرة الفرد حسب افتراضات نظرية الاستجابة للفقرة يكون مستقلاً إحصائياً عن عينة الفقرات، أي متحرر من تقدير صعوبة الفقرات التي اختبر بها المفحوص (Bock. Et al., 1997, p:192).

وهناك علاقة طردية بين علامة المجال، والقدرة (θ) بحيث يبدوان كأنهما شيء واحد، غير أن هناك اختلافان أساسيان بينهما هما:

١- إن علامة المجال تعرف على الفترة (1،0) بينما تعرف علامة القدرة على الفترة ($-\infty$ ، $+\infty$).

٢- إن تقدير قدرة الفرد مستقل عن عينة الفقرات المستعملة، بينما يعتمد تقدير علامة المجال على عينة الفقرات المستعملة (Hanson, 2002, p:221).

مشكلة الدراسة:

تهدف هذه الدراسة إلى التعرف على الخصائص السيكومترية للاختبار وفقراته التي يتم انتقاؤها لأغراض بناء اختبار محكي المرجع لتقدير علامة المجال (Domain score) ، وذلك من خلال الإجابة عن السؤالين الآتيين:

١- ما درجة مطابقة الاستجابة عن فقرات الاختبار لنموذج راش؟

٢- ما هي الخصائص السيكومترية للإختبار وفقراته التي يتم انتقاؤها لبناء اختبار محكي المرجع لتقدير علامة المجال وفق النظرية الحديثة في القياس؟

أهمية الدراسة:

لم يعد هدف التعليم مقتصراً على التمييز بين الطلاب في القدرة أو المستوى، وإنما يسعى إلى إكسابهم مهارات معينة، وتحقيق أهداف محددة، بل

وإتقان تلك المهارات. ولذا فإن على القائمين على التعليم بشكل عام، والمدرسة بشكل خاص تهيئة الظروف المناسبة التي تساعد على هذا النوع من التعليم، أو هذا النوع من الإتقان (Mastry Learning).

وقد جاءت هذه الدراسة لبيان كيف أن تقدير وتقييم الخصائص السيكومترية للاختبار وفقراته يساعد في بناء اختبارات محكية المرجع من خلال انتقاء فقرات لها خصائص سيكومترية محددة بهدف تقدير علامات الطلاب على مجال محدد من خلال انتقاء فقرات ممثلة لمجال سلوكي محدد، وبخصائص سيكومترية محددة، كما أن أهمية هذه الدراسة تنبع من أنها تساهم في تسليط الضوء على الاختبارات محكية المرجع، والتي لا تهتم بمقارنة أداء الطالب بأداء أقرانه، وإنما مقارنة أدائه بمستوى محدد من الإتقان، أو بتحقيق مجموعة من الأهداف الموضوعة مسبقاً.

مصطلحات الدراسة:

علامة المجال: هي العلامة التي يتوقع من الطالب تحصيلها على مجال دراسي محدد وهي تمثل قدرة الطالب في هذا المجال ويطلق عليها أحياناً العلامة الحقيقة.

(Hambleton & Swaminathan, 1985, p:61).

الاختبار محكي المرجع: هو الاختبار الذي يستخدم لتقدير أداء الطالب بالنسبة لمجال سلوكي محدد تحديداً مسبقاً أو هو الاختبار الذي يقارن فيه أداء الطالب مع مستوى محدد من الإتقان أو مع مدى تحقيقه لمجموعة من الأهداف.

الخصائص السيكومترية: يقصد بالخصائص السيكومترية معالم الصعوبة والتمييز والتخمين للفقرة والصدق والثبات للاختبار ،وفي هذه الدراسة فيقصد بها الصعوبة للفقرة والصدق والثبات للاختبار وذلك اعتمادا على نموذج راش

منهجية الدراسة وإجراءاتها:

مجتمع الدراسة:

تألف مجتمع الدراسة من طلاب وطالبات الصف الثامن الأساسي في المدارس التابعة لمديرية التربية والتعليم في عجلون، بلغ عددهم ٣٥٩٢ طالباً وطالبة موزعين على ٤٣ شعبة للذكور و ٤١ شعبة للإناث وذلك خلال الرجوع إلى التقرير الإحصائي التربوي لمديرية التربية .

عينة الدراسة:

تكونت عينة الدراسة من ٤٦٠ طالباً وطالبة موزعين على (١٢) شعبة من مدارس اساسيةللذكور والإناث حيث تم اختيار العينة بالطريقة العشوائية العنقودية والجدول(١) يبين توزيع أفراد عينة الدراسة على المدارس والشعب.

الجدول(١): توزيع أفراد المدارس وفق الجنس وأسماء المدارس والشعب

الجنس	المدرسة	عدد الشعب	عدد الطلبة
ذكور	صخرة الثانوية للذكور	٣	١٠٥
ذكور	عجلون الأساسية للذكور	٣	١٢٠
إناث	عجلون الأساسية للبنات	٣	١٣٢
إناث	الهاشمية الثانوية للبنات	٣	١٠٣
المجموع		١٢	٤٦٠

أداة الدراسة:

تم بناء اختبار محكي المرجع في مادة الرياضيات للصف الثامن الأساسي للفصل الدراسي الأول للعام الدراسي ٢٠١٠|٢٠١١ وفق الخطوات الآتية:

١- تم تحديد الغرض من الاختبار، وهو قياس تحصيل طلبة الصف الثامن في مادة الرياضيات.

٢- تم تحليل منهاج الرياضيات المقرر من وزارة التربية والتعليم الأردنية للصف الثامن للفصل الدراسي الأول، واشتقت منه الأهداف السلوكية كنتاجات تعلم يفترض أن تتحقق عند المتعلم في مستوى الإتقان التام. وقد

شملت عناصر البناء ألمفاهيمي لرياضيات الصف الثامن ، كما شملت القدرات الدنيا والعليا.

٣- تم إعداد جدول مواصفات يربط مستويات الأهداف بفئات المحتوى الدراسي موضوع الاختبار.

٤- عرض جدول المواصفات على ستة محكمين مختصين في الرياضيات؛ اثنان يعملان مشرفين تربويين ويحملان درجة الماجستير في أساليب تدريس الرياضيات، وثالث يحمل درجة دكتوراه في القياس والتقويم، وثلاثة يعملون مدرسين لمادة الرياضيات من مدة طويلة، وكان الهدف من الاحتكام إلى المحكمين إعطاء حكم على نتائج تحليل المحتوى والأهداف كما وردت في جدول المواصفات ومدى تغطيتها للمجال السلوكي المحدد في المنهاج والكتاب المقرر. وقد تم تعديل بعض الأهداف المصاغة بناء على رأي المحكمين واستقر عدد الأهداف السلوكية على(٢٠)هدف.

٥- تم كتابة فقرات عن كل هدف تم تحديده، واختلف عدد الفقرات حسب حجم المادة المقررة لكل هدف، وبلغ عدد هذه الفقرات (٥٤) فقرة وقد روعي أن تطابق الفقرة الهدف السلوكي الذي تقيسه من حيث المحتوى والمستوى المعرفي، كما روعي ألا تعتمد إجابة إحدى الفقرات على إجابة فقرة أخرى.

٦- للتأكد من صدق المحتوى تم عرض فقرات الاختبار على نفس المحكمين الستة ، وقد أخذت أحكام المحكمين حول الأمور الآتية:

أ. مدى التوافق بين الفقرة والهدف الذي يفترض أنها تقيسه.

ب. مدى وضوح الصياغة اللغوية للفقرة.

ج. مدى شمول الاختبار للمجال السلوكي.

وبناء على آراء هؤلاء المحكمين تم تعديل بعض الفقرات وحذفت أربع فقرات وبالتالي أصبح عدد الفقرات في صورتها الأولية (٥٠) فقرة أجمع المحكمون على أنها كافية وممثلة للمجال السلوكي الذي يقيسه الاختبار.

التجريب الأولي للاختبار:

تم تجريب الاختبار في صورته الأولية على شعبتين دراسيتين مؤلفة من (٧٨) طالبة من طالبات إحدى المدارس الأساسية من غير مجتمع الدراسة بغرض التعرف على مستوى صعوبة الفقرات وتمييزها والكشف عما يمكن أن يواجه الطلبة من صعوبات وكذلك إيجاد معامل الثبات ،ولم يكشف التجريب الأولى عن ملاحظات ذات أهمية فيما يخص وضوح الفقرات أو وجود صعوبات ومن ثم جمعت البيانات.

التحليل الإحصائي لبيانات التجريب الأولى:

استخدم برنامج التحليل الإحصائي (رزمة) SPSS لتحليل بيانات التجريب الأولى حيث تم حساب معامل الصعوبة لكل فقرة، وذلك بإيجاد نسبة الطالبات اللاتي أجبن عن الفقرة إجابة صحيحة من بين الطالبات اللاتي أجبن عن هذه الفقرات كما تم حساب معامل التمييز لكل فقرة، باستخدام معامل الارتباط الثنائي النقطي، والجدول (٢) يبين قيم الصعوبة والتمييز لفقرات الاختبار للعينة التجريبية. كما تم تقدير الثبات للاختبار بطريقة الاتساق الداخلي بمعادلة كرونباخ ألفا حيث بلغت قيمته ٠.٩٢ وتم حساب معامل الصدق بدلالة المحك التلازمي ممثلاً بمعامل الارتباط بين علامات الاختبار والتحصيل المدرسي في مادة الرياضيات لعينة التجريب الأولى حيث بلغ معامل الصدق ٠.٨٤.

جدول(٢): قيم الصعوبة والتمييز لفقرات الاختبار للعينة التجريبية

رقم الفقرة	معامل الصعوبة	معامل التمييز	رقم الفقرة	معامل الصعوبة	معامل التمييز
١	٠.٨٥	٠.٤٠	٣٠	0.45	0.36
٢	٠.٦٤	٠.٢٦	٣١	٠.٥٥	٠.٤٦
٣	٠.٤٣	٠.٣٥	٣٢	0.56	٠.٣٠
٤	0.42	٠.٢٨	٣٣	٠.٣٦	٠.١٢-
٥	٠.٧٩	٠.٣٧	٣٤	٠.٦٣	0.32
٦	٠.٣٢	٠.٢٤	٣٥	٠.٧٥	0.46
٧	0.61	٠.٣٨	٣٦	٠.٦٢	٠.٣٧
٨	0.56	0.60	٣٧	0.71	0.40
٩	٠.٩٢	٠.٤٦	٣٨	٠.٨٣	0.38
١٠	٠.٨٣	٠.٥٢	٣٩	٠.٦٠	٠.٤٠
١١	0.62	٠.٣١	٤٠	٠.٦٤	٠.٥٧
١٢	064	0.52	٤١	٠.٤١	٠.٣٠
١٣	0.71	٠.٣١	٤٢	٠.٨٠	٠.٤٤
١٤	0.94	٠.٤٣	٤٣	0.81	0.28
١٥	0.73	٠.٤١	٤٤	٠.٧٨	٠.٣٧
١٦	٠.٧٦	٠.٤٧	٤٥	٠.٥٥	٠.٤٦
١٧	٠.٧٥	٠.٤١	٤٦	٠.١٧	٠.٠٦-
١٨	0.56	٠.٦٥	٤٧	٠.٨٧	٠.٥١
١٩	٠.٦٤	٠.٤٤	٤٨	٠.٧٧	٠.٣٥
٢٠	٠.٧٣	٠.٤٦	٤٩	٠.٩٢	٠.٣٦
٢١	٠.٥٨	٠.٤٩	٥٠	٠.٦٦	٠.٤٠
٢٢	0.52	0.48	٥١	٠.٨١	٠.٤٦
٢٣	٠.٢٦	0.42	٥٢	٠.٦٦	٠.٤٦
٢٤	٠.٢٨	٠.٣٣	٥٣	٠.٦٠	٠.٤٥

٢٥	٠.٧٥	0.50	٥٤	0.58	0.54
٢٦	٠.٥٧	٠.٤٤			
٢٧	0.80	٠.٣٢			
٢٨	0.78	٠.٢٦			
٢٩	0.80	0.41			

وبناء على هذه التحليل تم حذف الفقرات الآتية(9، 14، 33، ٤٦) وذلك بسبب ارتفاع قيمة معامل صعوبتها ؛أي أن هذه الفقرات سهلة وهي الفقرتان ٩ و١٤ أو أن معامل التمييز لها سالب أو منخفض، وهي الفقرتان ٣٣ و٤٦. وبعد حذف هذه الفقرات أصبح عدد الفقرات بصورتها النهائية (٥٠) فقرة، وعند مراجعتها تبين أنها تمثل المحتوى المطلوب والأهداف السلوكية التي تغطي المجال السلوكي، وبعدها تم حساب معامل ثبات الاتساق الداخلي باستخدام معادلة كرونباخ ألفا فكانت قيمته ٠.٩٠ وهي قيمة مقبولة لأغراض هذه الدراسة.

إجراءات جمع البيانات:

بعد أن تم إعداد الاختبار بصورته النهائية ووضع التعليمات المناسبة له تم تطبيقه على ١٢شعبة دراسية في أربع مدارس (سبق بيانها في عينة الدراسة) ورغم وضوح تعليمات الاختبار، تمت الإجابة عن جميع استفسارات الطلاب والطالبات وتم تذكير الطلبة من التأكد من الإجابة عن جميع فقرات الاختبار، وبعد جمع

أوراق الإجابة تم تصحيح هذه الأوراق اعتماداً على مفتاح للإجابة سبق إعداده؛ حيث أعطيت الإجابة الصحيحة علامة واحدة والإجابة الخطأ صفراً.

عرض النتائج ومناقشتها:

أولاً: عرض النتائج المتعلقة بدرجة مطابقة الاستجابة عن فقرات الاختبار لنموذج راش (النموذج الأحادي المعلمة) ومناقشتها:

لتحديد مدى مطابقة الاستجابة عن فقرات الاختبار لنموذج راش، تم تصحيح استجابات الطلبة عن فقرات الاختبار بحيث أعطيت الإجابة الصحيحة للفقرة العلامة (١) والإجابة الخاطئة العلامة صفراً وبعد ذلك تم إدخال البيانات في ذاكرة جهاز الحاسوب وتم تحليل البيانات باستخدام برنامج (BIGSTEPS) والذي يحلل البيانات وفقاً للنموذج أحادي المعلمة (نموذج راش) وقبل عرض نتائج المطابقة لا بد من بيان النقاط الآتية والتي تتعلق ببرنامج BIGSTEPS (Linacre & Wright, 1993k p:34).

– يحذف البرنامج استجابات الأفراد الحاصلين على العلامة الكلية والأفراد الحاصلين على العلامة صفر، وكذلك يحذف الفقرات التي يجيب عنها جميع أفراد العينة إجابة صحيحة، أو تلك التي يخفق جميع أفراد العينة في الإجابة عنها،

وكذلك الأفراد الذين يزداد عدد استجاباتهم المتروكة كلما اتجها نحو نهاية الاختبار (والذي يعتبر مؤشراً على أن الوقت غير كاف).

– يحذف البرنامج الأفراد غير المطابقين مرة واحدة، وكذلك يحذف الفقرات غير المطابقة مرة واحدة؛ وذلك لأنه في كل مرة يعاد فيها التحليل يظهر هناك أفراد غير مطابقين وفقرات غير مطابقة.

– وبناء على ذلك لم يوجد فرد حصل على العلامة الكاملة أو أخفق في الإجابة عن جميع الفقرات، وكذلك لم توجد فقرة واحدة أجاب عنها إجابة صحيحة جميع أفراد العينة، وبناء على ذلك بلغ عدد أفراد العينة الذين تم تحليل استجاباتهم ٤٦٠ فرداً.

– قام البرنامج بتحديد الأفراد غير المطابقين للنموذج، ولمعرفة مؤشرات المطابقة الخاصة بالأفراد تم تقدير قدرة كل فرد بالإضافة الى الخطأ المعياري في قياس هذه القدرة وكذلك قيم إحصائي المطابقة الكلي (Tatal fit statistic) وله مؤشران:

المؤشر الأول هو مؤشر المطابقة الداخلية (التقاربي)
(In fit: the standardized information weighted fit statistics for person)

ويعبر عنه كذلك بإحصائي متوسط المربعات التقاربي (MNSQ) وهو مؤشر إحصائي للسلوكات غير المتوقعة والتي تؤثر في الاستجابات عن الفقرات التي تكون قريبة من مستوى قدرة الأفراد، وكذلك مؤشر المطابقة الخارجية (التباعدي (ZSTD)

the standardized information weighted fit statistics for person:) (out fit

ويعبر عنه كذلك بإحصائي متوسط المربعات التباعدي (MNSQ) وهذا المؤشر الإحصائي يعد بديلاً عن السابق، وله صفات مقاربة ومشابهة ولكنه أكثر حساسية للسلوكات غير المتوقعة من الأفراد عن الفقرات التي تبعد عن مستوى قدرة الأفراد، وذلك لكل تقدير من تقديرات القدرة. والجدول (٣) يبين المتوسط الحسابي والانحراف المعياري لكل من تقديرات القدرة والخطأ المعياري في قياس هذه القدرة، وقيم إحصائيات المطابقة الداخلية والخارجية،ومتوسطات المربعات للمطابقة الداخلية والخارجية.

جدول(٣): المتوسط الحسابي والانحراف المعياري لكل من تقديرات قدرات الأفراد والخطأ المعياري في القياس للقدرة وإحصاءات المطابقة الداخلية والخارجية (عدد الأفراد 460، عدد الفقرات 50)

	القدرة	الخطأ المعياري	إحصائي المطابقة الداخلية INFIT		إحصائي المطابقة الخارجية OUTDIT	
			متوسط المربعات NSQ	قيمة الإحصائي ZSTD	متوسط المربعات MNSQ	قيمة الإحصائي ZSTD
المتوسط الحسابي	٠.٩٩	٠.٣٧	٠.٩٧	٠.٠	١.٠٤	٠.١
الانحراف المعياري	١.٠٨	٠.٠٨	٠.١٧	٠.٩	٠.٥٨	١.٤

يتضح من الجدول (٣) اقتراب المتوسط الحسابي لمتوسطات المربعات الموزونة من الواحد وهو الوضع المثالي – كما يتوقعه النموذج – كما يلاحظ أن متوسط قيم إحصائي المطابقة الداخلية قد بلغ صفراً، والانحراف المعياري له ٠.٩، وهذه القيم تقترب من القيم المثالية التي يفترضها النموذج (١،٠) لكل من المتوسط الحسابي والانحراف المعياري على الترتيب، كما يلاحظ أن متوسط قيم إحصائي المطابقة الخارجية قد بلغ ٠.١ والانحراف المعياري له ١.٤، وهذه القيم تقترب من القيم المثالية التي يفترضها النموذج (صفر،١) على الترتيب.

وقد تبين وجود ٣٩ فرداً تبتعد استجاباتهم عن الاستجابات المتوقعة تبعاً لقدراتهم، أي غير مطابقين للنموذج ويعد الفرد غير مطابق للنموذج إذا كانت قيمة إحصائي المطابقة الخارجية له تزيد عن +٢ (Alastir & Hutchinson, 1987, p:79) ويعتبر هؤلاء الأفراد غير مطابقين للنموذج لأن استجاباتهم الملحوظة تبتعد

عن توقعات النموذج؛ وذلك كأن يجيب فرد ما فقرة إجابة صحيحة مع أن قدرته أدنى من صعوبة هذه الفقرة، والجدول (٤) يوضح أرقام الأفراد الذين تم استبعادهم نظراً لعدم مطابقتهم للنموذج.

جدول(٤): أرقام الأفراد غير المطابقين للنموذج وقيم متوسطات المربعات الموزونة وقيم لإحصائي المطابقة الخارجية المقابلة لقدراتهم

رقم الفرد	متوسط المربعات الموزونة	إحصائي المطابقة الكلي	رقم الفرد	متوسط المربعات الموزونة	إحصائي المطابقة الكلي	رقم الفرد	متوسط المربعات الموزونة	إحصائي المطابقة الكلي
٤٤٠	٥.٠٣	٥.٥	٤٤٢	٢.٤٥	٢.٣	٣٦٨	١.٦٧	٣.٠
٤٤١	٥.٠٣	٥.٨	٢٣٦	٢.٤٥	٤.٤	٢٩٨	١.٦٧	٢.٥
٣٩٠	٤.٨٥	٥.٧	٤٣١	٢.٤٢	٤.٨	٤١٣	١.٦٤	٢.٤
٢٢٠	٤.٤٦	٦.٣	٤٣٥	٢.٢٨	٤.٩	٢٤٢	١.٦٢	٢.٤
٢٧٧	٣.٩٦	٥.١	١٩٦	٢.١٣	٤.٦	٤٣٠	١.٦٠	٢.١
٤١٨	٣.٨١	٦.١	١٩٥	٢.٠٨	٢.١	١٥٧	١.٥٩	٢.٧
٣٩٦	٣.٥٧	٥.٨	٤١٢	٢.٠٢	٢.٤	١٥٥	١.٥٧	٢.٦
٤٠٨	٣.٤١	٥.٢	٢٣٧	٢.٠٢	٤.٠	٤٥٤	١.٥٠	٢.٣
٣٦٢	٣.١٦	٥.١	٣٩٥	١.٩٧	٢.٤	٤٢٦	١.٤٥	١.٧
١٦٨	٣.٠٢	٤.٩	١٠	١.٨٢	٢.١	٢٠٥	١.٤٢	٢.٠
٤١٤	٢.٧٤	٤.٦	٣٨٩	١.٧٧	٢.٧	٤٥٦	١.٣٢	١.٣
٤٦٠	٢.٦٧	٤.٧	٧٣	١.٧٦	٢.٧	٣٩٤	١.٢٧	٠.٩
١٦١	٢.٤٦	٤.٥	٤٢٤	١.٧٣	٢.٠	٤٣٦	١.٢٥	١.٠

بعد استبعاد الأفراد التسعة والثلاثين الذين لم تتطابق استجاباتهم مع توقعات النموذج وإعادة التحليل لاختبار مدى مطابقة الفقرات للنموذج، تم تقدير معلم الصعوبة لكل فقرة بالإضافة إلى الخطأ المعياري في قياس هذا المعلم وقيم إحصائي المطابقة الداخلية للفقرات Infit: (ZSTD) ومتوسط المربعات للفقرات (MNSQ) وكذلك قيم إحصائي المطابقة الخارجية للفقرات Infit: (ZSTD) ولإحصائي متوسط المربعات (MNSQ) لكل معلم من معالم الصعوبة ويبين الجدول ٥ المتوسط الحسابي والانحراف المعياري لكل من تقديرات معالم الصعوبة للفقرات والخطأ المعياري في قياس هذه الصعوبة، وكذلك قيم إحصائيات المطابقة الداخلية والخارجية لهذه المعالم.

جدول(٥): المتوسط الحسابي والانحراف المعياري لكل من معالم الصعوبة للفقرات والخطأ المعياري في القياس وإحصائيات المطابقة الداخلية والخارجية لهذه المعالم (عدد الأفراد = 421، عدد الفقرات= 50)

	القدرة	الخطأ المعياري	إحصائي المطابقة الداخلية INFIT		إحصائي المطابقة الخارجية OUTDIT	
			متوسط المربعات NSQ	قيمة الإحصائي ZSTD	متوسط المربعات MNSQ	قيمة الإحصائي ZSTD
المتوسط الحسابي	صفر	٠.١٣	١.٠	صفر	٠.٩٦	-٠.٤
الانحراف المعياري	٠.١٩	٠.٠٣	٠.١٦	٢.٦	٠.٣٠	٢.٨

يتبين من خلال الجدول ٥ اقتراب المتوسط الحسابي لمتوسطات المربعات الموزونة من الواحد،وهذا هو الوضع المثالي الذي يتوقعه النموذج، كما يلاحظ أن متوسط المطابقة الداخلية قد بلغ صفراً، وهذا وضع مثالي كذلك كما يفترضه النموذج، في حين بلغ الانحراف المعياري ٢.٦ وهو بعيد نوعاً ما عن القيمة المثالية التي يفترضها النموذج وهي الواحد.

وفي ضوء محكات المطابقة للفقرات المستخدمة تبين وجود تسع فقرات زادت قيم متوسط المربعات الموزونة لها عن الواحد الصحيح، وهذا مؤشر على أنها فقرات غير مطابقة وبعيدة عن توقعات النموذج وأنها قيم مضطربة ومشوشة للبيانات (Linacre & Wright, 1993, p:68). بالإضافة إلى أن فقرتين من هذه الفقرات التسع تم رفضها وذلك لأن قيم معامل التمييز (معامل الارتباط الثنائي النقطي) لهما جاءت سالبة؛ حيث أن القيم السالبة لمعامل التمييز تشير الى مطابقة سيئة أو تقدير في الاتجاه المعاكس هذه الفقرات هي: ٣٩، ٣٢، ٣٠، ٤٧، ١١، ٤٦، ٩، ١٩، ١٠ وقد تم حذف هذه الفقرات التسع لعدم مطابقتها لنموذج راش. بعد حذف الفقرات التي لم تطابق النموذج أحادي المعلمة (نموذج راش) أعيد التحليل للحصول على تقديرات متحررة من صعوبة الفقرات وقدرات الأفراد، والجدول ٦ يلخص نتائج التحليل للقيم المتحررة من قدرات الأفراد.

جدول(٦): نتائج التحليل للقيم المتحررة لقدرات الأفراد

(عدد الأفراد=421، عدد الفقرات=41)

	القدرة	الخطأ المعياري	المدى	الجذر التربيعي لمتوسط مربعات الأخطاء المعيارية في التقدير		الانحراف المعياري المعدل من الأخطاء		مؤشر الفصل		معد الثبات	
				الفعلي	المتوقع	الفعلي	المتوقع	الفعلي	المتوقع	الفعلي	المتوقع
المتوسط الحسابي	١.٢٤	٠.٤٦	(١.٤٦)- ٥.١٢-	٠.٥١	٠.٤٩	١.٢١	٢.٣٢	٢.٣٨	٢.٤٩	٠.٥٨	٠.٨٦
الانحراف المعياري	١.٣١	٠.١٧									
أقصى علامة خام	٤٦										
أدنى علامة خام	١٠										

من خلال الجدول٦ يلاحظ أن التقديرات النهائية المتحررة لقدرات الأفراد قد تراوحت بين العلامة الكلية ٤٦ كحد أقصى و ١٠ كحد أدنى، وقد بلغ متوسط توزيع القدرة ١.٢٤ وبمدى تراوح بين-١.٤٦ لوجيت للمستوى المنخفض من القدرة و٥.١٢ لوجيت للمستوى المرتفع من القدرة، وبلغ متوسط الأخطاء المعيارية لتقديرات قدرات الأفراد ٠.٤٦ وهذه قيمة متدنية، مما يشير إلى دقة تحديد مواقع الأفراد على متصل السمة. كما يلخص الجدول٧ نتائج التحليل للقيم المتحررة لصعوبة الفقرات.

جدول(٧): نتائج التحليل المتحررة لصعوبة الفقرات

(عدد الفقرات=41، عدد الأفراد= 421)

	القدرة	الخطأ المعياري	المدى	الجذر التربيعي لمتوسط مربعات الأخطاء المعيارية في التقدير		الانحراف المعياري المعدل من الأخطاء		مؤشر الفصل		معد الثبات	
				الفعلي	المتوقع	الفعلي	المتوقع	الفعلي	المتوقع	الفعلي	المتوقع
المتوسط الحسابي	صفر	٠.١٣	-١.٥٧ -٥.٣٨	٠.١٤	٠.١٤	١.٢٠	١.١٢	٨.٧١	٨.٨٦	٠.٩٩	٠.٩٩
الانحراف المعياري	١.٢١	٠.٠٣									

يلاحظ من الجدول٧ أن قيم التقديرات المتحررة لصعوبة الفقرات قد تراوحت من-١.٥٧ الى٥.٨٣ وحدة لوجيت، وتتوزع بمتوسط قدرة لوجيت وانحراف معياري مقداره ١.٢١ لوجيت، وقد بلغ متوسط الأخطاء المعيارية للفقرات ٠.١٣، وهذه قيمة متدنية مما يشير إلى دقة تقديرات الصعوبة للفقرات، وإذا قابلنا محوري تدريج الفقرات والقدرة نلاحظ وجود تطابق بينهما إلى حد ما، أي أنه يوجد نوع من الاتساق في تدرج صعوبة الفقرات، وأن الاختبار يقيس مدى واسعاً من القدرة ويقيس متغيراً أحادي البعد، وهذا يعكس صدق البناء للاختبار.

ثانيا: النتائج المتعلقة بالتحقق من أحادية البعد لفقرات الاختبار:

تم التحقق من أحادية البعد للاختبار (أداة الدراسة) بصعوبته النهائية بعدة مؤشرات؛ علماً أن أحادية البعد هو احد الافتراضات الأساسية لنظرية الاستجابة للفقرات التي يجب توفرها في البيانات، وقد تم الحصول عليها باستخدام التحليل العاملي(Factor analysis) باستخدام أسلوب تحليل المكونات الأساسية (Principal component analysis) لاستجابات الأفراد عن فقرات الاختبار بصورته النهائية، وقد تم حساب قيمة الجذر الكامن (Eigen value) ونسبة التباين المفسر (Explained variance) لكل عامل من العوامل، وقد جرى التدوير باستخدام طريقة التدوير المتعامد للعوامل التي كانت قيمة الجذر الكامن لها أكبر من واحد، وكان عددها ١٦ عاملاً فسرت ما نسبته ٥٦.٣٩% من التباين كما هو واضح من الجدول٨ والذي يبين قيمة الجذر الكامن ونسبة التباين المفسر لكل عامل من العوامل وكذلك نسبة التباين المفسر التراكمي المقابلة لكل عامل من العوامل.

جدول(٨): قيم الجذر الكامن ونسبة التباين المفسر

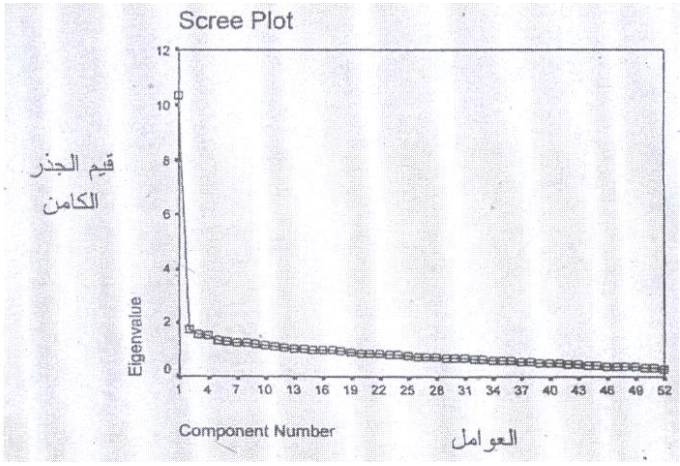
والتباين المفسر التراكمي المقابل لكل عامل في التحليل العاملي لفقرات الاختبار في استجابات ٤٢١ مفحوص

العامل	الجذر الكامن	نسبة التباين المفسر	نسبة التباين المفسر التراكمية
الأول	١٠.٣٧	١٩.٩٥	١٩.٩٥
الثاني	٠١.٧٦	٣.٣٨	٢٣.٣٣
الثالث	٠١.٥٧	٣.٠٢	٢٦.٣٤
الرابع	٠١.٥٣	٢.٩٥	٢٩.٢٩
الخامس	١.٣٩	٢.٦٧	٣١.٩٦
السادس	١.٣٣	٢.٥٦	٣٤.٥٢
السابع	١.٣٠	٢.٤٩	٣٧.٠٠
الثامن	١.٢٦	٢.٤٣	٣٩.٤٣
التاسع	١.٢٣	٢.٣٧	٤١.٨٠
العاشر	١.٢٠	٢.٣١	٤٤.١١
الحادي عشر	١.١٥	٢.٢٢	٤٦.٣٣
الثاني عشر	١.١١	٢.١٣	٤٨.٤٦
الثالث عشر	١.٠٤	٢.٠٠	٥٠.٤٦
الرابع عشر	١.٠٣	١.٩٩	٥٢.٤٦
الخامس عشر	١.٠٢	١.٩٧	٥٤.٤٢
السادس عشر	١.٠٠	١.٩٤	٥٦.٣٦

يتضح من الجدول(٨) أن هناك عاملاً واحداً فقط زادت قيمة الجذر الكامن له عن (٢) ويفسر ما نسبته ٢٠% من تباين علامات الأفراد على الصورة النهائية للاختبار؛ حيث بلغت قيمة الجذر الكامن لهذا العامل (١.٣٧) وهي قيمة مرتفعة

إذا ما قورنت مع قيم الجذور الكامنة لبقية العوامل الأخرى؛إذ أن نسبة الجذر الكامن للعامل الأول الى العامل الثاني كبيرة والمحك الآخر هو نسبة التباين المفسر للعامل الأول (٢٠%) في حين أن نسبة التباين المفسر لكل من العوامل المتبقية متقاربة جداً وصغيرة جداً مقارنة بالعامل الأول وهناك محك ثالث هو التمثيل البياني لقيم الجذور الكامنة للعوامل المكونة للاختبار بصورته النهائية أو ما يسمى باختبار الفرز (Scree Test)، كما في الشكل (٢) والذي يمثل بيانياً قيم الجذور الكامنة للعوامل المختلفة الناتجة من هذا التحليل، حيث أن هذا الشكل يثبت فرض أحادية البعد للاختبار.

الشكل: التمثيل البياني لقيم الجذور الكامنة للعوامل المكونة للاختبار.



ثالثاً: النتائج المتعلقة بالخصائص السيكومترية لاختبار محكي المرجع وفقراته التي تستخدم لتقدير علامة المجال وفق النظرية الحديثة في القياس ومناقشتها.

لمعرفة الخصائص السيكومترية للاختبار وفقراته التي تستخدم لبناء اختبار محكي المرجع لتقدير علامة المجال وفق النظرية الحديثة في القياس تم تحليل بيانات الدراسة باستخدام برنامج BIGSTEDPS وفيما يأتي عرض لهذه النتائج:

أ. النتائج المتعلقة بالخصائص السيكومترية للاختبار (الصدق والثبات):

تم التحقق من دلالات الصدق لاختبار المجال (مجتمع الفقرات) بطريقتين:

الطريقة الأولى: اعتمدت هذه الطريقة على الإجراءات التي اتبعت في تحديد مواصفات مجال التحصيل وتحليل الأهداف السلوكية والمحتوى وكفاية فقرات التي تمثل الأهداف كناتج تعليمي يراد تحقيقه والوصول به إلى مستوى إتقان تم تحديده مسبقاً، والتأكيد من صحة هذه الإجراءات من خلال الأحكام التي قدمها المختصون من ملائمة نتائج التحليل لمواصفات مجال التحصيل والتناظر بين محتوى الفقرات والأهداف التي أعدت لتمثيلها وقد سبق تفصيل ذلك عند الحديث عن أداة الدراسة.

الطريقة الثانية: (صدق المحك التلازمي) تم التحقق من دلالات الصدق لاختبار المجال وذلك بإيجاد معامل الارتباط بين العلامات المتحققة على جميع الفقرات (اختبار المجال) وبين علامات تحصيل طلبة عينة الدراسة في مادة

الرياضيات ممثلة بالعلامات التي حصلوا عليها في نهاية الفصل الدراسي الأول للعام الدراسي ٢٠١٠|٢٠١١ وقد حسب معامل الارتباط لجميع الطلبة الذين طبق عليهم اختبار المجال وعددهم (٤٦٠) طالباً وطالبة، وقد بلغت قيمة هذه المعامل ٠.٨٧ وهذه قيمة ومقبولة لأغراض هذه الدراسة.

ثبات الاختبار: بعد أن تم التحقق من مطابقة فقرات الاختبار لنموذج راش، واستبعدت الفقرات التي لم تحقق هذه المطابقة، وتم تحديد القيم التقديرية لكل من صعوبة الفقرات وقدرات الأفراد لا بد من الحصول على نوعين من معاملات الثبات: أحدهما يتعلق بثبات الأفراد والآخر يتعلق بثبات الفقرات، فمفهوم الثبات في النظرية الحديثة في القياس يشير إلى مدى الدقة في تقدير موقع كل من الفقرات والأفراد على متصل السمة ويمكن تحديد مدى دقة موقع الفقرات بإيجاد النسبة بين الانحراف المعياري للقيم التقديرية المتحررة من صعوبة الفقرات، ومتوسط الخطأ المعياري لهذه القيم، وتسمى هذه النسبة معامل الفصل بين الفقرات (Gi) (separation item index) والذي يجب أن تكون قيمته أكبر ما يمكن (أكبر من ٢)، فإذا لم تزد هذه القيمة عن(٢) فإنه يصعب قياس المتغير بواسطة هذه الفقرات، وبناء على التحليل الذي تم باستخدام برنامج BIGSTEDPS بلغ مؤشر الفصل بين الفقرات٨.٨٦ ومؤشر الفصل بين الأفراد ٢.٨٩، بمعنى أن معامل ثبات الاختبار بلغ ٠.٨٨،

ومعامل ثبات الأفراد ٠.٩٦، وهذه قيم مرتفعة جداً، وتشير إلى كفاية فقرات الاختبار في الفصل بين الأفراد، والتمييز بين مستويات القدرة المختلفة لديهم، وتشير كذلك إلى كفاية عينة الأفراد في الفصل بين الفقرات؛ وبمعنى أوضح فإن الأفراد يتوزعون بشكل مناسب على متصل السمة الممثل بفقرات الاختبار التي بنيت وفق نموذج راش.

ب. النتائج المتعلقة بالخصائص السيكومترية لفقرات الاختبار ومناقشتها:

لمعرفة قيم معالم فقرات الاختبار بصورته النهائية حسب النظرية الحديثة في القياس (نظرية الاستجابة للفقرة) اعتماداً على نموذج راش والذي يفترض أساساً أن الفقرات المطابقة للنموذج هي فقرات متقاربة في التمييز والتخمين فيها أقل ما يمكن(Hambleton & Swaminathan, 1985, P:45) تم تحليل استجابات أفراد عينة الدراسة، والبالغ عددهم (٤٢١) فرداً، وذلك بعد حذف الأفراد غير المطابقين للنموذج، حيث بلغ عدد الفقرات الاختبار في صورته النهائية (٤١) فقرة، بعد حذف (٩) فقرات لم تكن مطابقة للنموذج، وقد تم استخدام برنامج BIGSTEDPS في التحليل حيث تم تقدير صعوبات الفقرات والخطأ المعياري في تقدير هذه الصعوبات، علماً أن هذا البرنامج يقدر صعوبات الفقرات وقدرات الأفراد بطريقتين متتابعتين هما: طريقة (PROX) التقريبية أولاً ثم طريقة الأرجحية العظمى غير المشروطة (UCON) (Wright & stone, 1979, p:85) ويوضح الجدول(٩) قيم صعوبات

الفقرات والخطأ المعياري في تقدير هذه الصعوبات لكل فقرة من فقرات اختبار المجال وفق نموذج راش.

جدول(٩) قيم صعوبات الفقرات وأخطاؤها المعيارية مقدرة باللوجيت مرتبة تنازلياً حسب الصعوبة لفقرات اختبار المجال بصورته النهائية (٤١ فقرة) وفق نموذج راش

رقم الفقرة	الصعوبة الفقرة	الخطأ المعياري	رقم الفقرة	الصعوبة الفقرة	الخطأ المعياري	رقم الفقرة	الصعوبة الفقرة	الخطأ المعياري
٢٩	٥.٨٣	٠.٣٣	٣٦	٠.٠٥	٠.١٢	١	-٠.٥٨	٠.١٣
٥٠	٢.٣٤	٠.١٢	٤٠	٠.٠٤	٠.١٢	٣٢	-٠.٦١	٠.١٣
٧	٢.٠٨	٠.١٢	٢٧	-٠.٠٢	٠.١٢	٤٩	-٠.٦٤	٠.١٣
٣٧	١.٨٠	٠.١١	١٤	-٠.٠٦	٠.١٢	٢٤	-٠.٦٦	٠.١٣
٢٢	٠.٦٦	٠.١١	٢١	-٠.١٣	٠.١٢	١٢	-٠.٨٥	٠.١٤
٣٩	٠.٥٢	٠.١١	٣٤	-٠.١٣	٠.١٢	٣٥	-٠.٩٧	٠.١٤
٤٣	٠.٤٨	٠.١١	٤٥	-٠.١٥	٠.١٢	٤١	-٠.٩٩	٠.١٤
٨	٠.٣٩	٠.١١	٤	-٠.١٩	٠.١٢	٢٥	-١.٠١	٠.١٤
٣٨	٠.٣٨	٠.١١	١٥	-٠.٢٤	٠.١٢	٢	-١.٠٥	٠.١٥
٦	٠.٣١	٠.١٢	٥	-٠.٢٧	٠.١٢	١٧	-١.٢٣	٠.١٥
٢٣	٠.٣١	٠.١٢	٣٣	-٠.٢٧	٠.١٢	١٣	-١.٣٢	٠.١٦
١٦	٠.٢٠	٠.١٢	٤٢	-٠.٢٨	٠.١٣	٤٨	-١.٣٥	٠.١٦
٢٠	٠.١٩	٠.١٢	٤٤	-٠.٣٩	٠.١٣	٢٦	-١.٥٧	٠.١٤
٢٨	٠.١٥	٠.١٢	٣١	-٠.٤١	٠.١٣			
٣	٠.٠٧	٠.١٢	١٨	-٠.٤٣				
المتوسط	صفر	٠.١٣						
الانحراف المعياري	١.٢١	٠.٠٣						

يلاحظ من الجدول(٩) أن مدى صعوبة الفقرات لاختبار المجال يمتد من (-٠.٥٧) لوجيت إلى (٥.٨٣) لوجيت وبمتوسط حسابي قدره صفر لوجيت، وبخطأ معياري تراوح بين (٠.١١) و(٠.٣٣) بمتوسط قدرة (٠.١٣) فالعامل الرئيسي والمباشر الذي يؤثر في قيمة الخطأ المعياري هو حجم العينة، وفي العموم فإن زيادة حجم العينة سيقلل من قيمة الخطأ المعياري (Hambleton & Swaminathan, 1991, P:44) وأن أعلى قيمة لمعامل الصعوبة كانت للفقرة ٢٩ وإن أقل قيمة لمعامل الصعوبة كانت للفقرة ٢٦، وهذا وقد أشارت جمحاوي (٢٠٠٠) إلى أن الفقرات التي يزيد معامل صعوبتها عن (+١.٥٠) لوجيت تعد من الفقرات الصعبة، والفقرات التي يقل معامل صعوبتها عن (-١.٥٠) لوجيت تعد ضمن الفقرات السهلة أما الفقرات التي تقع قيم معامل صعوبتها بين (-.٥٠) و(+١.٥٠) لوجيت فإنها تعد ضمن المدى المتوسط لمعاملات الصعوبة، وبناء على ذلك فقد عدت الفقرة ٢٦ سهلة لأن معامل صعوبتها (-١.٥٧) لوجيت وأن الفقرات ٧، ٢٩، ٣٧، ٥٠ عدت صعبة لأن معامل صعوبتها زاد عن (+١.٥٠) لوجيت، وأما بقية الفقرات فإنها فقرات متوسطة الصعوبة لأن قيم صعوبتها تقع ضمن المدى المتوسط لمعاملات الصعوبة أي أن الفقرات التي تستخدم لتقدير علامة المجال ينبغي أن تكون بالمجمل متوسطة الصعوبة، كما أوضحت النتائج

أن قيم الأخطاء المعيارية للفقرات متوسطة الصعوبة كانت أقل منها للفقرات الصعبة أو السهلة، وهذا يدل على أن الفقرات متوسطة الصعوبة تكون أكثر دقة في قياس تحصيل الطلبة وتقدير علامة المجال بمعنى أن الفقرات متوسطة الصعوبة تقدم معلومات عن تحصيل الطلبة أكثر مما تقدمه الفقرات السهلة أو الفقرات الصعبة وهذا يتفق مع نتائج دراسة هالادياوروبد (Haladyna & Roid, 1983) ودراسة جمحاوي(٢٠٠١).

ويلاحظ أن هناك اتساقاً في تدرج صعوبة الفقرات ضمن المدى(-١.٥٧) إلى (٢.٣٤)لوجيت، وهذا يعني أن فقرات الاختبار تقيس سمة أحادية البعد، وهذا يؤكد ما جاء في التحليل العاملي من أن الاختبار يقيس سمة واحدة غير أن الفقرة ٢٩ متطرفة في صعوبتها، ويوجد مجال لإضافة فقرات صعبة بينها وبين الفقرة ٥٠ ذات معامل الصعوبة(٢.٣٤) لوجيت تعرّف المستويات العليا للسمة المقاسة.

والآن وبعد أن مرت عملية تكوين الاختبار بهذه المراحل من استبعاد للفقرات غير المطابقة للنموذج آحادي المعلمة (نموذج راش)، وبعد تحديد الخصائص السيكومترية للاختبار وفقراته، فإن هذا الاختبار والمؤلف من(٤١) فقرة يمثل مجال الرياضيات للصف الثامن الأساسي للفصل الدراسي الأول بشكل متوازن (متناسب) مع الأهداف، وبالتالي

يمكن استخدام هذا الاختبار لتقدير علامة المجال للطالب في الصف الثامن الأساسي.

المراجـع

١- جمحـاوي، إينـاس، (٢٠٠٠)، **مقارنـة خصـائص الفقـرات وفـق النظريـة التقليديـة ونظريـة اسـتجابة الفقـرة فـي مقيـاس القـدرة الرياضـية**، رسـالة ماجسـتير غيـر منشورة، جامعة اليرموك، إربد، الأردن.

2- Alastair, p. Hutchinson, G .(1987). **Calibrating graded assessment: Rasch partial credit analysis of performance in writing, language testing,** 4, pp73-92 .

3- Buck & Lawrence .(1978**). Is the federal government jumping on the criterion referenced testing bandwangan**. paper presented at the annual Meeting of the North eastern educational research association.

4- Crocker, L & Algian, J. (1986). **Introduction to classical and modern test theory**, clarification: cole publishing company.

5- Bock, R. D & Zimowski, M.F .(1997). IRT Estimation of domain scores. **Journal of educational Measurement**. Vol 34, 3, pp 197- 211.

6- Hanson, B.A; Beguin, A.A .(2002). Obtaining a common scale for item response theory item parameters using separate versus concurrent estimation in the common -item equating design, **Applied psychological measurement, vol 26, issel, 22.p**

7- Haldyna, T& Roid, G. (1983). A comparison of two approaches to criterion- referenced tests construction. **Journal of education measurement, 20, 271-282.**

8- Hambleton, R.K. & jonse. R.W. (1993). Comparison of classical test theory and item response theory and their applications to test development. **Educational measurement.** Issues and practice 4, 38-47.

9- Hambleton, R. K & Swaminathan, H. (1985). **Item response theory principles and applications.** Boston; Nijhoff publishing .

10- Linacre. J. M & wrigh, B. D.(1993). **Au ser's guide to BIGSTEPS. Computer program.** Chicago: MESA press .

11- Popham, w. j.(1980). **Modern educational measurement.** Englewood cliffs. NJ: Prentice- Hall, Inc.

12- Stocking. M. (1999). **Item response theory.** In. G. Masters, & J. Keevs (Eda). Advances in Measurement un educational research and Assessment (pp.322) the Boulevard, VK: Pergamon.

13- Wright. D. & stone, M. (1979). **Best test design: A handbook for research measurement,** Chicago: MESA press.