أثر نسبة فقرات اختبار الجذع المشترك في دقة المعادلة العمودية باستخدام نموذج راش لاختبار في الرياضيات

د. صبري حسن الطراونة
قسم علم النفس
كلية العلوم التربوية – جامعة مؤتة
sabri_razan@yahoo.com

أثر نسبة فقرات اختبار الجذع المشترك في دقة المعادلة العمودية باستخدام نموذج راش لاختبار في الرياضيات

د. صبري حسن الطراونة
قسم علم النفس
كلية العلوم التربوية – جامعة مؤتة

الملخص

هدفت الدراسة إلى دراسة أثر نسبة فقرات اختبار الجذع المشترك على دقة المعادلة العمودية باستخدام نموذج راش لاختبار في الرياضيات، تم بناء ثلاث صور من اختبار مكون من مستويين في الهندسة للصفين الخامس والسادس الأساسيين، كل مستوى مكون من (٢٠) فقرة، وتضمن كل اختبار فقرات مشتركة بين المستويين، الصورة الأولى تضمنت (٣) فقرات مشتركة، والثانية تضمنت (٦) فقرات مشتركة، والثالثة تضمنت (٩) فقرات مشتركة بين المستويين، وتكونت عينة الدراسة من (١٠٣١) طالباً وطالبةً موزعين على الصفين، تم التحقق من صدق الاختبار وثبات الاختبار، واستخدمت طريقة المعادلة العمودية باستخدام تصميم المجموعات العشوائية باختبار الجذع المشترك، باستخدام نموذج راش، تم استخدام محكي الصدق التقاطعي والخطأ المعياري للحكم على دقة المعادلة، أظهرت النتائج أن دقة المعادلة العمودية تزداد بزيادة نسبة فقرات اختبار الجذع المشترك، وأفضل نسبة هي (٤٥٪) من فقرات اختبار الجذع المشترك.

الكلمات المفتاحية: اختبار الجذع المشترك، المعادلة العمودية، نموذج راش والاختبار المتعددة المستويات.

* تاريخ تسلم البحث: ٢٠١٧/٩/٢٥م
* تاريخ قبوله للنشر: ٢٠١٨/١/١٦م

The Effect of Percent of Items in Anchor Test on the Accuracy of Vertical Equating Using Rasch Model for Mathematical Test

Dr. Sabri H. AL-Tarawneh
The College of Educational Sciences
Mutah University

Abstract

The study aimed at studying the effect of Percent of items in Anchor test on the accuracy of vertical equating using Rasch model for Mathematical Test. Three models of a two-level test in the Geometry subject were developed for fifth and sixth grades, each level consisted of (20), The first model included (3) common items, the second model included (6) common items, the third model included (9) common items between the two levels. The sample of the study consisted of (1031) students and the validity and reliability of the test were verified and the obtained results showed acceptable validity and reliability. The vertical equating was used by anchor test randomize groups using Rasch Model and Cross-validation and standard error were also used. The results showed that the accuracy of vertical equating increases by the increase of the anchor test items as the best obtained percentage was (45%) from the Items of Anchor Test.

Keywords: anchor test, vertical equating, rasch model, multi-ievel tests.

أثر نسبة فقرات اختبار الجذع المشترك على دقة المعادلة العمودية باستخدام نموذج راش لاختبار في الرياضيات

د. صبري حسن الطراونة
قسم علم النفس
كلية العلوم التربوية – جامعة مؤتة

المقدمـة

يمكن أن يأخذ الأفراد من المراحل العمرية أو الدراسية المختلفة مواقع ترتيبية على متصل نمائي من خلال أدائهم على الاختبارات المتعددة المستويات، والتي تقيس قدرات نمائية أو مجالات سلوكية تتضمن موضوعات دراسية يكون فيها التحصيل تراكمياً، لوجود تغيرات كمية تتماشى مع الارتقاء في المستوى العمري، ويوجد العديد من المباحث التي لها طبيعة تراكمية، مثل اللغة العربية واللغة الإنجليزية والرياضيات والعلوم (الطراونة، ٢٠٠٥).

ويمكن مقارنة اداء المفحوصين بهذه الاختبارات من خلال أحد تطبيقات نظرية استجابة الفقرة، وهي معادلة الاختبارات، ويتم من خلال معادلة الاختبارات تحويل العلامات على صورة من صور الاختبار إلى صورة أخرى له، وتكون هذه العلامات بعد عملية التحويل متكافئة، ويعرف كروكر والجينا (Croker and Algina) المعادلة بأنها عملية الحصول على علامات متكافئة لاختبارين، ويوجد نوعان للمعادلة، وهما: المعادلة الأفقية، وفيها يتم معادلة نماذج الاختبارات التي تقيس نفس السمة عند نفس المستوى من القدرة، والمعادلة العمودية، وفيها يتم معادلة اختبارات مختلفة تقيس نفس السمة صممت لمفحوصين عند مستويات قدرة مختلفة (Croker and Algina, 1986)

وهناك شروط يجب توافرها لإجراء المعادلة بين علامات الاختبارات، منها: أن تقيس هذه الاختبارات نفس السمة أو القدرة، واللاتباين (اللاتغاير) أي أن تحويل العلامات على هذه الاختبارات يبقى كما هو بغض النظر عن مجموعة المفحوصين، والتماثل، أي أن تحويل العلامات من صورة إلى أخرى للاختبارات قابل للانعكاس، أي أن الرسم البياني لتوزيع العلامات من الصورة (أ) للاختبار إلى الصورة (ب) يجب أن يكون نفس الانعكاس من الصورة (ب) إلى الصورة (أ) (Kolen and Brenan, 2004)، ومن الشروط كذلك، أن تكون الاختبارات متساوية في مستوى الثبات تقريباً، ويجب أن تكون الاختبارات المراد معادلتها

متكافئة، أي أن يكون المتوسط الحسابي والانحراف المعياري للأداء على الصورة (أ) ومستوى الصعوبة ومستوى التمييز للفقرات على الصورة (أ) من الاختبار تماثل نظيراتها في الصورة (ب) من الاختبار .(Croker and Algina, 1986)

وتبرز الحاجة إلى إجراء المعادلة العموديةعند تطبيق اختبارات مختلفة على مفحوصين مختلفين تقيس نفس السمة للمقارنة بينهم، وكذلك لتخطيط نمو الطالب عبر سنوات دراسته عند توافر مقياس مشترك لاختبار تحصيلي متعدد المستويات لقياس سمة نامية، وكذلك عند توافر بطارية اختبارات تتكون من عدة مستويات تناسب صفوفًا مختلفة، فمن الضروري في بعض الأحيان اختبار مفحوص في مستوى يختلف عن مستوى صفه، مثل اختبار مفحوص متميز في الصف الرابع باختبار يناسب طلبة الصف الخامس، أو اختبار مفحوص ضعيف في الصف الرابع بتطبيق اختبار عليه يناسب طلبة الصف الثالث، وهذا يتطلب معادلة الدرجات للاختبارات المعدة لمستويات الطلبة المختلفة للمقارنة بين أداء الطالب عندما يطبق عليه اختبار من خارج مستوى صفه مع الاختبار الذي يناسب صفه (الطراونة، ٢٠٠٥).

ويوجدعدة طرق للمعادلة، منها ما يعتمد على النظرية الكلاسيكية في القياس، مثل: طريقةالوحدات الصفية، وطريقة تكر الخطية، وطريقة براون هولد، والطريقة المئينية، ومنها ما يعتمد على نظرية استجابة الفقرة، مثل: طريقة معادلة العلامات الحقيقية وطريقة معادلة العلامات المشاهدة وطريقة المعادلة باستخدام نماذج نظرية استجابة الفقرة الثنائية والمتعددة (Crockerand and Algina, 1986; Hambleton, Swaminathan, and Roges, 1991; Kolen and Brennan, 2004).

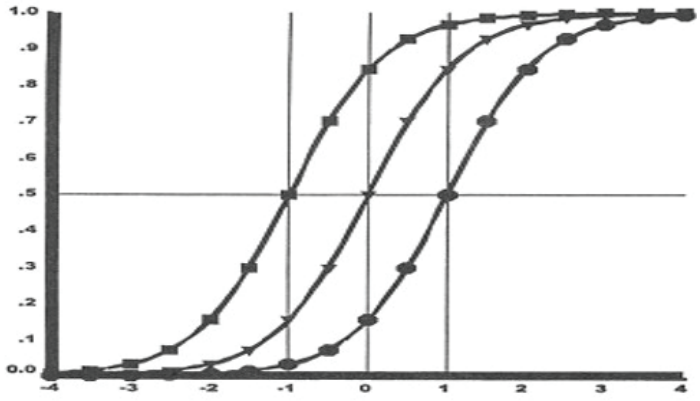
وتفترض نظرية استجابة الفقرة وجود سمة كامنة تفسر الاستجابة على فقرات الاختبار، ويستدل عليها بشكل غير مباشر من خلال أداء المفحوص على فقرات الاختبار(Croker and Algina, 1986)، ويوجد لنظرية استجابة الفقرة افتراضات تتمثل فيما يلي:

- أحادية البعد: يعني وجود سمة كامنة واحدة تفسر الاستجابة على فقرات الاختبار، أي أن الاختبار يقيس سمة واحدة.
- الاستقلال الموضعي: ويقصد به أن إجابة المفحوص على فقرة يجب أن لا تتأثر بإجابته على فقرات أخرى يتضمنها الاختبار، وعندما يتحقق هذا الافتراض، فإن ذلك يعني أن إجابة المفحوص على فقرة ما تمثل حدثا مستقلا عن إجابته على فقرة أخرى، ومن ثم فإن احتمال الحصول على نمط من أنماط الاستجابة على الفقرات يساوي حاصل ضرب احتمال الإجابة على كل فقرة.

- منحنى خصائص الفقرة: ويمثل هذا المنحنى العلاقة الرياضية التي تربط بين احتمالية إجابة الفرد إجابة صحيحة على الفقرة، وبين القدرة التي تقيسها فقرات الاختبار، وتختلف نماذج استجابة الفقرة بالصيغة الرياضية لمنحنى خصائص الفقرة.
- عامل السرعة: تفترض نماذج السمات الكامنة أن عامل السرعة في الأداء لا يلعب دورا في الإجابة عن الفقرات، أي أن إخفاق المفحوص في الإجابة عن فقرات الاختبار يعزى إلى انخفاض قدرته وليس إلى تأثير السرعة في الإجابة أو عدم وصوله لهذه الفقرات نتيجة قلة الوقت المخصص للاختبار (Hambleton and swaminathan, 1985).

ويوجد عدة نماذج ثنائية التدريج لنظرية استجابة الفقرة منها نموذج راش، ويفترض أن جميع الفقرات لها القدرة نفسها للتمييز بين الأفراد، ويفترض كذلك عدم وجود تأثير للتخمين في الإجابة عن الفقرات، وأن الفقرات تختلف في صعوبتها فقط (Baker, 2001).

والشكل (١) يمثل ثلاث فقرات كمثال على نموذج راش، ويبين أن الفقرات تختلف في معاملات الصعوبة فقط، وتتساوى في معاملات التمييز، ولا يوجد تخمين فيها.



شكل (١)
منحنى خصائص الفقرة لثلاث فقرات حسب نموذج راش

ويمكن إجراء المعادلة العمودية باستخدام نماذج نظرية استجابة الفقرة، ومنها نموذج راش، وتستخدم المعادلة الخطية للقدرة، لمعادلة الدرجات على مستويات الاختبار، وهي:

$$\theta_1 = \left[\frac{sd_1}{sd_2}\right]\theta_2 + \bar{\theta}_1 - \left[\frac{sd_1}{sd_2}\right]\bar{\theta}_2 \quad \text{.....(1)}$$

حيث:

θ_1: هي القدرة للاختبار المراد المعادلة إليه.

θ_2: هي القدرة للاختبار المراد معادلته.

sd_1: الانحراف المعياري لقدرة الأفراد للاختبار المراد المعادلة إليه.

sd_2: الانحراف المعياري لقدرة الأفراد للاختبار المراد معادلته.

θ_1: الوسط الحسابي لقدرة الأفراد للاختبار المراد المعادلة إليه.

θ_2: الوسط الحسابي لقدرة الأفراد للاختبار المراد معادلته (Thomas, 1998).

وأشار انجوف (Angoff, 1984) إلى أنه يوجد العديد من التصميمات للمعادلة، منها:

١- تصميم المجموعة الواحدة: في هذا التصميم يوجد عدة صور للاختبار مثل (أ، ب، ج)، ويوجد مجموعة واحدة من الأفراد، يتم تطبيق جميع صور الاختبار على المجموعة، ثم يتم معادلتها.

٢- تصميم المجموعات العشوائية باختبار لكل مجموعة: في هذا التصميم يتم اختيار عينة غير متجانسة من حيث التحصيل ذات حجم كبير، وتقسم هذه العينة إلى مجموعتين متكافئتين بشكل عشوائي، ويوجد صورتان للاختبار مثل (أ، ب)، يتم تطبيق الصورة الأولى (أ) على المجموعة الأولى وتطبق الصورة (ب) على المجموعة الثانية ثم يتم معادلتها.

٣- تصميم المجموعات العشوائية باختبارين لكل مجموعة: في هذا التصميم يتم اختيار عينة غير متجانسة من حيث التحصيل ذات حجم كبير، وتقسم هذه العينة إلى مجموعتين متكافئتين بشكل عشوائي، ويوجد صورتان للاختبار مثل (أ، ب)، يتم تطبيق الصورة الأولى (أ) على المجموعة الأولى ثم تتبع بالصورة (ب)، وتطبق الصورة (ب) على المجموعة الثانية ثم تتبع بالصورة (أ)، ثم يتم معادلتها.

٤- تصميم المجموعات غير العشوائية باختبار الجذع المشترك: في هذا التصميم يتم اختيار عينة مكونة من مجموعتين من الأفراد بشكل غير عشوائي، ويوجد صورتان للاختبار مثل (أ، ب) واختبار الجذع المشترك (فقرات مشتركة تعطى للمجموعتين)، تطبق الصورة (أ) على المجموعة الاولى وتطبق الصورة (ب) على المجموعة الثانية، ويطبق اختبار الجذع المشترك على المجموعتين، ثم يتم معادلتها.

٥- تصميم المجموعات العشوائية باختبار الجذع المشترك: في هذا التصميم يتم اختيار مجموعتين من الأفراد بشكل عشوائي، ويوجد صورتان للاختبار مثل (أ، ب) واختبار الجذع المشترك (فقرات مشتركة تعطى للمجموعتين)، تطبق الصورة (أ) على المجموعة

الاولى وتطبق الصورة (ب) على المجموعة الثانية، ويطبق اختبار الجذع المشترك على المجموعتين، ثم يتم معادلتها.

ويوجد محكات ومعادلات مختلفة تستخدم للتحقق من دقة فاعلية المعادلة العمودية، منها: الخطأ المعياري للمعادلة والذي يحسب من خلال برامج إحصائية باستخدام الحاسوب، وتحسب قيمته بمعادلات رياضية تختلف باختلاف التصميم المستخدم وطريقة المعادلة، ومحك الصدق التقاطعي، الذي يحسب بمعادلة واحدة، وهي المذكورة في كولن ووتني (Kolen, and Whiteny, 1981) وهي:

$$C= \sum (x_1-x_2)*2/NK \ldots\ldots\ldots\ldots (2)$$

x_1: العلامات المعادلة في العينة الأولى

x_2: العلامات المعادلة في العينة الثانية

N: عدد العلامات الملاحظ في توزيع الصدق التقاطعي

K: عدد الفقرات في الاختبار المتعدد

وقام يانغ وهويانغ (Yang and Houang, 1997) بدراسة هدفت إلى معرفة أثر طول اختبار الجذع المشترك في دقة معادلة الاختبارات الأفقية باستخدام طريقة تكر (tucker) الخطية، وطريقتين من طرق نظرية استجابة الفقرة، تم استخدام صورتين من اختبار للكفاءة، الأول مكون من (١٩٧) فقرة والثاني مكون من (٢٠٠) فقرة، استخدمت ثلاث صور من اختبار الجذع المشترك تختلف في عدد فقراتها، وأظهرت النتائج وجود دقة بدرجة متوسطة بغض النظر عن طريقة المعادلة، وأظهرت النتائج كذلك أن دقة المعادلة تميل لتكون أكثر عند زيادة فقرات اختبار الجذع المشترك.

وقام الطراونة (٢٠٠٥) بدراسة هدفت إلى تطوير اختبار رياضيات متعدد المستويات في العمليات الحسابية الأربع للصفوف الأساسية (الثالث، والرابع، والخامس والسادس) بفقرات متعددة التدريج، تكونت عينة الدراسة من (١٢٧٩) طالباً وطالبةً من طلبةالصفوف الأربعةالأساسية في محافظة الكرك، تكونت أداة الدراسة من اختبار من أربعة مسستويات في العمليات الحسابية الأربع في الرياضيات، تم التحقق من صدقه من خلال الصدق التمييزي، والصدق المرتبط بمحك والصدق العاملي، وتم التحقق من ثباته من خلال طريقة الإعادة، والطريقة النصفية ومعادلة كرونباخ ألفا، وتم استخدام الطريقة الخطية والطريقة المئينية والنموذج الأحادي المعلمة في المعادلة، تم استخدام محكي الخطأ المعياري والصدق التقاطعي للتحقق من فاعلية المعادلة، أظهرت نتائج الدراسة أن أكثر الطرق فاعلية باستخدام محك الخطأ المعياري كانت الطريقة الخطية في النظرية الكلاسيكية، وأن أكثر الطرق فاعلية

باستخدام محك الصدق التقاطعي كانت طريقة النموذج الأحادي المعلمة.

وقام السحيمات (٢٠٠٧) بدراسة هدفت إلى بناء اختبار متعدد المستويات في مبحث الرياضيات لطلبة الصفوف الأساسية (الثامن والتاسع والعاشر) في الأردن، وتم بناء اختبار لكل صف يتكون من (٣٠) فقرة، ويشترك كل اختبارين لكل صفين متجاورين بـ (١٥) فقرة كاختبار جذع مشترك، تم التحقق من خصائصها السيكومترية باستخدام النظرية الكلاسيكية ونظرية استجابة الفقرة، وقد استخرجت معادلات المعادلة التي يمكن من خلالها إيجاد العلامة في كل صف بما يقابلها في الصف الآخر في النظرية الكلاسيكية بحساب مقدار إزاحة الوسط الحسابي لكل صفين متجاورين، وفي نظرية استجابة الفقرة تم استخدام نموذج راش، أظهرت النتائج دلالات صدق وثبات مقبولة للاختبار.

كما قام سينهاري وهولاند (Sinharay and Holland, 2007) بدراسة هدفت إلى التحقق من أن اختبارات الجذع المشترك هي اختبارات مصغرة للمحتوى وللخصائص السيكومترية للاختبارات التي تجري عليها المعادلة، تم استخدام بيانات حقيقية تم الحصول عليها من دراسات سابقة وبيانات مولدة، تم بناء اختبارات الجذع المشترك مختلفة في مدى صعوبة فقراتها، ويختلف هذا المدى عن مدى صعوبة فقرات الاختبار المراد إجراء المعادلة له، تم استخدام انحرافات المعادلة والخطأ المعياري للتحقق من فاعلية المعادلة، أظهرت نتائج الدراسة أن اختبار الجذع المشترك يجب أن يتصف بنفس الخصائص التي يتصف بهاالاختبار المراد إجراء معادلة له.

وقام حمادنة (٢٠١٢) بدراسة هدفت إلى الكشف عن أثر تجانس المحتوى وطول الجذع المشترك علىدقة المعادلة لاختبار ثنائي التدريج ومتعدد المستوى في الرياضيات باستخدام تصميم الجذع المشترك للمجموعات غير المتكافئة، تم إجراء المعادلة المئينية العمودية في النظرية الكلاسيكية، وتم التحقق من دقة المعادلة باستخدام الخطأ المعياري للمعادلة، تكونت عينة الدراسة من (٧٧٦) طالباً وطالبةً من الصفين التاسع والعاشر الأساسيين التابعين لمديرية التربية والتعليم في محافظة المفرق، أظهرت النتائج أن دقة المعادلة المئينية العمودية كانت أعلى ما يمكن عندما كان محتوى اختبار الجذع المشترك متجانسا، وعدد فقراته أكثر من (٢٠) فقرة، إذ بلغت قيمة الخطأ المعياري للمعادلة العمودية (٠,٨٥١).

كما قام المدانات (٢٠١٢) بدراسة هدفت إلى مقارنة فاعلية طريقتي تكر (tucker) الخطية وليفين في النظرية الكلاسيكية في معادلة الاختبارات الأفقية عند استخدام التصميم القائم على اختبار الجذع المشترك ومجموعات غير عشوائية، تم بناء صورتين متكافئتين لاختبارفي مبحث الفيزياء كل اختبار مكون من (٢٠) فقرة، بالإضافة إلى (١٠) فقرات

استخدمت كاختبار جذع مشترك، تكونت عينة الدراسة من (١٩٩) طالباً وطالبةً، تم استخدام أربع مجموعات من فقرات اختبار الجذع المشترك (١، ٤، ٧، ١٠)، أظهرت النتائج أن الفروق بين علامات الطلبة المعادلة باستخدام طريقتي المعادلة ولجميع مجموعات اختبار الجذع هي فروق غير جوهرية، وفيما يتعلق بعدد فقرات اختبار الجذع المشترك، أظهرت النتائج أن الفروق كانت دالة إحصائيا عند مستوى الدلالة(α = ٠,٠٥) في متوسطات القيم المعادلة بطريقة تكر وطريقة ليفين عندما كانت عدد فقرات اختبار الجذع المشترك (٤) و(٧) فقرات، ولصالح طريقة ليفين، وكانت الفروق لصالح طريقة تكر عندما كانت عدد فقرات الجذع المشترك فقرة واحدة.

وقام الرواجفة (٢٠١٦) بدراسة هدفت إلى تطبيق المعادلة العمودية في اختبار متعدد المستويات للصفوف: العاشر الأساسي، والأول الثانوي العلمي والثاني الثانوي العلمي، تم استخدام تصميم الجذع المشترك للمجموعات المتكافئة، وتم بناء اختبار في الرياضيات مكون من ثلاثة مستويات، تم إجراء المعادلة الخطية العمودية باستخدام النموذج الأحادي المعلمة في نظرية استجابة الفقرة، وتم التحقق من دقة المعادلة باستخدام الخطأ المعياري للمعادلة، تكونت عينة الدراسة من (٨٠٠) طالبٍ وطالبة، من طلبة مديرية التربية والتعليم لمحافظة الطفيلة، أظهرت نتائج الدراسة دقة المعادلة العمودية فقد كانت قيمة الصدق التقاطعي متدنية.

يتضح من استعراض الدراسات السابقة أن بعضها درس المقارنة بين طرق المعادلة العمودية المنبثقة عن النظرية الكلاسيكية في القياس ونظرية استجابة الفقرة في دقة المعادلة، وبعضها درس بناء اختبارات متعددة المستويات (الرواجفة، ٢٠١٦، السحيمات، ٢٠٠٧، الطراونة، ٢٠٠٥)، وبعض الدراسات درست التحقق من أن اختبارات الجذع المشترك هي اختبارات مصغرة للمحتوى وبعضها درس تجانس محتوى اختبار الجذع المشترك (حمادنه، ٢٠١٢، Sinharay and Holland, 2007)، وبعض الدراسات درست أثر طول اختبار الجذع المشترك على دقة المعادلة الأفقية والمعادلة العمودية باستخدام النظرية الكلاسيكية (حمادنه، ٢٠١٢، المدانات، Yang and Houang, 1997,2012)، ولم تبحث أي دراسة في أفضل نسبة للفقرات في اختبار الجذع المشترك على دقة المعادلة العمودية لاختبار في موضوع الهندسة في الرياضيات باستخدام نموذج راش، لذلك هدفت هذه الدراسة إلى دراسة أثر نسبة الفقرات في اختبار الجذع المشترك على دقة المعادلة العمودية باستخدام نموذج راش لاختبار في الرياضيات.

مشكلة الدراسة

من مشكلات القياس النفسي عدم تساوي وحدات القياس، ولهذا فمن الصعوبة بمكان المقارنة بين أداء الطلبة في الاختبارات المختلفة، وبما أن معادلة الاختبارات من التطبيقات العملية لنظرية استجابة الفقرة، والتي تساعد على التوصل إلى وحدات قياس معيارية، فيمكن من خلالها المقارنة بين علامات المفحوصين على اختبارات مختلفة تقيس نفس السمة لمستويات مختلفة من القدرة، ولندرة الدراسات التي بحثت في نسبة فقرات اختبارات الجذع المشترك على دقة المعادلة باستخدام نماذج نظرية استجابة الفقرة، ولما يترتب على نتائج المعادلة من قرارات مهمة تتعلق بالمفحوصين، فمن الضروري تحديد أفضل نسبة لفقرات اختبار الجذع المشترك من فقرات الاختبار، ولهذا جاءت هذه الدراسة للكشف عن أثر نسبة الفقرات في اختبار الجذع المشترك على دقة المعادلة العمودية باستخدام نموذج راش، وتتحدد مشكلة الدراسة في الإجابة عن الأسئلة الآتية:

السؤال الأول: هل تتطابق الاستجابات على فقرات الاختبار بمستوييه مع افتراضات نموذج راش؟

السؤال الثاني: هل تتأثر دقة المعادلة الخطية العمودية لقدرات الطلبة في اختبار الرياضيات باستخدام نموذج راش في تصميم المجموعات العشوائية مع الجذر المشترك باختلاف عدد فقرات اختبار الجذع المشترك؟

السؤال الثالث: ما أفضل نسبة لفقرات اختبار الجذع المشترك في المعادلة العمودية؟

أهمية الدراسة

تنبع أهمية هذه الدراسة من كونها تطرقت إلى معادلة الاختبارات العمودية والتي هي إحدى تطبيقات نظرية استجابة الفقرة، وتنبع أهمية الدراسة كذلك من أهمية الاختبارات المتعددة المستويات، والتي تساعد على التوصل إلى وحدات قياس معيارية، تمكن القائمين على الشؤون التعليمة من مقارنة أداء الطلبة في المدارس والجامعات، وتكمن أهمية الدراسة كذلك في تناولها لموضوع نسبة عدد فقرات اختبار الجذع المشترك على دقة فاعلية المعادلة العمودية باستخدام نموذج راش، وتحديد أفضل نسبة لعدد فقرات اختبار الجذع المشترك المناسبة للمعادلة، فضلاً عن تناولها لموضوع الهندسة في الرياضيات، فالهندسة تعمل على تنمية القدرة على التخيل والرسم الثلاثي الأبعاد عند الطالب.

محددات الدراسة

اقتصرت الدراسة على:

- طريقة المعادلة العمودية الخطية باستخدام نموذج راش في نظرية استجابة الفقرة.
- تصميم المجموعات العشوائية باختبار الجذع المشترك.
- طلبة الصفين الخامس والسادس الأساسيين، في مديرية التربية والتعليم في لواء المزار الجنوبي على الفصل الدراسي الثاني من العام الدراسي (٢٠١٦/٢٠١٧).
- موضوع الهندسة في كتب الرياضيات للصفين الخامس والسادس الأساسيين.
- محكي الصدق التقاطعي والخطأ المعياري للتحقق من فاعلية المعادلة العمودية.

تعريف المصطلحات

منهجية الدراسة وإجراءاتها: المعادلة العمودية: هي عملية تحويل العلامات على نماذج الاختبار التي صممت عند مستويات مختلفة للقدرة للحصول على مقياس موحد يمتد عبر عدد من الصفوف الدراسية أو الأعمار المختلفة للطلبة.

اختبار الجذع المشترك: هو اختبار مكون من مجموعة فقرات مشتركة بين نماذج الاختبار المختلفة، والهدف منه معادلة الدرجات على نماذج الاختبار والمقارنة بينها.

معايير الفاعلية: هي معايير إحصائية تستخدم للتحقق من دقة وفاعلية المعادلة، وتستخدم للمقارنة بين طرق المعادلة، وفي هذه الدراسة تم استخدم محكي الصدق التقاطعي والخطأ المعياري للمعادلة.

نموذج راش: اقترحه العالم الدنماركي راش، ويهتم بتحديد موقع الفقرة الاختبارية على سلم الصعوبة لجميع الفقرات التي تشكل الاختبار، ويهتم بتدريج مستويات قدرة الفرد باختبار معين على نفس متصل الفقرات (علام، ٢٠٠٢).

الاختبارات متعددة المستويات: هي اختبارات تقيس التحصيل في موضوع معين عند مفحوصين من مستويات قدرة مختلفة، وتستخدم هذه الاختبارات في الغالب للمقارنة بين أداء المفحوصين من مستوى معين من القدرة مع مستوى أدنى منهم أو أعلى من القدرة.

مجتمع الدراسة

تكون مجتمع الدراسة من جميع طلبة الصفين الخامس والسادس الأساسيين، في مديرية التربية والتعليم في لواء المزار الجنوبي في المملكة الأردنية الهاشمية للعام الدراسي

٢٠١٧/٢٠١٦ والبالغ عددهم (٢٥٧٧) طالباً وطالبةً، والجدول (١) يبين عدد أفراد مجتمع الدراسة في كل صف.

جدول (١)
عدد أفراد مجتمع الدراسة في كل صف

الصف	المجموع
الخامس	١٢٦٩
السادس	١٣٠٨
المجموع	٢٥٧٧

عينة الدراسة

تكونت عينة الدراسة من (١٠٣١) طالباً وطالبةً، من طلبة الصفين الخامس والسادس الأساسيين في مديرية التربية والتعليم في لواء المزار الجنوبي في المملكة الاردنية الهاشمية، تم اختيارهم بطريقة العينة العشوائية الطبقية، وكانت وحدة الاختيار هي الشعبة، وبنسبة (٤٠٪) تقريبا من مجتمع الدراسة، والجدول (٢) يبين عدد أفراد عينة الدراسة حسب الصف.

جدول (٢)
عدد أفراد عينة الدراسة في كل صف

الصف	المجموع
الخامس	٥٠٨
السادس	٥٢٣
المجموع	١٠٣١

أداة الدراسة

تكون الاختبار بشكله النهائي من ست صور، ومن مستويين كل مستوى له ثلاث صور، المستوى الأول للصف الخامس والمستوى الثاني للصف السادس، وعدد الفقرات في جميع الصور كان (٢٠) فقرة، والصورة الأولى من المستويين فيها (٣) فقرات مشتركة، ترتبط بالمحتوى المشترك بين منهاجي الصفين الخامس الأساسي والسادس الاساسي، وفي الصورة الثانية فيها (٦) فقرات مشتركة ترتبط بالمحتوى المشترك بين منهاجي الصفين وفي الصورة الثالثة فيها (٩) فقرات مشتركة ترتبط بالمحتوى المشترك بين منهاجي الصفين، والجدول (٣) يبين ذلك.

جدول (٣)
مستويا الاختبار وصوره وعدد الفقرات المشتركة

المستوى (الصف)	الصورة	عدد الفقرات المشتركة	نسبة الفقرات المشتركة
الأول (الصف الخامس)	الأولى	٣	٪١٥
	الثانية	٦	٪٣٠
	الثالثة	٩	٪٤٥
الثاني (الصف السادس)	الأولى	٣	٪١٥
	الثانية	٦	٪٣٠
	الثالثة	٩	٪٤٥

وتم اتباع الخطوات الآتية لبناء الاختبار بمستوييه وبصوره الست:

أولاً) تحديد مادة الاختبار: وهي المحور الرئيسي من الاطار العام والنتاجات العامة الخاصة بالهندسة من منهاج الرياضيات للصفين الخامس والسادس الأساسيين (وزارة التربية والتعليم، ٢٠١٥أ، وزارة التربية والتعليم، ٢٠١٥ب)، والجدول (٤) يبين ذلك.

جدول (٤)
مواضيع الهندسه للصفين الخامس والسادس

الصف	المحتوى
الخامس	- المضلعات - رسم الدائرة - قياس الزاوية - رسم الزاوية
السادس	- مضلعات رباعية - مجموع قياسات زوايا المثلث - مجموع قياسات زوايا الشكل الرباعي - مجموع قياسات الزوايا حول نقطة - رسم المثلث - رسم متوازي الاضلاع - التماثل - استخدام الشبكات في بناء المجسمات

ثانياً) تحليل المحتوى لوحدات الهندسة في الصفين، وتم صياغة (٢٠) هدفًا لكل صف.

ثالثاً) تم صياغة (٢٠) فقرة، من نوع الاختيار من متعدد من أربعة بدائل لكل صف، كل فقرة تقيس هدفًا واحدًا، وتم وضع الاختبار بمستوييه في ست صور، كل مستوى له ثلاث صور، وتم تحديد المحتوى الذي تقيسه الفقرات في كل صورة من خلال استشارة (١٠) معلمين رياضيات يدرسون الصفين الخامس والسادس الأساسيين، وعلى النحو الآتي:

١) المستوى الأول: وهو اختبار الصف الخامس يتكون من ثلاث صور، وهي:

- الصورة الأولى: مكونة من (٢٠) فقرة للصف الخامس، منها (٣) فقرات مشتركة، وبنسبة (١٥٪) من فقرات الاختبار، مع اختبارالصف السادس كاختبار الجذع.
- الصورة الثانية: مكونة من (٢٠) فقرة للصف الخامس، منها (٦) فقرات مشتركة، وبنسبة (٣٠٪) من فقرات الاختبار، مع اختبار الصف السادس كاختبار الجذع المشترك.
- الصورة الثالثة: مكونة من (٢٠) فقرة للصف الخامس، منها (٩) فقرات مشتركة، وبنسبة (٤٥٪) من فقرات الاختبار، مع اختبار الصف السادس كاختبار الجذع المشترك.

٢) المستوى الثاني: وهو اختبار الصف السادس يتكون من ثلاث صور، وهي:

- الصورة الأولى: مكونة من (٢٠) فقرة للصف السادس، منها (٣) فقرات مشتركة، وبنسبة (١٥٪) من فقرات الاختبار، مع اختبار الصف الخامس كاختبار الجذع المشترك.
- الصورة الثانية: مكونة من (٢٠) فقرة للصف السادس، منها (٦) فقرات مشتركة، وبنسبة (٣٠٪) من فقرات الاختبار، مع اختبار الصف الخامس كاختبار الجذع المشترك.
- الصورة الثالثة: مكونة من (٢٠) فقرة للصف السادس، منها (٩) فقرات مشتركة، وبنسبة (٤٥٪) من فقرات الاختبار، مع اختبار الصف الخامس كاختبار الجذع المشترك.

رابعاً) عرض الاختبار بمستوييه وبصوره الست على (١٢) محكماً من أساتذة القياس والتقويم في الجامعات الأردنية ومعلمين ممن يدرسون الصفين ومشرفين للرياضيات في وزارة التربية والتعليم، لإبداء آرائهم حول فقرات كل مستوى من حيث سلامة اللغة، ووضوحها وملاءمتها لقياس المحتوى، وتدرج الفقرات في قياس المحتوى عبر الصفين، واقتراح أية تعديلات على الفقرات، وفي ضوء آراء المحكمين تم تعديل (٥) فقرات من فقرات اختبار الصف الخامس الأساسي، و(٤) فقرات من فقرات اختبار الصف السادس الأساسي، ولم يتم حذف أو إضافة أي فقرة.

خامساً) تطبيق الاختبار بمستوييه وبصوره الست على عينة استطلاعية مكونة من (١٤٨) طالبًا وطالبة، لتحديد الزمن المناسب لكل مستوى والتعرف على إن كان هناك أية ملاحظات على الفقرات، وتم تحديد (٦٠) دقيقة كوقت مناسب للاختبار لكل صورة من صور مستويي الاختبار.

سادساً) التحقق من صدق الاختبار بمستوييه من خلال:

- مؤشر التحليل العاملي: تم إجراء التحليل العاملي على العينة الكلية التي دخلت في التحليل للتحقق من أن الاختبار بمستوييه (اختبار الصف الخامس واختبار الصف السادس) يقيس سمة واحدة، وتم العمل على الصورة الثالثة التي عدد فقرات اختبار الجذع المشترك فيها (٩)

فقرات، لأنه في نظرية استجابة الفقرة تقدير قدرة الفرد يكون مستقلاً عن عينة الفقرات التي تطبق عليه، وتقدير خصائص الفقرات يكون مستقلاً عن عينة الأفراد، وتم إجراء التدوير المتعامد بطريقة الفاريمكس لكل مستوى للاختبار، والجدول (٥) يبين قيم الجذور الكامنة ونسبة التباين المفسر ونسبة التباين التراكمية.

الجدول (٥)
القيم المميزة ونسبة التباين المفسر للعوامل لمستويي الاختبار

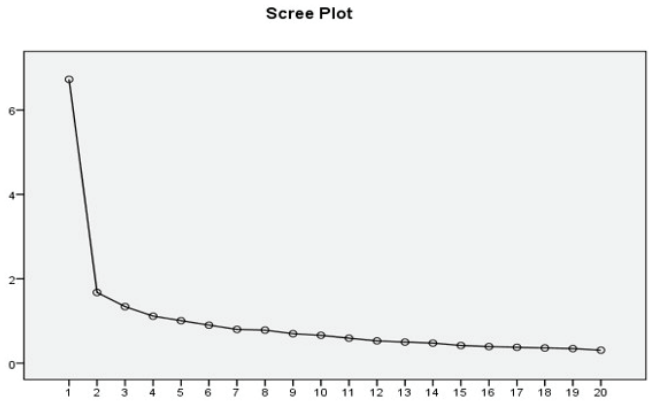
العامل	اختبار الصف الخامس		العامل	اختبار الصف السادس	
	الجذر الكامن	نسبة التباين المفسر		الجذر الكامن	نسبة التباين المفسر
١	٦,٧٢٨	٣٣,٦٣٩	١	٥,٩٨٦	٢٩,٩٢٨
٢	١,٦٧٦	٨,٣٧٨	٢	١,٨٩٢	٩,٤٥٨
٣	١,٣٤١	٦,٧٠٣	٣	١,٦٤٨	٨,٢٤١
٤	١,١١٤	٥,٥٧٢	٤	١,١٧٠	٥,٨٥١
٥	١,٠٠٦	٥,٠٣١	٥	٠,٩٨٠	٤,٩٠٢
٦	٠,٩٠٠	٤,٥٠٢	٦	٠,٩٤٤	٤,٧١٩
٧	٠,٨٠٠	٣,٩٩٨	٧	٠,٨٨٤	٤,٤٢٠
٨	٠,٧٨٣	٣,٩١٣	٨	٠,٧٤٠	٣,٧٠٢
٩	٠,٦٩٨	٣,٤٩٠	٩	٠,٧٠١	٣,٥٠٦
١٠	٠,٦٦٠	٣,٣٠٢	١٠	٠,٦٦٦	٣,٣٣٠
١١	٠,٥٩٢	٢,٩٥٩	١١	٠,٥٦٧	٢,٨٣٦
١٢	٠,٥٢٧	٢,٦٣٧	١٢	٠,٥٢٦	٢,٦٢٩
١٣	٠,٤٩٩	٢,٤٩٦	١٣	٠,٥٠٤	٢,٥٢١
١٤	٠,٤٧٥	٢,٣٧٥	١٤	٠,٤٩٥	٢,٤٧٣
١٥	٠,٤١٩	٢,٠٩٣	١٥	٠,٤٥٧	٢,٢٨٧
١٦	٠,٣٩٣	١,٩٦٧	١٦	٠,٤٣٧	٢,١٨٦
١٧	٠,٣٧٦	١,٨٧٩	١٧	٠,٤٢٠	٢,١٠١
١٨	٠,٣٦٠	١,٨٠٢	١٨	٠,٣٥١	١,٧٥٥
١٩	٠,٣٤٥	١,٧٢٧	١٩	٠,٣٢٥	١,٦٢٧
٢٠	٠,٣٠٧	١,٥٣٦	٢٠	٠,٣٠٦	١,٥٢٨

يتبين من الجدول (٥) ما يأتي:

- بالنسبة لاختبار الصف الخامس: كانت نسبة التباين التي فسرها العامل الأول تساوي (٣٣,٦٣٩٪)، وهي أكبر من (٢٠٪)، وهذا يدل على أن اختبار الصف الخامس يقيس سمة واحدة، أي تحقق افتراض أحادية البعد في نظرية استجابة الفقرة.
- بالنسبة لاختبار الصف السادس: كانت نسبة التباين التي فسرها العامل الأول تساوي (٢٩,٩٢٨٪)، وهي أكبر من (٢٠٪)، وهذا يدل على أن اختبار الصف السادس يقيس

سمة واحدة، أي تحقق افتراض أحادية البعد في نظرية استجابة الفقرة.

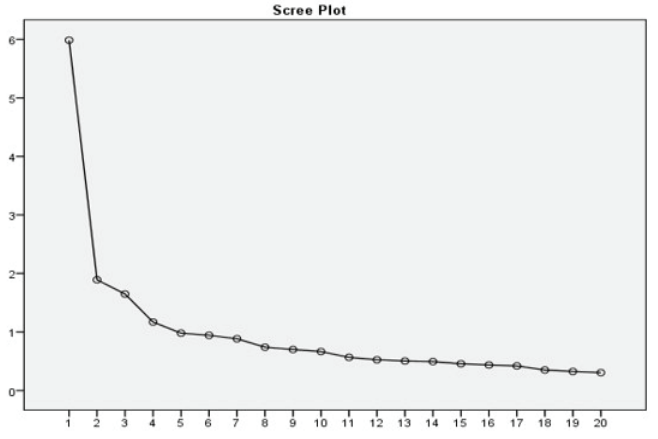
والشكلان (٢، ٣) يبينان رسم Scree Plot للجذور الكامنة لاختباري الصفين الخامس والسادس الاساسيين.



شكل (٢)

رسم Scree Plot للجذور الكامنة لمستوى الاختبار للصف الخامس

يتضح من الشكل (٢) ارتفاع قيمة الجذر الكامن للعامل الأول مقارنة ببقية العوامل، وهذا مؤشر آخر على تحقق افتراض أحادية البعد لمستوى الاختبار للصف الخامس.



شكل (٣)

رسم Scree Plot للجذور الكامنة لمستوى الاختبار للصف السادس

يتضح من الشكل (٣) ارتفاع قيمة الجذر الكامن للعامل الأول مقارنة ببقية العوامل، وهذا مؤشر آخر على تحقق افتراض أحادية البعد لمستوى الاختبار للصف السادس.

ب) التحقق من أداء الطلبة عبر الصفوف الدراسية:

للتحقق من أن الاختبار يقيس نفس السمة في مستوييه، تم حساب معاملات الارتباط لأداء أفراد العينة الاستطلاعية على الفقرات الخاصة بكل مستوى والفقرات المشتركة بالتتابع بين مستويي الاختبار للصفين، والجدول (٦) يبين نتائج ذلك.

جدول (٦)
معاملات الارتباط بين الأداء على الفقرات المشتركة والفقرات الخاصة بكل مستوى للاختبار

الأداء	معامل الارتباط للصورة		
	الأولى	الثانية	الثالثة
أداء طلبة الصف الخامس على اختبار صفهم وأداؤهم على الفقرات المشتركة مع الصف السادس	٠,٧٥**	٠,٧٩**	٠,٨٣**
أداء طلبة الصف السادس على اختبار صفهم وأداؤهم على الفقرات المشتركة مع الصف الخامس	٠,٧٣**	٠,٧٨**	٠,٨١**

**: دالة إحصائياً عند مستوى الدلالة ($\alpha \leq$ ٠,٠١).

يتبين من الجدول (٦) أن جميع قيم معاملات الارتباط كانت دالة إحصائياً عند مستوى الدلالة ($\alpha \leq$ ٠,٠١)، وتراوحت بين (٠,٧٣) وبين (٠,٨٣)، ومنها يتبين أن التباين المشترك تراوح بين (٠,٥٣) وبين (٠,٦٩)، وهذا يدل على أن فقرات الاختبار في مستوييه تقيس نفس السمة.

ج) التحقق من الصدق المرتبط بمحك: تم حساب معاملات الارتباط بين أداء أفراد العينة الاستطلاعية على الاختبار لكل مستوى وعلاماتهم في مبحث الرياضيات المدرسية في نهاية الفصل الدراسي الثاني من العام الدراسي ٢٠١٧/٢٠١٦ لكل صف، والجدول (٧) يبين نتائج ذلك.

جدول (٧)
معاملات الارتباط بين أداء أفراد العينة الاستطلاعية على الاختبار لكل مستوى وعلاماتهم في مبحث الرياضيات لكل صف

الصف	معامل الارتباط		
	الصورة الأولى	الصورة الثانية	الصورة الثالثة
الخامس	٠,٦٨**	٠,٧٣**	٠,٦٩**
السادس	٠,٧٤**	٠,٧١**	٠,٧٧**

يتبين من الجدول (٧) أن جميع قيم معاملات الارتباط كانت دالة إحصائياً عند مستوى

الدلالة ($\alpha \leq$ ٠,٠١)، وموجبة وتراوحت بين (٠,٦٨) وبين (٠,٧٧)، وهذا يدل على أن العلاقة طردية بين أداء أفراد العينة على صور الاختبار وتحصيلهم الدراسي في مبحث الرياضيات.

د) التحقق من أداء الطلبة عبر الصفوف الدراسية: تم حساب قيم معاملات الصعوبة لفقرات الصور الست للاختبار بمستوييه بعد تطبيقه على العينة الاستطلاعية مقدرة بمتوسط الاداء على الفقرة، والجدول (٨) يبين قيم معاملات الصعوبة للفقرات المشتركة بين كل مستويين وللصور الثلاث لاختبار.

جدول (٨)
قيم معاملات الصعوبة للفقرات المشتركة بين كل مستويين وللصور الثلاث

رقم الفقرة المشتركة	الفقرات المشتركة بين اختبار الصف الخامس واختبار الصف السادس للصورة الأولى (٣ فقرات)		الفقرات المشتركة بين اختبار الصف الخامس واختبار الصف السادس للصورة الثانية (٦ فقرات)		الفقرات المشتركة بين اختبار الصف الخامس واختبار الصف السادس للصورة الثالثة (٩ فقرات)	
	الصف الخامس	الصف السادس	الصف الخامس	الصف السادس	الصف الخامس	الصف السادس
١	٠,٣١	٠,٥٢	٠,٥٢	٠,٦٩	٠,٣٩	٠,٤٣
٢	٠,٣٢	٠,٤٩	٠,٣٠	٠,٤٨	٠,٤٥	٠,٦٤
٣	٠,٤٦	٠,٥٣	٠,٥٥	٠,٥٨	٠,٥١	٠,٥٩
٤			٠,٦٠	٠,٧٢	٠,٤٤	٠,٤٩
٥			٠,٣٣	٠,٤٧	٠,٤٨	٠,٥٤
٦			٠,٤٠	٠,٦٢	٠,٤٥	٠,٦٨
٧					٠,٦١	٠,٦٤
٨					٠,٤٥	٠,٥٥
٩					٠,٤٥	٠,٥٢

يتضح من الجدول (٨) أن قيم معاملات الصعوبة لفقرات الاختبار المشترك بين المستويين تزداد كلما ارتفع المستوى الصفي، أي أن قيم معاملات الصعوبة للفقرات المشتركة لاختبار الصف السادس كانت أكبر من قيم معاملات الصعوبة للفقرات المشتركة لاختبار الصف الخامس ولكل صور الاختبار، وهذه الزيادة في قيم معاملات الصعوبة تعبر عن قدرة الفقرات على التمييز بين المستويات الصفية المختلفة في الأداء عليها، ويعتبر هذا مؤشرًا على صدق الاختبار.

سابعاً) التحقق من ثبات الاختبار بمستوييه من خلال ما يأتي:

باستخدام النظرية الكلاسيكية: تم التحقق من الثباتمن خلال تطبيق الاختبار بمستوييه وبصوره الست مرتين وبفارق زمني مدته ثلاثة أسابيع على أفراد العينة الاستطلاعية والبالغ عددهم (١٤٨) طالبًا وطالبة، منهم (٧٥) طالباً وطالبةً من طلبة الصف الخامس الأساسي، و(٧٣) طالباً وطالبةً من طلبة الصف السادس الأساسي، وتم حساب معامل الثبات بطريقة

الإعادة، ومعادلة كودر- ريتشاردسون (٢٠) والطريقة النصفية، والجدول (٩) يبين قيم معاملات الثبات لكل صورة من صور الاختبار ولكل مستوى.

جدول (٩)
قيم معاملات الثبات لكل صورة من صور الاختبار ولكل مستوى

الصف	الصورة	معادلة كودر- ريتشاردسون (٢٠)	الطريقة النصفية	طريقة الاعادة
الخامس	الأولى	٠,٩٣	٠,٨٨	٠,٨١
	الثانية	٠,٨٣	٠,٨٧	٠,٨٦
	الثالثة	٠,٨٤	٠,٨٥	٠,٨١
السادس	الأولى	٠,٩٠	٠,٨٢	٠,٨٠
	الثانية	٠,٨٧	٠,٨٩	٠,٩٠
	الثالثة	٠,٨٠	٠,٩٢	٠,٨٩

يتبين من الجدول (٩) أن قيم معاملات الثبات مرتفعة ومقبولة.

باستخدام نظرية استجابة الفقرة: تم تقدير قيم معامل الفصل بين الأفراد، ومعامل الثبات للأفراد، وتم العمل على الصورة الثالثة التي عدد فقرات اختبار الجذع المشترك فيها (٩) فقرات، ويوضح الجدول (١٠) نتائج ذلك.

الجدول (١٠)
معامل الفصل بين الفقرات بين الأفراد ومعامل الثبات للاختبار للأفراد

المستوى	معامل الفصل بين الفقرات	معامل الثبات للاختبار	معامل الفصل بين الأفراد	معامل الثبات للأفراد
الأول	٢,٢٨	٠,٩١	٣,١٩	٠,٩٠
الثاني	٢,٠٩	٠,٨٣	٢,١٥	٠,٨٠

يتبين من الجدول (١٠) أن قيمة معامل الفصل للفقرات لمستويي الاختبار كانت أكبر من (٢)، مما يدل على أن فقرات الاختبار كافية لتعريف متصل السمة المراد قياسها، ويتبين كذلك أن قيمة معامل الفصل بين الأفراد لمستويي الاختبار كانت أكبر من (٢)، وهذا يدل على أن عينة الأفراد كافية للفصل بين الفقرات وتعريف متصل السمة الذي يقيسه الاختبار.

ثامناً) تم تصحيح الاختبار، حيث تم إعطاءعلامة واحدة لكل فقرة للإجابة الصحيحة والعلامة (٠) للإجابة غير الصحيحة لكل مستوى، وبهذا تتراوح درجات كل مستوى من مستويات الاختبار من (صفر) إلى (٢٠) درجة وللصورة الثلاث لكل مستوى.

ولتحليل نتائج الدراسة تم استخدام البرامج الإحصائية الآتية:

- SPSS
- BILOG-MG3
- NOHARM
- LDID

نتائج الدراسة

النتائج المتعلقة بالسؤال الأول: هل تتطابق الاستجابات على فقرات الاختبار بمستوييه مع افتراضات نموذج راش؟

للإجابة عن السؤال تم عمل الآتي:

اولاً) التحقق من مطابقة الأفراد والفقرات للنموذج:

بالنسبة للأفراد: لم يجب أي مفحوص عن جميع الفقرات بشكل صحيح أو بشكل خاطئ، وكذلك كانت قيم مستوى الدلالة لقيمة اختبار التوزيع الاحتمالي الهامشي أقل من (٠,٠١)، لجميع المفحوصين على مستويي الاختبار وفي الصور الست، وبذلك لم يتم حذف أي فرد، مما يعني مطابقة جميع أفراد العينة والبالغ عددهم (١٠٣١) لنموذج راش.

بالنسبة للفقرات: لم يوجد أي فقرة أجاب عنها جميع المفحوصين بشكل صحيح أو بشكل خاطئ، وكذلك لم تقل قيمة مستوى الدلالة لقيمة اختبار كاي تربيع لأي فقرة عن (٠,٠١)، وبذلك لم يتم حذف أي فقرة، مما يعني مطابقة جميع فقرات الاختبار وبمستوييه وبصوره الست لنموذج راش.

ثانياً) التحقق من افتراضات نظرية استجابة الفقرة وهي:

أحادية البعد: وتم التحقق منه من خلال:

أ-التحليل العاملي: تم الحديث عنه عند التحقق من صدق الاختبار.

ب- استخدام برنامج NOHARM (Normal Ogive Harmonic Analysis Robust Method)، وتم حساب مؤشرين، وهما: مؤشر (TANAKA)، وقد بلغت قيمته (٠,٩٨٢١) و(٠,٩٧٩٧) لمستويي الاختبار على التوالي، وهذه القيم تدل على تحقق افتراض أحادية البعد للاختبار بمستوييه، فإذا كانت قيمتة قريبة من (١) فإن ذلك يعتبر مؤشرًا على تحقق أحادية البعد (Jasper, 2010)، واستخدام مؤشر الجذر التربيعي لمتوسط مربعات البواقي، وإذا كانت قيمتة قريبة من الصفر وأقل من القيمة الحرجة، هذا يعتبر مؤشراً على تحقق افتراض أحادية البعد، وقد بلغت قيمته (٠,٠٠٥٢) و(٠,٠٠٦١) لمستويي الاختبار على التوالي، وهي

أقل من القيمة الحرجة، التي كانت تساوي (٠,١٧٩٢) و (٠,١٨٤٩)لمستويي الاختبار على التوالي، وبهذا فقد تحقق افتراض أحادية البعد للاختبار بمستوييه.

٢- الاستقلال الموضعي: يتحقق فرض الاستقلال الموضعي إذا تحقق فرض أحادية البعد، وإذا لم يتحقق فرض أحادية البعد، فإن فقرات المقياس ستكون مرتبطة عند مستوى قدرة معين، لأن فقرات المقياس تقيس قدرة ثانيه أو أكثر، وبما أنه تحقق فرض أحادية البعد للاختبار بمستوييه، فإن فرض الاستقلال الموضعي تحقق للاختبار بمستوييه (Hambleton & Swaminathan, 1985).

وتم التحقق من افتراض الاستقلال الموضعي كذلك باستخدام برنامج (LDID) (Local Dependence Indices for Dichotomous Items)، فقد تم حساب نسبة ارتباطات البواقي بين كل زوج من فقرات الاختبار التي حققت الاستقلال الموضعي والتي لم تحققه وفقاً لمؤشر Q3 للاستقلال الموضعي لمستويي الاختبار، وكانت تساوي (٨٠,٨١٪) و(٧٧,٤٨٪) لمستويي الاختبار الأول والثاني على التوالي، وهذا يدل على تحقق افتراض الاستقلال الموضعي، إذ يشترط البرنامج أن تزيد هذه النسبة عن (٥٠٪) حتى يتحقق الاستقلال الموضعي (Kim, Cohen & Lin, 2005).

ولمزيد من الدقة كذلك تم حساب معاملات الارتباط بين فقرات كل اختبار على العينة بأكملها، وحساب معاملات الارتباط بين فقرات كل اختبار على المجموعة ذات القدرات العليا من الطلاب (٢٧٪) من العينة الكلية، وكذلك حساب معاملات الارتباط بين مفردات كل اختبار على المجموعة ذات القدرات الدنيا لعينة الدراسة (٢٧٪) من العينة الكلية، ثم حساب متوسطات معاملات الارتباط في الحالات الثلاث، والجدول (١١) يبين نتائج ذلك

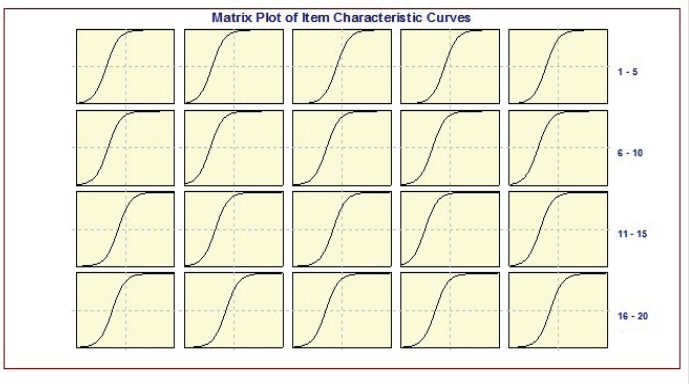
جدول(١١)
متوسط معاملات الارتباط بين فقرات كل مستوى من مستويي الاختبار على العينة ككل والمجموعتين العليا والدنيا

الصف	المجموعة	العدد	متوسط معاملات الارتباط بين كل فقرتين	متوسط أعلى معامل ارتباط بين كل فقرتين	متوسط أقل معامل ارتباط بين كل فقرتين
الخامس	الكلية	٥٠٨	٠,٠٣٢	٠,٠٢٤	٠,٠١١
	العليا	١٣٧	٠,٠٢١	٠,٠٢٠	٠,٠٠٩
	الدنيا	١٣٧	٠,٠١٣	٠,٠١٧	٠,٠٠٦
السادس	الكلية	٥٢٣	٠,٠٤٥	٠,٠٣١	٠,٠١٧
	العليا	١٤١	٠,٠٢٧	٠,٠٢٣	٠,٠٠٨
	الدنيا	١٤١	٠,٠١٧	٠,٠١١	٠,٠٠٧

يتبين من الجدول (١١) أن قيمة متوسط معاملات الارتباط في حالة العينة ككل كانت أكبر من متوسط معاملات الارتباط في حالة المجموعة العليا والمجموعة الدنيا، وجميع قيم متوسطات معاملات الارتباط للعينة ككل، وللمجموعة العليا، وللمجموعة الدنيا كانت قريبة من الصفر، وهذا يدل على أن فقرات الاختبار بمستوييه لا يوجد من بينها فقرات تتأثر الإجابة عنها بفقرات أخرى، وهذا ما يشار إليه بالاستقلال الموضعي (Onder, 2007).

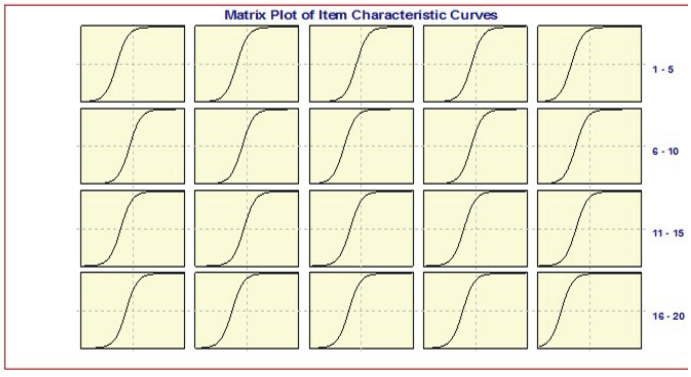
٣- التحرر من السّرعة: خلال تطبيق الدراسة: أنهى جميع أفراد عينة الدراسةالإجابة عن مستويي الاختبار خلال الوقت المحدد للإجابة، أي أن إخفاق بعض الطلبة في الإجابة عن بعض فقرات الاختبار يعزى إلى عامل القدرة ولا يعزى إلى عامل السرعة، فقد أكمل (٧٥٪) من المفحوصين الإجابة عن فقرات الاختبار خلال الوقت المحدد للاختبار (٦٠ دقيقة)، وأجاب المفحوصون عن (٨٠٪) من عدد الفقرات خلال الوقت المحدد للاختبار (٦٠ دقيقة)، وتحقق هذا للاختبار بمستوييه، وهذا يدل على تحقق افتراض التحرّر من السرعة للاختبار.

٤- افتراض منحنى خصائص الفقرة: للتحقق من هذا الافتراض تم رسم منحنى خصائص الفقرة لجميع فقرات الاختبار بمستوييه، والشكلان (٥، ٤) يبينان نتائج ذلك.



شكل (٤)
منحنى خصائص الفقرة لجميع فقرات المستوى الأول للاختبار

يتبين من الشكل (٤) تحقق افتراض اطرادية منحنى خصائص الفقرة لجميع فقرات المستوى الأول للاختبار.



شكل (٥)
منحنى خصائص الفقرة لجميع فقرات المستوى الثاني للاختبار

يتبين من الشكل (٥) تحقق افترض اطرادية منحنى خصائص الفقرة لجميع فقرات المستوى الثاني للاختبار.

ثالثا) التحقق من افتراضات نموذج راش، وهي:

أ- تساوي معاملات التمييز: لكي يتحقق هذا الافتراض يجب أن تكون قيم معاملات التمييز المحسوبة بالنظرية الكلاسيكية للفقرات واقعه ضمن الحدود، حسب المعادلة:

مدى معاملات التمييز = (متوسط معاملات التمييز $\pm$ ٠,١٥) ..... (٤)

حسب ماذكره كل من هامبلتون وسومیناثان (Hambleton & Swaminathan, 1985)، وتم حساب قيم معاملات التمييز لاستجابات العينة الاستطلاعية باستخدام النظرية الكلاسيكية، والجدول (١٢) يبين قيم المتوسط الحسابي وأعلى قيمة، وأدنى قيمة والمدى لمعاملات التمييز لكل مستوى.

جدول (١٢)
المتوسط الحسابي وأعلى وأدنى قيمة والمدى لمعاملات التمييز لكل مستوى

المستوى (الصف)	الصورة	المتوسط الحسابي	أدنى قيمة	أعلى قيمة	المدى
الخامس	الأولى	٠,٤٩	٠,٤٢	٠,٦١	٠,٣٤ -٠,٦٤
	الثانية	٠,٤٨	٠,٤١	٠,٦٠	٠,٦٣ -٠,٣٣٠
	الثالثة	٠,٥٨	٠,٤٥	٠,٧٢	٠,٤٣ -٠,٧٣
السادس	الأولى	٠,٥٣	٠,٤٦	٠,٦٥	٠,٣٨ -٠,٦٨
	الثانية	٠,٥١	٠,٤٢	٠,٦٣	٠,٣٦ -٠,٦٦
	الثالثة	٠,٥٦	٠,٤٠	٠,٧١	٠,٧١ -٠,٤١٠

يتبين من الجدول (١٢) أن جميع قيم معاملات التمييز لفقرات المستويين للصور الست وقعت ضمن المدى المحسوب بالمعادلة (١) لكل صورة، وهذا يبين تحقق افتراض تساوي معاملات التمييز لفقرات الاختبار بمستوييه وبصوره الست.

ب) التخمين: تم التحقق من عدم التخمين أو انخفاض مؤشر التخمين من خلال فحص أداء (١٠٪) من الطلبة الذين حصلوا على أقل العلامات على الفقرات الأكثر صعوبة (Hambleton & Swaminathan, 1985)، لمقارنة نسبة هؤلاء الطلبة الأقل قدرة والذين أجابوا إجابة صحيحة عن الفقرات الصعبة بالقيمة النظرية للتخمين العشوائي في حالة البدائل الأربع، التي تساوي (٠,٢٥)، تم أخذ أكثر (٤) فقرات صعبة من كل مستوى، والجدول (١٣) يوضح نسبة الطلبة منخفضي القدرة الذين أجابوا إجابة صحيحة عن الفقرات الصعبة في الاختبار بمستوييه.

الجدول (١٣)
نسبة الطلاب منخفضي القدرة الذين أجابوا إجابة صحيحة على الفقرات الصعبة في الاختبار بمستوييه

المستوى	رقم الفقرة	نسبة الطلبة الذين أجابوا إجابة صحيحة
الأول	١	٠,٢٠
	٣	٠,١٦
	١٠	٠,١٨
	١١	٠,٢١
الثاني	٤	٠,١٤
	٧	٠,١٩
	٩	٠,٢٣
	١٢	٠,١١

يتبين من الجدول (١٣) أن نسبة الطلبة منخفضي القدرة الذين أجابوا إجابة صحيحة عن الفقرات الصعبة هي نسب منخفضة وأقل من (٠,٢٥)، وبهذا يتحقق افتراض التخمين لفقرات الاختبار بمستوييه.

ج- معاملات الصعوبة : تم حساب معامل الصعوبة لكل فقرة من فقرات اختبار الصف الخامس للصورة الثالثة، واختبار الصف السادس للصورة الثالثة فقط للعينة الكلية باستخدام برنامج (Bilog-mg3)، ولم يتم عرض نتائج الصورتين الأولى والثانية للصفين الخامس والسادس الأساسيين، لأنه في نظرية استجابة الفقرة تقدير قدرة الفرد يكون مستقلاً عن عينة الفقرات التي تطبق عليه، وهذا ما يعرف بـ (Item Free)، وتقدير خصائص الفقرات يكون مستقلا عن عينة الأفراد التي طبقت عليهم الفقرات، وهذا ما يعرف بـ (Person Free)، والجدول (١٤) يبين نتائج ذلك.

جدول (١٤)
قيم معاملات الصعوبة لكل فقرة من مستويي الاختبار باستخدام نظرية استجابة الفقرة

رقم الفقرة	معامل الصعوبة لفقرات الصف الخامس	معامل الصعوبة لفقرات الصف السادس
١	٠,٨٥٧	١,٢١٧-
٢	٠,٤٢٥	١,٣٦٩-
٣	٠,١٢٤-	٠,٣٥٦-
٤	٠,٠٤٧-	٠,٣٧٦-
٥	١,٠٣٩	٠,٧٩٢
٦	٠,٠٠٢-	٠,٢١٣
٧	٠,١٢٥-	٠,٦٨٥
٨	١,٠٠٥-	٠,٩٧٤-
٩	٠,٠٩٤	٠,٤٢٤
١٠	٠,٦٨٥-	١,٢٩٠
١١	٠,٧٥٠	٠,٤٧٢-
١٢	١,٣٣٢	١,٠٥٨-
١٣	٠,٨٨٤	٠,٥٦٨
١٤	٠,٤٤٤-	٠,٥٨٤-
١٥	٠,٦٧٨	١,١٥٨
١٦	٠,٩٦٣	٠,٨٤٢-
١٧	١,٠٧٤	٠,٥٠١-
١٨	٠,٨٣٢	١,٠٢٦-
١٩	١,٢١٣	٠,٩٤٣
٢٠	١,٥٣١	٠,٥٣٠-

يتبين من الجدول (١٤) أن قيم معاملات الصعوبة تراوحت بين (-١,٠٠٥) (١,٥٣١) للمستوى الأول، وبين (-١,٣٦٩) وبين (١,٢٩٠) للمستوى الثاني، وهي قيم مقبولة حسب نظرية استجابة الفقرة.

نتائج السؤال الثاني: هل تتأثر دقة المعادلة الخطية العمودية لقدرات الطلبة في اختبار الرياضيات باستخدام نموذج راش في تصميم المجموعات العشوائية مع الجذر المشترك باختلاف عدد فقرات اختبار الجذع المشترك؟

للإجابة عن السؤال تم معادلة الدرجات لمستويي الاختبار (اختبار الصف الخامس واختبار الصف السادس) بالطريقة الخطية باستخدام نموذج راش كما في المعادلة (١) باستخدام تصميم المجموعات العشوائية باختبار الجذع المشترك، ثم تم حساب قيمة الصدق التقاطعي باستخدام المعادلة (٢) وقيمة خطأ المعادلة باستخدام برنامج (Bilog-mg3) للتحقق من

فاعلية المعادلة العمودية باستخدام نموذج راش لكل مستوى على كل صورة للاختبار، والجدول (١٥) يبن نتائج ذلك.

جدول (١٥)
قيمة الصدق التقاطعي لصور الاختبار

صورة الاختبار	قيمة الصدق التقاطعي	قيمة الخطأ المعياري
الأولى (عدد فقرات الجذع المشترك ٣)	٠,٢٦٣	٠,٨٧٢
الثانية (عدد فقرات الجذع المشترك ٦)	٠,٢٢٢	٠,٧٧٤
الثالثة (عدد فقرات الجذع المشترك ٩)	٠,٠٩٩	٠,٣٠١

يتبين من الجدول (١٥) أن قيمة معامل الصدق التقاطعي تراوحت بين (٠,٠٩٩) (٠,٢٦٣)، وكانت أقل قيمة للصورة الثالثة (عدد فقرات الجذع المشترك ٩)، وهذا يدل على أن المعادلة العمودية كانت أكثر دقة للصورة الثالثة، أي عندما كان عدد فقرات الجذع المشترك (٩)، ويتبين من الجدول (١٥) أن قيمة الخطأ المعياري للمعادلة تراوحت بين (٠,٣٠١) وبين (٠,٨٧٢)، وكانت أقل قيمة للصورة الثالثة كذلك، وهذا يدل على أن المعادلة العمودية كانت أكثر دقة للصورة الثالثة، أي عندما كان عدد فقرات الجذع المشترك يساوي (٩)، وهذا يدل على أن المعادلة العمودية تكون اكثر دقة عندما تزيد عدد الفقرات المشتركة بين الاختبارات، أي كلما زاد عدد فقرات اختبار الجذع المشترك كلما زادت دقة المعادلة العمودية، وقد يعزى ذلك إلى أن نسبة تمثيل المحتوى للاختبار تكون أكثر عندما تكون الفقرات المشتركة أكثر، وتتفق هذه النتيجة مع نتيجة دراسة كل من (الحمادنه، ٢٠١٢; Sinharay and Holland 2007;Yang and Houang, 1997)، وقد يعزى كذلك إلى طبيعة المحتوى الذي بني الاختبار في مستوييه لقياسه، فلا يوجد اختلاف في نوع المحتوى، ولكن الاختلاف كان في المستوى (الصف)، فالاختبار في مستوييه يقيس القدرة نفسها (الهندسة) لكن بشكل متدرج في الصعوبة عبر الصفين الدراسيين.

نتائج السؤال الثالث: ما أفضل نسبة لفقرات اختبار الجذع المشترك في المعادلة العمودية؟

للإجابة عن السؤال تم حساب الفروق بين قيم الصدق التقاطعي وقيم الخطأ المعياري لصور الاختبار الثلاث، والجدول (١٦) يبين ذلك.

جدول (١٦)
الفروق بين قيم الصدق التقاطعي وقيم الخطأ المعياري بين صور الاختبار الثلاث

الفقرق بين قيم الصورتين	الصدق التقاطعي	الخطأ المعياري
الصورة الأولى التي نسبة فقرات اختبار الجذع المشتركفيها (١٥٪) والصورة الثانية التي نسبة فقرات اختبار الجذع المشترك فيها (٣٠٪)	٠,٠٤١	٠,٠٩٨
الصورة الأولى التي نسبة فقرات اختبار الجذع المشترك فيها (١٥٪) والصورة الثالثة التي نسبة فقرات اختبار الجذع المشترك فيها (٤٥٪)	٠,١٦٤	٠,٥٧١
الصورة الثانية التي نسبة فقرات اختبار الجذع المشترك فيها (٣٠٪) والصورة الثالثة التي نسبة فقرات اختبار الجذع المشترك فيها (٤٥٪)	٠,١٢٣	٠,٤٧٣

يتبين من الجدول (١٦) أن أقل قيمة للفرق بين قيم الصدق التقاطعي وقيم الخطأ المعياري لصور الاختبار الثلاث كانت بين الصورة الأولى التي نسبة فقرات اختبار الجذع المشترك فيها (١٥٪) وبين الصورة الثانية التي نسبة فقرات اختبار الجذع المشترك فيها (٣٠٪)، وكانت تساوي (٠,٠٤١) للصدق التقاطعي و(٠,٠٩٨) للخطأ المعياري، وأن أكبر قيمة للفرق بين قيم الصدق التقاطعي وقيم الخطأ المعياري لصور الاختبار الثلاث كانت بين الصورة الأولى التي نسبة فقرات اختبار الجذع المشترك فيها (١٥٪) الصورة الثالثة التي نسبة فقرات اختبار الجذع المشترك فيها (٤٥٪)، وكانت تساوي (٠,١٦٤) للصدق التقاطعي و(٠,٥٧١) للخطأ المعياري، وبذلك تكون أفضل نسبة لفقرات اختبار الجذع المشترك في المعادلة الخطية العمودية هي (٤٥٪)، وتتفق هذه النتيجة مع ما اقترحه انجوف (Angoff, 1984) بأن لا تقل نسبة فقرات اختبار الجذع المشترك عن (٢٠٪) من عدد فقرات الاختبارات المراد معادلتها، وقد يعزى ذلك إلى أن اختبار الجذع المشترك يمثل نسبة أكبر من المحتوى المشرك بين الصفين، وأن صدق وثبات الاختبارات تزيد قيمتاهما بزيادة عدد فقراتها، وبهذا يكون اختبار الجذع المشترك صورة مصغرة عن اختبار كل صف.

التوصيات

توصي الدراسة بما يأتي:

- ألا تقل نسبة الفقرات في اختبار الجذع المشترك عن (٤٥٪) من عدد فقرات الاختبارات التي يراد معادلتها.
- زيادة عدد الفقرات في اختبار الجذع المشترك في الاختبارات المراد معادلتها، سواء في المعادلة الأفقية أو المعادلة العمودية.

- إجراء المزيد من الدراسات على المعادلة بنوعيها الأفقية والعمودية باستخدام طرق أخرى للمعادلة.
- إجراء المزيد من الدراسات على المعادلة بنوعيها الأفقية والعمودية واستخدام معايير أخرى للتحقق من فعاليتها.

المراجع

حمادنة، إياد (٢٠١٢). أثر تجانس المحتوى وطول الجذع المشترك على دقة معادلة اختبار ثنائي التدريج ومتعدد المستوى في مبحث الرياضيات.جامعة آل البيت، مجلة المنارة للبحوث والدراسات، ٨(٤)، ٩- ٤٣.

السحيمات، منير (٢٠٠٧). بناء اختبار تحصيلي في الرياضيات متعدد المستويات لطلبة الصفوف (الثامن والتاسع والعاشر الأساسي) في الأردن. رسالة ماجستير غير منشورة، جامعة مؤتة، الكرك.

الطراونة، صبري(٢٠٠٥). تطوير اختبار رياضيات متعددة المستويات للصفوف الأساسية (٣- ٦) بفقرات متعددة التدريج. اطروحة دكتوراه غير منشورة، جامعة عمان العربية، عمّان الأردن.

علام، صلاح الدين(٢٠٠٢). القياس والتقويم التربوي والنفسي: أساسياته وتطبيقاته وتوجيهاته المعاصرة. القاهرة: دار الفكر العربي.

المدانات، رائد (٢٠١٢). مقارنة فاعلية طريقة تكر الخطية وطريقة ليفين في معادلة الاختبارات عند استخدام التصميم القائم على اختبار جذع مشترك ومجموعات غيرعشوائية.مجلة العلوم النفسية والتربوية، ١٣(٢)، ٣٦٦-٣٩٤.

وزارة التربية والتعليم (٢٠١٥). الرياضيات للصف الخامس.الجزء الثاني، الطبعة الثانية، عمان: وزارة التربية والتعليم.

وزارة التربية والتعليم (٢٠١٥). الرياضيات للصف السادس. الجزء الثاني، الطبعة الثانية، عمان: وزارة التربية والتعليم.

Angoff, W. (1984). *Scales, Norms, and Equivalent Scores*. Educational Testing Service. Princton: New Jersey.

Baker, F. (2001). *Thebasicofitemresponsetheory*. Maryland: University of Maryland.

Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. NewYork: Holt Rinehart and Winston.

Hambleton, R & Swaminathan, H. (1985). *Item Response Theory: Principle and Application*. Boston: Kluwer Nijhoff Publishing.

Hambleton, R. Swaminathan, H. & Roges, H. (1991). *Fundamentals of item response theory*. London, Sage Publications.

Jasper, F. (2010). Applied dimensionality and test structure assessment with the START-M mathematics test. *The International Journal of Educational and Psychological Assessment, 6*(1), 104-125.

Kim, S., Cohen, A. & Lin, Y. (2005). *LDID: acomputer program for local dependence indices for dichomous item,* Georgia, University of Georgia.

Kolen, M. & Brennan, L. (2004). *Test equating, scaling, and linking methods and practices*. Second edition, New York: Springer.

Kolen, M. & Whiteny, R. (1981). Comparison of four procedures for equating the test of general education development. *Journal of Educational Measurment, 9,* 279-293.

Onder, I. (2007). An investigation of goodness of model data fit model Veri Uyumunun Araştirilmasi. *Hacettepe Universitesi Egitim Fakultesi Dergisi, 32,* 210-220.

Sinharay, S. & Holland, P. (2007). *The correlation between the scores of atest and an anchor test*. Research Report, Ets, March, by Educational Testing Service (ETS).

Thomas, A. W. (1998). *Aprimerof Item item response theory*. NewYork: Oklahoma City.

Yang, W. & Houang, R. (1997). *The effect of anchor length and equating method on the accuaracy of test equating: comparisons of linear and irt-based equating using anchor test design*. Paper presented at the annual meeting of the Amerrican educational research assosiation, New York, April 8-12, 1996.