المجلة الأردنية في العلوم التربوية، مجلد 8، عدد 2، 2012، 151- 166

أثر اختلاف شكل توزيع القدرة على معالم الفقرة ودالة المعلومات للاختبار

زايد بني عطا* ونضال الشريفين*

تاريخ تسلم البحث 2011/5/16 تاريخ قبوله 2012/5/15

ملخص: هدفت هذه الدراسة إلى التحقق من أثر اختلاف شكل توزيع القدرة على معالم الفقرة ودالة المعلومات للاختبار. ولتحقيق الغرض من الدراسة تم توليد استجابات أربعة أشكال من توزيعات القدرة وفق النموذج اللوجستي ثلاثي المعلمة من خلال استخدام برنامج WINGEN على 60 فقرة ثنائية الاستجابة. وباستخدام برنامج BILOG-MG 3 تم تحليل استجابات الأفراد لكل شكل من أشكال توزيع القدرة وفق النموذج اللوجستي ثلاثي المعلمة. كشفت نتائج تحليل التباين الأحادي للقياسات المتكررة عن وجود فروق ذات دلالة إحصائية عند مستوى دلالة (0.05) بين متوسطات معالم الصعوبة للفقرات المعايرة من أشكال توزيع القدرة، وكذلك أظهرت النتائج وجود فروق ذات دلالة إحصائية بين متوسطات معالم التمييز للفقرات، وعلى الرغم من تباين معالم الفقرات فقد أظهرت النتائج أنّ قيم معاملات الارتباط بين تقديرات معالم الفقرات المعايرة تبعا لشكل توزيع القدرة كانت دالة إحصائيا كمؤشر لاستقرار معالم الفقرات عبر التوزيعات المختلفة للقدرة. وكشفت النتائج كذلك بأنّ دالة المعلومات للاختبار قد تباينت باختلاف شكل توزيع القدرة، إلا أنها جميعاً قد أعطت أعلى قيم لدالة المعلومات ضمن مدى القدرة 0.5- و0.63 حيث كانت قيم الخطأ المعياري أقل ما يمكن. **(الكلمات المفتاحية:** المحاكاة، شكل توزيع القدرة، النموذج الثلاثي المعلمة، معالم الفقرة، دالة المعلومات للاختبار).

The Effect of Different Ability Distribution on Item Parameters and Test Information Function

Zaid Bani Ata and **Nedal Al Shraifin,** *Faculty of Education, Yarmouk University, Irbid, Jordan.*

***Abstract**: This study aimed at verifying the effect of the ability (θ) distribution on item parameters and Test Information Function (TIF). To achieve the aim of the study, four ability distributions' were generated according to three parameter logistic model using WINGEN program on 60 dichotomous items. The responses of the examinees were analyzed by BILOG– MG 3 programs for each ability distributions according to the three parameter logistic model. The results of the repeated measurement ANOVA revealed that there were statistical significant differences between difficulty means and discrimination means of the scaled items from ability distribution. In spite of the variance of item parameters, the results revealed that the correlation between the scaled items with different ability were statistically significant as indicators of item parameters stability among different distributions of ability. The results also revealed that the Test Information Function varied with the variance of the ability distribution, but they all gave the highest values of the items function within the range of ability (0.5-0.63), where the measurement's standard errors were the least . (**Keywords**: Simulation, Ability Distribution, Three- Parameter Logistic Model, Item Parameter, Test Information Function (TIF)).*

خلفية الدراسة: تعد نظرية الاستجابة للفقرة (Item Response Theory) من التطورات الحديثة في مجال القياس النفسي والتربوي بسبب ما قدمته من طرق سيكومترية ذات فعالية كبيرة في بناء المقاييس النفسية والتربوية، وطريقة تفسير الدرجات على هذه المقاييس مقارنة بالنظرية التقليدية في القياس (Mislevy & Bock, 1990). وتقوم نظرية الاستجابة للفقرة على افتراض وجود متصل للسمة، بحيث يمكن تقدير احتمال الإجابة الصحيحة للفرد عن فقرة ما إذا علم موقعه(θ) على هذا المتصل، وأن العلاقة بين أداء الفرد على الفقرة وقدرته يمكن أن تحدد من خلال ما يسمى منحنى خصائص الفقرة (Item Characteristic Curve)، كما تفترض أن مقدار الاحتمال يكون دالة متزايدة مطردة (Monotonically Increasing) لموقع الفرد على متصل السمة، مما يعني أن احتمال الإجابة الصحيحة يزداد بزيادة قدرة الفرد (Hambleton, 1994).

وقد بُنيت هذه النظرية على افتراضات ينبغي تحققها في البيانات، لكي تؤدي الى نتائج يمكن الوثوق بها، ومن أهم هذه الافتراضات افتراض أحادية البعد (Unidimensionality) الذي يعني بأن يقيس الاختبار سمة واحدة تفسر أداء الفرد عليها، ويرى ورم (Warm, 1978) أن افتراض أحادية البعد أكثر الافتراضات تعقيداً في نظرية الاستجابة للفقرة بسبب تدخل عوامل أخرى لها علاقة بظروف تطبيق الاختبارالتي تؤثر في الأداء على الاختبار مثل: الدافعية والقلق والعوامل الشخصية، ولذلك فان مثل هذا الافتراض لا يتحقق بشكل مؤكد دائماً، وعلى الرغم من وجود هذه العوامل فقد رأى هامبلتون وسوامينيثان وروجرز (Hambleton, Swaminathan & Rogers, 1991) بأنه يتطلب وجود عامل أساسي في الأداء وهو ما يشار إليه بالقدرة التي يقيسها الاختبار. أما الافتراض الثاني فيتمثل بالاستقلال الموضعي (Local Independence) ويقصد به أن تكون استجابات الفرد على فقرات الاختبار مستقلة إحصائياً عند مستوى قدرة معينة، أي أن استجابة الفرد عن فقرة ما يجب أن لا تؤثر سلباً أو إيجاباً على استجابته لفقرة أخرى (Croker & Algina, 1986; Hambleton & Swaminthan, 1985).

* كلية التربية، جامعة اليرموك، إربد، الأردن.

© حقوق الطبع محفوظة لجامعة اليرموك، اربد، الأردن.

إن أحد أهم المزايا المرتبطة بنظرية الاستجابة للفقرة استقلالية القياس، ويعني أن تقدير معالم الفقرات يكون مستقلاً عن خصائص الأفراد التي استخدمت في تقدير هذه المعالم (Sample-Free)، وان تقدير قدرة الفرد يكون مستقلاً عن عينة الفقرات التي تطبق عليه (Item-Free) (Hambleton, 1994). وتعد استقلالية القياس بمثابة النقطة المفصلية بين النظرية التقليدية ونظرية الاستجابة للفقرة، وهو ما عبر عنه لورد (Lord, 1980) بخاصية اللاتغاير (Invariance). فإذا ما تحققت مثل هذه الخاصية يصبح بالإمكان القيام ببعض التطبيقات مثل: معادلة الاختبارات، بنوك الأسئلة، الكشف عن تحيز الفقرات، القياس التكيفي (Rup & Zumbo, 2006).

وحتى يتحقق القياس الموضوعي أو ما يسمى باستقلالية القياس، فقد اهتم البحث السيكومتري المتعلق بنظرية الاستجابة للفقرة بتطوير النماذج الاحتمالية لتحديد العلاقة بين أداء الفرد على فقرات الاختبار وبين السمات أو القدرات الكامنة وراء هذا الأداء، ويمكن تصنيف هذه النماذج إلى فئتين، الفئة الأولى تسمى بالنماذج ثنائية التدريج (Dichotomous Models) والفئة الثانية تسمى بالنماذج متعددة التدريج (Polytomous Models). وتعد النماذج ثنائية التدريج من أشهر النماذج استخداما في بناء الاختبارات والمقاييس (De Gruijter & Van Der Kamp, 2005; Embreston & Rise, 2000). ولم يقتصر البحث السيكومتري بالبحث عن النماذج الرياضية بل تعدى ذلك إلى إيجاد طرق التقدير العددي التي تستخدم في تقدير معالم الفقرات وقدرة الفرد للوصول إلى أفضل التقديرات (Hambleton & Sawminthan, 1985).

وقد بينت ستوكنغ (Stocking, 1990) أن تقدير معالم الفقرة يعد قضية مهمة عند استخدام نظرية الاستجابة للفقرة خصوصا في التطبيقات العملية التي تعتمد على تلك التقديرات، وقد أشارت إلى أن التخطيط لعينات المعايرة يتم اختيارها في معظم التطبيقات بالأسلوب العشوائي. ويفترض غالبية مطوري الاختبارات والمقاييس بأن السمات تتوزع توزيعا طبيعيا في المجتمع، وكلما كان حجم العينة كبيرا كلما كانت العينة ممثلة، وبالتالي يكون توزيع السمة في العينة قريبا من التوزيع الطبيعي، وقد أثبتت الدراسات (الثوابية، 2004؛ Hambleton & Cook,1980; Hambleton & Jones, 1994) بأن حجم العينة يلعب دوراً كبيراً في دقة تقديرات معالم الفقرات وقدرة الأفراد عند استخدام نظرية الاستجابة للفقرة.

والمتصفح للأدب السيكومتري المتعلق بنظرية الاستجابة يجد بأنها لا تشترط التوزيع الطبيعي على غرار النظرية التقليدية في القياس(علام، 2005). ولأن غالبية الدراسات كانت تعتمد على التوزيع الطبيعي، ولا يوجد أدلة امبريقية على أنّ توزيع القدرة المسبق طبيعي، فإن مثل هذا قد لا ينسجم مع بعض المواقف الاختبارية خصوصا إذا تم التعامل مع عينات غير عشوائية، فقد يتم التعامل مع افرد ذوي قدرة عالية أو متدنية مما يعني أن توزيعا ملتويا قد يحدث، ومن هنا يمكن طرح أسئلة كثيرة منها:مدى تأثر معالم الفقرات وقدرة الفرد بشكل توزيع القدرة؟ لذا برزت الحاجة لإجراء هذه الدراسة لتقصي أثر اختلاف شكل توزيع القدرة على معالم الفقرات وقدرة الأفراد.

وعلى الرغم من كثرة الدراسات التي اهتمت بدراسة العوامل المؤثرة على دقة تقديرات معالم الفقرات وقدرة الأفراد، إلا أن المتصفح للدراسات التي اهتمت بدراسة أثر شكل توزيع القدرة على معالم الفرات وقدرة الأفراد لم تحظى بشكل كبير. فمن الدراسات المبكرة في هذا السياق دراسة سيبرس (Cypress, 1973) التي هدفت لفحص أثر التواء العلامات وحجم العينة على تقديرات معلمة الصعوبة وقدرة الأفراد، ولتحقيق الهدف من الدراسة استخدم الباحث خمس أحجام من العينات (75، 150، 300، 600، 1200) و35 توزيعا للعلامات ميزها بسبع مستويات للالتواء وهي: موجب عال،موجب متوسط، موجب متدن، سالب عالي، سالب متوسط، سالب متدن ومجموعة محكية توزيعها طبيعي، وتم إخضاع الجميع لاختبار في الرياضيات مكون من 90 فقرة. أشارت نتائج الدراسة بأنّ الفروق بين معلمة القدرة للأفراد ذات التوزيع الملتوي ومعلمة القدرة للمجموعة المحكية (توزيع طبيعي) تزيد كلما زادت درجة الالتواء باتجاه السالب أو الموجب، وبينت النتائج كذلك بأنّ الأخطاء المعيارية لمعلمة الصعوبة كانت متدنية عند التوزيعات المتوسطة وعالية عند التوزيعات المتطرفة بالالتواء.

وكشفت دراسة سونج (Seong, 1990) التي هدفت للتحقق من مدى حساسية تقديرات معالم الفقرة لتوزيع القدرة باستخدام النموذج اللوجستي ثنائي المعلمة بأنّ تقديرات معلمة التمييز تكون أكثر دقة إذا كان شكل توزيع القدرة الحقيقة ملتويا التواءً سالباً وتوزيع شكل القدرة المسبق طبيعيا عندما يكون حجم العينة صغيراً، في حين كانت تقديرات معلمة الصعوبة أكثر دقة عندما كان شكل توزيع القدرة الحقيقة ملتوياً التواءً موجباً وحجم العينة صغيراً.

وأجرى ستون (Stone, 1992) دراسة هدفت لتقييم التقديرات الناتجة عن استخدام طريقة الارجحية العظمى الهامشية (Maximum Likelihood Estimaation) تحت شروط حجم العينة وطول الاختبار وتوزيع طبيعي ومفلطح للقدرة باستخدام النموذج اللوجستي تنائي المعلمة. كشفت نتائج الدراسة أنّ تقديرات معلمة الصعوبة بقيت مستقرة عبر توزيعات القدرة ولكن تغيرت تقديرات معلمة التمييز عندما كان التوزيع المسبق للقدرة غير طبيعي وبقيت مستقرة عند التوزيع الطبيعي للقدرة.

واهتمت دراسة ساس وسشمت ووكر (Sass, Schmitt & Walker, 2004) بدراسة أثر شكل التوزيع الملتوي للقدرة وحجم العينة وطريقة التقدير على معالم الفقرة ومعلمة القدرة للفرد وفق النموذج اللوجستي ثلاثي المعلمة، ولتحقيق الهدف من الدراسة استخدم الباحثون شكلين من توزيع القدرة ملتو(موجب، سالب) وطبيعي وحجم عينة (500، 1000) وست طرق تقدير. بينت النتائج بشكل عام بأنّ التقديرات في معالم الفقرات كانت أقل دقة

عندما كان توزيع القدرة ملتوي وأكثر دقة عندما كان توزيع القدرة طبيعي، وعند مقارنة الأخطاء المعيارية بين معلمة الصعوبة والتمييز لوحظ بأنها الأعلى عند معلمة التمييز.

وعلى صعيد الدراسات التي اهتمت بالتحقق من افتراضات نظرية الاستجابة للفقرة خصوصا فيما يتعلق بخاصية اللاتغاير(*Invariance*) في معالم الفقرات والأفراد التي اعتبرها لورد(*Lord, 1980*) بأنها من أكثر الصفات المهمة لنظرية الاستجابة للفقرة مقارنة مع النظرية التقليدية، فقد تباينت نظرة الباحثين حول إمكانية التحقق من افتراض اللاتغاير في تقديرات الفقرات وقدرة الأفراد، ولا زال هذا الجدل واضحا من خلال نتائج الدراسات التي اهتمت بذلك. فقد قام ميلر ولين (*Miller & Linn, 1988*) بدراسة هدفت إلى التحقق من مدى استقرار معالم الفقرات المستمدة من نظرية الاستجابة للفقرة، ولتحقيق الهدف من الدراسة استخدم الباحثان بيانات اختبار الرياضيات الوطني الثاني للعام 1985 بعينة حجمها 6799 طالبا وطالبة. أظهرت نتائج الدراسة أن معالم الفقرات تتغير بتغير حجم العينة.

وأجرى فان (*Fan, 1998*) دراسة هدفت لمقارنة النظرية التقليدية والنماذج ثنائية التدريج من حيث معالم الفقرات وقدرات الأفراد ومدى استقرار معالم الفقرات عبر العينات المختلفة. أشارت نتائج الدراسة إلى أنّ درجة استقرار تقديرات النظرية التقليدية ونظرية الاستجابة للفقرة كانت الأكثر عبر العينات العشوائية واقل استقراراً عبر عينات القدرات المرتفعة والمتدنية، وبينت النتائج كذلك مع أن النموذج أحادي المعلمة اقل النماذج مطابقة للبيانات إلا أن تقديرات هذا النموذج كانت أكثر استقراراً من تقديرات النموذجين الثنائي والثلاثي المعلمة.

وقامت جمحاوي (2000) بدراسة هدفت للمقارنة بين النظرية التقليدية ونظرية الاستجابة للفقرة، ولتحقيق الهدف من الدراسة استخدمت اختبار للقدرة العقلية، وطبقت الاختبار على عينة مؤلفة من 1061 طالبا وطالبة من طلبة الصف التاسع الأساسي في مديرتي اربد الأولى والثانية. أشارت نتائج الدراسة إلى وجود اتفاق عالي بين النظريتين في تقدير صعوبة الفقرات، وأشارت كذلك إلى عدم وجود فروق ذات دلالة إحصائية بين معاملي ثبات الاختبار الناتج من النظرية التقليدية والنموذج ثنائي المعلمة، بينما وجدت فروق ذات دلالة إحصائية بين معامل الثبات للاختبار الناتج من النظرية التقليدية والناتج من نموذج راش والنموذج ثلاثي المعلمة.

وفي دراسة كيكار واويتمان ولوتيش (Kelkar, Wightman & Lueech, 2000) التي هدفت للتحقق من خاصية اللاتغاير في معالم الفقرات باستخدام النموذج أحادي وثنائي وثلاثي المعلمة، ولتحقيق الهدف من الدراسة تم استخدام اختبار القبول لكلية الطب. توصلت الدراسة بأن صعوبة الفقرات استقرت باختلاف عينات الطلبة، وأنّ تقديرات مستويات قدرة الطلبة قد استقرت رغم اختلاف صعوبة الفقرات ونسخ الاختبار.

وقام عبابنه (2008) بدراسة هدفت الى استقصاء اللاتغاير في تقدير إحصائيات الفقرة المعايرة باستخدام نظرية الاستجابة للفقرة والنظرية التقليدية، ولتحقيق الهدف من الدراسة تم استخدام اختبار قدرة عقلية مكون من 71 فقرة تم تطبيقه على عينة مؤلفة من 1000طالبا وطالبة من طلبة الصف السابع الاساسي في مديرية التربية والتعليم لمنطقة اربد الأولى. اظهرت نتائج الدراسة تحقق افتراض اللاتغاير لتقديرات معلم الصعوبة ومعلم التمييز عند استخدام عينات معايرة ذات قدرة متقاربة، وتقع في مدى صعوبة الفقرة(عينة ذوي القدرة العليا، العينة العشوائية) في النموذج ثلاثي المعلمة، في حين أن افتراض اللاتغاير لتقديرات معلم الصعوبة والتمييز لم يتحقق عند استخدام النظرية الكلاسيكية، كما لم يتحقق افتراض اللاتغاير لتقديرات معلمة التخمين للفقرة المعايرة باستخدام النموذج ثلاثي المعلمة، ومع ذلك فقد توصل الباحث إلى أن افتراض اللاتغاير يعد افتراضاً نظرياً يصعب تحقيقه في الواقع العملي.

على الرغم من كثرة الدراسات التي اهتمت بنظرية الاستجابة للفقرة، واختلاف أهدافها خصوصا فيما يتعلق بفحص افتراضاتها ودراسة العوامل المؤثرة على دقة التقديرات، إلا أنّ المتصفح لها يجد غالبيتها اهتم بدراسة أثر حجم العينة وطول الاختبار على دقة التقديرات، وبنفس الوقت التحقق من افتراض اللاتغاير مقارنة مع النظرية التقليدية، وقليل منها اهتم بدراسة أثر شكل توزيع القدرة على معالم الفقرات وقدرة الأفراد، والمتصفح لهذه الدراسات يجد غالبيتها اهتم بدراسة التوزيعات الملتوية بمصاحبة متغيرات مثل حجم العينة وطول الاختبار وطريقة التقدير. لذا جاءت هذه الدراسة لتتناول أثر اختلاف شكل توزيع القدرة على معالم الفقرات ودالة المعلومات للاختبار كمحاولة لتقييم دقة وكفاءة نظرية الاستجابة للفقرة في تقدير معالم الفقرات وأخطاء القياس مستخدمة النموذج ثلاثي المعلمة لمناسبته للاختبارات الموضوعية.

مشكلة الدراسة وأسئلتها

تعد دقة تقدير معالم الفقرات والأفراد من القضايا المهمة في نظرية الاستجابة للفقرة، حيث تعتمد هذه الدقة على كثير من العوامل الذي اهتم البحث السيكومتري في دراستها. فقد تباينت وجهة نظر الباحثين حول العوامل المؤثرة على دقة التقديرات، فمنهم من يرى لحجم العينة وطول الاختبار أثر على دقة التقدير، حيث يتضح بأنّ العينات المختلفة قد تولد تقديرات مختلفة، وربما يكون بسبب انتهاك افتراضات النموذج. وكذلك تباينت وجهة النظر في انتهاك الافتراضات التي تستند عليها النظرية وأثر ذلك على التقديرات، ولكن هناك مواقف كثيرة قد تتطلب تقييم أفراد مختلفين في مستوى القدرة، خصوصا في الاختبارات التكيفية والمقننة التي قد تستخدم لأكثر من مستوى من مستويات القدرة، والذي يؤثر على خصائص الفقرات مما يستدعي دراسة أثر اختلاف توزيع القدرة على معالم الفقرات تمشيا مع التوجه السيكومتري في تقييم مدى دقة وكفاءة نظرية الاستجابة للفقرة في تقدير معالم الفقرات

ودقة القياس، وفي التحقق من افتراض اللاتغاير الذي يعد نقطة ارتكاز استقلالية القياس.لذا ستحاول هذه الدراسة على وجه التحديد الإجابة عن الأسئلة التالية :-

1. ما تقديرات معالم الفقرات (الصعوبة، التمييز، التخمين) المعايرة من كل شكل من أشكال توزيع القدرة المختلفة (طبيعي؛ ملتو نحو اليمين؛ ملتو نحو اليسار؛ وثنائي المنوال)؟
2. هل تختلف تقديرات معالم الفقرات(الصعوبة، التمييز، التخمين) باختلاف شكل التوزيع للقدرة (طبيعي؛ ملتو نحو اليمين؛ ملتو نحو اليسار؛ وثنائي المنوال)؟
3. هل تختلف دالة المعلومات للاختبار والخطأ المعياري للقياس باختلاف شكل التوزيع للقدرة (طبيعي؛ ملتو نحو اليمين؛ ملتو نحو اليسار؛ وثنائي المنوال)؟

أهمية الدراسة

إن لخاصية استقلالية القياس أهمية كبيرة عند مطوري الاختبارات والمقاييس النفسية والتربوية الذين يستخدمون نظرية الاستجابة للفقرة، حيث ساهمت هذه الخاصية في حل الكثير من الصعوبات والمشكلات السيكومترية التي تواجه النظرية التقليدية في القياس في إعطاء تفسيرات حول خصائص الأفراد والفقرات معتمدة في ذلك على افتراض التوزيع الطبيعي، ولأن التوزيع الطبيعي ليس شرطاً مسبقاً لاستخدام نظرية الاستجابة للفقرة، لذا تكمن أهمية هذه الدراسة من الناحية النظرية في دراسة استقلالية القياس من خلال معرفة الفروق بين معالم الفقرة المقدرة الناتجة من استخدام أشكال مختلفة من توزيع القدرة، وانعكاس تلك الفروق على دالة معلومات الاختبار، وبتوفير أدلة للتحقق من استقلالية القياس. أما من الناحية العملية فيمكن أن تقدم هذه الدراسة معلومات وإرشادات تساعد مطوري الاختبارات النفسية والتربوية للحصول على تقديرات دقيقة لمعالم الفقرة بناء على توزيعات مختلفة لأشكال القدرة.

مصطلحات الدراسة

النموذج اللوجستي ثلاثي المعلمة: هو أحد نماذج نظرية الاستجابة للفقرة ثنائية التدريج، حيث يمكن لهذا النموذج تقدير ثلاث معلمات هي: (معلمة الصعوبة، معلمة التمييز، ومعلمة التخمين للفقرة).

معالم الفقرات : وهي معالم الصعوبة، والتمييز، والتخمين المنبثقة عن النموذج اللوجستي ثلاثي المعلمة.

معلمة الصعوبة: وهي نقطة على متصل القدرة تقابل احتمال $(1+c_i/2)$ للإجابة على الفقرة i إجابة صحيحة، حيث يشير الرمز c_i إلى معلمة التخمين.

معلمة التمييز: وهي ميل منحنى خصائص الفقرة الذي يحدث عنده تغيير في اتجاه المنحنى(inflexion) الذي يقابل الصعوبة على متصل القدرة.

معلمة التخمين: وهي احتمال إجابة فرد ذو قدرة متدنية على فقرة ما إجابة صحيحة.

شكل توزيع القدرة: وهي الخصائص المتعلقة بالتوزيع التكراري للقدرة من حيث شكل التوزيع ووسطه الحسابي وإنحرافه المعياري والتوائه وتفلطحه.

الطريقة والإجراءات

توليد البيانات

جرى توليد البيانات باستخدام برنامج *WINGEN* وهو من تصميم وإنتاج هان وهامبيلتون (Han & Hambleton, 2007) حيث يمكن من خلال هذا البرنامج توليد استجابات لاختبارات أحادية البعد سواء كانت ثنائية التدريج أم متعددة التدريج، بالإضافة إلى اختبارات متعددة الأبعاد. وتجدر الإشارة كذلك إلى أنّ البرنامج يؤهل المستخدم 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؛ وأخيراً توزيع بيتا (Beta Distribution) الذي من خلاله يمكن توليد توزيعات ملتوية سواء كانت ملتوية التواءً موجباً أم سالباً. وقد تمت عملية توليد البيانات وفق المراحل الآتية:

المرحلة الأولى: توليد اختبار طوله 60 فقرة وفق النموذج اللوجستي ثلاثي المعلمة، وتم اختيار هذا الطول للاختبار بالاعتماد على ما أشار إليه دراسجو (Drasgow, 1982) بأنّ هذا الطول مناسب للوصول إلى دقة في التقديرات وفق النموذج اللوجستي ثلاثي المعلمة، بالإضافة إلى ذلك مناسبة هذا الطول للاختبارات التي تعد في المدارس والجامعات والاختبارات الموضوعية المقننة.وقد تم توليد معالم الفقرات وفق الشروط التالية :

- توليد معلمة التمييز للفقرات وفق توزيع Log normal ~ (0 , 0.25) حيث كان الوسط الحسابي والانحراف المعياري لمعلمة التمييز 1.024، 0.286 على الترتيب، وتعد قيم معلمة التمييز الحقيقية التي تم توليدها مماثلة للقيم الحقيقية التي استخدمها هامبلتون وسوامينيثان (Hambleton & Swaminthan, 1985) حيث أكدا على أن تكون قيم معلمة التمييز تتراوح ما بين 0 و2 لوجت.
- توليد معلمة الصعوبة للفقرات وفق التوزيع الطبيعي Normal~ (0 , 1)، حيث كان الوسط الحسابي والانحراف المعياري لقيم معلمة الصعوبة الحقيقية 0.053، 1.024 على الترتيب.

- 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، حيث كان الوسط الحسابي والانحراف المعياري لقيم معلمة التخمين 0.191، 0.061 على الترتيب. ويلخص الجدول (1) قيم معالم الفقرات التمييز والصعوبة والتخمين التي تم توليدها.

جدول 1: قيم معالم التمييز والصعوبة والتخمين التي تم توليدها وفق النموذج ثلاثي المعلمة

رقم الفقرة	التمييز a_i	الصعوبة b_i	التخمين c_i	رقم الفقرة	التمييز a_i	الصعوبة b_i	التخمين c_i
1	1.027	0.061	0.252	**31**	1.397	1.066	0.253
2	0.819	0.461	0.139	**32**	1.097	1.173	0.140
3	1.288	-0.851	0.188	**33**	0.726	-0.003	0.144
4	0.701	01.16	0.224	**34**	1.248	0.746	0.201
5	0.739	-0.064	0.281	**35**	0.906	-0.828	0.150
6	1.080	0.943	0.175	**36**	1.275	-0.067	0.139
7	0.826	0.689	0.139	**37**	1.008	0.456	0.085
8	0.796	-0.461	0.259	**38**	0.891	0.414	0.121
9	1.172	1.120	0.138	**39**	1.533	0.651	0.154
10	0.645	-0.575	0.192	**40**	1.053	-1.625	0.131
11	1.176	0.656	0.118	**41**	1.038	0.405	0.155
12	0.817	-2.815	0.170	**42**	1.514	0.665	0.271
13	1.276	0.188	0.153	**43**	0.924	-1.024	0.117
14	0.946	-0.414	0.238	**44**	0.868	-1.394	0.181
15	1.045	1.055	0.199	**45**	1.207	0.849	0.181
16	0.641	-1.398	0.212	**46**	1.018	-0.420	0.195
17	0.941	-0.542	0.397	**47**	0.955	-0.689	0.146
18	0.869	-0.214	0.188	**48**	1.375	-0.609	0.193
19	2.064	0.235	0.116	**49**	1.135	2.009	0.220
20	0.798	-0.439	0.147	**50**	1.451	-0.654	0.153
21	0.895	0.007	0.164	**51**	1.096	0.266	0.271
22	1.500	-0.669	0.206	**52**	0.698	1.484	0.174
23	0.974	-0.850	0.178	**53**	1.442	0.180	0.217
24	1.354	-0.982	0.180	**54**	0.522	1.477	0.189
25	1.218	0.949	0.115	**55**	1.013	-0.583	0.215
26	0.674	2.288	0.237	**56**	0.720	-2.322	0.137
27	1.390	-0.917	0.258	**57**	0.849	0.026	0.280
28	0.952	-0.925	0.392	**58**	0.765	-0.542	0.169
29	1.103	0.127	0.177	**59**	0.794	0.775	0.256
30	1.384	1.430	0.206	**60**	1.114	2.039	0.162

المرحلة الثانية: توليد استجابات 10000 مفحوص باستخدام القيم نفسها للمعالم الحقيقة للفقرات التي تم توليدها في المرحلة الأولى لكل توزيع من توزيعات القدرة وفق الظروف الآتية:

(1) توليد استجابات 10000 فرد وفق التوزيع الطبيعي Normal (0 , 1) ~، حيث كان الوسط الحسابي والانحراف المعياري لهذا التوزيع 0.0005، 1.01 على الترتيب، وكانت درجة الالتواء والتفلطح له 0.02 -، 0.07 على الترتيب.

(2) توليد استجابات 10000 فرد وفق توزيع بيتا Beta ~ (2 , 3) حيث كان توزيع القدرة ملتوياً نحو اليمين بوسط حسابي وانحراف معياري 0.60-، 1.20، وكانت درجة الالتواء والتفلطح 0.263، - 0.67 على الترتيب.

(3) توليد استجابات 10000 فرد وفق توزيع بيتا Beta ~ (3 , 2)، وكان توزيع القدرة ملتويا نحو اليسار بوسط حسابي وانحراف معياري 0.61 ، 1.19 على الترتيب، وبلغت قيمة درجة الالتواء والتفلطح لهذا التوزيع 0.29 –، 0.66 – على الترتيب.

(4) اختيار استجابات 5000 فرد عشوائياً من توزيع القدرة الملتوي نحو اليمين، وكذلك اختيار استجابات 5000 فرد عشوائياً من توزيع القدرة الملتوي نحو اليسار، وقد تم دمج

الاستجابات لتشكل استجابات 10000 فرد وذلك لعمل توزيع ثنائي المنوال، وبلغت قيمة الوسط الحسابي والانحراف المعياري لهذا التوزيع 0.006، 1.34 على الترتيب، وكانت درجة الالتواء والتفلطح له 0.01-، 0.86- على الترتيب.

وللتحقق من شكل توزيع الاستجابات التي تم توليدها في المرحلة الثانية بأنها تأخذ شكل التوزيع الطبيعي أم لا. فقد قام الباحثان باختبار الأشكال السابقة لتوزيع القدرة من خلال استخدام اختبار كولموغروف-سمرنوف (Kolmogorove-Smirnov)، وبينت نتائج اختبار كولموغروف-سمرنوف بأن قيمه كانت دالة إحصائياً عند مستوى دلالة ($\alpha = 0.01$) لكل من التوزيع الملتوي نحو اليمين واليسار وثنائي المنوال، مما يعني ابتعادها عن التوزيع الطبيعي في حين لم تكن قيمته دالة إحصائيا للتوزيع الطبيعي.

تحليل البيانات

لتحقيق الهدف من الدراسة، قام الباحثان باختيار استجابات 1000 فرد عشوائيا من كل توزيع من التوزيعات الأربعة للقدرة بحيث تم تشكيل أربع عينات، وقد تم الحفاظ على طول الاختبار وحجم العينات لما لها من أثر على دقة التقديرات، وفي الوقت نفسه اختار الباحثان طول الاختبار 60 فقرة وحجم العينات 1000 فرد بالاستناد على دراسة دراسجو (Drasgow, 1982) للوصول إلى دقة في التقديرات عند استخدام النموذج اللوجستي ثلاثي المعلمة، وكذلك أوصى لورد (Lord, 1980) بأن يكون طول الاختبار 50 فقرة وعدد الأفراد 1000 للحصول على أفضل التقديرات. وقد استخدم الباحثان برامج SPSS, BILOG-MG 3, NOHARM لتحليل البيانات المولدة، وقد تمت عملية التحليل وفق المراحل الآتية:-

أولا: التحقق من افتراضات النموذج

للتحقق من افتراضات النموذج ثلاثي المعلمة ,وهي: أحادية البعد لما له من أثر على دقة التقديرات، وذلك باستخدام التحليل العاملي الاستكشافي لكل توزيع من التوزيعات الأربعة باستخدام طريقة المكونات الرئيسة (Principle Component Analysis) ومن ثم تدوير العوامل باستخدام طريقة (Varimax). والجدول (2) يوضح ذلك.

جدول 2: قيم الجذر الكامن ونسب التباين المفسر للعامل الأول والثاني وناتج قسمة قيمة الجذر الأول على العامل الثاني في التوزيعات المختلفة

شكل التوزيع		العامل الأول	العامل الثاني	ناتج القسمة
طبيعي	الجذر الكامن	12.122	1.96	6.172
	التباين المفسر	20.203	3.273	
ملتو نحو اليمين	الجذر الكامن	12.863	1.988	6.472
	التباين المفسر	24.446	3.313	
ملتو نحو اليسار	الجذر الكامن	16.063	2.150	7.472
	التباين المفسر	26.771	3.584	
ثنائي المنوال	الجذر الكامن	17.148	2.127	8.060
	التباين المفسر	28.580	3.546	

يتضح من النتائج الواردة في الجدول (2) بأن جميع نسب التباين المفسر من العامل الأول كانت أعلى من 20 لجميع التوزيعات، وأن ناتج قسمة الجذر الكامن للعامل الأول على العامل الثاني اكبر من 2، مما يشير إلى وجود سمة سائدة للاختبار المولدة فقراته؛ وهذا يعني تحقق افتراض أحادية البعد للاختبار (Hatti, 1984 ; Reckase, 1997).

كما تم استخدام برنامج NOHARM بغرض التحقق من أحادية البعد للبيانات من خلال إجراء تحليل البواقي Residual Analysis، حيث يتم فحص المواءمة بين عدد الأبعاد التي يتم تحديدها للاختبار والبيانات من خلال حساب مصفوفة البواقي Residual Matrix لإيجاد قيمة الجذر التربيعي لمتوسطات مربعات البواقي Root Mean Square of Residuals (RMSR). وبين ميلر (Miller, 1991) أنه إذا كانت قيمة مؤشر RMSR قريبة من الصفر فهذا دليل على تحقق أحادية البعد، وبالوقت نفسه إذا كانت قيمة مؤشر تناكا (Tanaka Index) قريبة من الواحد يعد دليلاً آخر على أحادية البعد. ويلخص الجدول (3) قيم RSMR وقيم مؤشر تناكا للبيانات في التوزيعات المختلفة.

جدول 3: قيم الجذر التربيعي لوسط مربعات البواقي RMSR **ومؤشر تناكا للبيانات في التوزيعات المختلفة**

شكل التوزيع	RMSR	مؤشر تناكا
طبيعي	0.006	0.982
ملتو نحو اليمين	0.008	0.978
ملتو نحو اليسار	0.003	0.995
ثنائي المنوال	0.006	0.991

يتضح من النتائج الواردة في الجدول (3) بأن قيم الجذر التربيعي لمتوسط مربعات البواقي تقترب من الصفر، وبالوقت نفسه جاءت قيم مؤشر تناكا للمطابقة قريبة من الواحد لجميع التوزيعات مما يبين تحقق افتراض أحادية البعد في البيانات. ونظراً لتحقق افتراض أحادية البعد في البيانات فهذا يعني تحقق افتراض الاستقلال الموضعي الذي يعد مؤشراً لأحادية البعد إذا كان النموذج المستخدم يحدد قيمة تقديرية لقدرة الفرد على بعد أحادي (Emberston & Reise, 2000; Hambleton & Swaminthan, 1985).

ثانياً : مطابقة البيانات للنموذج

بهدف التحقق من مطابقة الأفراد والفقرات للنموذج اللوجستي ثلاثي المعلمة، تم استخدام برنامج BILOG-MG 3 لتحليل البيانات لكل توزيع من التوزيعات الأربعة، وأجريت عملية التحليل لكل توزيع كما يلي:

1. حللت بيانات التوزيع الطبيعي (1000 فرد) حيث أظهرت نتائج التحليل بأن جميع الفقرات جاءت متطابقة مع النموذج باستثناء الفقرة رقم 56 حيث كانت قيمة مربع كاي χ^2 لها ذات دلالة إحصائية عند مستوى دلالة ($\alpha = 0.01$). وكانت قيمة χ^2 (0.0041). وبينت نتائج التحليل بأنّ جميع استجابات الأفراد جاءت متطابقة مع توقعات النموذج باستثناء مفحوصين (705، 353) حيث كانت قيم مربع كاي ذات دلالة إحصائية عند مستوى دلالة ($\alpha = 0.01$).
2. حللت بيانات التوزيع الملتوي نحو اليمين (1000 فرد) بينت نتائج التحليل بأنّ جميع الفقرات قد تطابقت مع النموذج، وبالوقت نفسه تطابقت استجابات الأفراد باستثناء فرد واحد (453).
3. حللت بيانات التوزيع الملتوي نحو اليسار(1000 فرد) أظهرت نتائج التحليل بأن جميع الفقرات جاءت متطابقة مع النموذج باستثناء الفقرة 52 حيث كانت قيمة مربع كاي دالة إحصائياً عند مستوى الدلالة ($\alpha = 0.01$) وبلغت قيمته 0.0054، وبينت النتائج كذلك بأنّ جميع استجابات الأفراد متطابقة مع النموذج باستثناء ثلاثة أفراد (63، 414، 994).
4. حللت بيانات التوزيع ثنائي المنوال (1000 فرد) أظهرت نتائج التحليل بأن جميع الفقرات جاءت متطابقة مع النموذج باستثناء الفقرة 40 حيث كانت قيمة مربع كاي دالة إحصائياً عند مستوى الدلالة ($\alpha = 0.01$) وبلغت قيمته 0.0006، وبينت النتائج كذلك بأنّ جميع استجابات الأفراد قد تطابقت مع النموذج باستثناء فردين (219، 356).

بالاعتماد على نتائج الخطوة الثانية، تم حذف الفقرات التي لم تكن متطابقة مع النموذج اللوجستي ثلاثي المعلمة وهي الفقرات التي تحمل الأرقام 40، 52، 56 وبذلك تم الاحتفاظ بـ 57 فقرة مطابقة للنموذج، أما بالنسبة للأفراد فقد ارتأى الباحثان إبقاء حجم العينات 1000 فرد نظراً لقلة عدد الأفراد وعددهم (8) أفراد الذين لم تتطابق استجاباتهم مع النموذج لإجراء عملية المقارنة بين معالم الفقرات الناتجة من كل توزيع من التوزيعات المختلفة للقدرة.

المعالجة الإحصائية

من أجل الإجابة عن أسئلة الدراسة، تم استخدام المعالجات الإحصائية الآتية:-

- إيجاد معاملات الصعوبة، التمييز، التخمين لكل شكل من أشكال توزيعات القدرة باستخدام النموذج اللوجستي ثلاثي المعلمة.
- إيجاد معامل ارتباط بيرسون بين معالم الفقرات الناتجة من كل توزيع من توزيعات القدرة.
- استخدام تحليل التباين للقياسات المتكررة واختبار أدنى فرق Least Significant Differences (LSD) للكشف عن الفروق في معالم الفقرات.

عرض النتائج ومناقشتها

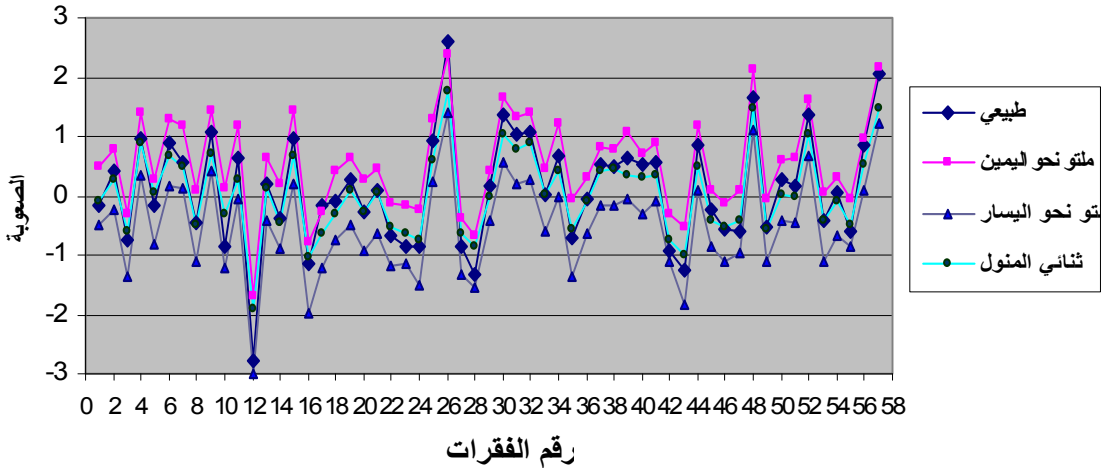
للإجابة عن السؤال الأول والمتعلق بإيجاد قيم معالم الفقرات (الصعوبة، التمييز، التخمين) لكل توزيع من توزيعات القدرة المختلفة باستخدام النموذج ثلاثي المعلمة، فقد تم استخدام برنامج BILOG-MG 3 لإيجاد قيم معالم الفقرات. ويوضح الجدول (4) قيم معالم الفقرات لكل توزيع من توزيعات القدرة.

جدول 4: قيم معالم الفقرات (الصعوبة، التمييز، التخمين) لكل توزيع من توزيعات القدرة

رقم الفقرة	شكل التوزيع: طبيعي			ملتو نحو اليمين			ملتو نحو اليسار			ثنائي المنوال		
	الصعوبة	التمييز	التخمين	الصعوبة	التمييز	التخمين	الصعوبة	التمييز	التخمين	الصعوبة	التمييز	التخمين
1	-0.147	1.584	0.131	0.504	1.799	0.225	-0.504	2.072	0.263	-0.077	2.259	0.268
2	0.421	1.195	0.109	0.797	1.613	0.107	-0.255	1.577	0.102	0.259	1.911	0.167
3	-0.738	2.426	0.237	-0.309	2.070	0.162	-1.378	2.384	0.233	-0.589	2.994	0.205
4	0.950	1.289	0.179	1.399	1.235	0.193	0.332	1.252	0.203	0.874	1.623	0.226
5	-0.180	1.287	0.262	0.269	1.449	0.220	-0.834	1.426	0.215	0.058	2.295	0.329
6	0.897	2.112	0.185	1.298	2.311	0.200	0.157	2.036	0.123	0.661	2.574	0.166
7	0.549	1.501	0.140	1.198	2.150	0.205	0.118	1.784	0.172	0.480	2.057	0.155
8	-0.470	1.551	0.268	0.092	1.531	0.230	-1.124	1.680	0.216	-0.492	1.768	0.219
9	1.080	2.661	0.147	1.425	2.284	0.121	0.413	2.563	0.165	0.724	2.744	0.142
10	-0.839	0.907	0.208	0.127	1.404	0.232	-1.215	1.102	0.184	-0.323	1.668	0.270
11	0.618	1.967	0.121	1.164	2.843	0.165	-0.059	2.624	0.140	0.281	2.616	0.091
12	-2.795	1.472	0.218	-1.696	1.762	0.191	-2.995	1.832	0.197	-1.912	2.230	0.199
13	0.195	2.274	0.164	0.635	2.480	0.171	-0.427	2.838	0.155	0.136	2.737	0.155
14	-0.368	1.518	0.268	0.203	1.750	0.231	-0.897	1.718	0.273	-0.467	1.962	0.170
15	0.976	1.902	0.212	1.429	2.009	0.200	0.198	2.028	0.192	0.679	2.805	0.220
16	-1.138	1.100	0.303	-0.785	1.253	0.191	-1.964	1.289	0.234	-1.019	1.461	0.233
17	-0.152	2.197	0.444	-0.278	1.531	0.212	-1.200	1.876	0.202	-0.626	2.039	0.333
18	-0.084	1.907	0.227	0.431	2.347	0.254	-0.734	1.981	0.209	-0.297	1.783	0.176
19	0.284	3.690	0.124	0.646	4.221	0.112	-0.507	4.435	0.071	0.101	5.039	0.117
20	-0.289	1.519	0.212	0.271	1.717	0.156	-0.924	1.721	0.144	-0.270	1.669	0.145
21	0.106	1.700	0.214	0.456	1.639	0.163	-0.643	1.608	0.151	0.039	2.248	0.198
22	-0.671	2.617	0.205	-0.121	2.971	0.170	-1.188	3.065	0.203	-0.529	3.271	0.201
23	-0.860	1.637	0.228	-0.170	1.979	0.195	-1.138	1.929	0.205	-0.630	2.322	0.186
24	-0.853	3.168	0.238	-0.247	2.720	0.201	-1.498	2.303	0.209	-0.762	3.416	0.136
25	0.914	2.562	0.109	1.309	2.768	0.145	0.254	2.480	0.132	0.616	2.399	0.121
26	2.590	1.192	0.244	2.375	1.562	0.259	1.397	1.126	0.212	1.748	1.630	0.263
27	-0.841	2.385	0.256	-0.389	2.456	0.187	-1.322	2.908	0.211	-0.648	3.579	0.264
28	-1.345	1.388	0.208	-0.675	1.504	0.273	-1.545	2.163	0.218	-0.871	1.848	0.319
29	0.161	2.44	0.213	0.426	1.883	0.112	-0.427	2.123	0.180	-0.036	2.348	0.136
30	1.367	2.489	0.217	1.670	3.292	0.204	0.563	2.471	0.193	1.040	2.516	0.217
31	1.029	2.833	0.254	1.319	2.309	0.222	0.193	2.562	0.232	0.769	3.263	0.238
32	1.076	2.033	0.132	1.409	2.584	0.148	0.273	2.218	0.117	0.873	2.421	0.174
33	0.005	1.258	0.139	0.465	1.588	0.143	-0.613	1.605	0.154	0.021	1.808	0.147
34	0.676	2.503	0.200	1.202	2.606	0.208	-0.014	2.568	0.213	0.434	2.733	0.176
35	-0.698	1.661	0.177	-0.065	1.745	0.186	-1.380	1.766	0.870	-0.566	2.379	0.208
36	-0.066	2.725	0.147	0.326	2.116	0.107	-0.632	2.788	0.111	-0.077	2.883	0.144
37	0.540	1.917	0.135	0.805	2.046	0.099	-0.158	2.293	0.103	0.403	2.814	0.150
38	0.474	1.713	0.193	0.799	1.885	0.126	-0.165	1.862	0.199	0.443	2.452	0.179
39	0.650	2.767	0.139	1.060	3.902	0.172	-0.056	3.309	0.160	0.345	3.491	0.150
40	0.528	2.141	0.226	0.716	1.950	0.147	-0.303	1.795	0.195	0.313	2.564	0.170
41	0.574	2.538	0.228	0.906	3.055	0.247	-0.087	3.230	0.244	0.338	3.974	0.289
42	-0.937	1.694	0.199	-0.304	2.133	0.193	-1.126	2.378	0.304	-0.743	1.905	0.156
43	-1.272	1.476	0.162	-0.533	2.216	0.227	-1.839	1.781	0.175	-1.008	2.050	0.177
44	0.858	2.586	0.199	1.179	2.444	0.170	0.085	2.448	0.217	0.505	2.983	0.158
45	-0.240	1.919	0.303	0.105	1.696	0.195	-0.869	2.274	0.304	-0.421	2.168	0.169
46	-0.554	1.495	0.211	-0.116	1.861	0.150	-1.093	2.023	0.190	-0.523	2.476	0.149
47	-0.587	2.300	0.204	0.094	2.823	0.224	-0.976	3.112	0.245	-0.428	2.811	0.222
48	1.669	1.651	0.205	2.133	3.338	0.243	1.120	1.820	0.206	1.470	2.731	0.228
49	-0.511	3.071	0.187	-0.043	2.932	0.151	-1.124	3.398	0.179	-0.557	3.114	0.119
50	0.263	2.290	0.310	0.615	2.277	0.230	-0.425	2.173	0.268	0.018	1.888	0.196
51	0.162	2.398	0.178	0.624	3.088	0.223	-0.453	2.988	0.214	-0.002	3.049	0.178
52	1.348	1.137	0.209	1.620	1.694	0.226	0.657	1.319	0.220	1.033	1.416	0.210
53	-0.416	1.851	0.320	0.056	2.189	0.248	-1.126	2.016	0.171	-0.406	2.214	0.270
54	0.055	1.522	0.311	0.320	1.533	0.199	-0.665	1.568	0.192	-0.092	1.943	0.307
55	-0.599	1.387	0.185	-0.044	1.500	0.139	-0.839	1.762	0.319	-0.491	1.669	0.148
56	0.853	1.425	0.304	0.982	1.509	0.216	0.098	1.827	0.263	0.527	1.828	0.273
57	2.060	1.555	0.162	2.148	2.172	0.155	1.203	2.051	0.166	1.459	2.441	0.151
الوسط الحسابي	0.111	1.952	0.209	0.548	2.171	0.187	-0.520	2.164	0.196	0.031	2.444	0.196
الانحراف المعياري	0.926	0.596	0.063	0.791	0.643	0.043	0.815	0.634	0.051	0.695	0.675	0.057

يتضح من النتائج الواردة في الجدول (4) تبايناً ملحوظاً في مدى صعوبة الفقرات، حيث كان المدى في صعوبة الفقرات هو الأعلى في التوزيع الطبيعي الذي تراوحت قيمه من 2.795 – إلى 2.590، في حين كان الأقل في التوزيع ثنائي المنوال حيث تراوحت قيمه من 1.912 – إلى 1.748. بينما كان متقارباً في التوزيعين الملتوي نحو اليمين واليسار، ويوضح الشكل (1) هذا التباين في مدى صعوبة الفقرات.

ويلاحظ من الشكل (1) بأنّ التوزيعيْن الملتويْين نحو اليمين واليسار قد أنتجا فقرات صعبة وسهلة، وهذه النتيجة تتفق مع ما أشار إليه ستوكينغ (Stocking, 1990) بأنّ الأفراد ذوي القدرة العالية ينتج عن تقديرات تدل على فقرات سهلة، والأفراد ذوي القدرة المتدنية تنتج عن تقديرات تدل على فقرات صعبة، وبالوقت نفسه تعارضت مع نتائج دراسة ستون(Stone, 1992) بأنّ معالم الصعوبة بقيت ثابتة، وهذا يدل على أن معلمة الصعوبة قد تأثرت بخصائص القدرة للأفراد. ونظراً لتأثر صعوبة الفقرة فقد أثر ذلك على معلمة التمييز أيضا، حيث كان الوسط الحسابي لمعلمة التمييز الأعلى في التوزيع ثنائي المنوال (2.444)، ويعود السبب في ذلك لوجود قدرات متطرفة ناتجة عن دمج أفراد ذوي قدرة عالية وأفراد ذوي قدرة متدنية مما جعل معاملات الصعوبة لأن تكون صعبة وسهلة وبذلك انتج معاملات تمييز عالية، يلي ذلك التوزيعين الملتويين نحو اليمين واليسار(2.171، 2.164). بينما أشارت نتائج الجدول (4) بأنّ التوزيع الطبيعي كان الأقل في الوسط الحسابي لمعلمة التمييز (1.952)، والسبب في ذلك يعود للتباين في قدرة الأفراد مما أعطى مدى واسعا في صعوبة الفقرات، ومع ذلك فإنها تعتبر قيم عالية.



شكل (1) : مخطط توضيحي لقيم معاملات الصعوبة لكل توزيع من توزيعات القدرة

أما بالنسبة لمعلمة التخمين، فقد أشارت نتائج الجدول (4) بأنّ الأوساط الحسابية للتخمين كانت متقاربة لجميع توزيعات القدرة، وقد جاءت متفقة مع الاتجاه العام لصعوبة الفقرة حيث كانت معلمة التخمين للفقرة الصعبة عالية، مما يعني لجوء الأفراد ذوي القدرة المتدنية للإجابة عن مثل هذه الفقرة إلى التخمين.

وللإجابة عن السؤال الثاني حول الاختلافات بين معالم الفقرات عند استخدام أشكال مختلفة لتوزيع قدرة الأفراد. فقد تم استخدام تحليل التباين الأحادي للقياسات المتكررة (one way ANOVA for repeated measures)، ويوضح الجدول (5) نتائج تحليل التباين للقياسات المتكررة لمعرفة دلالة الفروق في معالم الفقرات تبعا لشكل توزيع القدرة.

جدول 5 : نتائج تحليل التباين للقياسات المتكررة لأثر شكل توزيع القدرة على تقدير معالم الفقرات

معالم الفقرات	مصدر التباين	درجات الحرية	مجموع المربعات	متوسط المربعات	قيمة f	الدلالة الاحصائية
الصعوبة	عدد الفقرات	56	144.161	2.574	581.838	0.0000*
	شكل التوزيع	3	32.884	10.961		
	الخطأ	168	3.165	0.019		
	الكلي	227	180.210			
التمييز	عدد الفقرات	56	79.293	1.446	32.597	0.0000*
	شكل التوزيع	3	6.931	2.130		
	الخطأ	168	11.907	0.071		
	الكلي	227	98.130			
التخمين	عدد الفقرات	56	0.141	0.007	3.210	0.02400*
	شكل التوزيع	3	0.014	0.005		
	الخطأ	168	0.236	0.001		
	الكلي	227	0.664			

* دال إحصائيا عند مستوى الدلالة $\alpha = 0.01$

جدول 6: نتائج المقارنات الثنائية بين الأوساط الحسابية لمعلمتي الصعوبة والتمييز لكل توزيع من توزيعات القدرة

معالم الفقرة	شكل التوزيع	المتوسط الحسابي	ملتو نحو اليمين	ملتو نحو اليسار	ثنائي المنوال
الصعوبة	طبيعي	0.111	0.437*	0.631*	0.080
	ملتو نحو اليمين	0.548		1.068*	0.517*
	ملتو نحو ليسار	- 0.520			0.551*
	ثنائي المنوال	0.031			
التمييز	طبيعي	1.952	0.219*	0.212*	0.492*
	ملتو نحو اليمين	2.171		0.007	0.273*
	ملتو نحو اليسار	2.164			0.280*
	ثنائي المنوال	2.444			

* دال إحصائيا عند مستوى الدلالة $\alpha = 0.01$

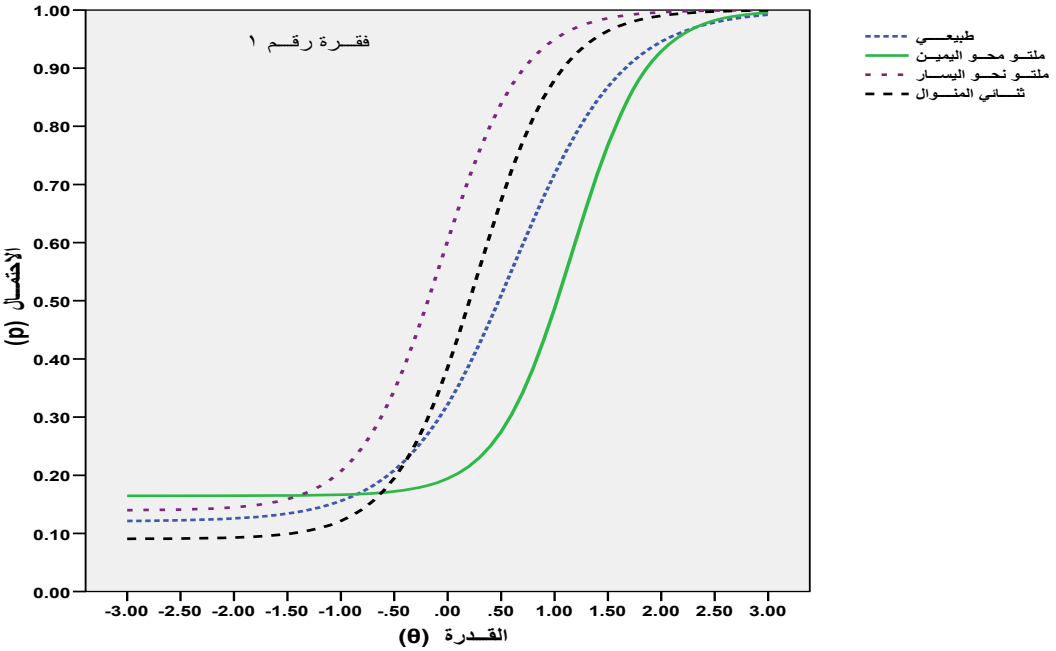
تشير النتائج الواردة في الجدول (5) إلى أنه توجد فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha = 0.01$) بين متوسطات معالم الصعوبة للفقرات وكذلك بين متوسطات معالم التمييز للفقرات، في حين لم تكن دالة إحصائياً في معلمة التخمين الذي يشير بالتالي إلى تأثر معالم الفقرات بشكل توزيع القدرة. وللكشف عن مواقع الفروق بين المتوسطات، تم إجراء المقارنات البعدية باستخدام أدنى فرق Least Significant (LSD) Differences. ويوضح الجدول (6) مواقع هذه الفروق.

يكشف الجدول (6) أن هناك فروق ذات دلالة إحصائية ($\alpha = 0.01$) بين الأوساط الحسابية لمعاملات الصعوبة المقدرة تبعا لكل شكل من أشكال توزيعات القدرة باستثناء الفرق بين الوسطين الحسابين لصعوبة الفقرات الناتجة من التوزيع الطبيعي وثنائي المنوال، حيث أظهرت هذه الفروق أن أعلى فرق كان بين الوسطين الحسابين لصعوبة الفقرات الناتجة من التوزيع الملتوي نحو اليمين ونحو اليسار كانت عالية، إذ بلغت قيمة الفرق 1.068، وقد عزا الباحثان ذلك للتطرف بين قدرات المفحوصين والتي أظهرت خاصيتي اثر السقف Ceiling effect واثر الحد الأدنى Floor effect. في حين أظهرت الفروق بان اقل فرق كان بين الوسطين الحسابين لصعوبة الفقرة الناتج من التوزيع الطبيعي وثنائي المنوال حيث بلغت قيمته 0.080 ولم يكن دالا إحصائيا، مما يعني تقارب معالم صعوبة الفقرات لكلا التوزيعين.

أما بالنسبة لمعاملات التمييز، فقد أشارت نتائج الفروق الواردة في الجدول (6) بأن هناك فروقا ذات دلالة إحصائية بين متوسطات معالم التمييز الناتجة من كل توزيع من توزيعات القدرة باستثناء الفرق بين وسطي معلمتي التمييز الخاصة بالتوزيعين الملتو نحو اليمين والملتوي نحو اليسار، ويعود السبب في ذلك إلى تجانس القدرة في التوزيعين والذي انعكس بالتالي على صعوبة الفقرات وتقارب معاملات التمييز، بينما أشارت نتائج الفروق بأنّ قيمة الفرق بين الوسط الحسابي لمعلمة التمييز في التوزيع الطبيعي

والتوزيع ثنائي المنوال كان أعلى قيمة حيث بلغت 0.492، وقيمة هذا الفرق كانت عالية ويعود الى أن التوزيع ثنائي المنوال أنتج أعلى معاملات لتمييز الفقرات.

وتجدر الإشارة كذلك بأنّ هذه النتائج المتعلقة بتباين معالم الفقرات خصوصا فيما يتعلق بالصعوبة والتمييز جاءت متفقة مع نتائج دراسة ميلر ولين (Miller & Linn, 1988) بوجود اختلافات كبيرة بين المنحنيات المميزة للفقرة عبر المجموعات المختلفة، والذي ربما بدوره يؤدي إلى تغير معالم الفقرات عبر المجموعات المختلفة وعدم استقرارها، وهذا يجعل من افتراض اللاتغاير في معالم الفقرات اقرب إلى الافتراض النظري منه إلى الافتراض العملي. وهذا ما اتضح عند متابعة منحنيات خصائص الفقرة عبر التوزيعات المختلفة، فعلى سبيل المثال يوضح الشكل (2) منحنى خاصية الفقرة للفقرة رقم 1 عبر التوزيعات المختلفة.



شكل (2) : منحنى خصائص الفقرة 1 لكل توزيع من توزيعات القدرة المختلفة.

يتضح من الشكل (2) بأنّ منحنيات خصائص الفقرة 1 أظهرت تحيزاً منتظماً عبر التوزيعات المختلفة، فقد بين الشكل بأنّ احتمالات الإجابة الصحيحة يكون أكبر دائماً عندما كان توزيع القدرة ملتو نحو اليسار، والأقل دائما عندما يكون توزيع القدرة ملتو نحو اليمين، في حين كانت احتمالات الإجابة الصحيحة للتوزيعين الطبيعي وثنائي المنوال كانت متقاربة عندما كان توزيع القدرة في المدى -1 و1 واختلفت عند الأطراف، وهذا نتج بسب اقتراب صعوبة الفقرات من قدرات الأفراد. ويتضح من الشكل كذلك بأن عامل التخمين كان الأعلى دائما في التوزيع الملتوي نحو اليسار وذلك بسب صعوبة الفقرات، بينما كان الأقل دائما عندما كان توزيع القدرة طبيعي.

وعلى الرغم من تباين معالم الفقرات عبر التوزيعات المختلفة ودلالة هذا التباين، فقد تم إيجاد معاملات الارتباط بين معالم الفقرات لكل توزيع من توزيعات القدرة لملاحظة مدى التوافق فيما بينها عبر التوزيعات المختلفة للقدرة، ويبين الجدول (7) قيم هذه المعاملات

جدول 7: قيم معاملات الارتباط بين تقديرات معالم الفقرات لكل توزيع من توزيعات القدرة

معالم الفقرة	شكل التوزيع	ملتو نحو اليمين	ملتو نحو اليسار	ثنائي المنوال
الصعوبة	طبيعي	* 0.981	* 0.984	* 0.983
	ملتو نحو اليمين		* 0.988	* 0.988
	ملتو نحو اليسار			* 0.986
التمييز	طبيعي	* 0.772	* 0.867	* 0.841
	ملتو نحو اليمين		* 0.834	* 0.793
	ملتو نحو اليسار			* 0.862
التخمين	طبيعي	* 0.492	* 0.518	*.620
	ملتو نحو اليمين	1	* 0.527	* 0.650
	ملتو نحو اليسار		1	* 0.360

* دال إحصائيا عند مستوى الدلالة $\alpha = 0.01$

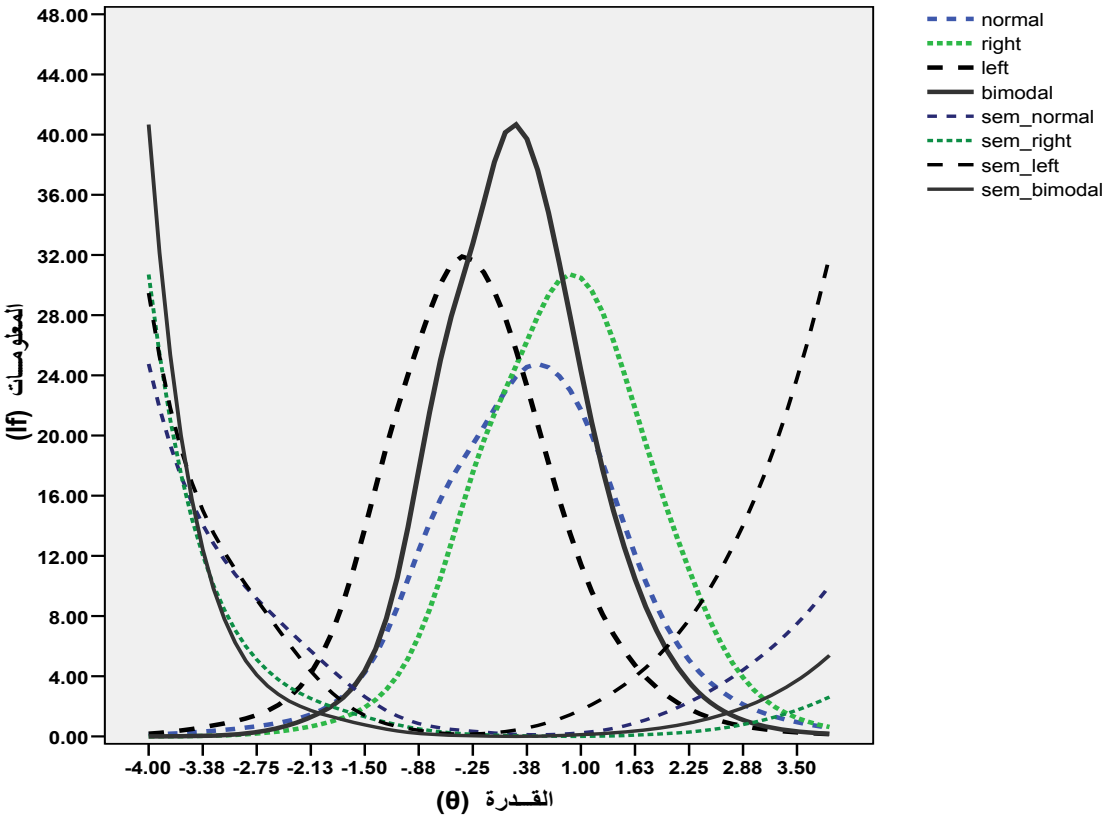
أظهرت النتائج الواردة في الجدول (7) أن جميع معاملات ارتباط معلم الصعوبة للفقرات المعايرة لكل توزيع من توزيعات القدرة كانت عالية ودالة إحصائيا، حيث تراوحت هذه القيم ما بين 0.981 و0.988. وبالنظر إلى هذه القيم يلاحظ بأن جميعها كانت متقاربة. وبينت النتائج كذلك بأن قيم معاملات الارتباط معلم التمييز للفقرات المعايرة لكل من التوزيعات الأربعة تراوحت قيمها ما بين 0.772 و0.867 وهي دالة إحصائيا إلا أنها اقل من قيم معاملات الارتباط لمعلم الصعوبة. وأظهرت النتائج كذلك بأنّ قيم معاملات الارتباط لمعلمة التخمين جاءت دالة إحصائيا وتراوحت قيمها ما بين 0.360 و0.650.

إن النتائج التي وردت في الجدول (7) جاءت متفقة مع نتائج الدراسات التي تناولت قيم معاملات الارتباط بين تقديرات معالم الفقرات باختلاف عينات المعايرة بحسب القدرة من مثل دراسة ميلر ولين(Miller & Linn, 1988)،دراسة كيكار وايتمان ولويتش(Kelkar, Wightman & Liuech, 2000)، دراسة جمحاوي (2000) ودراسة عبابنة (2008) التي اعتمدت على قيمة معامل الارتباط كمؤشر لفحص سلوك تقديرات معالم الفقرات عند استخدام عينات مختلفة في القدرة للتحقق من افتراض اللاتغاير للفقرات استنادا لما أشار إليه هامبلتون وسوامينيثان (Hambleton & Swaminthan, 1985).

وعلى الرغم من التوافق الذي أظهره معامل الارتباط بين التقديرات المختلفة لمعالم الفقرات المعايرة بتوزيعات مختلفة للقدرة كمؤشر إحصائي يعتمد عليه في فحص الثبات النسبي لتقديرات معالم الفقرات عبر العينات المختلفة، إلا أن ذلك قد تعارض مع فحص دلالة الفروق بين هذه المعالم كما أظهرها تحليل التباين للقياس المتكرر، مما يعني أن العلاقة الارتباطية لم تعكس حالة عدم الاستقرار في معلمتي الصعوبة والتمييز، وهذا يقود إلى أن خصائص الفقرات قد تتأثر بخصائص الأفراد، مما يجعل تحرر أدوات القياس من خصائص الفقرات وتحرر قدرة الأفراد من خصائص أدوات القياس اقرب إلى الافتراض النظري منه إلى الافتراض العملي لعدم توفر المؤشرات الإحصائية الكافية للتحقق من ذلك، ويعبر عن هذا الافتراض بما يسمى بخاصية اللاتغاير عبر العينات المختلفة، وهذا يفتح مجالا خصبا للدراسات الامبريقية للتحقق من ذلك الافتراض من خلال استخدام الخطأ المعياري للتنبؤ(Standard error of estimate) لتقييم نتائج دقة التقدير لمعالم الفقرات عبر توزيعات مختلفة للقدرة مع الأخذ بعين الاعتبار حجم العينات وطول الاختبار كعوامل مؤثرة على تقدير الفقرات.

أما بالنسبة للسؤال الثالث والمتعلق بأثر شكل توزيع القدرة على دالة المعلومات والخطأ المعياري للفقرات والمقياس ككل، حيث تعد دالة معلومات الاختبار Test Information Function (TIF) من المؤشرات التي يستدل منها على معامل الثبات للمقياس في نظرية الاستجابة للفقرة، ويعمل منحنى دالة المعلومات الاختبار عكس منحنى الخطأ المعياري للقياس، وبالتالي فان زيادة كمية المعلومات يؤدي إلى نقصان الخطأ المعياري للقياس. وقد بين ريفي (Reeve, 2004) بأن دالة المعلومات للاختبار ترتبط مع ثبات المقياس من خلال العلاقة التالية $r = 1 - \frac{1}{\sum_{i=1}^{I} I(\theta)}$ ، فيتوقع أنه كلما زادت دالة المعلومات للاختبار عند مستوى معين من القدرة يزيد الثبات بمعنى أنه يقل الخطأ المعياري للقياس الذي يتيح الفرصة لتقدير الخطأ المعياري عند كل مستوى من مستويات القدرة ومعرفة مدى مساهمة كل فقرة في تحديد دقة القياس (Hambleton, 1994).

ولدراسة أثر شكل توزيع القدرة على دالة معلومات الاختبار، تم استخدام منحنيات دوال المعلومات للاختبار والخطأ المعياري للقياس لكل توزيع من توزيعات القدرة من برنامج BILOGMG-MG 3، ولتسهيل عملية المقارنة بين منحنيات دوال المعلومات وضعت تلك المنحنيات في شكل واحد كما يظهر في الشكل (3).



شكل (3): دالة معلومات المقياس والخطأ المعياري للقياس لكل توزيع من توزيعات القدرة المختلفة

يتضح من الشكل (3) بأن اكبر كمية لدالة المعلومات تكون للاختبار المعاير من التوزيع ثنائي المنوال للقدرة، في حين كانت أقل كمية لدالة المعلومات الاختبار عندما تمت معايرته بالتوزيع الطبيعي للقدرة، وربما يعود ذلك إلى معلمة التمييز للفقرات، حيث كانت قيم التمييز للفقرات عندما تم معايرتها بتوزيع ثنائي المنوال للقدرة هي الأعلى بينما كانت الأقل عندما تم معايرتها بالتوزيع الطبيعي، فكلما زاد معلم التمييز للفقرة زادت كمية المعلومات التي تسهم بها الفقرة حول معلم الصعوبة.

أما بالنسبة لدالة معلومات الاختبار المعاير من التوزيع الملتوي نحو اليمين للقدرة، فقد كانت كمية المعلومات هي الأقصى عند مستويات القدرة العالية، بينما كانت كمية المعلومات التي يسهم بها المقياس عندما تمت معايرته من التوزيع الملتوي نحو اليسار هي الاقصى عند مستويات القدرة المنخفضة، وهذا ينسجم مع ما أشار إليه هامبلتون وسوامينيثان (Hambleton & Swaminthan, 1985) إلى أن الاختبار السهل من المتوقع أن يقدم معلومات أكبر عند مستويات القدرة المنخفضة، بينما الاختبار الصعب متوقع أن يقدم معلومات اكبر عند مستويات القدرة العالية.

ويظهر الشكل (3) كذلك تناقص قيم منحنى الخطأ المعياري للاختبار المعاير من التوزيع ثنائي المنوال في مدى القدرة -1 و، 1.50، بينما تناقصت قيم منحنى الخطأ المعياري للاختبار المعاير من التوزيع الطبيعي للقدرة في المدى -0.13 و1.13. ويبين الشكل كذلك بانّ قيم الخطأ المعياري للاختبار المعاير من التوزيع الملتوي نحو اليمين تناقصت عند المستويات العليا للقدرة على عكس قيم الخطأ المعياري للاختبار المعاير من التوزيع الملتوي نحو اليسار التي كانت قيمها الأقل عند مستويات القدرة المنخفضة. ومما تجدر الإشارة إليه بأن القيم الدنيا للخطأ المعياري للاختبار كانت في مدى القدرة -0.5 و0.63 لجميع التوزيعات المختلفة للقدرة وبالتالي فان فقرات المقياس أكثر دقة في هذا المدى بغض النظر عن شكل التوزيع لقدرات الأفراد، وبذلك فهي تعطي أقصى كمية من المعلومات في هذا المدى لارتباط الخطأ المعياري ارتباطا عكسياً مع دالة المعلومات كما أشار بيكر(Baker, 2001).

ولارتباط دالة معلومات الاختبار ككل بدوال معلومات الفقرات، ولتحديد أثر اختلاف شكل توزيع القدرة على دالة معلومات الفقرة تم تلخيص القيم القصوى لدالة معلومات الفقرة لكل شكل من أشكال توزيع القدرة للأفراد كما في الجدول (8).

جدول 8: القيم القصوى لدالة معلومات الفقرة لكل شكل من أشكال التوزيعات المختلفة للقدرة

رقم الفقرة	القيم القصوى لدالة معلومات الفقرة — شكل التوزيع				رقم الفقرة	القيم القصوى لدالة معلومات الفقرة — شكل التوزيع للقدرة			
	طبيعي	نحو اليمن	نحو اليسار	ثنائي المنوال		طبيعي	نحو اليمن	نحو اليسار	ثنائي المنوال
1	0.483	0.526	0.647	0.762	**30**	1.022	1.832	1.055	1.044
2	0.289	0.529	0.510	0.663	**31**	1.232	0.871	1.051	1.688
3	0.933	0.785	0.909	1.513	**32**	0.801	1.257	0.980	1.049
4	0.295	0.263	0.266	0.427	**33**	0.303	0.478	0.479	0.616
5	0.250	0.344	0.337	0.695	**34**	1.067	1.139	1.096	1.331
6	0.781	0.909	0.818	1.205	**35**	0.491	0.532	0.549	0.951
7	0.430	0.781	0.572	0.786	**36**	1.399	0.910	1.570	1.576
8	0.359	0.189	0.467	0.513	**37**	0.708	0.865	1.077	1.484
9	1.334	0.377	1.195	1.434	**38**	0.507	0.697	0.592	1.066
10	0.138	0.316	0.213	0.413	**39**	1.466	2.739	2.013	2.286
11	0.766	1.472	1.315	1.436	**40**	0.743	0.717	0.554	1.187
12	0.357	0.538	0.575	0.849	**41**	1.040	1.453	1.634	2.262
13	0.943	1.108	1.495	1.390	**42**	0.489	0.785	0.787	0.672
14	0.344	0.491	0.437	0.695	**43**	0.399	0.794	0.568	0.749
15	0.602	0.687	0.712	1.289	**44**	1.143	1.079	0.988	1.644
16	0.169	0.272	0.265	0.341	**45**	0.513	0.495	0.719	0.850
17	0.497	0.390	0.597	0.545	**46**	0.373	0.649	0.711	1.152
18	0.588	0.846	0.657	0.567	**47**	0.894	1.298	1.512	1.290
19	2.679	3.590	4.283	5.064	**48**	0.460	1.748	0.559	1.205
20	0.384	0.546	0.562	0.526	**49**	1.648	1.608	2.048	1.929
21	0.480	0.491	0.484	0.865	**50**	0.720	0.834	0.706	0.612
22	1.156	1.594	1.592	1.821	**51**	1.022	1.554	1.481	1.651
23	0.433	0.673	0.628	0.943	**52**	0.217	0.465	0.285	0.335
24	1.589	1.259	0.889	2.245	**53**	0.461	0.743	0.732	0.729
25	1.329	1.450	1.194	1.140	**54**	0.317	0.401	0.425	0.521
26	0.223	0.371	0.210	0.401	**55**	0.338	0.430	0.419	0.524
27	0.870	1.053	1.411	1.926	**56**	0.282	0.377	0.504	0.494
28	0.323	0.335	0.770	0.460	**57**	0.443	0.875	0.765	1.114
29	0.839	0.713	0.798	1.061					

يتضح من نتائج الجدول (8) بأن معظم الفقرات المعايرة من التوزيع الثنائي للقدرة كانت القيم القصوى لدالة المعلومات لها أعلى من القيم القصوى لدالة معلومات الفقرة لبقية توزيعات القدرة. وبينت النتائج بأن الفقرة 19 كانت القيمة القصوى لها هي الأعلى لكل توزيعات القدرة بينما كانت الفقرة 16 هي الأقل لجميع التوزيعات، وبتتبع القيم القصوى لدالة المعلومات للفقرات المعايرة من التوزيع الطبيعي للقدرة يلاحظ تدني القيم لها مقارنة مع بقية التوزيعات حيث تراوحت قيم دالة المعلومات للفقرة ما بين 0.169 ولغاية 2.679، ويعود السبب في ذلك بأنّ التوزيع الطبيعي للقدرة قد أنتج مدى واسعا لمعلمة الصعوبة أكثر من غيره من التوزيعات الأخرى.

ولأن مفهوم الثبات في نظرية الاستجابة للفقرة يرتبط بدالة معلومات الفقرة والاختبار، فقد تم إيجاد معاملات الثبات الامبريقي

المنبثق من نظرية الاستجابة للفقرة ومعاملات ثبات الاتساق الداخلي كرونباخ ألفا حسب النظرية التقليدية في القياس لكل توزيع من توزيعات القدرة المختلفة، ويوضح الجدول (9) هذه القيم.

جدول 9: قيم معاملات الثبات (كرونباخ الفا) والامبريقي لكل توزيع من توزيعات القدرة المختلفة

شكل التوزيع	الثبات (α)	الثبات (TF)
طبيعي	0.93	0.95
ملتو نحو اليمين	0.93	0.96
ملتو نحو اليسار	0.95	0.96
ثنائي المنوال	0.95	0.97

يتضح من النتائج الواردة في الجدول (9) تباين قيم معاملات الثبات المحسوبة وفق النظرية الكلاسيكية، حيث كانت هذه القيم متماثلة بالنسبة للتوزيع الطبيعي والتوزيع الملتوي نحو اليمين، وأقل مما هي علية في التوزيعين الملتوي نحو اليسار وثنائي المنوال، بينما أشارت النتائج بان قيمة معامل الثبات الامبريقي كانت الأقل بالنسبة للتوزيع الطبيعي والأعلى بالنسبة للتوزيع ثنائي المنوال، وهذا يتفق مع ما ورد في الشكل (3)، حيث كان الاختبار يقدم اقصى معلومات عندما تمت معايرته من التوزيع ثنائي المنوال للقدرة، في حين قدم أقل المعلومات عندما تم معايرته من التوزيع الطبيعي للقدرة. وتجدر الإشارة كذلك بان قيم معاملات الثبات الامبريقي للاختبار عندما تم معايرته من التوزيع الملتوي نحو

اليسار واليمين للقدرة قد تماثلت، وبالنظر إلى الشكل (3) نجد أن كمية المعلومات للاختبار المعاير من التوزيعين متماثل.

ولمعرفة دلالة الفروق في معاملات الثبات المحسوبة وفق النظرية التقليدية ونظرية الاستجابة للفقرة، فقد تم استخدام اختبار (M) المقترح من قبل هاكستين وولن (Hakstin & Whalen, 1976) الذي يتبع توزيع مربع كاي بدرجات حرية (عدد المعاملات – 1) وتحسب قيمة الإحصائي M وفق المعادلة:-

$$M = \frac{J-1}{18\ J}\left\{\sum_{k=1}^{k} B_k - \frac{\left[\sum_{k=1}^{k} B_k\left(1-r_k\right)^{\frac{-1}{3}}\right]^2}{\sum_{k=1}^{k} B_k\left(1-r_k\right)^{\frac{-2}{3}}}\right\}$$

حيث تحسب B_k من المعادلة $B_K = \frac{\left(9\,n_k - 11\right)^2}{n_k - 1}$ ، وتشير الرموز الواردة في المعادلة كما يلي:

J: عدد فقرات الاختبار r_k : معامل الثبات الطريقة n_k : عدد أفراد العينة في الطريقة المستخدمة

وقد كشفت نتائج التقديرات الخاصة بقيم الثبات التقليدي عن وجود فروق ذات دلالة إحصائية بين معاملات الثبات التقليدي عند مستوى الدلالة ($\alpha = 0.01$)، حيث كانت قيمة الإحصائي M المحسوبة (55.24) وهي أعلى من القيمة الحرجة لمربع كاي بدرجات حرية(3)، وبالوقت نفسه أشارت النتائج الخاصة بتقديرات قيم معاملات الثبات الامبريقي بوجود فروق ذات دلالة إحصائية بينها، حيث بلغت قيمة الإحصائي M (65.28) وهي أعلى من القيمة الحرجة لمربع كاي بدرجات حرية (3). إن السبب في اختلاف قيم معاملات الثبات ربما يعود لاختلاف تباين قدرة الأفراد على الرغم من أنّ الفقرات نفسها كانت لجميع الأفراد، وهذا يتفق مع أشارت إليه نتائج تحليل التباين الأحادي للقياس المتكرر باختلاف الصعوبة والتمييز لكل توزيع من توزيعات القدرة خصوصا وان تقدير تباين الخطأ في الثبات الامبريقي يعتمد على مقلوب دالة المعلومات.

لقد تبين من النتائج التي توصلت إليها الدراسة الحالية بأنه يوجد أثر لاختلاف شكل توزيع القدرة على معالم الفقرات ودالة المعلومات للاختبار، حيث تباينت معالم الفقرات عبر التوزيعات المختلفة لقدرات الأفراد، مما يشير إلى عدم استقرار معالم الفقرات عبر التوزيعات المختلفة، على الرغم من وجود علاقة خطية قوية بين هذه المعالم والذي يعد احد المؤشرات لفحص اللاتغاير في معالم الفقرات. ولأن النتائج التي تم التوصل إليها تتأثر بمجموعة من العوامل مثل: طول الاختبار، حجم العينة، وطريقة التقدير تصبح الحاجة ملحة لإجراء دراسات مماثلة تتناول أثر اختلاف شكل توزيع القدرة على معالم الفقرات ودالة المعلومات في سياق هذه العوامل لإخراج الافتراض المتعلق بخاصية اللاتغاير للفقرات من الإطار النظري إلى التطبيق العملي من خلال توفير مؤشرات أكثر قوة من معامل الارتباط للكشف عن ذلك لأهميته في التطبيقات العملية عند استخدام نظرية الاستجابة للفقرة.

المراجع

الثوابية، احمد محمود (2010). اثر حجم العينة على تقدير صعوبة الفقرة والخطا المعياري في تقديرها باستخدام نظرية الاستجابة للفقرة. **مجلة جامعة دمشق**، *26*، 225-556.

جمحاوي، ايناس(2000). **مقارنة خصائص الفقرات وفق النظرية التقليدية ونظرية استجابة الفقرة في مقياس للقدرة العقلية**. رسالة ماجستير غير منشورة، جامعة اليرموك، الأردن.

عبابنة، عماد غصاب (2008). استقصاء اللاتغير في تقدير إحصائيات الفقرة باستخدام نظرية الاستجابة للفقرة والنظرية الكلاسيكية في القياس. WWW.ULUM.NL، **مجلة علوم إنسانية**، *39*، 1-26.

علام، صلاح الدين (2005). **نماذج الاستجابة للمفردات الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي**، *(ط1)*. القاهرة: دار الفكر العربي.

Baker, F.B.(2001). *The basics of item response theory(2nd ed)*. College Park, MD: ERIC Clearing House on Assessment and Evaluation.

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

Cypress, k.(1973).*The effect of diverse test score distribution characteristics on the estimation of Rash measurement model*. Paper presented at the Annual Meeting of Americans Educational Research Association, Orleans.

De Gruijter, D. M., & Van der Kamp, L. J. Th.(2005). *Statistical Test Theory for Education and*

Miller, T. R.(1991). *Empirical estimation of standard errors of compensatory MI model parameters obtained from the NOHARM estimation program.*(ACT Research Report No. onr91-2). Iowa City IA: ACT Inc.

Miller, M.D., & Linn, R.L.(1988). Invariance of item characteristic function with variations in instructional coverage. *Journal of Educational Measurement, 25***,** 305-319.

Mislevy, R.J., & Bock, R.D.(1990). *BILOG3: Item analysis and test scoring with binary logistic models(2nd ed)*. Scientific software, Inc.

Reckase, M.D.(1997). The past and future of multidimensional item response theory. *Applied Psychological Measurement,1***,** 25-36.

Reeve, B. (2004). *Applications of item response theory (IRT) modeling for building and evaluating questionnaires measuring patient-reported outcomes*[On-Line]:http://outcomes.caner. gov/conference/irt/reeve.pdf.

Rup, A.A. & Zumbo, B.D. (2006). Understanding parameter invariance in unidimensional IRT models. *Educational and Psychological Measurement, 66*, 63-84.

Sass, D. A., Schmitt, T. A., & Walker, C. M.(2004). *An evaluation of BILOG-MG with skewed theta distribution using various estimation procedures : A simulation study*. Poster presented at the National Council on Measurement in Education. San Diego, California.

Seong, T. J.(1990). Sensitivity of Marginal Maximum Likelihood Estimation of item and ability parameter to the characteristics of the prior ability distribution. *Applied Psychological Measurement ,14*, 299-311.

Stocking, M. L.(1990). Specifying optimum examines for item response theory. *Psychometrika,55,* 461-475.

Stone,C. A.(1992). Recovery of Marginal Maximum Likelihood Estimation in Two Parameter Logistic Response Model: An evaluation of MULTILOG. *Applied Psychological Measurement ,16*, 1-16.

Warm, T.A.(1978). *A Primer of Item Response Theory***.** Technical Report No, 940279, Oklahoma City: U.S, Coast Guard Institute.

Psychology. Retrieved, from: www.leidenuniv,n1/ griuijterdnme.

Drasgow, F.(1982). Biased test items and differential validity**.** *Psychological Bulletin ,92***,**526-531.

Embretson, S.E., & Reise, S. P.(2000). *Item Response Theory for Psychologists*. New Jersey: Lawrence Erlbaum Associates, Publishers.

Fan, X.(1998). Item response theory and classical test theory. An empirical comparison of their item /person statistics. *Educational and Psychological Measurement, 58,* 257-381.

Hambleton, R., K.(1994). Item Response Theory: A broad psychometric frame work for measurement advances. *Psicothema,6,* 535-556.

Hambleton, R.K, and Cook, L.(1980). The Robustness of Latent Trait Models and effect of test length and sample size on the precision of ability estimate. In D. J. Weiss(Ed), *Proceding of the 1979 Computerized Adaptive Testing Conferenc***,** 36-52. Minneapolis, Minnesota: University of Minnesota Department of Psychology. Psychometric Methods Program , Computerized Adaptive Testing Laboratory.

Hambleton , R.K & Jones, R. W.(1994). Item parameter estimation errors and their influence on test information function. *Applied Measurement in Education, 7,* 171-186.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory : Principles and applications.*Boston : Kluwer-Nijhoff.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory.* London : Sage Publications, Inc.

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, Amherst , MA: University of Massachusetts Center for Educational Assessment.

Hatti, J. 1984). An empirical study of various indices for determining unidimensionality. *Multivariate Behavioral Research*, 19,49-78.

Hkstain, A.R & Whalen, T.E.(1976). A K-sample significance test for independent alpha coefficients. *Pyschometrika, 41***,**219-231

Kelkar, V., Wightman, L.P. & Luecht, R.M.(2000). *Evaluation of the IRT parameter invariance property for the MCAT.* Paper presented at the Annual Meeting of National Council on Measurement in Education, New Orleans, LA (ERIC Document Reproduction Service No ED. 443830.

Lord, F. M.(1980**).** *Application of item response theory to practical testing problems.* New Jersey: Lawrence Erlbaum Associates, Publishers.