أثر ترتيب فقرات اختبار الاختيار من متعدد وفقاً لمعالم الفقرة على تقدير قدرة المفحوصي ودالة المعلومات للاختبار وفقاً للنموذج ثلاثي المعلمة

د. حابس سعد الزبون
جامعة الحسين بن طلال – كلية العلوم التربوية
قسم المناهج وأساليب التدريس.

د. راجي عوض الصرايرة
جامعة مؤتة – كلية العلوم التربوية
قسم علم النفس التربوي

الملخص: هدفت هذه الدراسة إلى التحقق من أثر ترتيب فقرات اختبار الاختيار من متعدد وفقاً لمعالم الفقرة على تقدير قدرة المفحوصين ودالة المعلومات للاختبار وفقاً للنموذج ثلاثي المعلمة، ولتحقيق أهداف الدراسة قام الباحثان بالاعتماد على برنامج (WINGEN3) لتوليد البيانات حيث تم توليد معالم الفقرات (الصعوبة، التمييز ، التخمين) لاختبار مكون من (٦٠) فقرة و تم الاعتماد على معالم الفقرات الناتجة لتوليد القدرة لـ (٥٠٠٠) مفحوص باستخدام على النموذج ثلاثي المعلمة. تلا ذلك ترتيب الفقرات حسب معاملات الصعوبة (تنازلي، تصاعدي، عشوائي) وكذلك ترتيب الفقرات حسب معاملات التمييز (ai) ومعاملات التخمين(ci) وتم توليد قدرات المفحوصين بناء على ترتيب معاملات الصعوبة والتمييز والتخمين لتكوين ثلاث صور للاختبار. كما تم استخدام برنامج (BILOG-MG3) من أجل فحص جودة مطابقة النموذج من خلال مطابقة الأفراد والفقرات للنموذج اللوجستي ثلاثي المعلمة و إيجاد دالة المعلومات لفقرات الصور الثلاث للاختبار والاختبار ككل.كشفت نتائج الدراسة عن عدم وجود فروق ذات دلالة إحصائيّة عند مستوى الدلالة (α=٠,٠٥) بين المتوسطات الحسابية لتقديرات معلمة القدرة تعزى لاختلاف مؤشر الصعوبة والتمييز والتخمين للفقرة (عشوائي، تصاعدي، تنازلي). كما كشفت نتائج الدراسة أن أعلى قيمة لدالة معلومات الاختبار كانت عندما تم ترتيب فقرات الاختبار حسب معاملات التمييز تنازلياً والترتيب التصاعدي لفقرات الاختبار حسب معاملات الصعوبة.

الكلمات المفتاحية: ترتيب فقرات الاختبار، دالة المعلومات ، النموذج ثلاثي المعلمة

The effect of multiple-choice test items order as per item parameters on estimating examinees' ability and test information function according to the three-parameter logistic model

Abstract This study aimed to verify the effect of multiple choice test item order according to item parameters on estimating examinees' ability and test information function as per to the three-parameters logistic model. In order to achieve the objectives of the study, WINGEN3 program was utilized to generate data parameters including difficulty, discrimination, and guessing of a 60-item test. The resultant item parameters resulted were used to generate the ability of (5000) examinees through using three-parameter logistic model.Items were ordered according to difficulty coefficients descendingly, ascendingly, and randomly. Furthermore, items were sorted according to discrimination (ai) and guessing (ci) coefficients. Examinees' abilities were generated based on discrimination and guessing difficulty coefficients order to build three forms for the test. BILOG-MG3 was used to examine the extent of the model congruity quality between examinees and items against the three-parameter logistic model and to find the information coefficient of the three forms of the test and the test as a whole. The results of the study revealed that there were not any significant statistical differences at (α = 0.05) between the arithmetic means of ability parameter estimates which are attributed to difference in difficulty discrimination, and guessing coefficients of the items (randomly, ascendingly, descendingly). The results also demonstrated that the highest value for

Keywords: test item order, information function, three-parameter logistic model

المقدمة يعرف الاختبار التحصيلي بأنه أداة قياس يتم إعدادها وفق طريقة منظمة من عدة خطوات تتضمن مجموعة من الإجراءات التي تخضع لشروط محددة، بغرض تحديد درجة امتلاك الفرد لسمة أو قدرة معينة من إجابته عن الفقرات التي تمثل السمة أو القدرة المرغوب قياسها.(عوده، ٢٠١٠)

وهناك أخطاء ترافق عملية القياس النفسي، قد تؤدي هذه الأخطاء في تقدير الدرجة الحقيقية للسمة المقاسة ، إذ إن الدرجة الملاحظة التي نحصل عليها من تطبيق الاختبار تتضمن أخطاء، تزداد قيمتها أو تقل تبعاً لمصدر هذه الأخطاء، ومن أهم هذه المصادر :المقياس المستخدم نفسه، والعينة، والظروف التجريبية. التي أشارت إليها النظرية الكلاسيكية في القياس إذ تقوم هذه النظرية على مجموعة من الافتراضات الأساسية مفادها: أن أي عملية قياس يتخللها هامش للخطأ وذلك حسب الصيغة التالية :

$$X=T+E$$

حيث X: الدرجة الملاحظة، T: الدرجة الحقيقية، E: قيمة الخطأ. ويُشكل الاختبار منطق النظرية الكلاسيكية، ولذلك أُطلق عليها نظريـة الاختبار (Test Theory) (Crocker and Algina, 1986).

وتُعتبر الاختبارات وسيلة من أهم الوسائل التي يُعتمد عليها في قياس وتقويم قدرات الطلاب؛ وذلك من أجل تحديد مستواهم التحصيلي، ومعرفة مدى تحقيق الأهداف السلوكية أو النواتج التعليمية، وما يقدمه المعلم من نشاطات تعليمية مختلفة تساعد على رفع الكفايات التحصيلية لدى الطلاب؛ لذلك حرص المشرفون التربويون على أن تكون هذه الاختبارات ذات كفاءة عالية في عملية القياس والتقويم (النجار، ٢٠١٠).

وتتنوع الاختبارات تبعاً للأهداف التي وضعت من أجلها، فما يصلح لمادة دراسية قد لا يصلح لأخرى، وما يصلح لهدف قد لا يصلح لآخر، وما يصلح لمستوى من الطلاب قد لا يصلح لمستوى آخر، فعلى المدرس أن يختار النوع المناسب، ومن أهم هذه الأنواع اختبارات التحصيل (Achievement Tests) التي تهدف لقياس تحصيل الطلبة (الكيلاني وعدس، ١٩٨٦).

ويعتمد الاتجاه المعاصر في القياس النفسى والتربوى على نظرية استجابة الفقرة أو نظرية السمات الكامنة التى تفترض أنه يمكن التنبؤ بأداء الأفراد أو يمكن تفسير أدائهم في اختبار نفسي أو تربوي معين بناء على خاصية أو خصائص معينة مميزة لهذا الأداء تسمى السمات Traits ، وتحاول هذه النظرية تقدير درجات الأفراد في هذه السمات . ونظراً لصعوبة ملاحظة هذه السمات بطريقة مباشرة أو قياسها بصورة مباشرة ، فإنه يجب تقديرها أو الاستدلال عليها من خلال استجابات الأفراد الملاحظة على الفقرات الاختبار الذى يقيس السمة أو القدرة المطلوبة ولهذا أطلق عليها " نظرية السمات الكامنة " أو نظرية الاستجابة للمفردة (عماد عبد المسيح ، ١٩٩١: ٤٤٧).

وتهدف نظرية الاستجابة للفقرة والنماذج الرياضية المرتبطة بها إلى تقدير جميع إحصاءات الفقرة والقدرة ، وكلما كان هناك جودة مطابقة (Goodness of Fit) بين خصائص النموذج المستخدم وخصائص البيانات أدى ذلك إلى الحصول على تقديرات دقيقة لهذه الإحصاءات ، حيث يجب وضع كل من الأفراد والفقرات على مقياس للقدرة ، ويتم ذلك من خلال عمليات التقدير طالما يوجد علاقة تقارب ممكنة بين الاحتمالات المتوقعة للممتحنين والاحتمالات الواقعية الفعلية لأدائهم في كل مستوى من مستويات القدرة ، مع الأخذ في الاعتبار أن تقديرات إحصاءات الفقرة وتقديرات قدرة الممتحنين يجب أن تعدل وتراجع بصفة مستمرة حتى يمكن الحصول على أقصى اتساق ممكن بين التنبؤات على أساس تقديرات القدرة ومتغيرات الفقرة وبين البيانات الواقعية الفعلية للاختبار ويتم ذلك من خلال برامج الحاسب الآلي المعدة لذلك (Hambleton &Swaminathan , 1989) . ونتج عن نظرية استجابة الفقرة مجموعة من النماذج، تعرف باسم نماذج السمات الكامنة (Latent Trait Models) ، تهدف إلى تحديد العلاقات بين أداء الفرد في الاختبار، والسمة التي يمكن وراء هذا الأداء وتفسره، وتعد النماذج اللوجستية ذات المعلمة الواحد (نموذج راش) و ذات المعلمتين والثلاثة معالم من أهم النماذج واسعة الانتشار ، وتختلف هذه النماذج عن بعضها في عدد معالم الفقرة ، ويمثل النموذج اللوجستي ثلاثي المعلمة الشكل العام للنماذج اللوجستية؛ لأنه يحتوي على المعالم الثلاثة الممكنة للفقرة، وهي الصعوبة، والتمييز، التخمين (

bi,ai,ci) على الترتيب، ويعبر عنه من خلال المعادلة التالية ، التي تقيس احتمال إجابة المفحوص ذي القدرة (Θ) على الفقرة (i).

$$P_i(\Theta) = C_i + \frac{1 - C_i}{1 + e^{-1.7a_i(\theta_i - b_i)}}$$

وتفترض نماذج نظرية استجابة الفقرة عدد من الافتراضات في البيانات التي ستطبق عليها ومن هذه الافتراضات :

١- أحادية القياس (البعد) Unidimensionality :

ولا تعنى أحادية القياس بساطة المتغير موضوع القياس ، أى بساطة ما تقيسه الفقرة ، وإنما تعنى أن فقرات المقياس تكون متجانسة فيما بينها وتقيس في أساسها نفس السمة ومعنى ذلك أن أي فقرة من الفقرات المتدرجة الصعوبة يتطلب في حلها نفس النوع من الإجراءات والعمليات السلوكية ، ولكنها تختلف فيما بينها من حيث تدرج صعوبتها فقط .

ويرى "هامبلتون وسوامينثان" (١٩٨٩) أنه يصعب تحقيق هذا الافتراض ، نظراً لوجود بعض العوامل التى تؤثر في أداء الأفراد على الاختبار مثل مستوى الدافعية ، وقلق الاختبار ، والقدرة على الإجابة بسرعة والحكمة الاختبارية ، والتخمين في إجابة بعض فقرات الاختبار.

٢- استقلالية القياس (الاستقلال الموضعي) Local Independence :

ويقصد بهذا الافتراض أن تكون استجابات الفرد للفقرات المختلفة في الاختبار مستقلة استقلالاً إحصائياً، وهذا يعنى ألا تؤثر استجابة الفرد لإحدى الفقرات في استجاباته للفقرات الأخرى . ويتضح هذا في :

أ- تحرر القياس من توزيع العينة المستخدمة Sample Free وهذا يعنى ثبات تقدير معالم الفقرة واستقرارهما بالرغم من اختلاف عينة الأفراد المستخدمة في تدريج المقياس طالما أنها عينة ملائمة .

ب- تحرر القياس من مجموعة الفقرات المستخدمة Item Free وهذا يعنى ثبات تقدير قدرة الفرد واستقرارهما بالرغم من اختلاف مجموعة الفقرات المستخدمة في القياس ، طالما أنها فقرات ملائمة ، وطالما أن هذه المجموعات المختلفة من الفقرات تقع على ميزان تدرج واحد ، أي أنها تعرف متغيراً واحداً .

٣- منحنى خصائص الفقرة :

فالمنحنيات المميزة للفقرة هى دوال رياضية تربط بين احتمال نجاح الفرد في الإجابة على فقرة ما والقدرة التى تقيسها مجموعة الفقرات التى يشتمل عليها الاختبار ، أو هو انحدار الدرجة التى يحصل عليها الفرد في إحدى الفقرات على قدرته .

وتُعدّ دالة المعلوماتية الفقرة (item information function, IIF) من الاحصاءات المهمة في النظرية الحديثة في القياس، حيث يمكن من خلالها تحديد الخطأ المعياري في التقدير، حيث إنه يمكن من إيجادها تحديد الخطأ المعياري في التقدير' وذلك بالاعتماد على الأرجحية العظمى لمعلم القدرة، حيث إن تباين الخطأ في تقدير القدرة يساوي معكوس دالة المعلومات، وكما أن مصفوفة التباين المشترك (variance covariance matrix) للتقديرات تساوي معكوس مصفوفة المعلومات(information matrix) لتقديرات معالم الفقرة (Hambleton & Swaminathan,1993) .

وتلعب دالة معلومات الاختبار دوراً مهماً في النظرية الحديثة في القياس، إذ يمكن من خلالها تحديد الخطأ المعياري في التقدير، وتتمتع دالة معلومات الاختبار، التي تمثل مجموع دوال معلومات الفقرات عند مستوى معين من القدرة بميزة، وهي كون دالة معلومات الاختبار مستقلة عن عينة المفحوصين، وبذلك تقدم النظرية الحديثة في القياس مميزات إضافية، فيما يتعلق بزيادة القدرة على تقدير أخطاء القياس(Brannick, 2003) . كما أن الخطأ المعياري في التقدير يرتبط عكسياً مع الجذر التربيعي لدالة معلومات الاختبار وفق العلاقة الآتية:

$$SE(\theta) = \frac{1}{\sqrt{I(\theta)}}$$

وهذا يعني أن الخطأ المعياري في التقدير، يكون أقل ما يمكن عند مستويات القدرة التي تناظر أقصى معلوماتية.

ومن الاستخدامات المفيدة لدالة معلومات الفقرة، إمكانية معرفة مدى مساهمة كل فقرة في دالة معلومات الاختبار بشكل مستقل عن الفقرات الأخرى للاختبار، فإذا كان لدينا فكرة جيدة عن قدرات مجموعة المفحوصين، فيمكن انتقاء فقرات الاختبار التي تعظم المعلومات التي يقدمها الاختبار في المدى الذي تتوزع فيه قدرات المفحوصين في الاختبار.

وتستخدم دالة معلومات الفقرة لاختيار الفقرات عند بناء الاختبارات، بالاعتماد على النظرية الحديثة في القياس، على افتراض انّ معلومات الفقرة تتغير عبر مستويات السمة المختلفة؛ وعليه فإنه من الممكن اختيار فقرات تقدم دقة قياس مرتفعة عند نقطة معينة على متصل السمة، كما أن الفقرات التي تتضمن معالم تمييز كبيرة، تقدم معلومات أكبر عن قدرة المفحوصين؛ وبالتالي الحصول على دقة أكبر، وتعطى دالة المعلومات للاختبار من خلال المعادلة الآتية:

$$I(\theta)=\sum_{i=1}^{n}\frac{P_i'(\theta)^2}{P_i(\theta)Q_i(\theta)}$$

حيث إنّ:

I(θ) : دالة معلومات الاختبار.

(θ) : معلم القدرة للمفحوصين.

$P_i(\theta)$: دالة استجابة الفقرة $Q_i(\theta)=1-P_i(\theta)$

$P_i'(\theta)$: المشتقة الأولى لدالة استجابة الفقرة.

وعليه يمكن انتقاء فقرات الاختبار اعتماداً على كمية المعلومات التي تساهم بها الفقرات في كمية المعلومات الكلية للاختبار.

وهناك عدد من الخطوات المهمة التي من الممكن الاعتماد عليها عند بناء الاختبار وهي:

وصف شكل دالة المعلومات المستهدفة في الاختبار (target information).

اختيار فقرات تغطي دوال معلوماتها مستويات الصعوبة المناظرة لدالة المعلومات المستهدفة.

بعد إضافة كل فقرة إلى الاختبار يتم حساب دالة معلومات الاختبار لفقرات الاختبار التي يتم اختبارها.

الاستمرار في اختيار فقرات الاختبار، حتى تقترب دالة معلومات الاختبار من دالة المعلومات المستهدفة.

وتتأثر دالة المعلومات بمعالم الفقرة، ففي النماذج اللوجستية الأحادية والثنائية تكون أعلى كمية معلومات عند معلم الصعوبة للفقرة (عند تساوي القدرة مع الصعوبة)؛ وذلك لأن شكل دالة معلومات الفقرة بشكل عام قريب من الشكل الجرسي (bell shape) ، بينما في حالة النموذج اللوجستي الثلاثي يمكن ملاحظة تأثير معالم الفقرة على دالة معلومات الاختبار من خلال تبسيط المعادلة الرياضية لدالة معلومات الاختبار على النحو الآتي (Hambleton & Swaminathan,1993).

$$I_i(\theta)=\frac{2.89a_i^{\,2}(1-c_i)}{(c_i+e^{(1.7a_i(\theta-b_i))})(1+e^{(1.7a_i(\theta-b_i)})^2}$$

وبالاعتماد على المعادلة الرياضية السابقة يمكن ملاحظة ما يأتي:

تكون كمية المعلومات أكبر بشكل عام عندما تكون معلمة التمييز كبيرة.

تزداد كمية المعلومات كلما اقتربت معلمة التخمين من الصفر.

تكون كمية المعلومات عندما تكون (θ) قريبة من (b_i) أكبر من كمية المعلومات عندما تكون (θ) بعيدة من (b_i) .

وقد بين ستوكنج (stocking,1990) أن تقديرات الفقرة من القضايا المهمة في النظرية الحديثة في القياس، وأنه من خلال إيجاد القيم القصوى لتقديرات القدرة من خلال دالة المعلومات لمعالم الفقرة باستخدام إجراءات المشتقة الأولى للنماذج اللوجستية الأحادية والثنائية والثلاثية يمكن الحصول على تقدير لمعالم الفقرة يؤدي إلى تعظيم دالة المعلومات لمعلم الفقرة.

ويشير هامبلتون وسوامينثان Hambleton & (Swaminathan, 1985) أن إحدى الطرق التي تمكننا من خلال نماذج السمات الكامنة بناء اختبارات أفضل، بحيث تعطي

معلومات أكثر حول السمة الكامنة المقدرة هو استخدام اقتران المعلومات، ويمكن استخدامها أيضا للتأكد من دقة تقدير معالم الفقرة ومعلمة القدرة، وتأتي أهمية اقتران معلومات الفقرة من كون مساهمة كل فقرة باقتران معلومات الاختبار يتحدد بشكل مستقل عن باقي الفقرات في الاختبار، أما في النظرية التقليدية، فإنّ مساهمة الفقرة بثبات الاختبار وصدقه لا يتحدد بشكل مستقل عن باقي فقرات الاختبار. واقتران معلومات الاختبار هي مقدار يتناسب عكسياً مع الخطأ المعياري في التقدير، و بذلك تعتبر تقديراً لثبات الاختبار لأنه بزيادة الخطأ المعياري في التقدير يقل الثبات والعكس صحيح . وفي حالة نموذج راش يمكن كتابة اقتران المعلومات للفقرة كالآتي:

$$Ii(\theta) = Pi(\theta)\,(1 - Pi(\theta))$$

حيث:

$Pi(\theta)$:احتمال الإجابة عن الفقرة (i)إجابة صحيحة

أما اقتران معلومات الاختبار، فهي عبارة عن ناتج جمع اقترانات المعلومات للفقرات جميعها و يمكن التعبير عنها بالصيغة *الآتية*:

$$I(\theta) = \Sigma\, Ii(\theta)$$

حيث:

$Ii(\theta)$: اقتران معلومات الفقرة(i) .

$I(\theta)$: اقتران معلومات الاختبار .

والمعلومات التي يقدمها الاختبار ليس من الضروري أن تكون وهي عند مستوى قدرة معين تعتمد على ، (θ)ثابتة باختلاف قيمة فعلى سبيل المثال (θ) خصائص الفقرات المستخدمة لتقدير المعلومات التي تزودنا بها عند المستويات العليا من القدرة تكون قليلة باستخدام فقرات سهلة جداً لأن كل ذوي القدرة المرتفعة تقريباً سيجيبون عنها إجابة صحيحة، ويمكن للفقرات نفسها أن تزودنا بكمية أكبر من المعلومات عن القدرة لذوي القدرة المنخفضة .(Hulin, Drasgow & Parsons,1983).

وقد أجريت بعض الدراسات التي تناولت أثر ترتيب فقرات الاختبار على بعض خصائص الفقرات والاختبار وقدرات المفحوصين ومن هذه الدراسات دراسة كندال (Kendall, 1964) والتي هدفت إلى بحث إمكانية زيادة أو تخفيض صدق الاختبار نتيجة تغيير طول زمن الاختبار، فقد طور كندال اختباراً في الذكاء مُكوَّناً من (٧٥) فقرةً اختيار من متعدد، مرتبة حسب صعوبتها، ويستعمل للتنبؤ بكفاءة المتطوعين للخدمة العسكرية، ولقد أعد منه صورتين متكافئتين: الأول باللغة الإنجليزيّة، والثاني بالفرنسية، وكان المحك اختبار كفاءة (M. test)، قُدِّم الاختبار التنبؤي تحت الأوقات (١٥، ١٨، ٢٠، ٢٢، ٢٥، ٣٠) دقيقة لكل مجموعة، وحددت هذه الأوقات بناءً على جلسات سابقة، ووجدت معاملات الصدق التلازمي لكل مجموعة وذلك بإيجاد معامل الارتباط بين علامات كل جلسة وعلامات الاختبار المحكي، وكان أكبرها تحت الوقت (٢٥) دقيقةً في اختبار اللغة الإنجليزيّة، وتحت الوقت (٢٢) دقيقة في اختبار اللغة الفرنسية، ووجد أن الفروق في معاملات الصدق للمجموعات المتعددة من المفحوصين تكون أقل ما يمكن تحت هذه الأوقات (٢٢، ٢٥ دقيقة)، وتتزايد هذه الفروق في حالة تقصير الوقت، وتكون الفروق أكبر ما يمكن في الوقت القصير جداً.

كما قام ننللي (Nannally, 1978) بإجراء دراسة هدفت إلى معرفة أثر طول وقت الاختبار، وصعوبة وتجانس فقراته على خصائص الاختبار السيكومترية، وأظهرت النتائج أن سبب اختلاف الخصائص السيكومترية للاختبار تحت الأوقات المختلفة الطول إلى ارتفاع معاملات صعوبة فقراته أو إلى الترتيب التصاعدي لهذه الفقرات حسب صعوبتها، وذلك لأن تقصير زمن الاختبار يحرم معظم المفحوصين من إجابة القسم الأكثر صعوبة من الفقرات، هذا الأمر سيجعل البدائل تقيس عوامل مختلفة، وعليه ستختلف الخصائص السيكومترية للاختبار.

وفي دراسة لجودي وآخرون (Jodi & Claudio,2007) والتي هدفت إلى دراسة أثر ترتيب الفقرات في تقدير النشاط البدني للمفحوصين باستخدام الاستبانة العالمية (IPAQ)، وذلك في توزيع الاستبانة ثم عكس فقراتها، حيث أبرزت نتائج الدراسة فروقاً في استجابات العينة وفقاً لطريقة عرض الفقرات علما بأن الفقرات كانت أصلاً مرتبة بطريقة عشوائية في المقياس.

و دراسة العمري(٢٠١١) التي هدفت إلى الكشف عن ترتيب الفقرات في المقاييس النفسية على خصائص المقياس وخصائص فقراته، حيث تم استخدام مقياس للثقة بالنفس، تم إعادة ترتيب فقراته بأربع طرق ، الثلاث الأولى هي ترتيب الفقرات

تنازلياً حسب موشر كل من تمييز الفقرات، ثبات الفقرة، صدق الفقرة، إضافة إلى الترتيب الرابع، وهو ترتيب الفقرات عشوائياً.

وقد أشارت نتائج الدراسة إلى تفوق أسلوب ترتيب الفقرات اعتماداً على مؤشر ثباتها، حيث أعطى هذا الترتيب أعلى معامل ثبات للمقياس بالإضافة إلى أن هذا الترتيب قد تقدم على الطرق الأخرى في تحسين متوسطات الخصائص السيكومترية للفقرات.

مما سبق يتضح أنّ الدراسات السابقة – التي في حدود علم الباحثين – قد تناولت ترتيب الفقرات وأثره في الخصائص السيكومترية للمقاييس النفسية، كما تناولت أثر طول الاختبار وزمن الاختبار على خصائصه السكومترية، ولأن ترتيب الفقرات حسب معالمها يعد من أهم الطرق المستخدمة في ترتيب فقرات الاختبار، جاءت هذه الدراسة للتعرف على أثر ترتيب فقرات اختبار الاختيار من متعدد وفقاً لمعالم الفقرة في تقدير قدرة المفحوصين ودالة المعلومات للاختبار وفقاً للنموذج ثلاثي المعلمة.

مشكلة الدراسة :

تعتبر إجراءات بناء الاختبارات التي منها ترتيب فقرات الاختبار من الأمور الهامة التي قد تؤثر في خصائص الاختبار السيكومترية وخصائص الفقرات التي قد تؤدي إلى خطأ في أداة القياس والذي يعتبر أحد مصادر أخطاء القياس، فأخطاء القياس قد تأتي من عيوب في المقياس وقد تأتي من أخطاء المعاينة لفقرات المقياس من حيث الحجم والتمثيل المناسب، وقد تأتي أيضاً من طبيعة الظروف التجريبية المؤثرة في الاستجابة أو في المستجيب (Lord,1960)

ويرى بعض علماء القياس أن الأخطاء المتأتية من المقياس تعد من أكثر الأخطاء تأثيراً في الدرجة الملاحظة، وبالتالي التأثير على الدرجة الحقيقية، وهذا يتضح من المعادلة التالية :

$$T=X\pm E$$

ويمكن تقليل الأخطاء المتأتية من الظروف التجريبية أو الظروف المحيطة بعملية القياس من خلال تثبيت هذه الظروف أو عزل تأثيرها تجريبيا أو احصائياً.

وتعد إجراءات بناء الاختبار أحد هذه الظروف التجريبية لذلك جاءت هذه الدراسة لمحاولة التعرف على أثر ترتيب فقرات الاختبار الذي يعد أحد الظروف التجريبية الخاصة ببناء الاختبارات في تقدير دالة المعلومات حيث إن ترتيب الفقرات حسب معاملات الصعوبة أو التمييز أو التخمين قد يؤثر في الخطأ المعياري للقياس مما يؤثر في دالة المعلومات للفقرة وللاختبار.

ويمكن أن تتمثل مشكلة الدراسة في الإجابة عن الأسئلة التالية:

١. هل توجد فروق ذات دلالة إحصائيّة في تقدير قدرة المفحوص باختلاف مؤشر تمييز الفقرة(تصاعدي، تنازلي، عشوائي)؟
٢. هل توجد فروق ذات دلالة إحصائيّة في تقدير قدرة المفحوص باختلاف مؤشر صعوبة الفقرة(تصاعدي، تنازلي، عشوائي)؟
٣. هل توجد فروق ذات دلالة إحصائيّة في تقدير قدرة المفحوص باختلاف مؤشر التخمين الفقرة(تصاعدي، تنازلي، عشوائي)؟
٤. هل توجد فروق ذات دلالة إحصائيّة في تقدير دالة المعلومات باختلاف مؤشر تمييز الفقرة(تصاعدي، تنازلي، عشوائي)؟
٥. هل توجد فروق ذات دلالة إحصائيّة في تقدير دالة المعلومات باختلاف مؤشر صعوبة الفقرة(تصاعدي، تنازلي، عشوائي)؟

أهمية الدراسة:

تأتي أهمية هذه الدراسة من الموضوع الذي تناقشه إذ إنها تتحدث عن موضوع غاية في الأهمية وهو موضوع ترتيب الفقرات في الاختبار التحصلي حيث إن ما يتوصل إليه البحث سيرشد واضعي الاختبارات إلى الطريقة المناسبة لترتيب فقرات الاختبار في الاختبار للحصول على اختبار يتمتع بخصائص جيدة ويعطي دالة معلومات عالية مما يقلل من الخطأ المعياري في القياس مما يزيد من دقة القياس.

مصطلحات الدراسة :

معالم الفقرة(item parameter) :هي معلمة الصعوبة، والتمييز، ومعلمة التخمين.

معلمة الصعوبة (Threshold): هي نقطة على متصل القدرة تقابل احتمال 2/(1+ci) للإجابة على الفقرة (i) إجابة صحيحة، ويشير الرمز (ci) إلى معلمة التخمين.

معلمة التمييز (Slope): نسبة ميل منحنى خصائص الفقرة الذي يقابل النقطة التي تكون فيها علامة القدرة تساوي صعوبة الفقرة.

معلمة التخمين (Asymptote) :هي خط المقاربة الأدنى Lower Asymptote من منحنى خصائص الفقرة ويمثل احتمالية إجابة المفحوصين ذوي القدرة المتدنية على الفقرة إجابة صحيحة.

دالة معلومات **الاختبار (test information function)**

هو اقتران رياضي يعبر عن مجموع دوال المعلومات لفقرات الاختبار، وفي هذه الدراسة تعرف دالة معلومات الاختبار بالقيمة التي يتم الحصول عليها من تطبيق المعادلة الآتية:

$$I(\theta) = \sum_{i=1}^{n} \frac{P_i'(\theta)^2}{P_i(\theta)Q_i(\theta)}$$

توليد البيانات:

تم الاعتماد على برنامج (WINGEN3) لتوليد البيانات وهو من تصميم وإنتاج هان و هامبلتون (Han & Hambleton,2007) حيث يمكن بالاعتماد على هذا البرنامج توليد استجابات لاختبارات أحادية البعد سواء كانت ثنائية التدريج أم متعددة التدريج، بالإضافة إلى اختبار متعدد الأبعاد، ويمكن الاعتماد على هذا البرنامج لتوليد استجابات لأكثر من مجموعة من المفحوصين بأنواع مختلفة من التوزيعات، حيث يستخدم ثلاثة توزيعات لتوليد قدرة الأفراد وهي : التوزيع الطبيعي (Normal Distribution) ، التوزيع المنتظم (Uniform Distribution) ، وتزيع بيتا (Beta Distribution) الذي من خلاله يمكن توليد توزيعات ملتوية سواء أكانت ملتوية التواء موجب أم التواء سالب.

مراحل توليد البيانات :

أولا: توليد الاختبار

١. توليد اختبار طوله (٦٠) فقرة بالاعتماد على النموذج اللوجستي ثلاثي المعلمة(Three – Parameter Logistic Model) ، وتم اختيار هذا الطول للاختبار بالاعتماد على ما أشار اليه (Drasgow,1982) بأن هذا الطول المناسب بالاعتماد على النموذج الثلاثي للوصول إلى دقة في التقديرات وفقاً لنموذج اللوجستي ثلاثي المعلمة .

٢. توليد معلمة التمييز للفقرات وفق توزيع Log normal (0,0.25)~ وبعد توليد البيانات تبين أن الوسط الحسابي لقيم لمعلمة التمييز(١,٠٧) والإنحراف المعياري كان (٠,٢٤) ، وقد اعتبرت هذه القيم جيده بمقارنتها بالمعيار الذي تم تحديده من قبل هاملبتون وسوامينيثان (Hambleton & swaminthan,1985) والذي ينص على أن قيم معلمة التمييز الحقيقية تتروح ما بين (٠،٢) لوجت.

٣. توليد معلمة الصعوبة للفقرات وفق التوزيع الطبيعي Normal ~ (0,1) حيث كان الوسط الحسابي لقيم لمعلمة الصعوبة (0.45) والانحراف المعياري (0.67) على الترتيب.

٤. توليد معلمة التخمين للفقرات وفق توزيع بيتا Beta (8,32)~ وهذا التوزيع ينتج قيم لمعلمة التخمين، تماثل قيم التخمين للاختبار الموضوعي (ثنائي الاستجابة) المؤلف من خمسة بدائل، وكان الوسط الحسابي لمعلمة التخمين (0.21) والانحراف المعياري (0.06) على الترتيب.

ويلخص الجدول (١) قيم معاملات الصعوبة والتمييز والتخمين التي تم توليدها .

جدول (١) : قيم معالم الصعوبة والتمييز والتخمين التي تم توليدها بالاعتماد على النموذج ثلاثي المعلمة

رقم الفقرة	التمييز ai	الصعوبة bi	التخمين Ci	رقم الفقرة	التمييز ai	الصعوبة bi	التخمين ci
١	٠,٨٤	٠,٥٦	٠,٢٣	٣١	١,٠١	٢,٤٩	٠,١٤
٢	١,١٤	١,٥٠-	٠,٢٤	٣٢	١,٣٥	١,٠٧	٠,٢٠
٣	٠,٩٦	٠,١٩	٠,١١	٣٣	٠,٩٥	٠,٠٢	٠,١٦
٤	١,٢٥	١,٦٩	٠,١٥	٣٤	١,١١	٠,٨٦-	٠,٢٣
٥	٠,٩٣	١,٢١-	٠,٢٧	٣٥	١,٢٥	٠,٠٧-	٠,٢٢
٦	٠,٧٩	٠,٢٦-	٠,٢٣	٣٦	١,١٧	٠,١١-	٠,٢٢
٧	١,٣٠	١,٧٧	٠,١٤	٣٧	١,٣٩	٠,٩٣-	٠,١٩
٨	١,٣١	٠,١٦-	٠,١٧	٣٨	٠,٩٤	٠,٣٣	٠,٠٩
٩	٠,٥٨	١,٦٤	٠,٢٣	٣٩	٠,٩٧	٠,٩٦	٠,١٨
١٠	٠,٨١	٠,٢٩	٠,٢٠	٤٠	٠,٧٠	٠,٩١-	٠,٢١
١١	٠,٩٧	٠,٣٤-	٠,٣١	٤١	١,١٤	١,٤١-	٠,١٥
١٢	٠,٩٣	٢,٠٢	٠,١٨	٤٢	١,٠٠٦	١,٠٥	٠,١٢
١٣	٠,٩٠	٠,٠٢-	٠,١٦	٤٣	٠,٩٣	١,٠٥-	٠,١٦
١٤	١,٠٦	٠,٨٠	٠,١٨	٤٤	١,٠٣	٠,٢١	٠,٢٣
١٥	٠,٥٦	٠,٦٣	٠,١٧	٤٥	١,٨٦	١,٢٥-	٠,١٨
١٦	٠,٨٥	٠,٨١-	٠,١٥	٤٦	٠,٩٣	٠,٨٨-	٠,٣٢
١٧	٠,٨٨	٠,٥٢-	٠,١٧	٤٧	٠,٩٩	٠,٩٢-	٠,٢٠
١٨	١,٠٣	١,١٧	٠,٢٢	٤٨	١,٠٦	٢,٠٨	٠,٢٢
١٩	١,٠٩	٠,٥٣-	٠,١٦	٤٩	١,٣٧	٠,٢٩	٠,٢٥
٢٠	٠,٩٨	٠,٠٤-	٠,٣٦	٥٠	١,٤٣	٠,٥٩	٠,٢٢
٢١	١,٢٥	٠,٦٣	٠,١٣	٥١	١,٢٠	٠,٧٩	٠,١٩
٢٢	١,٠٤	١,٠٠٥	٠,٢٧	٥٢	١,٠٨	٠,٩٥-	٠,٢٥
٢٣	١,١٤	٠,٩٦-	٠,٣٧	٥٣	٠,٩١	١,٣٦	٠,١٨
٢٤	١,٠٧	١,٨٧-	٠,٢٢	٥٤	٠,٩٩	١,٧٦-	٠,٢٢
٢٥	١,٥٦	٠,٢٦	٠,٢٤	٥٥	٠,٧٢	٠,٥٣-	٠,١٩
٢٦	١,٣١	٠,٠٦	٠,١٢	٥٦	١,٢٤	٠,٦٥-	٠,٢٥
٢٧	١,٢١	١,٥٤	٠,١٣	٥٧	٠,٩٣	٠,٦٧-	٠,٢٣
٢٨	١,٠٧	١,٨١-	٠,٣٦	٥٨	٠,٨٧	٠,٨٦-	٠,١٦
٢٩	١,٣٩	١,٤٤	٠,١٩	٥٩	٠,٨٥	٠,٦٢-	٠,٢١
٣٠	١,٠٠٩	١,٢٦	٠,٢٩	٦٠	١,٦٩	٠,٩٨-	٠,٢٠

ثانيا: توليد الاستجابات

تم توليد استجابات (٥٠٠٠) مفحوص باستخدام القيم نفسها للمعالم الحقيقية للفقرات التي تم توليدها في السابق بالاعتماد على التوزيع الطبيعي Normal ~ (0,1) ، حيث كان الوسط الحسابي والانحراف المعياري لهذا التوزيع (٠,٠١ ، ٠,٩٩) على الترتيب وكانت درجة الالتواء والتفلطح (٠,٠٣ ، ٠,٠٥) على الترتيب.

تحليل البيانات:

لتحقيق أهداف الدراسة قام الباحثان بالاعتماد على برنامج (WINGEN3) لتوليد البيانات حيث تم توليد معالم الفقرات (الصعوبة ، التمييز ، التخمين) لاختبار مكون من (٦٠) فقرة و تم الاعتماد على على معالم الفقرات الناتجة لتوليد القدرة لـ (٥٠٠٠) مفحوص بالاعتماد على النموذج ثلاثي المعلمة.

بعدها تم ترتيب الفقرات حسب معاملات الصعوبة (تنازلي، تصاعدي، عشوائي) وكذلك ترتيب الفقرات حسب معاملات التمييز (ai) ومعاملات التخمين(ci)؛ ليتم تكوين ثلاث صور للاختبار وتم توليد قدرات المفحوصين بناء على ترتيب معاملات الصعوبة والتمييز والتخمين للصور الثلاث للاختبار.

مطابقة البيانات للنموذج

تم الاعتماد على برنامج (BILOG-MG3) وذلك من أجل مطابقة الأفراد والفقرات للنموذج اللوجستي ثلاثي المعلمة، حيث تم تحليل بيانات (٥٠٠٠) مفحوص وقد أشارت نتائج التحليل بأن جميع الفقرات كانت مطابقة للنموذج حيث كانت قيمة مربع كاي لها غير دالة إحصائياً عند مستوى دلالة (α=٠,٠٥) كما تبين من نتائج التحليل أن جميع استجابات المفحوصين مطابقة مع توقعات النموذج باستثناء خمسة مفحوصين حيث كانت قيمة مربع كاي ذات دلالة إحصائيّة عند مستوى دلالة(α=٠,٠٥) .

المعالجة الإحصائيّة

للإجابة عن أسئلة الدراسة تم الاعتماد على المعالجات الإحصائيّة التالية.

١. الاعتماد على برنامج (WINGEN 3) لتوليد اختبار مكون (٦٠) فقرة وتوليد معاملات الصعوبة والتمييز والتخمين بالاعتماد على النموذج اللوجستي ثلاثي المعلمة
٢. الاعتماد على برنامج (WINGEN 3) لتوليد استجابات (٥٠٠٠) مفحوص بناء بناء على معالم الفقرة التي تم توليدها.
٣. إعادة ترتيب معالم الفقرات (تصاعدي، تنازلي) وتوليد استجابات للمفحوصين بالاعتماد على هذه المعالم.
٤. استخدام تحليل التباين الأحادي للكشف عن الفروق في قدرات المفحوصين باختلاف معالم الفقرات.
٥. الاعتماد على برنامج ((BILOG-MG3) لإيجاد دالة المعلومات لفقرات الصور الثلاث للاختبار والاختبار ككل.

النتائج ومناقشتها:

للإجابة عن السؤال الأول والذي ينص على " هل توجد فروق ذات دلالة إحصائيّة في تقدير قدرة المفحوص باختلاف مؤشر صعوبة الفقرة(تصاعدي، تنازلي، عشوائي)؟.

للإجابة عن هذا السؤال تم استخدام برنامج (WINGEN 3) لتوليد معالم الفقرات لاختبار مكون من (٦٠) فقرة وبالاعتماد على قيم معالم الفقرات، تم توليد استجابات (٥٠٠٠) مفحوص، وبالاعتماد على معالم الفقرات التي تم توليدها تم إعادة ترتيب الفقرات بالاعتماد على قيم معاملات الصعوبة تصاعدياً وتنازلياً وبالاعتماد على هذه القيم تم توليد استجابات (٥٠٠٠) مفحوص لصورتي الاختبار.

وبالاعتماد على هذه القيم تم إيجاد المتوسطات الحسابية والانحرافات المعيارية لتقديرات معلمة القدرة للأفراد حسب متغير الصعوبة: (عشوائي، تصاعدي، تنازلي)، ويبين الجدول (٢) هذه القيم

جدول ٢ : الأوساط الحسابية والانحرافات المعيارية لتقديرات معلمة القدرة للأفراد (θ) تبعاً لمتغير صعوبة الفقرة (عشوائي، تصاعدي، تنازلي)

صورة الاختبار	الوسط الحسابي	الانحراف المعياري
الأولى: عشوائي	-٠,٠١١	٠,٩٩٤
الثانية: تصاعدي	٠,٠٠٤	١,٠٠٢
الثالثة: تنازلي	٠,٠٢١	٠,٩٨٩

يتبين من الجدول (٢) وجود فروق ظاهرية بين المتوسطات الحسابية لتقديرات معلمة القدرة للأفراد تبعاً لمتغير صعوبة الفقرة، ولمعرفة الدلالة الإحصائيّة لتلك الفروق الظاهرية تم الاعتماد على تحليل التباين الأحادي (One Way ANOVA) ويبين الجدول (٣) نتائج تحليل التباين الأحادي.

جدول ٣ : نتائج تحليل التباين الأحادي (One Way ANOVA) للمتوسطات الحسابية لتقديرات معلمة القدرة تبعاً لمتغير صعوبة الفقرة (عشوائي، تصاعدي، تنازلي)

مصادر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	F	Sig الدلالة
بين المجموعات	٢,٧٢	٢	١,٣٥٩	١,٣٧	٠,٢٥٤
داخل المجموعات	١٤٨٥١,١٤	١٤٩٩٧	٠,٩٩		
الكلي	١٤٨٥٣,٨٦	١٤٩٩٩			

يتبين من الجدول (٣) عدم وجود فروق ذات دلالة إحصائيّة عند مستوى الدلالة (α=٠,٠٥) بين المتوسطات الحسابية لتقديرات معلمة القدرة تعزى لاختلاف مؤشر صعوبة الفقرة (عشوائي، تصاعدي، تنازلي). ويمكن تفسير ذلك بأن تقدير قدرات المفحوصين تكون مستقلة عن فقرات الاختبار التي أجاب عليها المفحوصون (invariant)، أي أن تقدير قدرات المفحوصين باستخدام مجموعة مختلفة من الفقرات سيكون كما هو.

للإجابة عن السؤال الثاني: الذي ينص على " هل توجد فروق ذات دلالة إحصائيّة في تقدير قدرة المفحوص باختلاف مؤشر صعوبة تمييز الفقرة (تصاعدي، تنازلي، عشوائي)

للإجابة عن هذا السؤال الاعتماد على برنامج (WINGEN 3) لتوليد معالم الفقرات لاختبار مكون من (٦٠) فقرة وبالاعتماد على قيم معالم الفقرات تم توليد استجابات (٥٠٠٠) مفحوص، وبالاعتماد على معالم الفقرات التي تم توليدها تم إعادة ترتيب الفقرات حسب معاملات التمييز تصاعدياً وتنازلياً وبالاعتماد على هذه القيم تم توليد استجابات (٥٠٠٠) مفحوص لصورتي الاختبار.

وبالاعتماد على هذه القيم تم إيجاد المتوسطات الحسابية والانحرافات المعيارية لتقديرات معلمة القدرة للأفراد حسب متغير التمييز(عشوائي، تصاعدي، تنازلي)، ويبين الجدول (٤) هذه القيم

جدول (٤): الأوساط الحسابية والانحرافات المعيارية لتقديرات معلمة القدرة للأفراد (θ) تبعاً لمتغير تمييز الفقرة (عشوائي، تصاعدي، تنازلي)

صورة الاختبار	الوسط الحسابي	الانحراف المعياري
الأول: عشوائي	-0.012	٠,٩٩
الثاني: تصاعدي	-0.148	١,٠٢
الثالث: تنازلي	0.012	٠,٩٨

يتبين من الجدول (٤) وجود فروق ظاهرية بين المتوسطات الحسابية لتقديرات معلمة القدرة للأفراد تبعاً لمتغير تمييز الفقرة، ولمعرفة الدلالة الإحصائيّة لتلك الفروق الظاهرية تم الاعتماد على تحليل التباين الأحادي (One Way ANOVA) ويبين الجدول (٥) نتائج تحليل التباين الأحادي.

دول ٥ : نتائج تحليل التباين الأحادي (One Way ANOVA) للمتوسطات الحسابية لتقديرات معلمة القدرة تبعاً لمتغير تمييز الفقرة (عشوائي، تصاعدي، تنازلي)

مصادر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	F	SIG الدلالة
بين المجموعات	٢,٠٥٣	٢	١,٠٧٢	١,٠٤	٠,٣٥٥
داخل المجموعات	١٤٨٧٢,٢٩	١٤٩٩٧	٠,٩٩٢		
الكلي	١٤٨٧٤,٣٤	١٤٩٩٩			

يتبين من الجدول (٥) عدم وجود فروق ذات دلالة إحصائيّة عند مستوى الدلالة (α=٠,٠٥) بين المتوسطات الحسابية لتقديرات معلمة القدرة تعزى لاختلاف مؤشر تمييز الفقرة (عشوائي، تصاعدي، تنازلي). ويمكن تفسير ذلك بأن تقدير قدرات المفحوصين تكون مستقلة عن فقرات الاختبار التي أجاب عليها المفحوصون (invariant)، أي أن تقدير قدرات المفحوصين باستخدام مجموعة مختلفة من الفقرات سيكون كما هو.

للإجابة عن السؤال الثالث: الذي ينص على " هل توجد فروق ذات دلالة إحصائيّة في تقدير قدرة المفحوص باختلاف مؤشر تخمين الفقرة (تصاعدي، تنازلي، عشوائي)

فقد تم الاعتماد على برنامج (WINGEN 3) لتوليد معالم الفقرات لاختبار مكون من (٦٠) فقرة وبالاعتماد على قيم معالم الفقرات تم توليد استجابات (٥٠٠٠) مفحوص، وبالاعتماد على معالم الفقرات التي تم توليدها تم إعادة ترتيب الفقرات حسب معاملات التخمين تصاعدياً وتنازلياً وبالاعتماد على هذه القيم تم توليد استجابات (٥٠٠٠) مفحوص لصورتي الاختبار.وبالاعتماد على هذه القيم تم إيجاد المتوسطات الحسابية والانحرافات المعيارية لتقديرات معلمة القدرة للأفراد حسب متغير التخمين(عشوائي، تصاعدي، تنازلي)، ويبين الجدول (٦) هذه القيم

جدول ٦ : الأوساط الحسابية والانحرافات المعيارية لتقديرات معلمة القدرة للأفراد (θ) تبعاً لمتغير تخمين الفقرة (عشوائي، تصاعدي، تنازلي)

صورة الاختبار	الوسط الحسابي	الانحراف المعياري
الأول: عشوائي	-0.012	0.99
الثاني: تصاعدي	-0.002	1.002
الثالث: تنازلي	0.013	0.99

يتبين من الجدول (٦) وجود فروق ظاهرية بين المتوسطات الحسابية لتقديرات معلمة القدرة للأفراد تبعاً لمتغير تخمين الفقرة، ولمعرفة الدلالة الإحصائيّة لتلك الفروق الظاهرية تم الاعتماد على تحليل التباين الأحادي (One Way ANOVA) ويبين الجدول (٧) نتائج تحليل التباين الأحادي.

جدول ٧ : نتائج تحليل التباين الأحادي (One Way ANOVA) للمتوسطات الحسابية لتقديرات معلمة القدرة تبعاً لمتغير تخمين الفقرة (عشوائي، تصاعدي، تنازلي)

مصادر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	F	SIG الدلالة
بين المجموعات	1.505	٢	0.752	0.759	0.468
داخل المجموعات	14869.48	١٤٩٩٧	0.99		
الكلي	14870.98	١٤٩٩٩			

يتبين من الجدول (٧) عدم وجود فروق ذات دلالة إحصائيّة عند مستوى الدلالة (α=٠,٠٥) بين المتوسطات الحسابية لتقديرات معلمة القدرة تعزى لاختلاف مؤشر تخمين الفقرة (عشوائي، تصاعدي، تنازلي). ويمكن تفسير ذلك بأن تقدير قدرات المفحوصين تكون مستقلة عن فقرات الاختبار التي أجاب عليها المفحوصون (invariant)، أي أن تقدير قدرات المفحوصين باستخدام مجموعة مختلفة من الفقرات سيكون كما هو.

للإجابة عن السؤال الرابع: الذي ينص على " هل توجد فروق ذات دلالة إحصائيّة في تقدير دالة المعلومات باختلاف مؤشر تمييز الفقرة(تصاعدي، تنازلي، عشوائي)

للإجابة على هذه السؤال تم الاعتماد على المقارنة بين المنحنيات الخاصة بدالة المعلومات والتي تم إيجادها بالاعتماد على الخطوات التالية.

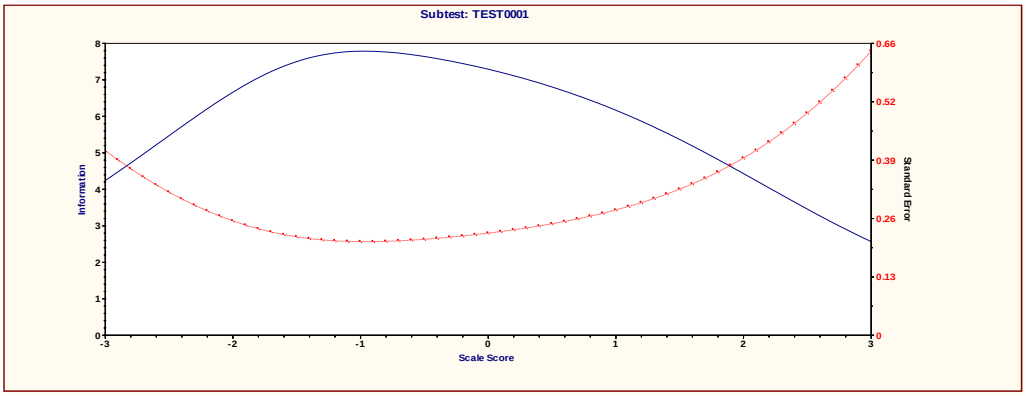
١. الاعتماد على برنامج (WINGEN 3) لتوليد معالم الفقرات لاختبار مكون من (٦٠) فقرة

٢. توليد استجابات (٥٠٠٠) مفحوص بالاعتماد على قيم معالم الفقرات .

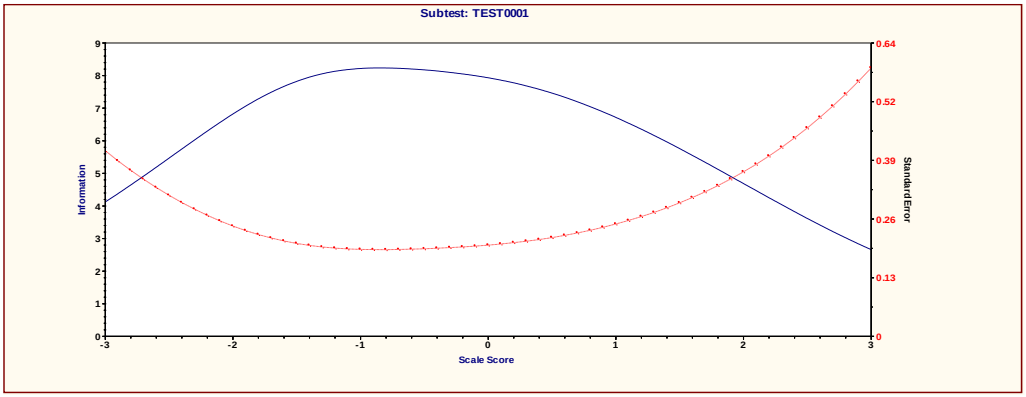
٣. بالاعتماد على معالم الفقرات التي تم توليدها تم إعادة ترتيب الفقرات حسب معاملات تمييز الفقرة تصاعدياً وتنازلياً وبالاعتماد على هذه القيم ثم توليد استجابات (٥٠٠٠) مفحوص لصورتي الاختبار.

٤. إيجاد دالة المعلومات بالاعتماد على برنامج (BILOG-MG) وباستخدام هذه القيم تم إيجاد منحنيات دالة المعلومات الخاصة بكل صورة من صور الاختبار والأشكال التالية توضح ذلك

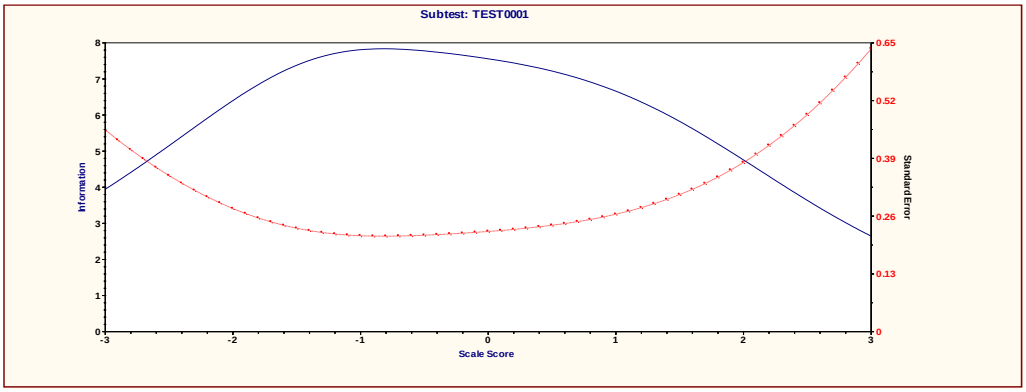
منحى دالة المعلومات وفقاً للترتيب التنازلي لمعلمة التمييز



منحى دالة المعلومات وفقاً للترتيب التصاعدي لمعلمة التمييز



منحى دالة المعلومات وفقاً للترتيب العشوائي لمعلمة التمييز



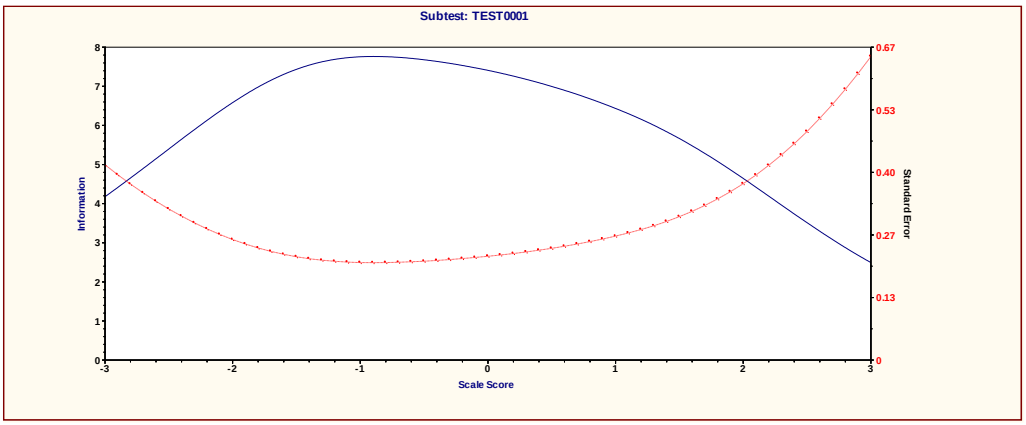
يتبين من الأشكال السابقة أن أعلى دالة معلومات للاختبار كانت للصورة الأولى من الاختبار التي تم ترتيب معاملات التمييز فيها تنازلياً أي من الأعلى إلى الأقل ، ويمكن تفسير ذلك أن الفقرات ذات التمييز العالي سوف تعطي دالة معلومات أعلى للفقرات، وبالتالي للاختبار ككل وهذا أمر منطقي، حيث إن الفقرة ذات التمييز العالي لديها قدرة عالية على التمييز بين المجموعتين العليا والدنيا من الطلبة حيث إنه كلما كان معامل التمييز للفقرة عالي ، أشار ذلك إلى قدرة الفقرة على تصنيف المفحوصين وتدريجهم حسب ما لديهم من قدرة، مما يقلل من الخطأ

المعياري في التقدير مما يزيد من دالة المعلومات، يليها الصورة الثالثة عندما يتم ترتيب الفقرات بشكل عشوائياً، وأخيراً الصورة الثانية في حال ترتيب الفقرات تصاعدياً حسب معاملات التمييز.

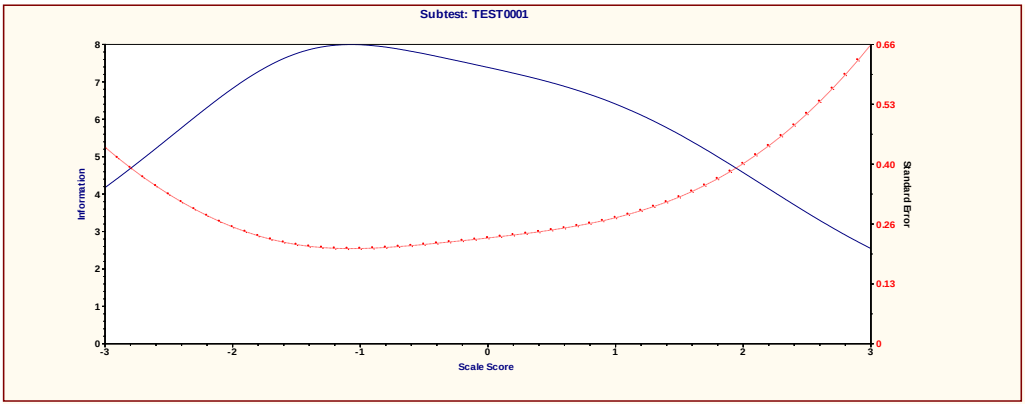
للإجابة عن السؤال الخامس : الذي ينص على " هل توجد فروق ذات دلالة إحصائيّة في تقدير دالة المعلومات باختلاف مؤشر صعوبة الفقرة (تصاعدي، تنازلي، عشوائي).

للإجابة على هذه السؤال تم الاعتماد على المقارنة بين المنحنيات الخاصة بدالة المعلومات والتي تم إيجادها بالاعتماد على الخطوات السابقة والأشكال التالية توضح ذلك.

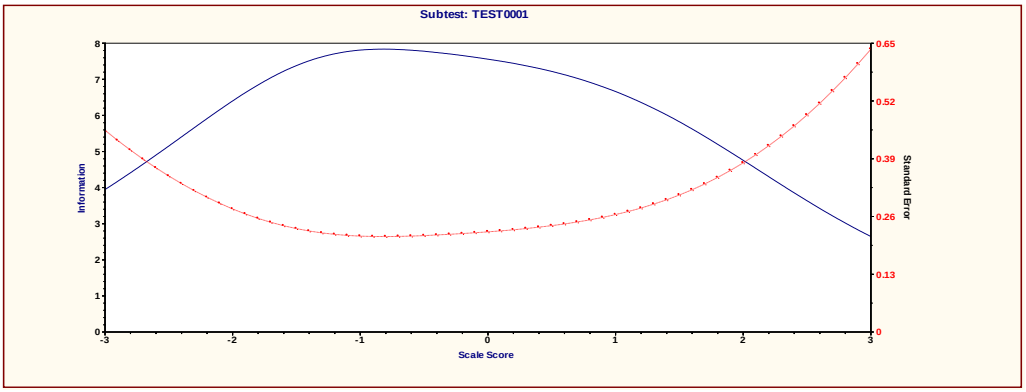
منحى دالة المعلومات وفقاً للترتيب التنازلي لمعامل الصعوبة



منحى دالة المعلومات وفقاً للترتيب التصاعدي لمعلمة الصعوبة



منحى دالة المعلومات وفقاً للترتيب العشوائي لمعلمة الصعوبة



يتبين من الأشكال السابقة أن أعلى دالة معلومات للاختبار كانت للصورة الثانية من الاختبار التي تم ترتيب معاملات الصعوبة فيها تصاعدياً؛ أي من الأسهل إلى الأصعب ، ويمكن تفسير ذلك في أن هذا الترتيب يوفر الدافعية عند الطلبة للاستمرار في محاولاتهم

للإجابة عن الفقرات؛ لأنه سوف يشعر بالتعزيز الفوري بسبب قدرته على الإجابة عن الأسئلة الأولى من الاختبار، التي تسمى بالأسئلة التشجيعية أو الأسئلة الماصة للصدمة، كما إنه عندما تكون معاملات الصعوبة للفقرة منخفضة ربما يكون مستوى قدرة معظم الطلبة أكبر من مستوى صعوبة الفقرة مما يساعدهم في الإجابة عن فقرات الاختبار بشكل أفضل مما يقلل من الخطأ المعياري في التقديرمما يزيد من دالة معلومات الفقرة، بالتالي زيادة دالة معلومات الاختبار، كما أنه في حالة ترتيب الفقرات حسب درجة الصعوبة للفقرة تصاعدياً سيؤدي ذلك بأن تكون تتزايد درجة الصعوبة للفقرات بشكل تدريجي حتى تصل قيمة مساوية إلى قدرة المفحوص (θ) أو أعلى بقليل منها مما يؤدي إلى زيادة دالة المعلومات وهذا يمكن توضيحه من المعادلة التالية:

$$I_i(\theta) = \frac{2.89a_i^{\,2}(1-c_i)}{(c_i + e^{(1.7a_i(\theta-b_i))})(1+e^{(1.7a_i(\theta-b_i)})^2}$$

من المعادلة السابقة يتضح أنّ أقصى كمية معلومات تكون عندما تناظر قيمة (θ) مع درجة الصعوبة (b) كما أنّ كمية المعلومات تتزايد كلما زادت درجة الصعوبة للفقرة عن قدرة المفحوص، أي إنّه عندما نتدرج في معاملات الصعوبة للفقرات سوف تزداد كمية المعلومات للفقرة بشكل تدريجي مما يعظم دالة معلومات الاختبار.

التوصيات :

- ترتيب فقرات الاختبار تصاعدياً حسب درجة الصعوبة أو تنازلياً حسب درجة تمييز الفقرة.
- عدم وضع الفقرات بشكل عشوائي في الاختبار.

المراجع:

- عبابنة،عماد.(٢٠٠٤).أأثر حجم العينة وطريقة انتقائها وعدد الفقرات وطريقة انتقائها على دقة تقدير معالم الفقرة والقدرة لاختبار قدرة عقلية باستخدام نظرية الاستجابة للفقرة. رسالة دكتوراه غير منشورة. جامعة عمان العربية،الاردن
- العمري، حسان . (٢٠١١). أثر ترتيب الفقرات في المقاييس النفسية على التجانس الداخلي للمقياس وخصائص فقراته، مجلة جامعة دمشق، المجلد (٢٧)، العدد (٣) ، ص ص ١١١-١٢٩.
- علام، صلاح الدين .(٢٠٠٢) . القياس والتقويم التربوي والنفسي أساسياته وتطبيقاته وتوجيهاته. القاهرة :دار الفكر العربي.
- عوده، أحمد (٢٠١٠). القياس والتقويم في العملية التدريسية.(ط٤) إربد: دار الأمل.
- الكيلاني، عبدالله وعدس، عبد الرحمن (١٩٨٦). القياس والتقويم في التربية وعلم النفس. مركز الكتاب الأردني، الأردن، عمان.
- النجار، نبيل (٢٠١٠). **القياس والتقويم منظور تطبيقي مع تطبيقات برمجية SPSS**. الأردن، عمان: دار الحامد للنشر والتوزيع.

Brannick, M. (2003). **Basics of IRT one –** linefile : //a:/ item response theory.htm.

- Crocker, L ,&. Algina ,J. (1986) . Introduction to classical and modern test theory , University of Florida,U.S.A.
- Drasgow,F.(1982). Biased test items and differential validity. Psychological Bulletin, 92, 526-531.
- Embretson ,S. & Reiase , S. (2000). Item Response Theory for Psychologists.New jersey: Lawrence Erlbaum Associates,Inc.
- Hambleton,R.k., & Swaminathan,H, &Rogers. H.j.(1991). Fundamentals of Item Response Theory :International Educational and Professional. Publisher Newbury park.
- Hambleton,R.k., & Swaminathan, H.(1985). Item Response Theory :Principles and applications. Boston MA:Kluwer-Nyjhoff.
- Han, K. T; & Hambleton, R.K.(2007). User's Manual for WinGen : Windows Software that Generated IRT Model Parameter and Item Response. Center for Educational Assessment Research Report No 642, Amherts . MA : University of Massachusetts Center for Educational Assessment.
- Hamleton , R.H & Jones , R , W. (1994). Item Parameter estimation errors and their Influence on test information function. Applied Measurement in education. 7 (3), 171 - 186.
- Jodi .B, Claudio. N, Ilse De. B, Christine. M, Jay .M.(2007). The Effect of Item Order on

Physical Activity Estimates Using the IPAQ. Californian Journal of Health Promotion V 5, Issue 1, 23-29

- Lazarsfeld,P ,F .(1950). The Logical and Mathematical Foundation of Latent structure Analysis , measurement and prediction, Presentation University Press
- Stocking, M.K. (1990). Specifying optimum examinees for item parameters estimation in item response theory . Psyehometrika,vol.55, N03,361-475.

أثر التدريس باستخدام القصة التعليمية المصورة في اكتساب المفاهيم العلمية لدى طلبة المرحلة الأساسية الدنيا واتجاهاتهم نحوالعلوم*

سوسن جمال العمري. عبد السلام موسى العديلي**

كلية العلوم التربوية/ جامعة آل البيت

الملخص هدفت الدراسة إلى الكشف عن أثر تدريس العلوم باستخدام القصة التعليمية المصورة في اكتساب المفاهيم العلمية لدى طلبة المرحلة الأساسية الدنيا واتجاهاتهم نحو العلوم، لتحقيق أهداف الدراسة، استخدم المنهج شبه التجريبي، وتكون أفراد الدراسة من(٤٤) طالباً وطالبة من طلبة الصف الثالث الأساسي في مدرسة الأبرار النموذجية الأساسية التابعة لمديرية قصبة المفرق في الفصل الدراسي الثاني من العام الدراسي ٢٠١٥/٢٠١٦. توزعوا على شعبتين اعتبرت إحداهما تجريبية، ودرست بطريقة القصة التعليمية المصورة، فيما اعتبرت الأخرى ضابطة ودرست بالطريقة الاعتيادية، تم استخدام اختبار لاكتساب المفاهيم، ومقياس للاتجاهات نحو العلوم، بعد التحقق من صدقهما وثباتهما، أظهرت نتائج الدراسة أن هناك فروقاً ذات دلالة إحصائية عند مستوى الدلالة (α=0.05) بين المتوسطات الحسابية لدى أفراد الدراسة على اختبار اكتساب المفاهيم البعدي ومقياس الاتجاهات نحو العلوم تعزى لطريقة التدريس لصالح القصة التعليمية المصورة. وفي ضوء النتائج أوصت الدراسة ببعض التوصيات منها، دعوة المعلمين والمعلمات إلى استخدام طريقة القصة التعليمية المصورة في تدريس مادة العلوم للمرحلة الأساسية الدنيا.

الكلمات المفتاحيه: القصة التعليمية المصورة، اكتساب المفاهيم العلمية، الاتجاهات نحو العلوم، تدريس العلوم