أثر حجم العينة وشكل توزيع قدرة المفحوصين على معالم الفقرة عند استخدام النموذج اللوغارتيمي الثلاثي

د. طه عقلة الخرشه
قسم علم النفس
كلية التربية والآداب-جامعة تبوك
taha_oglah@yahoo.com

أثر حجم العينة وشكل توزيع قدرة المفحوصين على معالم الفقرة عند استخدام النموذج اللوغارتيمي الثلاثي

د. طه عقلة الخرشه
قسم علم النفس
كلية التربية والآداب-جامعة تبوك

الملخص

هدفت الدراسة الحالية للكشف عن أثر حجم العينة واختلاف شكل توزيع القدرة (طبيعي، ملتوي نحو اليمين، ملتوي نحو اليسار) على معالم الفقرة. ولتحقيق الغرض من الدراسة تم توليد استجابات تسع مجموعات من المفحوصين وفق النموذج اللوجستي ثلاثي المعلمة من خلال استخدام برنامج WINGEN-3 على (٥٠) فقرة ثنائية الاستجابة، وباستخدام برنامج Bilog-mg3 تم تحليل استجابات الأفراد، كشفت نتائج تحليل التباين الثنائي عن وجود فروق ذات دلالة إحصائية عن مستوى دلالة بين متوسطات معالم الفقرة للفقرات تبعا لحجم العينة، وكذلك أظهرت النتائج وجود فروق ذات دلالة إحصائية بين متوسطات معالم الفقرة تبعا لشكل التوزيع.

الكلمات المفتاحية: شكل توزيع القدرة، المحاكاة، نظرية الاستجابة للفقرة.

* تاريخ تسلم البحث: ٢٠١٦/٥/٢٥م
* تاريخ قبوله للنشر: ٢٠١٧/١٢/٢٨م

The Effect of Sample Size and Different Ability Distribution on Item Parameters Utilizing Three Logistic Model

Dr. Taha O. AL-Kursheh
Faculty of Education and Arts
University of Tabuk

Abstract

This study aimed at verifying the effect of Sample Size and the ability distribution (normal, positive and negative skewed) on item parameters, To achieve the aim of the study, generating responses to the nine groups of subjects according to three parameter logistic model using WINGEN program on 50 dichotomous items, The responses of the examinees were analyzed by BILOG– MG 3 programs for each group, The results of the repeated measurement ANOVA revealed that there were statistical significant differences between difficulty means and discrimination means attributed to Sample Size, The results also revealed that there were statistical significant differences between difficulty means and discrimination means attributed to ability distribution.

Keywords: ability distribution, simulation, item response theory.

أثر حجم العينة وشكل توزيع قدرة المفحوصين على معالم الفقرة عند استخدام النموذج اللوغارتيمي الثلاثي

د. طه عقلة الخرشه
قسم علم النفس
كلية التربية والآداب-جامعة تبوك

المقدمة

لقد مر علم القياس والتقويم بمراحل عديدة خلال العقود القليلة الماضية شملت تغيرات في مفاهيمه ومبادئه الأساسية وأساليبه وأدواته ومرجعياته، وقد أسهمت هذه التطورات في إحداث تغييرات تربوية شاملة في مختلف مكونات المنظومة التعليمية. وكل هذه التغيرات والتطورات تسعى لجعل القياس النفسي والتربوي أكثر موضوعيه. أي تحرر أدوات القياس من خصائص الأفراد الذين تطبق عليهم هذه الأدوات وتحرر قياس قدرات وسمات الأفراد من خصائص عينة الأسئلة التي يشتمل عليها الاختبار الذي طبق عليهم.

وموضوعية القياس هذه ارتبطت بمدخل جديد يطلق عليه نظرية استجابة الفقرة التي تفترض أن هناك علاقة بين مقدار ما يمتلكه المفحوص من السمة ويعبر عنها بالرمز θ_j ومعامل صعوبة فقرة ما، ويعبر عنها بالرمز b_i واحتمال حصول المفحوص على الإجابة الصحيحة عند مستوى قدرة معين ويعبر عنها بالرمز p_i (θ_j) وأن هذه العلاقة تتخذ شكل المنحنى الرياضي الذي يفترض أن يكون شكله كحرف S حيث يتم تدريج قدرات المفحوصين وصعوبة الفقرات على متصل واحد. والذي يحدد مقدار احتمال الاستجابة الصحيحة أو الخطأ على الفقرة هو الفرق بين قدرة الفرد θ_j وصعوبة الفقرة b_i . وتتمتع هذه النظرية بمجموعة من الميزات التي أشار إليها هامبلتون وسوامينثان (Hambleton & Swaminathan,1985) مثل:

١- بافتراض وجود مجموعة كبيرة من الفقرات التي تقيس نفس السمة فإن تقدير قدرات المفحوصين مستقلة عن عينة المفحوصين.(Item-Free).

٢- بافتراض وجود مجتمع كبير من المفحوصين، فإن تقدير إحصائيات الفقرة لفقرات الاختبار كالصعوبة والتمييز مستقلة عن عينة المفحوصين المختارة بهدف معايرة الفقرات (Person-Free)

يوجد إحصائي يشير إلى الدقة في تقدير قدرة كل فرد مثل الخطأ المعياري للتقدير وهذا الإحصائي مختلف من فرد لآخر.

وتقوم هذه النظرية كغيرها من النظريات على عدد من الافتراضات التي تميزها عن غيرها.

١- افتراض أُحادية البعد Assumption of unidirictionality

ان التعريف العملي لمفهوم أُحادية البعد في الاختبارات العقلية يتمثل في امكانية ترتيب فقرات اختبار عقلي تبعاً لصعوبتها، فإذا كانت نماذج استجابات الأفراد متدرجة فإن الاختبار سيكون أُحادي البعد، وهو ما يسمى بأنموذج الاستجابة لجتمان Guttman، إذ إن الذي يفشل في الفقرة الأولى حتماً سيفشل في بقية الفقرات، ومن ينجح في الفقرة الاخيرة لا بد أن ينجح في كل الفقرات الأقل صعوبة (Lumsden, 1961: 65).

وإن فرض أحادية البعد يصعب تحقيقه بشكل تام، إذ إن هناك العديد من العوامل التي تؤثر في استجابة الافراد كالدافعية، والقلق، والناحية المعرفية وغيرها. (Anastasi & Urbina, 1997: 192)

ويتطلب تحقيق هذا الفرض وجود عامل مسيطر واحد على الاختبار وهو القدرة المقيسة (Lord, 1980: 19). لذا فلا بد أن نتحقق من هذا الفرض عند بناء الاختبار من خلال أربعة مؤشرات مهمة وهي:

- **نمط الاستجابة:** وذلك باستعمال أنموذج جتمان في الاستجابة عن الفقرات بترتيب الفقرات بحسب صعوبتها، فالمفحوص الذي يفشل على فقرة بمستوى صعوبة معين لا بد أن يفشل في اجابته عن فقرة بمستوى صعوبة اعلى.
- **الاتساق الداخلي:** ويعد معامل ارتباط الفقرة بالدرجة الكلية مؤشراً على اتساق ما تقيسه هذه الفقرات ومن ثَم فهو مؤشر على أن الفقرات تقيس سمة واحدة. فالفقرات ذات أعلى معامل ارتباط بوينت بايسريال (r_{pa}) معامل الارتباط ثنائي التسلسل) تستبقى في الاختبار لأن مثل هذه الفقرات عند تجميعها تؤلف تدريجاً بأعلى اتساق داخلي، وهذا يعطي مؤشراً على ان مجموعة الفقرات تقيس سمة واحدة أو قدرة واحدة (Nunnally, 1978: 77).
- **المكونات الرئيسة:** يقدم لومسدن بناء اختبار بافتراض أحادية البعد اولاً ثم إجراء التحليل العاملي بطريقة المكونات الرئيسة Principal componerts وحذف الفقرات التي لا تقيس العامل العام ومن ثم إعادة التحليل وحذف الفقرات غير المنسجمة مع العامل العام .وتكرار الفقرات التي تقيس العامل العام فقط. وأما ريكس Reckase فقد اقترح أن تكون نسبة ٢٠٪ من التباين المفسر بالعامل العام مؤشراً كافياً لأحادية البعد (Reckase 1979: 209)،

وقد اقترح لورد Lord دلالة احصائية تقريبية في حساب الجذور الكامنة Latent roots في التحليل العاملي لتحقيق أُحادية البعد لمجموعة فقرات الاختبار وهي كون قيمة الجذر الأول كبيرة مقارنة بالجذر الثاني والجذر الثاني قريباً في قيمته من الجذور الاخرى الذي يدل على أن الفقرات أحادية البعد (Lord, 1980: 19).

- مطابقة البيانات لأحد نماذج السمات الكامنة: يمكن اعتماد الاستقلال المحلي Local Independence لكل زوج من الفقرات ولأي مجتمع متجانس من المفحوصين كمؤشر لأحادية البعد، إذ يترتب على الاستقلال المحلي وجود سمة واحدة يقيسها الاختبار (Crocker& Algina, 1986: 345).

٢. افتراض الاستقلال المركزي Assumption of local Independence

ويقصد بهذا الافتراض أن تكون استجابات الفرد للبنود المختلفة في الاختبار مستقلة استقلالاً احصائياً. وهذا يعني ألا تؤثر استجابة الفرد لإحدى المفردات على استجاباته للمفردات الاخرى، ويتضح هذا في:

- تحرر القياس من توزيع العينة المستعملة sample free ويعني ثبات تقدير كل من قدرة الفرد وصعوبة البند واستقرارها بالرغم من اختلاف عينة الافراد المستعملة في تدريج المقياس طالما أنها عينة ملائمة.
- تحرر القياس من مجموعة المفردات المستعملة Item free ويعني ثبات تقدير كل من قدرة الفرد وصعوبة المفردة واستقرارها بالرغم من اختلاف مجموعة المفردات المستعملة في القياس، طالما أنها مفردات ملائمة، وطالما أن هذا المجموعات المختلفة من المفردات تقع على ميزان تدرج واحد أي أنها تعرف متغيراً واحداً (Ranadall 51 ,1998:).

٣.افتراض المنحني المميز للفقرة Assumption of Item characteristic

ان المنحني المميز للفقرة (I.C.C Item Characteristic Curve) عبارة عن انحدار غير خطي لدرجة الفقرة على القدرة أو السمة المقيسة من خلال المقياس او الاختبار، ويرجع الاختلاف بين نماذج السمات الكامنة إلى اختلاف صور دوالها الرياضية ومن ثم اختلاف واشكال منحنياتها. ويعد هذا الفرض دالة رياضية تربط بين احتمال نجاح الفرد في الإجابة عن الفقرة والقدرة التي يقيسها الاختبار أو مجموعة الفقرات التي يشتمل عليها المقياس. (Allen & Yen, 1979: 128).

وبشكل عام فإن شكل المنحني المميز للفقرة يأخذ الشكل نفسه– شكل (S S- shaped)، ويعد هذا مؤشراً على أن الافتراضات المتنوعة تطابق المواقف الاختبارية (Anastasi & Urbina, 1997: 190). وتعتمد احتمالية الإجابة الصحيحة لمفحوص معين على فقرة ما

على شكل المنحني فقط، أي تعتمد على الصعوبة فقط، وهذا يدلل على استقلاليتها عن توزيع القدرة لمجتمع المفحوصين المطبق عليهم الاختبار، أي أن احتمالية الاجابة الصحيحة لا تعتمد على عدد المفحوصين ذوي مستوى القدرة نفسها. ولا على الأفراد عند مستويات القدرة الاخرى، ولان منحني الانحدار لا يتباين فإن الخط التقاربي الادنى ونقطة الانعطاف والميل عند هذه النقطة تبقى نفسها بغض النظر عن توزيع القدرة في مجتمع المفحوصين. أي ان شكل المنحني المميز للفقرة لا يعتمد على توزيع القدرة في مجتمع المفحوصين، ويطلق على هذه الخاصية عدم تباين المنحني المميز للفقرة (Hambelton & Eignon& Gifford, 1978: 47) (Lord, 1980: 35).

وقد انبثق عنها مجموعة من النماذج منها:

النموذج اللوجستي أحادي المعلمة (نموذج راش) Model Rasch الذي استخدم في تحليل استجابات أفراد عينة الدراسة. يعد هذا النموذج ابسط نماذج استجابة الفقرة ثنائية التدريج، ويفترض هذا النموذج أن جميع الفقرات لها القدرة التمييزية نفسها للتمييز بين المفحوصين، أي تكون جميع المنحنيات المميزة لفقرات الاختبار متوازية (متساوية في الميل)، وتختلف عن بعضها البعض في نقاط التقائها بالمحور الأفقي، الذي يمثل متصل القدرة أو السمة الكامنة، كما يفترض عدم لجوء المفحوصين للتخمين العشوائي عند إجابتهم عن فقرات الاختبار، وهذا يعني أن الفقرات تختلف فقط في صعوبتها، وتكون معادلة هذا النموذج كما يأتي:

$$P_{i\,(\theta)} = \frac{e^{D(\theta - b_i)}}{1 + e^{D(\theta - b_i)}} \quad , (i = 1 \dots \dots \dots \dots \dots \dots \dots \dots n)$$

حيث $P_{i\,(\theta)}$ احتمال إجابة الفقرة (i) إجابة صحيحة من قبل مفحوص له القدرة θ تم اختياره عشوائياً من بين المفحوصين الذين لهم القدرة نفسها.

D: ثابت مقداره ١,٧، (Sampling Factor) وهو عامل تدريج يجعل القدرة المستنتجة من استعمال هذا النموذج مساوية للقدرة المستنتجة من استعمال النموذج، الذي يعتمد على المنحنى الطبيعي التراكمي.

b_i : معلمة الصعوبة للفقرة i

e: الأساس اللوغاريتمي الطبيعي (٢,٧١٨)

النموذج اللوغارتيمي ثنائي البارامتر .(Lord Model) كان لورد ١٩٥٢ أول من طور النموذج ثنائي البارامتر معتمدا على التوزيع الاعتدالي التراكمي. ثم قام بيرنبوم ١٩٦٨ باستبدال دالة اللوغاريتم التراجيحي لمعلمين بدلا من دالة التوزيع الاعتدالي كصيغة للدالة

المميزة للمفردة. وتتميز دالة اللوغاريتم الترجيحي بأنها أكثر إقناعا للملاءمة مع الدالة الاعتدالية. ويعتبر النموذج اللوغاريتمي سهل التطويع رياضيا أكثر من النموذج الاعتدالي لأنه يحتوي بداخلة بعض صيغ التكامل، بينما يعتبر النموذج الاعتدالي دالة صريحة لمعلمي المفردة والقدرة ولديه خصائص إحصائية هامة.

الصورة الرياضية للنموذج اللوغارتيمي ثنائي البارامتر

تمثل المعادلة التالية المنحنى المميز للمفردة للنموذج اللوغاريتمي ثنائي المعلم والذي طوره بيرنبوم (Birnbanm)

$$p_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}}$$

حيث إن:

$p_i(\theta)$:احتمال إجابة المفحوص الذي اختير عشوائيا من مستوى القدرة (q) على المفردة (i) إجابة صحيحة (b_i)، معلم الصعوبة، (a_i) : معلم التمييز، (q) (θ) معلمة القدرة، (D) عامل التدريج ويساوي (١,٧).

من الواضح أن النموذج اللوغاريتمي ثنائي البارامتر يشبه النموذج اللوغاريتمي أحادي البارامتر فيما عدا وجود عنصرين جديدين هما:

العنصر الأول: عامل D وهو عامل القياس أو التدريج وتم إدخاله للمعادلة لكي يجعل الدالة اللوغاريتمية أقرب ما يكون للدالة الاعتدالية. وقد وجد أنه حينما تكون (D =١,٧) فإن قيم $p_i(\theta)$ للنموذج اللوغاريتمي ثنائي البارامتر والنموذج الاعتدالي ثنائي البارامتر تختلف بقيمة مطلقة أقل من (٠,٠١) لجميع قيم (θ).

العنصر الثاني الإضافي للنموذج اللوغاريتمي ثنائي البارامتر هو البارامتر هو a_i والذي يسمى معلم تمييز المفردة. ويتناسب معلم التمييز a_i مع ميل منحنى خاصية المفردة عند النقطة على مقياس القدرة. فالمفردات التي تتميز منحنياتها بميل شديد الانحدار تكون أكثر فائدة في تصنيف المفحوصين إلى مستويات مختلفة القدرة مقارنة بالمفردات الأقل انحداراً في ميل منحنياتها، وفي الحقيقة فإن فائدة المفردة تكمن في قدرتها على التمييز بين المفحوصين القريبين من القدرة (θ) وهذا يتناسب مع ميل منحنى خاصية المفردة عند نقطةq.

النموذج اللوغارتيمي ثلاثي البارامتر .(Birnbaum Model) في هذا النموذج أضاف بيرنيوم معلم ثالث هو التخمين، لأن بعض مفردات الاختبار تسمح أحيانا لبعض المفحوصين من ذوي القدرة المنخفضة جداً، بالتوصل إلى الإجابة الصحيحة عن طريق التخمين، والقيمة

التقديرية لهذا المعلم، تمثل احتمال توصل هؤلاء المفحوصين إلى الإجابة الصحيحة عن المفردة، مع وجوب أخذ ذلك بعين الاعتبار، عند مطابقة البيانات المستمدة من الاختبار للنموذج. وهنا تكون المنحنيات المميزة للمفردات متقاطعة؛ أي تختلف عن بعضها البعض في الجزء الذي تلتقي فيه بالمحور الأفقي الممثل لمتصل السمة الكامنة.

الصورة الرياضية للنموذج اللوغارتيمي ثلاثي البارامتر

التعبير الرياضي للنموذج اللوغارتيمي ثلاثي البارامتر كمايلي:

$$P\left(X=\frac{1}{\theta_s},\beta_i,\alpha_i,\gamma_i\right)=\gamma_i+(1-\gamma_i)\frac{e^{[\alpha_i(\theta_s-\beta_i)]}}{1+e^{[\alpha_i(\theta_s-\beta_i)]}},X=0,1$$

وقد تكتب بصيغة أخرى هي:

$$P_i(\theta)=C_i+(-C_i)\frac{e^{Da_i(\theta-b_i)}}{1+e^{Da_i(\theta-b_i)}},i=1,2,\dots\dots,n$$

حيث:

احتمال اجابة المفحوص ذي القدرة (θ) إجابة صحيحة على المفردة i.

$p_i(\theta)$: معلم صعوبة المفردة.

a_i: معلم تمييز المفردة.

C_i: معلم تخمين المفردة الذي يمثل المقطع الرأسي (الصادي) لخط التقارب الأدنى لمنحنى خاصية المفردة، كما يعبر عن احتمالية أن يجيب المفحوص ذو القدرة المنخفضة إجابة صحيحة عن المفردة بالتخمين، وان هناك علاقة طردية بين احتمالية أن يجيب المفحوص إجابة صحيحة عن المفردة، وبين قيمة مؤشر التخمين، حيث تزداد هذه الاحتمالية بزيادة معامل التخمين.

وأهم ما يميز هذه النظرية، توفر الموضوعية في القياس النفسي التربوي الذي يمكننا من الحصول على إحصائيات الفقرة التي لا تعتمد على خصائص المفحوصين، ودرجات تعبر عن قدرة المفحوصين لا تعتمد على خصائص الفقرات.

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). 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 (الشرفين وزياد، ٢٠١٢)

وقد أجريت العديد من الدراسات لفحص افتراضات نظرية الاستجابة للفقرة، ولتحديد العوامل المؤثرة في تقدير معالم الفقرات وقدرات الأفراد ومن هذه الدراسات الدراسة التي أجرتها بارنس ووايز (Barnes and Wise, 1991) بعنوان استخدام النموذج أحادي المعلمة من نظرية الاستجابة للفقرة مع العينات الصغيرة. فقد استخدمت عينات صغيرة الحجم (٥٠، ١٠٠، ٢٠٠) وقد جرى تطبيق اختبارين: الأول مكون من (٢٥) فقرة والثاني مكون من (٥٠) فقرة، واستخدم نموذج راش في تقدير معلمة صعوبة الفقرة وقدرات المفحوصين، وتوصلت الدراسة إلى أن تقدير قدرات المفحوصين يتأثر بعدد الفقرات ولا يتأثر بحجم العينة، وأن تغير معلمة الصعوبة كان تغيراً بسيطاً عند تغير حجم العينة، وقد يعود ذلك إلى أن الفرق بين حجم العينات كان صغيراً.

كما توصلت دراسة ستون ويوموتو (Stone & Yamoto, 2004) والتي كانت بعنوان أثر حجم العينة على تقديرات معالم الفقرات ثنائية التدريج باستخدام نموذج راش ونماذج نظرية استجابة الفقرة، حيث استخدم نتائج اختبار Knox's Cube Test Revised في سحب (٣٠) عينة عشوائية لتقدير معالم الفقرات، وقد وجدا بأن نموذج راش يعطي أقل تقدير لمعلمة الصعوبة، وأن العينات الصغيرة تكون أقل في مطابقة نموذج حسن المطابقة.

وفي دراسة عبابنة (٢٠٠٤) بعنوان أثر حجم العينة وطريقة انتقائها وعدد الفقرات وطريقة انتقائها على دقة تقدير معالم الفقرة والقدرة لاختبار قدرة عقلية باستخدام نظرية الاستجابة للفقرة، ولتحقيق أهداف هذه الدراسة تمّ إعداد اختبار قدرة عقلية مؤلف من أربعة اختبارات فرعية هي اختبار المفردات، اختبار المتشابهات، اختبار المتضادات، واختبار الحساب، حيث بلغ عدد فقرات الاختبار النهائية (٧١) فقرة، طبقت فقرات الاختبار على عينة مكونة من طلاب وطالبات الصف السابع من مدارس التربية والتعليم، وقد بلغ عدد أفراد عينة الدراسة (١٠٠٠) طالب وطالبة، وقد توصلت الدراسة إلى أنه تزداد الدقة في تقدير معلمة الصعوبة والقدرة عندما يكون مدى القدرة للمفحوص متوافقاً مع مدى صعوبة الفقرات، وكذلك تزداد الدقة في تقدير معلمة التخمين عند استخدام عينة من ذوي القدرة

المتدنية في معايرة الفقرات، كما أظهرت النتائج بأن تقديرات معالم الفقرة لعينة مختارة من الفقرات تتسم بالاستقرار النسبي عند معايرتها ضمن المقياس ككل أو بصورة مستقلة باستخدام العينة نفسها من الأفراد، وأن الدقة في تقدير معالم الفقرة تزداد بزيادة حجم عينة المفحوصين.

وقام 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 ملتويًا (موجبًا، سالبًا)، وطبيعي وحجم عينة (٥٠٠، ٦٠٠) وست طرق تقدير. بينت النتائج بشكل عام بأنَّ التقديرات في معالم الفقرات كانت أقل دقة عندما كان توزيع القدرة ملتويًا وأكثر دقة عندما كان توزيع القدرة طبيعيًا، وعند مقارنة الأخطاء المعيارية بين معلمة الصعوبة والتمييز لوحظ بأنها الأعلى عند معلمة التمييز.

كما أجرى دي لا توري ويوان (De La Torre & Yuan, 2010) دراسة هدفت إلى التعرف على أثر حجم العينة على دقة تقدير معلمة القدرة ومعالم الفقرة في اختبارات مطورة حسب نماذج نظرية الاستجابة للفقرة (IRT). واستخدمت الدراسة نموذج (HO-IRT) في توليد مجموعة من البيانات ضمن ظروف اختبار مختلفة. حيث جرى توليد البيانات باستخدام طريقة مونتي كارلو من أجل التعرف إلى العلاقات بين معلمة القدرة ومعالم الفقرة في الاختبار، وأثر حجم العينة في دقة تقدير معلمة القدرة ومعالم الفقرة. وأشارت النتائج إلى أن حجم العينة يؤثر على قدرة الاختبار في تقدير معلمة القدرة ومعالم الفقرة، والى وجود علاقة ارتباطية بين تقدير معلمة القدرة ومعالم الفقرة.

وأجرى الشرفين وبني عطا (٢٠١١) دراسة هدفت إلى التحقق من أثر اختلاف شكل توزيع القدرة على معالم الفقرة ودالة المعلومات للاختبار، وجود فروق ذات دلالة إحصائية عند مستوى دلالة (٠,٠٥) 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 أنَّ قيم معاملات الارتباط بين تقديرات معالم الفقرات المعايرة تبعا لشكل توزيع القدرة كانت دالة إحصائيا كمؤشر لاستقرار معالم الفقرات عبر التوزيعات المختلفة للقدرة.

في ضوء ما تَقدَّم من الدراسات السابقة، يُمكن ملاحَظَة أن بعض تلك الدراسات قد تم إجراؤها باستخدام البيانات المُولَّدة، والتي يجري توليدها بوساطة برامج التوليد المختلفة، وبعضها الآخر قد تم أجري باستخدام البيانات الفعلية. وكذلك نلاحظ أن غالبيتها اهتم

بدراسة أثر حجم العينة وطول الاختبار على دقة التقديرات، وقليل منها اهتم بدراسة أثر شكل توزيع القدرة على معالم الفقرات بمعزل عن العوامل الاخرى المؤثرة في دقة تقدير معالم الفقرات مثل حجم العينة. لذا تأتي هذه الدراسة لتدرس أثر تغير حجم العينة واختلاف شكل توزيع القدرة على معالم الفقرات.

مشكلة الدراسة وفرضياتها

تُعد عملية تقدير معالم الفقرات وقدرات الأفراد وفق نظرية الاستجابة للفقرة من الخطوات الأساسية في تطبيق هذه النظرية، وقد تباينت وجهات النظر حول الطريقة المُثلى في تقدير معالم الفقرات، فبعضها يرى أن التقديرات التي يتم الحصول عليها تكون متشابهة على الرغم من اختلاف المتغيرات، وبعضها الآخر يرى أن تلك التقديرات تكون مختلفة تبعا لاختلاف عدد فقرات الاختبار، أو لاختلاف توزيع قدرات الأفراد، أو لاختلاف حجم العينة حيث يتضح بأنَّ العينات المختلفة قد تولد تقديرات مختلفة، 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، ومن هنا نشأت الحاجة إلى مثل هذه الدراسة والتي اعتمدت على بيانات مُولَّدة، وهَدَفَت لبيان أثر ثلاث أحجام للعينة هي (١٠٠٠، ٤٠٠٠، ٨٠٠٠) على دقة تقدير معالم الفقرات وذلك باستخدام ثلاث أشكال لتوزيع القدرة لمعالم الأفراد هي (التوزيع الطبيعي، التوزيع الملتوي نحو اليمين، والتوزيع الملتوي نحو اليسار)، وباستخدام النموذج ثلاثي المعلمة (PL3) لمناسبته للاختبارات الموضوعية. حيث وجد الباحث نُدرة في الدراسات التي تناولت هذين الجانبين معاً.

وسوف تقوم الدراسة الحالية بفحص الفرضيات الإحصائية الآتية:

- لا توجد فروق ذات دلالة إحصائية عند مستوى الدلالة بين المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم الصعوبة للفقرات ، تعزى لمتغيري حجم العينة وشكل توزيع القدرة للأفراد، والتفاعل بينهما.

- لا توجد فروق ذات دلالة إحصائية عند مستوى الدلالة بين المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم التمييز للفقرات، تعزى لمتغيري حجم العينة وشكل توزيع القدرة للأفراد، والتفاعل بينهما.
- لا توجد فروق ذات دلالة إحصائية عند مستوى الدلالة بين المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم التخمين للفقرات ، تعزى لمتغيري حجم العينة وشكل توزيع القدرة للأفراد، والتفاعل بينهما.

أهمية الدراسة

لقد قدَّمت نظرية الاستجابة للفقرة طريقة فعالة لاختيار فقرات الاختبار، حيث تكون من خلالها معالم كل فقرة من الفقرات ثابتة، إضافة إلى كون صعوبة الفقرة وقدرة المفحوص يتم قياسهما على نفس التدريج، الأمر الذي يوفر إمكانية اختيار الفقرات التي تكون أكثر فائدة ضمن مدى معين على متصل القدرة، ومن مميزات نظرية الاستجابة للفقرة أيضا ، إمكانية اختيار الفقرة وفقا لما تُسهم به من معلومات ضمن المعلومات الكلية التي نحتاجها للوصول إلى خصائص الاختبار المرغوب، وبما أن كمية المعلومات ترتبط بدقة القياس فإن من الممكن اختيار الفقرات للوصول إلى اختبار يتمتع بدرجة جيدة من الدقة عند أيّ مستوى من مستويات القدرة .لذا تكمن أهمية هذه الدراسة من الناحية النظرية في دراسة استقلالية القياس من خلال معرفة الفروق بين معالم الفقرة المقدرة الناتجة من استخدام أحجام مختلفة للعينة وأشكال مختلفة من توزيع القدرة، 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 احجام مختلفة للعينة وتوزيعات مختلفة لأشكال القدرة.

حدود الدراسة

تتحدد نتائج هذه الدراسة بالأسلوب الذي استخدمه الباحث في توليد البيانات في الدراسة الحالية وهو استخدام بيانات المونت كارلو للمحاكاة الحاسوبية Monte Carlo (MC) باستخدام برنامج WinGen-3.

اقتصرت هذه الدراسة على استخدام النموذج ثلاثي المَعلمة.

مصطلحات الدراسة

النموذج اللوغاريتمي ثلاثي المعلم (three-Parameter Logistic Model): أحد نماذج نظرية استجابة المفردة، ويعتمد على ثلاثة معالم للمفردة وهي معلم الصعوبة، والتمييز، والتخمين، إضافة إلى تقديرات قدرة المفحوصين.

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

معلمة الصعوبة : Item Difficulty: نقطة تُمثل موقع الفقرة على متصل القدرة تقابل احتمال ($\frac{c_i}{2}+1$) للإجابة عن الفقرة إجابة صحيحة، حيث (c_i) تمثل معلمة التخمين، ويُرمز لمَعلمة صعوبة الفقرة بالرمز p_i .

معلمة التمييز Item Discrimination: نسبة ميل (Slope) منحنى خصائص الفقرة (ICC) والذي يقابل النقطة التي تكون فيها معلمة القدرة على متصل السمة مساوية لصعوبة الفقرة، وتُعرَّف أيضا بأنها قدرة الفقرة على التمييز بين مستويات المفحوصين المختلفة على متصل القدرة، ويُرمز لها بالرمز a_i.

معلمة التخمين: هي عبارة عن خط المقاربة الأدنى (Lower Asymptote) من منحنى خصائص الفقرة، ويُمَثِّل احتمال إجابة المفحوصين ذوي القدرة المتدنية، إجابة صحيحة عن الفقرة، عن طريق التخمين ويُدعى (Pseudo Chance Level) ويُرمز لمَعلَمة التخمين بالرمز c_i.

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

منهجية الدراسة وإجراءاتها

اعتمدت هذه الدراسة على توليد بيانات باستخدام المحاكاة (هذا الاسلوب يمكننا من ضبط المتغيرات الخارجية) من خلال استخدام الحاسوب Computer Simulation Methods لتوليد البيانات اللازمة للدراسة الحالية، والمسماة طرق مونت كارلو (Monte Carlo Methods (MCM)، بحجوم عينات وأشكال توزيع مختلفة وذلك باستخدام برنامج (WINGEN-3) بحيث تتراوح أحجام العينات من (١٠٠٠ إلى ٨٠٠٠) مفحوص، واشكال توزيع (طبيعي، ملتوٍ نحو اليمين، ملتوق نحو اليسار)، وقد جرت عملية توليد البيانات وفقا للخطوات الآتية:

الخطوة الأولى: توليد الفقرات Items Generation

ولتحقيق هدف الدراسة، فقد تمَّ توليد اختبار مكون من (٥٠) فقرة ثنائية الاستجابة (Dichotomous) بحيث تكون هذه الفقرات مناسِبة للنموذج اللوغاريتمي ثلاثي المعلمة (PL3)) حيث تم توليد معلمة الصعوبة للفقرات وفق التوزيع الطبيعي (Normal~ 0,1) ومعلمة التمييز للفقرة وفق توزيع (Log normal~ 0,0.25) وتوليد معلمة التخمين للفقرات وفق توزيع بيتا(8,32) ~ Beta) وهذا التوزيع ينتج قيم لمعلمة التخمين تماثل قيم التخمين للاختبار الموضوعي (ثنائي الاستجابة) المؤلف من خمسة بدائل. والجدول (١) يُمثل ملخصا للإحصاءات الوصفية لمعالم الفقرات المُكوِّنة للاختبار.

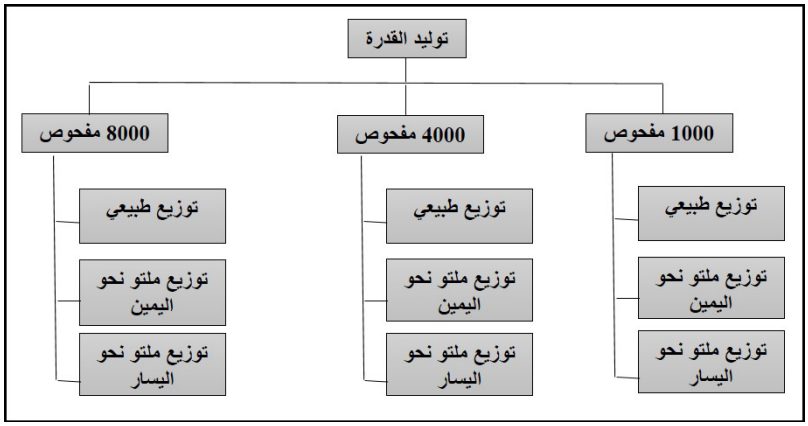
جدول (١)

الإحصاءات الوصفية لمعالم فقرات الاختبار

الإحصائي	الصعوبة (b)	التمييز (a)	التخمين (c)
الوسط الحسابي	٠,٠٢٧	١,٠٦١	٠,١٩٢
الانحراف المعياري	١,٠٤٤	٠,٢٧٦	٠,٠٧٤
القيمة الصغرى	-٢,٢٤٣	٠,٥٨٧	٠,٠٦٥
القيمة العظمى	٢,٤٥٧	١,٩٥٨	٠,٤١٩

الخطوة الثانية: توليد القدرات Abilities Generation

اعتمد الباحث بتوليد قدرات المفحوصين على ثلاثة أحجام للعينة (١٠٠٠، ٤٠٠٠، ٨٠٠٠) وثلاثة أشكال لتوزيع القدرات وهي: (توزيع طبيعي (Normal~ 0,1)، توزيع ملتوٍ نحو اليمين (Beta ~ 2,3)، توزيع ملتوٍ نحو اليسار (Beta ~ (3,2) والشكل رقم (١) يوضح توليد قدرات المفحوصين وفقا لمتغيري (حجم العينة، وشكل التوزيع للقدرة).



شكل (١)

المخطط البياني لتوليد قدرات المفحوصين حسب حجم العينة وشكل التوزيع

وللتحقق من اعتدالية شكل توزيع الاستجابات التي تم توليدها في المرحلة الثانية. قام الباحث باختبار الأشكال السابقة لتوزيع القدرة من خلال استخدام اختبار كولموغروف-سمرنوف، (Kolmogorove-Smirnov) وبينت نتائج اختبار كولموغروف-سمرنوف بأن قيمته كانت دالة إحصائياً عند مستوى دلالة لكل من التوزيع الملتوي نحو اليمين واليسار، مما يعني ابتعادها عن التوزيع الطبيعي في حين لم تكن قيمته دالة إحصائيا للتوزيع الطبيعي.

الخطوة الثالثة: توليد الاستجابات: Responses Generation

بعد أن جرى توليد فقرات الاختبار والمعالم لهذه الفقرات، وكذلك توليد قدرات الافراد المفحوصين، تم توليد استجابات هؤلاء الافراد على الفقرات، بناء على معلمة القدرة، ومعالم الفقرات، وذلك باستخدام برنامج التوليد (WINGEN). حيث جرى إعداد تسع ملفات تحتوي على استجابات المفحوصين وحسب متغيرات الدراسة وكما في الشكل رقم (١) السابق.

المعالجات الإحصائية

قبل معالجة البيانات إحصائياً كان لا بد من فحص الافتراضات الأساسية التي يجب توفرها في البيانات عند تطبيق النموذج اللوغاريتمي ثلاثي المعلم، وأهم هذه الافتراضات أحادية البعد للسمة المقيسة. وقد جرى ذلك من خلال التحليل العاملي للبيانات وقد أُجرى تحليل عاملي للبيانات باستخدام طريقة المكونات الأساسية .(Principal Components Analysis) وفي التحليل العاملي ، تعتمد أحادية البعد على أن نسبة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني تكون نسبة كبيرة لا تقل عن (٢)، وأن ما يفسره العامل الأول من التباين أكثر من ٢٠٪ تقريبا. (Hambleton and Swaminathan, 1985) ويبين الجدول (٢) قيم الجذر الكامن، ونسب التباين المفسر للعامل الأول والعامل الثاني، ونواتج قسمة قيمة الجذر الكامن للعامل الأول على العامل الثاني للبيانات وبحسب متغيرات الدراسة.

جدول (٢)
قيم الجذر الكامن للعامل الأول والثاني وناتج قسمة قيمة الجذر الأول على العامل الثاني في التوزيعات المختلفة

شكل التوزيع	حجم العينة		العامل الأول	العامل الثاني	نسبة الجذر الكامن للعامل الأول إلى الثاني
التوزيع الطبيعي	١٠٠٠	الجذر الكامن	١٠,٦١٢	٢,٤٤٢	٤,٣٤٥
	٤٠٠٠	الجذر الكامن	١٠,٦٣	٢,٢٥	٤,٧٢٤
	٨٠٠٠	الجذر الكامن	١٠,٥٨٨	٢,٣٢	٤,٥٦٣

تابع جدول (٢)

شكل التوزيع	حجم العينة		العامل الأول	العامل الثاني	نسبة الجذر الكامن للعامل الأول إلى الثاني
توزيع ملتوي نحو اليمين	١٠٠٠	الجذر الكامن	١٢,٥٢٣	١,٣٢٨	٩,٤٣
	٤٠٠٠	الجذر الكامن	٩,٢٣٥	٣,٤٢٥	٢,٦٩٦
	٨٠٠٠	الجذر الكامن	١٠,٨٥٦	٢,٤٥٦	٤,٤٢
توزيع ملتوي نحو اليسار	١٠٠٠	الجذر الكامن	١٨,٩٧٥	٣,١٢٣	٤,٣٤٥
	٤٠٠٠	الجذر الكامن	١٢,٤٢٣	٢,٣٤٦	٥,٢٩٥
	٨٠٠٠	الجذر الكامن	١٢,٤٢٣	٣,٤٢٣	٣,٦٢٩

ويتبين من جدول (٢) تحقق افتراض أحادية البعد للاختبار، حيث أن نسبة الجذر الكامن للعامل الأول إلى العامل الثاني كانت كبيرة وتزيد على (٢)، وأن نسبة التباين المفسر من العامل الأول في جميع التوزيعات هي نسبة عالية (أكبر أو تساوي ٢٠ ٪ تقريباً). مما يشير إلى وجود سمة سائدة للاختبار المولدة فقراته. وفيما يتعلق بفحص افتراض الاستقلال المحلّي وافتراض التحرر من السرعة، فإن الاستجابات المولّدة لأفراد افتراضيين، وكذلك من الصعب اختبار مدى فحص ارتباط محتوى كل فقرة بمحتوى الفقرة الأخرى لأن ليس لها محتوى محدد، وعليه فإن هذا الإجراء يكون إجراءً غير ذي معنى. واكتفى الباحث بالتحقق فقط من افتراض أحادية البعد.

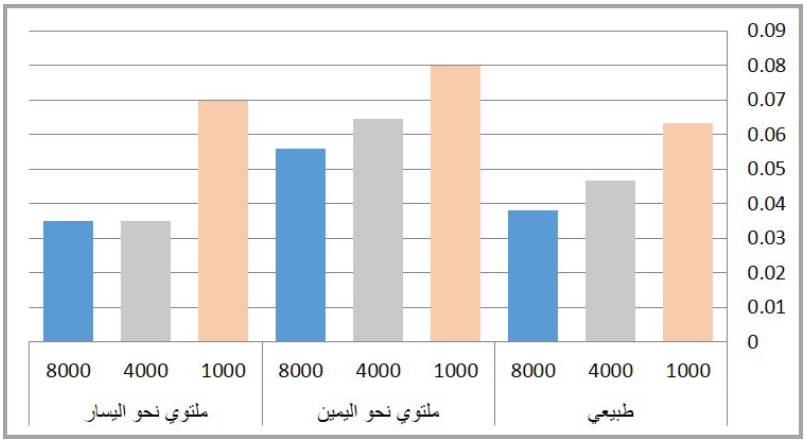
عرض النتائج ومناقشتها

وللإجابة عن الفرضية الأولى حول دقة تقدير معلمة الصعوبة في النموذج اللوغاريتمي الثلاثي عند استخدام أشكال مختلفة لتوزيع قدرة الأفراد واحجام مختلفة للعينة، حُسِبت المتوسطات الحسابية، والانحرافات المعيارية للأخطاء المعيارية لتقديرات معالم صعوبة الفقرة، وحسب متغيرًا؛ شكل توزيع القدرة، وحجم العينة، وذلك باستخدام برنامج BILOG-MG 3؛ والجدول رقم (٣) يبين ذلك.

جدول (٣)
المتوسطات الحسابية والانحرافات المعيارية للأخطاء المعيارية لتقديرات معالم صعوبة الفقرات، وفقاً لمتغيري؛ شكل التوزيع وحجم العينة

شكل التوزيع	حجم العينة	الوسط الحسابي	الانحراف المعياري
طبيعي	١٠٠٠	٠,٠٦٣٢	٠,٠٣٨٢٦
	٤٠٠٠	٠,٠٤٦٨	٠,٠٢٧٨٨
	٨٠٠٠	٠,٠٣٨١	٠,٠١٩٧٤
ملتوي نحو اليمين	١٠٠٠	٠,٠٧٩٩	٠,٠٦٥٨٣
	٤٠٠٠	٠,٠٦٤٦	٠,٠٣٣٠٧
	٨٠٠٠	٠,٠٥٥٩	٠,٠١٩٦٥
ملتوي نحو اليسار	١٠٠٠	٠,٠٦٩٩	٠,٠٣١٩٥
	٤٠٠٠	٠,٠٣٥٠	٠,٠١٣٩٩
	٨٠٠٠	٠,٠٣٤٩	٠,٠١٤٠٢

يُلاحظ من بيانات الجدول (٣) وجود فروق ظاهرية بين المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم صعوبة الفقرة، وحسب متغيري؛ شكل التوزيع لقدرات الأفراد، وحجم العينة. والشكل (٢) يوضح ذلك.



شكل (٢)
المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم صعوبة الفقرة

ولمعرفة الدلالة الإحصائية لتلك الفروق الظاهرية وفقا لمتغيري شكل التوزيع، وحجم العينة والتفاعل بينهما؛ جرى استخدام تحليل التباين الثنائي للقياسات المتكررة على عاملين (شكل التوزيع لقدرات الأفراد، وحجم العينة) وكما هو موضح في الجدول (٤).

جدول (٤)
نتائج تحليل التباين الثنائي (قياسات متكررة)، للمتوسطات الحسابية لمعالم صعوبة الفقرة، وحسب متغيري؛ شكل التوزيع لقدرات الأفراد، وحجم العينة، والتفاعل بينهما

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	الدلالة
شكل التوزيع	٠,٠٣٥٩٦	٢	٠,٠١٧٩٨	١٦,٤١٤	٠,٠٠٠
حجم العينة	٠,٠٦٥٦٣	٢	٠,٠٣٢٨١	٢٩,٩٥٨	٠,٠٠٠
العينة ×التوزيع	٠,٠٠٦١٣	٤	٠,٠٠١٥٣	١,٣٩٨	٠,٢٣٤
الخطأ	٠,٤٨٣٠٣	٤٤١	٠,٠٠١١٠		
المجموع	١,٩١٦	٤٥٠			

أظهرت النتائج المتعلقة بمعالم صعوبة الفقرات الاختبار، وكما يتبين من الجدول (٤) ما يلي:

وجود فروق ذات دلالة إحصائية عند مستوى الدلالة الاحصائية) بين المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم صعوبة الفقرات تُعزى لمتغير شكل التوزيع لقدرات الأفراد (طبيعي، ملتوي نحو اليمين، ملتوي نحو اليسار). حيث بلغت قيمة ف (١٦,٤١٤) ولمعرفة أيٍّ من أشكال التوزيع لقدرات الأفراد كانت لصالحها هذه الفروق، أُجريت المقارنات باستخدام أدنى فرق (LSD Least Significant Differences) للمقارنات البَعدية على المتوسطات الحسابية المقدرة للأخطاء المعيارية لتقديرات معالم صعوبة الفقرات وِفقا لمتغير شكل التوزيع لقدرات الأفراد، وكما هو مبين في الجدول (٥).

جدول (٥)
نتائج المقارنات الثنائية بين المتوسطات الحسابية المُقدَّرة للأخطاء المعيارية لتقديرات معالم صعوبة الفقرات وفقا لمتغير شكل التوزيع

شكل التوزيع	الوسط الحسابي	ملتوي نحو اليمين	ملتوي نحو اليسار
طبيعي	٠,٠٤٩٤	٠,٠١٧٤*	٠,٠٠٢٨
ملتوي نحو اليمين	٠,٠٦٦٨		٠,٠٢٠٢*
ملتوي نحو اليسار	٠,٠٤٦٦		

يلاحظ من الجدول (٦) وجود فروق ذات دلالة إحصائية عند مستوى الدلالة بين الأوساط الحسابية المقدرة للأخطاء المعيارية لمعالم الصعوبة تبعا لكل شكل من أشكال توزيعات القدرة ، حيث أظهرت هذه الفروق أن أعلى فرق كان بين الوسطين الحسابين المقدرين للأخطاء المعيارية لمعالم صعوبة الفقرات الناتجة من التوزيع الملتوي نحو اليمين ونحو اليسار، إذ بلغت قيمة الفرق (٠,٠٢٠٢) ويمكن تفسير ذلك على أساس أنه عندما لا يكون هناك تطابق بين مدى القدرة

للمفحوصين (ذوي القدرات المتدنية و ذوي القدرات العالية) ومدى صعوبة الفقرات فمن المتوقع أن تظهر فروق ذات دلالة إحصائية بين تقديرات معالم الفقرة .حيث تظهر خاصيتا اثر السقف Ceiling effect واثر الحد الأدنى Floor effect في حين أظهرت الفروق بأن أقل فرق كان بين الوسطين الحسابين المقدرين للأخطاء المعيارية لمعالم صعوبة الفقرات الناتج من التوزيع الطبيعي و الملتوي نحو اليسار مما يعني تقارب معالم صعوبة الفقرات لكلا التوزيعين. وهذا يتفق مع ما أشارت إليه دراسة الشريفين وبني عطا (٢٠١١) ودراسة ساس وسشمت ووكر (Sass, Schmitt & Walker, 2004).

وجود فروق ذات دلالة إحصائية عند مستوى الدلالة الاحصائية) بين المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم صعوبة الفقرات تُعزى لمتغير حجم العينة. حيث بلغت قيمة ف (٢٩,٩٥٨) ولمعرفة أيٍّ مِن حجم العينة كانت لصالحها هذه الفروق، تم أجريت المقارنات باستخدام أدنى فرق (LSD Least Significant Differences) للمقارنات البَعدية على المتوسطات الحسابية المقدرة للأخطاء المعيارية لتقديرات معالم صعوبة الفقرات وِفقا لمتغير حجم العينة، وكما هو مبين في الجدول (٦).

جدول (٦)
نتائج المقارنات الثنائية بين المتوسطات الحسابية المُقدَّرة للأخطاء المعيارية لتقديرات معالم صعوبة الفقرات وفقاً حجم العينة

حجم العينة	الوسط الحسابي	٤٠٠٠فرد	٨٠٠٠ فرد
١٠٠٠ فرد	٠,٠٧١٠	٠,٠٢٢٢*	٠,٠٢٨٠*
٤٠٠٠ فرد	٠,٠٤٨٨	-	٠,٠٠٥٩
٨٠٠٠ فرد	٠,٠٤٣٠		-

*دال احصائيا عند ($\alpha \leq$ ٠,٠٥)

يلاحظ من الجدول (٦) وجود فروق ذات دلالة إحصائية عند مستوى الدلالة بين الأوساط الحسابية المقدرة للأخطاء المعيارية لمعالم الصعوبة تبعا لحجم العينة، وتشير نتائج وكما هو موضح بالجدول رقم (٦) بأن الخطأ المعياري في تقدير معلمة صعوبة الفقرة يتناقص بزيادة حجم العينة؛ وهذا يشير إلى أهمية مراعاة حجم العينة عند بناء الاختبارات وفق نظرية الاستجابة للفقرة، وذلك لأن الخطأ المعياري في التقدير يتناسب عكسياً مع دقة القياس. وهذا يتفق مع ما أشارت إليه دراسة عبابنة (٢٠٠٤) ودراسة ستون ويوموتو (Stone & Yamoto, 2004) ودراسة دي لا توري ويوان (De La Torre & Yuan, 2010).

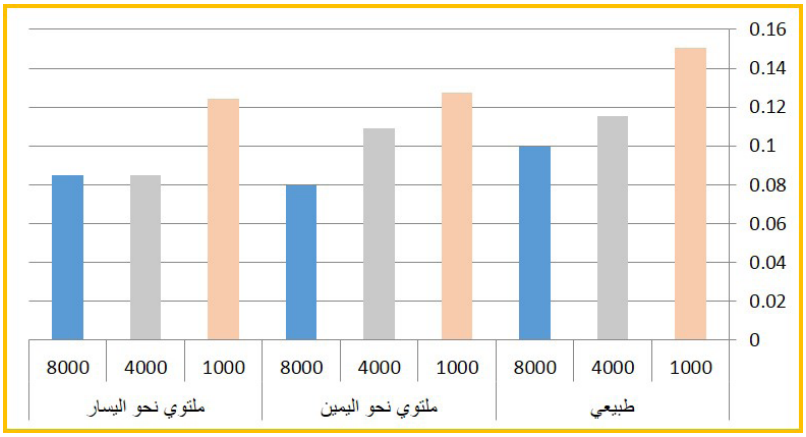
عدم وجود فروق ذات دلالة إحصائية عند مستوى الدلالة) بين المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم صعوبة الفقرات تُعزى للتفاعل الثنائي بين متغيري (حجم العينة ،شكل التوزيع لقدرات الأفراد).

وللإجابة عن الفرضية الثانية حول دقة تقدير معلمة التمييز في النموذج اللوغاريتمي الثلاثي عند استخدام أشكال مختلفة لتوزيع قدرة الأفراد وأحجام مختلفة للعينة، حُسِب المتوسطات الحسابية والانحرافات المعيارية للأخطاء المعيارية لتقديرات معالم تمييز الفقرات، وحُسِب متغيرًا شكل توزيع القدرة، وحجم العينة وذلك باستخدام برنامج BILOG-MG 3 والجدول رقم (٧) يبين ذلك.

جدول(٧)
المتوسطات الحسابية والانحرافات المعيارية للأخطاء المعيارية لتقديرات معالم تمييز الفقرات، وفقاً لمتغيري؛ شكل التوزيع وحجم العينة

شكل التوزيع	حجم العينة	الوسط الحسابي	الانحراف المعياري
طبيعي	١٠٠٠	٠,١٥٠٥	٠,٠٨٥٧٥
	٤٠٠٠	٠,١١٥٢	٠,٠٥٢٧٩
	٨٠٠٠	٠,٠٩٩٨	٠,٠٦٨٢٥
ملتوي نحو اليمين	١٠٠٠	٠,١٢٧٥	٠,٠٧٦٧٦
	٤٠٠٠	٠,١٠٩٢	٠,٠٧٣٠٣
	٨٠٠٠	٠,٠٧٩٥	٠,٠٤٠٢٧
ملتوي نحو اليسار	١٠٠٠	٠,١٢٤٢	٠,٠٥٤٨٩
	٤٠٠٠	٠,٠٨٥١	٠,٠٥٩١٢
	٨٠٠٠	٠,٠٨٥١	٠,٠٦١٢٥

يُلاحظ من بيانات الجدول (٧) وجود فروق ظاهرية بين المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم تمييز الفقرة، وحسب متغيري؛ شكل التوزيع لقدرات الأفراد، وحجم العينة. والشكل(٣) يوضح ذلك.



شكل (٣)
المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم تمييز الفقرة

ولمعرفة الدلالة الإحصائية لتلك الفروق الظاهرية وفقا لمتغيري شكل التوزيع، وحجم العينة والتفاعل بينهما؛ استُخدم التصميم العاملي للقياسات المتكررة على عاملين، ولذلك فقد تم استخدام تحليل التباين الثنائي للقياسات المتكررة على عاملين (شكل التوزيع لقدرات الأفراد، وحجم العينة) وكما هو موضح في الجدول (٨).

جدول(٨)
نتائج تحليل التباين الثنائي (قياسات متكررة)، للمتوسطات الحسابية لمعالم تمييز الفقرة، وحسب متغيري؛ شكل التوزيع لقدرات الأفراد، وحجم العينة، والتفاعل بينهما

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	الدلالة
شكل التوزيع	٠,٠٤٤١٩	٢	٠,٠٢٢١٠	٥,٢٤٧	٠,٠٠٦
حجم العينة	٠,١٦٤٧٨	٢	٠,٠٨٢٣٩	١٩,٥٦٢	٠,٠٠٠
العينة × التوزيع	٠,٠١٢٧٤	٤	٠,٠٠٣١٨	٠,٧٥٦	٠,٥٥٤
الخطأ	١,٨٥٧٣٨	٤٤١	٠,٠٠٤٢١		
المجموع	٧,٣٧٢٩١	٤٥٠			

أظهرت النتائج المتعلقة بمعالم تمييز فقرات الاختبار، وكما يتبين من الجدول (٨) ما يلي: وجود فروق ذات دلالة إحصائية عند مستوى الدلالة الاحصائية) بين المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم تمييز الفقرات تُعزى لمتغير شكل التوزيع لقدرات الأفراد (طبيعي ، ملتو نحو اليمين ، ملتو نحو اليسار). حيث بلغت قيمة ف (٥,٢٤٧) ولمعرفة أيٍّ

مِن اشكال التوزيع لقدرات الأفراد كانت لصالحها هذه الفروق، أُجريت المقارنات باستخدام أدنى فرق (LSD Least Significant Differences) للمقارنات البَعدية على المتوسطات الحسابية المقدرة للأخطاء المعيارية لتقديرات معالم تمييز الفقرات وِفقا لمتغير شكل التوزيع لقدرات الأفراد، وكما هو مبين في الجدول (٩).

جدول (٩)
نتائج المقارنات الثنائية بين المتوسطات الحسابية المُقدَّرة للأخطاء المعيارية لتقديرات معالم تمييز الفقرات وفقاً لمتغير شكل التوزيع

شكل التوزيع	الوسط الحسابي	ملتوي نحو اليمين	ملتوي نحو اليسار
طبيعي	٠,١٢١٨	٠,٠١٦٤*	٠,٠٢٣٧*
ملتوي نحو اليمين	٠,١٠٥٤	-	٠,٠٠٧٣
ملتوي نحو اليسار	٠,٠٩٨١		-

*دال احصائيا عند ($\alpha \leq$ ٠,٠٥)

يلاحظ من الجدول (٩) وجود فروق ذات دلالة إحصائية عند مستوى الدلالة بين الأوساط الحسابية المقدرة للأخطاء المعيارية لمعالم التمييز تبعا لكل شكل من أشكال توزيعات القدرة ، باستثناء الفرق بين التوزيعين (الملتوي 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 بأنَّ قيمة الفرق بين الوسط الحسابي لمعلمة التمييز في التوزيع الطبيعي والتوزيع الملتوي نحو اليسار كان أعلى قيمة حيث بلغت(٠,٠٢٣٧). وهذا يتفق مع ما أشارت إليه دراسة الشريفين وبني عطا (٢٠١١) ودراسة ساس وسشمت ووكر (Sass, Schmitt & Walker, 2004).

وجود فروق ذات دلالة إحصائية عند مستوى الدلالة الاحصائية) بين المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم تمييز الفقرات تُعزى لمتغير حجم العينة. حيث بلغت قيمة ف (١٩,٥٦٢) ولمعرفة اتجاه الفروق، تم إجراء المقارنات البعدية باستخدام أدنى فرق (LSD Least Significant Differences) للمقارنات البَعدية على المتوسطات الحسابية المقدرة للأخطاء المعيارية لتقديرات معالم تمييز الفقرات وِفقا لمتغير حجم العينة، وكما هو مبين في الجدول (١٠).

جدول (١٠)
نتائج المقارنات الثنائية بين المتوسطات الحسابية المُقدَّرة للأخطاء المعيارية لتقديرات معالم تمييز الفقرات وفقاً لحجم العينة

حجم العينة	الوسط الحسابي	٤٠٠٠ فرد	٨٠٠٠ فرد
١٠٠٠فرد	٠,١٣٤١	*٠,٠٣٠٩	*٠,٠٤٦٠
٤٠٠٠ فرد	٠,١٠٣٢	-	٠,٠١٥٠
٨٠٠٠ فرد	٠,٠٨٨١		-

*دال احصائيا عند ($\alpha \leq$ ٠,٠٥)

يلاحظ من الجدول (١٠) وجود فروق ذات دلالة إحصائية عند مستوى الدلالة بين الأوساط الحسابية المقدرة للأخطاء المعيارية لمعالم التمييز تبعا لحجم العينة، وتشير نتائج وكما هو موضح بالجدول رقم (١٠) بأن الخطأ المعياري في تقدير معلمة تمييز الفقرة يتناقص بزيادة حجم العينة ؛ وهذا يشير إلى أهمية مراعاة حجم العينة عند بناء الاختبارات وفق نظرية الاستجابة للفقرة، وذلك لأن الخطأ المعياري في التقدير يتناسب عكسياً مع دقة القياس. وهذا يتفق مع ما أشارت إليه دراسة عبابنة (٢٠٠٤) ودراسة ستون ويوموتو (Stone & Yamoto, 2004) ودراسة دي لا توري ويوان (De La Torre & Yuan, 2010).

عدم وجود فروق ذات دلالة إحصائية عند مستوى الدلالة) بين المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم تمييز الفقرات تُعزى للتفاعل الثنائي بين متغيري (حجم العينة، وشكل التوزيع لقدرات الأفراد).

وللإجابة عن الفرضية الثالثة حول دقة تقدير معلمة التخمين في النموذج اللوغاريتمي الثلاثي عند استخدام أشكال مختلفة لتوزيع قدرة الأفراد واحجام مختلفة للعينة، تم حساب المتوسطات الحسابية والانحرافات المعيارية للأخطاء المعيارية لتقديرات معالم التخمين للفقرات، وحسب متغيري شكل توزيع القدرة، وحجم العينة وذلك باستخدام برنامج BILOG-MG 3 والجدول رقم (١١) يبين ذلك.

جدول(١١)
المتوسطات الحسابية والانحرافات المعيارية للأخطاء المعيارية لتقديرات معالم التخمين للفقرات، وفقاً لمتغيري؛ شكل التوزيع وحجم العينة

شكل التوزيع	حجم العينة	الوسط الحسابي	الانحراف المعياري
طبيعي	١٠٠٠	٠,٠٢٤١	٠,٠٢٥٧٢
	٤٠٠٠	٠,٠٢٣٠	٠,٠٢٢٨٣
	٨٠٠٠	٠,٠١٦٥	٠,٠١٨٢٧

تابع جدول (١١)

شكل التوزيع	حجم العينة	الوسط الحسابي	الانحراف المعياري
ملتوي نحو اليمين	١٠٠٠	٠,٠٣١٥	٠,٠٢٦٢٣
	٤٠٠٠	٠,٠٢٥٢	٠,٠١٦٣٦
	٨٠٠٠	٠,٠١٨٩	٠,٠١٦٩٠
ملتوي نحو اليسار	١٠٠٠	٠,٠٢٦١	٠,٠٢٤٨٧
	٤٠٠٠	٠,٠٢٥٥	٠,٠٢٥٨٩
	٨٠٠٠	٠,٠١٦٧	٠,٠١٨٥٩

يُلاحظ من بيانات الجدول (١١) وجود فروق ظاهرية بين المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم التخمين للفقرة، وحسب متغيري؛ شكل التوزيع لقدرات الأفراد، وحجم العينة. ولمعرفة الدلالة الإحصائية لتلك الفروق الظاهرية وفقا لمتغيري شكل التوزيع، وحجم العينة والتفاعل بينهما؛ تم استخدام تحليل التباين الثنائي للقياسات المتكررة على عاملين (شكل التوزيع لقدرات الأفراد، وحجم العينة) وكما هو موضح في الجدول (١٢)

جدول (١٢)
نتائج تحليل التباين الثنائي (قياسات متكررة)، للمتوسطات الحسابية لمعالم تخمين الفقرة، وحسب متغيري؛ شكل التوزيع لقدرات الأفراد، وحجم العينة، والتفاعل بينهما

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	الدلالة
شكل التوزيع	٠,٠٠١٢١	٢	٠,٠٠٠٦٠	١,٢٣٩٢	٠,٢٩١
حجم العينة	٠,٠٠٧٨٤	٢	٠,٠٠٣٩٢	٨,٠٣٢٨	٠,٠٠٠
العينة × التوزيع	٠,٠٠٠٦٣	٤	٠,٠٠٠١٦	٠,٣٢١٦	٠,٨٦٤
الخطأ	٠,٢١٥٢١	٤٤١	٠,٠٠٠٤٩		
المجموع	٠,٤٦٤٠٩	٤٥٠			

أظهرت النتائج المتعلقة بمعالم تخمين فقرات الاختبار، وكما يتبين من الجدول (١٢) ما يلي:

عدم وجود فروق ذات دلالة إحصائية عند مستوى الدلالة) بين المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم التخمين للفقرات تُعزى لمتغير شكل التوزيع لقدرات الأفراد. حيث بلغت قيمة "ف" (١,٢٣٩) بدلالة إحصائية (٠,٢٩١) ويتبين وجود تقارب في قيم الأخطاء المعيارية لتقديرات معالم التخمين للفقرات الاختبار في اشكال التوزيع.

وجود فروق ذات دلالة إحصائية عند مستوى الدلالة الاحصائية) بين المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم تخمين الفقرات تُعزى لمتغير حجم العينة. حيث بلغت قيمة ف (٨,٠٣٣) ولمعرفة اتجاه الفروق، أُجريت المقارنات باستخدام أدنى فرق (LSD Least Significant Differences) للمقارنات البَعدية على المتوسطات الحسابية المقدرة للأخطاء المعيارية لتقديرات معالم تخمين الفقرات وِفقا لمتغير حجم العينة، وكما هو مبين في الجدول (١٣).

جدول (١٣)
نتائج المقارنات الثنائية بين المتوسطات الحسابية المُقدَّرة للأخطاء المعيارية لتقديرات معالم تخمين الفقرات وفقاً لحجم العينة

حجم العينة	الوسط الحسابي	٤٠٠٠	٨٠٠٠
١٠٠٠	٠,٠٢٧٢	*٠,٠٠٩٩	*٠,٠٠٧٢
٤٠٠٠	٠,٠٢٤٦	–	٠,٠٠٢٧
٨٠٠٠	٠,٠١٧٤		–

*دال احصائيا عند ($\alpha \leq$ ٠,٠٥)

يلاحظ من الجدول (١٣) وجود فروق ذات دلالة إحصائية عند مستوى الدلالة بين الأوساط الحسابية المقدرة للأخطاء المعيارية لمعالم التخمين تبعا لحجم العينة، وكما هو موضح بالجدول رقم (١٣) بأن الخطأ المعياري في تقدير معلمة تخمين الفقرة يتناقص بزيادة حجم العينة؛ وهذا يشير إلى أهمية مراعاة حجم العينة عند بناء الاختبارات وفق نظرية الاستجابة للفقرة، وذلك لأن الخطأ المعياري في التقدير يتناسب عكسياً مع دقة القياس.

عدم وجود فروق ذات دلالة إحصائية عند مستوى الدلالة) بين المتوسطات الحسابية للأخطاء المعيارية لتقديرات معالم تخمين الفقرات تُعزى للتفاعل الثنائي بين متغيري (حجم العينة، وشكل التوزيع لقدرات الأفراد).

الخلاصة والتوصيات

من خلال استعراض النتائج التي توصلت إليها الدراسة الحالية يتضح وجود أثر لحجم العينة على معالم الفقرات، حيث تبين عند تفحص قيم الخطأ المعياري في تقدير معالم الفقرات بأنها تتناقص بزيادة حجم عينة التطبيق، بمعنى إن عدد المفحوصين المستخدمين في تقدير معالم الفقرات يؤثر على استقرار ودقة معالم الفقرات .مما يشير إلى أهمية حجم العينة بوصفه عاملاً مؤثراً على دقة القياس وموضوعيته، وهذا يتفق مع ما أشارت إليه دراسة

عبابنة (٢٠٠٤) ودراسة ستون ويوموتو (Stone & Yamoto, 2004) ودراسة دي لا توري ويوان (De La Torre & Yuan, 2010). وتبين كذلك أثر لاختلاف شكل توزيع القدرة على معالم الفقرات، حيث يلاحظ أن تقديرات معالم الفقرات تتسم بعدم الاستقرار والثبات عبر التوزيعات المختلفة لقدرات الأفراد. وهذا يتفق مع ما أشارت إليه دراسة الشريفين وبني عطا (٢٠١١) ودراسة ساس وسشمت ووكر (Sass, Schmitt& Walker, 2004)، واستناداً إلى النتائج ومناقشتها فإن الدراسة توصي بضرورة مراعاة حجم العينة عند بناء الاختبارات بالاعتماد على النظرية الحديثة في القياس، وذلك بزيادة عددها قدر المستطاع للتقليل من أخطاء القياس وبالتالي زيادة الدقة في القياس. وتوصي كذلك بإجراء دراسات مماثلة تتناول أثر اختلاف شكل توزيع القدرة على معالم الفقرات في سياق طريقة التقدير.

المراجع

الشرفين، نضال؛ بني عطا، زياد (٢٠١٢). أثر اختلاف شكل توزيع القدرة على معالم الفقرة ودالة المعلومات للاختبار. المجلة الاردنية في العلوم التربوية، ٢، ١٥١ -١٦٦.

عبابنة، عماد (٢٠٠٤). "أثر 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

علام، صلاح الدين (٢٠٠٥). نماذج الاستجابة للمفردات الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي. (ط ١)، القاهرة: دار الفكر العربي.

Allen, M. J. & Yen, W. M. (1979). *Introducation temeasureme Theory California*. USA: Brooks /colepublishing compony.

Anastasie & Urbian (1997). *Psycholoical testing*. (7th Ed.), New York: prentices, Hall.

Barnes, L. and Wise, S. (1991). The utility of a modified one - parameter IRT model with small samples. *Journal of Applied Measurement in Education, 4*(2), 143 – 153.

Crocher, L. & Algian, J. (1986). *Introduction to classical and modern test theory*. New York G.B.S.: college publishing.

De Gruijter, D. M., & Van der Kamp, L. J. Th. (2005). *Statistical Test Theory for Education and Psychology*. Retrieved, from: www.leidenuniv,n1/griuijterdnme.

De la Torre, J. & Hong, Y .(2010). Parameter estimation with small sample size a higher-order IRT model approach. *Applied Psychological Measurement*. *34*(4), 267-285.

Hambelton, R. K. & Swaminathan H. (1985). *Item response theory: principles and application*. Boston, Dordrecht caster, Kluwer: Nijhoff publishing.

Lord, F. (1980). *Application of item response theory to practical testing problems*. New Jersey: Lawrence Erlbau Associats, Inc.

Lumsden, J. (1961). The construction of unidimensional test. *Psychological Bulletin, 58*(2), 122-131.

Nunnally, J. C. (1978). *Psychological theory*. (2nd Ed.), New York: McGraw – Hill.

Randall, S. (1998). *Comparing measurement theories*. Paper present at the annual meeting of the Amrican Educational Research associdion, San Diego CA. April Stage.

Reckase, M. D. (1979). Unifactor trait models applied to multifactor test: rsults and application, *Journal of Educational Statistics, 4*)3), 207-230.

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, 2004.

Stone, M. and Yumoto, F. (2004). The effect of sample size for estimating Rasch / IRT parameters with dichotomous items. *Journal of Applied Measurement, 5*(1), 48 - 61.