أثر عدد نقاط التربيع في القيمة القصوى لدالة معلوماتية الفقرة والثبات الامبريقي والنظري للاختبار

*راجي عوض الصرايرة**

ملخص

هدفت هذه الدراسة إلى فحص أثر عدد نقاط التربيع في القيمة القصوى لدالة معلوماتية الفقرة والثبات الامبريقي والنظري للاختبار تحت افتراضات النموذج اللوجستي ثلاثي المعلمة. وللإجابة عن أسئلة الدراسة تم توليد استجابات تحت افتراضات النموذج اللوجستي ثلاثي المعلمة على (50) فقرة ثنائية الإجابة، من خلال برمجية (WINGEN3) وتم تحليل الاستجابات المولدة تبعا لاختلاف عدد نقاط التربيع من خلال برمجية (BILOG-MG3). وأظهرت نتائج تحليل التباين الأحادي للمجموعات المعشاة (One Way ANOVA Randomize Block Design) وجود فروق ذات دلالة إحصائية عند مستوى دلالة ($\alpha \leq 0.05$) بين متوسطات القيم القصوى لدالة معلوماتية الفقرة تعزى إلى عدد نقاط التربيع، حيث أظهرت النتائج أن القيمة القصوى لدالة معلوماتية الفقرة تزداد بزيادة عدد نقاط التربيع، وكما أظهرت نتائج الدراسة عدم وجود فروق دالة إحصائيا عند مستوى دلالة ($\alpha \leq 0.05$) بين تقديرات الثبات الامبريقي والنظري تبعا لاختلاف عدد نقاط التربيع.

الكلمات الدالة: عدد نقاط التربيع، القيمة القصوى لدالة معلوماتية الفقرة، الثبات الامبريقي، الثبات النظري النموذج اللوجستي ثلاثي المعالم.

المقدمة

أسهمت نظرية استجابة الفقرة (Item Response Theory) بحلٍّ معظم جوانب القصور التي لم تستطع نظرية الاختبار الكلاسيكية (Classical Test Theory) معالجتها، مثل معادلة درجات الاختبار وتحيّز الفقرات وبنوك الأسئلة واللاتغير في تقدير خصائص الفقرات، إذ قامت هذه النظرية الجديدة على مجموعة من الافتراضات وهي افترض أحادية البعد (Unidimensionality) والذي يعني وجود سمة واحدة تُفسر أداء المفحوص على الاختبار، بمعنى أن درجة المفحوص على الاختبار تعكس السمة التي يقيسها الاختبار فقط. وهناك طرق إحصائية مختلفة تُستخدم لفحص تحقيق البيانات لهذا الافتراض، ومن أهمها إسلوب التحليل العاملي في حين تمثل الافتراض الثاني والذي يطلق عليه الاستقلال المحلي (Local Independence) وتحقيق هذا الافتراض يتطلب أن تكون استجابة المفحوصين على فقرات الاختبار مستقلة احصائياً عند مستوى قدرة معينة بمعنى آخر أن لا تؤثر إجابة المفحوص في فقرة ما إيجابا أو سلبا بإجابته على الفقرات الأخرى (Crocker and Algina, 1986) وهذا الافتراض يتحقق فقط في حالة أن الاختبار أحادي البعد والذي يعني أن الأفراد من مستوى قدرة معين لا يتأثر اداؤهم على فقرة ما بأدائهم على فقرة أخرى (Hambleton and Swaminathan, 1985) أما الافتراض الثالث والذي يسمى بمنحى خصائص الفقرة (ICC) (Item Characteristics Curve) حيث إن مفهوم منحنى نظرية استجابة الفقرة عبارة عن اقتران رياضي يربط بين احتمالية نجاح المفحوص على الفقرة والقدرة التي تقيسها مجموعة من الفقرات المكونة للاختبار، وأخيرا افتراض السرعة في الأداء والذي يفترض أن عامل السرعة لا يلعب دوراً في الإجابة عن فقرات الاختبار، بمعنى أن الإجابة الخاطئة على الفقرة سببها القدرة وليس الوقت المخصص للاختبار (علام، 2005) وتعتمد نظرية استجابة الفقرة على نماذج رياضية تهدف لتحديد العلاقة بين احتمالية إجابة المفحوص على فقرة ما إجابةً صحيحة وبين القدرة الكامنة التي تكمن وراء هذا الأداء سواء كانت ثنائية الإجابة (Dichotomous response) أو متعددة الإجابة (polytomous response) وهذه النماذج عبارة عن دوال رياضية احتمالية متنوعة تختلف الصيغة الرياضية للنموذج باختلاف عدد معالم الفقرة المكونة لبنائها الرياضي فنماذج نظرية استجابة الفقرة ثنائية الإجابة (Dichotomous IRT Models)

* قسم علم النفس، كلية العلوم التربوية، جامعة مؤتة، الكرك، الأردن.
تاريخ استلام البحث 2014/1/14، وتاريخ قبوله 2014/4/6.

©2016 عمادة البحث العلمي/ الجامعة الأردنية. جميع الحقوق محفوظة.

تقسم حسب عدد المعالم كما أوردها هامبلتون وسومينثان وروجرز (Hambelton, Swaminathan and Rogers, 1991) وكذلك أمبريتسون وريز (Embretson and Reise, 2000) إلى ثلاثة نماذج هي النموذج اللوجستي أحادي المعلمة (One Parameter Logistic Model- 1PLM) ويعد هذا النموذج من أوسع النماذج المستخدمة في نظرية استجابة الفقرة والذي يفترض أن جميع الفقرات لا تختلف عن بعضها إلا بمعلمة صعوبة الفقرة (b_i) ويفترض تساوي معلم التمييز (a_i) لجميع الفقرات في حين تقترب معلمة التخمين (c_i) للفقرات من الصفر ويتخذ النموذج الصيغة الرياضية كما في المعادلة (1) لتعبر عن احتمالية الإجابة الصحيحة على الفقرة والنموذج الثاني يسمى بالنموذج اللوجستي ثنائي المعلمة (Two Parameter Logistic Model- 2PLM) وهذا النموذج يفترض أن الفقرات تختلف في معلمتي الصعوبة والتمييز في حين تقترب معلمة التخمين من الصفر، وتعطى احتمالية الإجابة الصحيحة على الفقرة بالصيغة الرياضية كما في المعادلة (2) واخيرا النموذج اللوجستي ثلاثي المعلمة (Three Parameter Logistic Model- 3PLM) ويعد النموذج اللوجستي ثلاثي المعلمة الصيغة الرياضية الأكثر عموميه، وذلك بإضافته معلمة أخرى إلى معالم الفقرة وهي معلمة التخمين ويُعبَّر عن احتمالية الإجابة الصحيحة على الفقرة بالصيغة الرياضية كما في المعادلة (3).

$$p_i(\theta) = \frac{1}{1 + e^{-D(\theta - b_i)}} \qquad i = 1,2,3.....,n \qquad (1)$$

$$p_i(\theta) = \frac{1}{1 + e^{-D a_i(\theta - b_i)}} \qquad (2)$$

$$p_i(\theta) = c_i + (1 - c_i)\frac{1}{1 + e^{-D a_i(\theta - b_i)}} \qquad (3)$$

$P_i(\theta)$: احتمالية الإجابة الصحيحة على الفقرة (i) من قبل مفحوص تم اختياره عشوائيا من مستوى القدرة (θ).

b_i: معلمة الصعوبة للفقرة (i).

n: عدد فقرات المكونة للاختبار.

e: الأساس اللوغرتمي الطبيعي ويساوي (2.718)

θ: قيمة مستوى القدرة.

D: عامل التدريج ويساوي (1.7).

(c_i): معلمة التخمين للفقرة (i)

ويضاف ايضا لما قدمته نظرية استجابة الفقرة تلك الطرق الرياضية المختلفة لتقدير معلمة القدرة، إذ اختلفت هذه الطرق في إسلوبها الرياضي لتقدير معلمة القدرة، ونورد بعضاً منها والتي تستخدم بشكل خاص في برمجية (Bilog-MG3) ونعرضها بشكل موجز كما وضحها كل من أمبريتسون وريز (Embretson and Reise, 2000) و(التقي، 2000) ومنها طريقة الأرجحية العظمى (Maximum likelihood (ML)) وكذلك طريقة القيمة العظمى للتوزيع البعدي (Bayes or expected a posteriori (EAP)) اما الطريقة التي استخدمت في هذه الدراسة فهي طريقة توقُّع التوزيع البعدي (Bayes modal or maximum a posteriori (MAP)) اذ تقوم فكرة تقدير القدرة وفقاً لهذه الطريقة على تقسيم متصل السمة إلى (61) نقطة بطول (0.1) وتسمى كل نقطة تربيع (Q_r) ويحدد لكل نقطة تربيع وزن يرمز له بالرمز $(w(Q_r))$ ويتم تحيد القيمة العددية لاقتران الأرجحية العظمى عند كل نقطة (Q_r) وفي هذه الطريقة يتم تقدير القدرة بدون اجراء عمليات تقريب متتابع بل من خلال الصيغة الرياضية الآتية:

$$\theta = \sum_{r=1}^{61} \frac{\left[Q_r \times L(Q_r) \times w(Q_r)\right]}{\left[L(Q_r) \times w(Q_r)\right]}$$

(Q_r): قيم معلمة القدرة عند كل ربيعي (r)

$w(Q_r)$: وزن قيمة معلمة القدرة عند الربيعي (r)

$L(Q_r)$: اقتران الارجحية العظمى.

ومن المفاهيم الأساسية التي أدخلتها نظرية استجابة الفقرة في بناء الاختبارات مفهوم دالة معلوماتية الفقرة (IIF) Item Information Function) ودالة معلوماتية الاختبار (TIF) (Test Information Function)، والتي يعتمد عليها في بناء الاختبارات التكيفية (Adaptive Tests)، حيث أن استراتيجيات بناء الاختبار التكيفي تعتمد على مفهوم دالة المعلومات، إذ يتم تحديد مساهمة الفقرة في دالة معلومات الاختبار بشكل مستقل عن باقي الفقرات المكونة للاختبار، فإذا كان هناك معرفة مسبقة أولية عن قدرات المفحوصين فيمكن انتقاء الفقرات التي تعظم المعلوماتية التي يسهم بها الاختبار للوصول إلى تقديرات دقيقة للقدرة، ولمفهوم دالة المعلومات تعريف رياضي يعطى من خلال الصيغة الرياضية الآتية:

$$I_g(\theta) = \sum_{i=1}^{n} \left[\frac{(P_g'(\theta))^2}{(P_g(\theta))Q_g(\theta)} \right]$$

حيث إن $P_g^{\backslash}(\theta)$ المشتقة الأولى لمنحنى خصائص الفقرة عند النقطة (g) على متصل القدرة وفي علم المعادلات التفاضلية فإن المشتقة الأولى عند أي قيمة للقدرة تساوي ميل الخط المستقيم عند مماس منحنى خصائص الفقرة عند النقطة (g) على متصل القدرة وتكون قيمة دالة معلومات الفقرة أكبر ما يمكن عند القيمة (g) على متصل القدرة والتي تقابل أكبر ميل للمماس كما أن بسط المعادلة يمثل المشتقة الأولى، لذلك فإن دالة معلومات الفقرة ستكون أكبر ما يمكن عندما تكون قيمة المشتقة أكبر وبالتالي فإن الفقرة تزودنا بمعلوماتية عند النقطة (θ_g) التي تقابل أكبر ميل وتحديد قيمة (θ) التي يكون عندها للفقرة أكبر معلوماتية يقلل من الخطأ المعياري الذي يتعلق بتقدير القدرة للمفحوصين، وبالتالي يتم بناء فقرات الاختبار التكيفي بشكل أكثر دقة (Swaminathan, 1985) Hambleton and.

والمتفحص في الصيغة الرياضية لدالة معلوماتية الفقرة يجد أن معلم القدرة ومعالم الفقرة أساس تلك الصيغة، وهذا يقودنا إلى حديث عن إجراءات طريقة تقدير معالم الفقرات، إذ لا بد من الاشارة إلى مفهوم أساسي يعتمد عليه في تقدير معالم الفقرات وهو تربيع جاوس (Gaussian quadrature) إذ إن طريقة الأرجحية العظمى الهامشية المستخدمة في برمجية (BILOG-MG3) لتقدير معالم الفقرات تستخدم تكامل عددي (numerical integration) والذي يعرف بتربيع جاوس (Gaussian quadrature) من أجل تقريب وتسهيل عملية التكامل (Baker and Kim, 2004) ولتحقيق ذلك فإن برمجية (BILOG-MG3) واعتمادا على تربيع جاوس تستخدم إسلوب العرض البياني من خلال المدرج التكراري البسيط (Simple Histogram Technique) كإجراء رياضي لحساب التكامل، ومن أهم شروط تطبيق هذا الإجراء هو تحقيق افتراض أن المفحوصين تم سحبهم من المجتمع عشوائيا وكما أن الشكل النموذجي لتوزيع القدرة القبلي هو التوزيع الطبيعي المعياري $g(\theta/\tau)$ والذي يكون فيه مدى قيم معلم القدرة يتراوح من (4- إلى +4) مقسمة على فترات متساوية بمقدار (q) بحيث يتم تمثيل هذه النقاط بأعمدة تكرارية يتضمن كل عمود نقطة منتصف تسمى التربيع (quadrature) ويتكون متصل القدرة من (χ_q) قيم التربيع، $\chi_q = (q_1, q_2, ..., q_n)$ ويرتبط بكل نقطة تربيع (q) بوزن يرمز له بالرمز $A(\chi_q)$ وهذا الوزن يرتبط بارتفاع اقتران توزيع القدرة القبلي وكما أن وزن التربيع يتم حسابه من خلال حاصل ضرب عرض كل مستطيل في المدرج التكراري في ارتفاع المستطيل ومن الجدير بالذكر أن استخدام تربيع جاوس لا تقدر معالم الفقرة مباشرة من استجابات المفحوصين على فقرات الاختبار بل يتم تقديرها من بيانات افتراضية قبلية (artificial data) عند كل نقطة تربيع والبيانات الافتراضية القبلية عند كل نقطة تربيع تتضمن العدد المتوقع للمفحوصين (expected number of examinees) (Baker and Harwell, 2004)، ومن هنا فإن الدراسة الحالية تسعى إلى فحص أثر عدد نقاط التربيع التي تقسم متصل القدرة على القيمة القصوى لدالة معلوماتية الفقرة والثبات الامبريقي والنظري. وتأتي أيضا اهميتها من كون عدد نقاط التربيع من المكونات الاساسية لمعادلة تقدير القدرة المعتمدة بهذه الدراسة والتي بدورها تؤثر في دالة تقدير القيمة القصوى لدالة معلوماتية الفقرة والثبات الامبريقي والنظري.

ويوضح ديو تويت (Du Toit,2003) أن نظرية استجابة الفقرة وظفت مفهوم الثبات الكلاسيكي من خلال طريقتين فالثبات الكلاسيكي (Classical Reliability) بشكل عام يعرف وفق النظرية الكلاسيكية للاختبار بأنه نسبة تباين الدرجة الحقيقة إلى تباين الدرجة الملاحظة وتباين الدرجة الملاحظة هو مجموع تباين الدرجات الحقيقة وتباين الخطأ أما في سياق نظرية استجابة الفقرة فان الدرجات الحقيقة تتمثل في قيم القدرة غير الملاحظة للمفحوصين (unobservable theta) والتي يتم تقديرها بخطأ معياري اعتمادا على نمط الإجابة على الفقرة وتظهر قيم تقدير معلمة القدرة للمفحوصين والخطأ المعياري لتقديرها في الطور الثالث (phase 3) في برمجية BILOG-MG3) وكما اسلفنا أن الثبات الكلاسيكي تم توظيفه في نظرية استجابة الفقرة من خلال برمجية (BILOG-MG3) بطريقتين مختلفتين تبعا للكيفية التي تقدر بها تباينات الدرجة الحقيقة وتباين الخطأ إذ تسمى الطريقة الأولى بالثبات النظري (theoretical reliability) والطريقة الثانية تدعى بالثبات الامبريقي (empirical reliability) وفيما يلي توضيح مفهوم كل منهم.

الثبات النظري (theoretical reliability)

يتم تقدير معامل الثبات النظري اعتمادا على معالم الفقرة المكونة للاختبار ولا تدخل تقديرات معالم القدرة للمفحوصين في تقدير معامل الثبات النظري وبدلاً من ذلك يتم افتراض أن قيم القدرة للمفحوصين تتبع التوزيع الطبيعي بوسط صفر وانحراف معياري واحد، إذ يعتمد هذا الافتراض في تقدير معامل الثبات والصيغة الرياضية لحساب معامل الثبات النظري تعطى كالآتي:

$$theoreticl - reliability = \left[\frac{1}{1+\sigma^2}\right]$$

حيث أن (σ^2) تمثل تباين الخطأ.

ويتم تقدير تباين الخطأ في حالة النموذج اللوجستي ثلاثي المعلمة الذي اعتمدته الدراسة من خلال الصيغة الرياضية الآتية:

$$\sigma^2 = (S.E.(\hat{\theta}))^2 = \left\{\frac{1}{D^2\sum_{j=1}^{n} a^2\left[\frac{1-p(\hat{\theta})}{p(\hat{\theta})}\right]\left[\frac{p(\hat{\theta})-g_j}{1-g_j}\right]^2}\right\}$$

حيث أن:

$p(\hat{\theta})$: تمثل احتمالية الإجابة الصحيحة على الفقرة عند مستوى القدرة $(\hat{\theta})$.

D: عامل التدريج ويساوي (1.7).

g_j: اقتران الكثافة الاحتمالية.

الثبات الامبريقي (empirical reliability)

يتم تقدير معامل الثبات الامبريقي من خلال إيجاد النسبة بين تباين الدرجة الحقيقية إلى تباين الدرجة الملاحظة، حيث إن تباين الدرجة الملاحظة يتكون من تباين الدرجة الحقيقية وتباين الخطأ، ويتم تقدير تباين الدرجة الحقيقية بشكل مباشر من خلال حساب تباين تقديرات قيم القدرة للمفحوصين، وكما يتم تقدير تباين الخطأ من خلال إيجاد دقة القياس والذي يسمى الانحراف المعياري البعدي (posterior standard deviation) ويعبر عنه بالصيغة الرياضية الآتية:

$$PSD(\bar{\theta}) \cong \left[\frac{\sum_{k=1}^{q}(X_k-\bar{\theta})^2 P(x_i/X_k)A(X_k)}{\sum_{k=1}^{q}P(x_i/X_k)A(X_k)}\right]$$

$A(X_k)$: الوزن الموجب (positive weight) المعطى من خلال اقتران الكثافة الاحتمالية $g(x)$.

$P(x_i/X_k)$: احتمالية الإجابة الصحيحة على النمط (x_i) عند مستوى القدرة (X_k).

(X_k): قيمة القدرة عند النقطة (k) على متصل القدرة.

ويتم تقدير القيمة $(\bar{\theta})$ من خلال الصيغة الرياضية الآتية:

$$\bar{\theta} \cong \left[\frac{\sum_{k=1}^{q}(X_k)P(x_i/X_k)A(X_k)}{\sum_{k=1}^{q}P(x_i/X_k)A(X_k)}\right]$$

مشكلة الدراسة

تكمن مشكلة هذه الدراسة في الكشف عن أثر عدد قيم التربيع على القيمة القصوى لدالة معلوماتية الفقرة والثبات الامبريقي والثبات النظري للاختبار باستخدام نظرية استجابة الفقرة، وتحددت مشكلتها في الإجابة عن الأسئلة الآتية:

السؤال الأول: هل تختلف القيمة القصوى لدالة معلوماتية باختلاف عدد قيم التربيع؟

السؤال الثاني: هل تختلف تقديرات معامل الثبات الامبريقي والثبات النظري باختلاف عدد قيم التربيع؟

أهمية الدراسة

تبرز أهمية الدراسة في التحقق الامبريقي بالكشف عن أثر عدد قيم التربيع على القيمة القصوى لدالة معلوماتية الفقرة والثبات الامبريقي والثبات النظري باستخدام نظرية استجابة الفقرة، وذلك نظرًا لأهمية دالة معلوماتية الفقرة وثبات الاختبار في القياس النفسي والتربوي، إذ إن لبّ عملية القياس يكمن في دقة تقدير مَعْلمة القدرة للمفحوصين ودقة تقدير معالم الفقرة، حيث تتأثر دقة تقديرات معلمة القدرة ومعالم الفقرة بمجموعة من العوامل بينت بعضها نتائج الأبحاث مثل حجم العينة وطول الاختبار وشكل التوزيع لمعلم القدرة ونوع النموذج اللوجستي ولكن مجال البحث في المعالم المكونة للبناء

الرياضي للدوال الرياضية التي يتم التقدير من خلالها حظى باهتمام أقل مما يتطلب تقييم أثر هذه المعالم المكونة للبناء الرياضي على دقة التقديرات التي صمم لأجلها ولأن دالة المعلومات للفقرة (IIF) من الاحصاءات المهمة في نظرية استجابة الفقرة حيث يمكن من خلالها تحديد الخطأ المعياري للتقدير وجاءت الحاجة لتقييم أثر عدد نقاط التربيع على تقدير قيمة دالة معلومات الفقرة وكذلك على الثبات الامبريقي والنظري كمفاهيم كلاسيكية استخدمت من وجهة نظر النظرية الحديثة في القياس كون عدد نقاط التربيع اساس في بنائها الرياضي وكما جاءت هذه الدراسة أيضا استكمالا لدعم الجهود التي بذلت في هذا الجانب.

متغيرات الدراسة

المتغير المستقل: عدد قيم التربيع لمتصل القدرة وله خمسة مستويات (10، 15 ،25 ،35 ، 50) نقطة تربيع والذي يتم التحكم في عددها من خلال أوامر خاصة في برمجية BILOG-MG3).

المتغير التابع: القيم القصوى لدالة معلوماتية الفقرة،تقديرات معامل الثبات الامبريقي والثبات النظري.

مصطلحات الدراسة

عدد قيم التربيع: عدد الفترات التي تم تقسيم متصل القدرة اليها وفي الدراسة الحالية تم تقسيم متصل القدرة إلى (10،15 ، 25،35 ، 50) نقطة تربيع.

برنامج (Wingen3): وهي من تصميم هان وهامبلتون (Han and Hambleton, 2007) وتتكون من عدة نوافذ، منها نافذة تتعلق بخصائص المفحوصين اذ تتضمن ثلاثة أشكال لتوزيع القدرة وهي التوزيع الطبيعي (Normal distribution) والتوزيع المنتظم (Uniform distribution) وتوزيع بيتا (Beta distribution) والذي يتيح لنا توليد توزيعات موجبة الالتواء وسالبة بدرجات مختلفة ونافذة اخرى تتعلق بخصائص الفقرات حيث تتضمن عدد من الخيارات مثل عدد الفقرات (number of items) المراد توليد استجابات عليها، وأيضا تتضمن هذه النافذة عدة خيارات لنموذج استجابة الفقرة المستخدم في توليدها، سواء النماذج المتعلقة بالاستجابات الثنائية (Dichotomous responses) أو النماذج متعددة الإجابة (Polychromous response) كما تتضمن توزيعات مختلفة لمعالم الفقرة المراد توليد استجابات عليها مثل التوزيع الطبيعي (Normal distribution) والتوزيع المنتظم (Uniform distribution) وتوزيع لوغريتم الطبيعي (Lognormal).

برمجية (BILOG- MG3): هي إحدى البرمجيات الإحصائية المتعلقة بالتحليلات الخاصة بنظرية استجابة الفقرة ثنائية الإجابة، حيث تظهر مخرجات البرنامج والمتعلقة بالطور الثالث تقديرًا للقيم القصوى لدالة معلوماتية الفقرة وكذلك تقدير معاملات الثبات النظري والامبريقي، وكما يتضمن مخرجات الطور الثاني إحصائيات جودة مطابقة الفقرة للنموذج اللوجستي (Du Toit, 2003).

حدود الدراسة: تحددت الدراسة الحالية في النقاط الآتية:

- تعتمد الدراسة الحالية على استجابات تم توليدها بالمحاكاة باستخدام برمجية Wingen3.
- تعتمد الدراسة الحالية على أستخدام استجابات ثنائية التصحيح (0,1) لفقرات اختيار من متعدد بأربعة بدائل.
- تستخدم الدراسة التحليلات الخاصة بالنموذج اللوجستي ثلاثي المعلمة وطريقة توقع التوزيع البعدي (EAP) لتقدير معلمة القدرة .
- اقتصرت الدراسة على حجم عينة مقداره (1000) فرد و(50) فقرة.
- التوزيع القبلي للقدرة توزيعًا طبيعيًا.

الدراسات السابقة

بحثت العديد من الدراسات في موضوع دالة معلومات الفقرة، ومن هذه الدراسات دراسة الزبون (2013) التي هدفت إلى معرفة أثر حجم العينة على تقدير دالة معلومات الاختبار والخطأ المعياري في تقديرها باستخدام نظرية استجابة الفقرة، وللإجابة عن أسئلة الدراسة تم استخدام نتائج الاختبار الوطني لمادة الرياضيات، حيث تم اختيار عشوائيا اجابات (7500) طالبا وطالبة من طلبة الصف الثامن الأساسي على اختبار مكون من (40) فقرة من نوع الاختيار من متعدد، تم توزيعهم على خمس مجموعات بأحجام (500،1000،1500،2000،2500) مفحوص، وتم الاعتماد على برمجية (BILOG-MG3) وبرمجية (SPSS) لتحليل نتائج الاختبار الوطني لمادة الرياضيات، حيث تم إيجاد معالم الفقرات بالاعتماد على النموذج اللوجستي ثلاثي المعالم في نظرية استجابة الفقرة، وقد أظهرت النتائج أن مقدار دالة معلومات الاختبار تزداد بزيادة حجم العينة، كما أن الخطأ المعياري تقدير دالة معلومات الاختبار يتناقص بزيادة حجم العينة.

وفي دراسة أجراها بني عطا والشريفين (2012) هدفت إلى

التحقق من أثر اختلاف شكل توزيع القدرة على معالم الفقرة ودالة المعلومات للاختبار، ومن أجل الإجابة عن أسئلة الدراسة تم توليد بيانات ثنائية الاستجابة على (60) فقرة تبعا لأربعة أشكال من توزيعات القدرة وفق افتراضات النموذج اللوجستي ثلاثي المعلمة من خلال استخدام برنامج (WINGEN3)، وتم تحليل استجابات البيانات المولدة لكل شكل من أشكال توزيع القدرة وفق النموذج اللوجستي ثلاثي المعلمة، وباستخدام برنامج (BILOG-MG 3)، إذ كشفت نتائج تحليل التباين الأحادي للقياسات المتكررة عن وجود فروق ذات دلالة إحصائية بين متوسطات معالم الصعوبة للفقرات، وكذلك بين متوسطات معالم التمييز للفقرات تبعا لشكل توزيع القدرة، وعلى الرغم من تباين معالم الفقرات، فقد أظهرت النتائج أنَّ قيم معاملات الارتباط بين تقديرات معالم الفقرات (الصعوبة، التمييز) تبعا لشكل توزيع القدرة كانت دالة إحصائيا، والذي عدَّته الدراسة مؤشرًا لاستقرار معالم الفقرات عبر الأشكال المختلفة لتوزيع القدرة، وكما كشفت النتائج كذلك أنَّ دالة المعلومات للاختبار أعطت أعلى قيم لدالة المعلومات ضمن مدى القدرة (0.5− و0.63) حيث كانت قيم الخطأ المعياري أقل ما يمكن في جميع أشكال توزيع القدرة وكما كشفت عن وجود فروق بين معاملات الثبات الامبريقي تعزى إلى شكل توزيع القدرة.

واجرت دراسة جمحاوي (2000) والتي كانت بعنوان مقارنة خصائص الفقرات وفق النظرية التقليدية ونظرية استجابة الفقرة في مقياس للقدرة العقلية وللإجابة عن اسئلة الدراسة طبق اختبار للقدرة العقلية على عينة مؤلفة من (1061) طالباً وطالبة من طلبة الصف التاسع الاساسي في مديريات محافظة إربد، وتوصلت الدراسة إلى عدم وجود فروق دالة احصائيا بين معامل الثبات للاختبار الناتج من النظرية التقليدية ومعامل الثبات الناتج من نموذج راش والنموذج اللوجستي ثلاثي المعلمة. وكذلك من الدراسات التي اهتمت بدالة معلوماتية الفقرة دراسة اجراها هامبلتون وجونز (Hambleton and Jones, 1994) هدفت إلى فحص أثر الفقرات ذات القدرة التمييزية العالية على دقة تقدير دالة معلوماتية الفقرة ومن اجل الإجابة عن اسئلة الدراسة تم توليد استجابات افراد بحجمي عينة (2000,500) مفحوص وبنك اسئلة يتضمن (80) فقرة اختيار من متعدد وكشفت نتائج الدراسة ان حجم العينة يؤثر في دقة تقدير معلوماتية الفقرة أكثر من طول الاختبار والفقرات ذات القدرة التمييزية المرتفعة تبالغ في قيم دالة معلوماتية الفقرة.

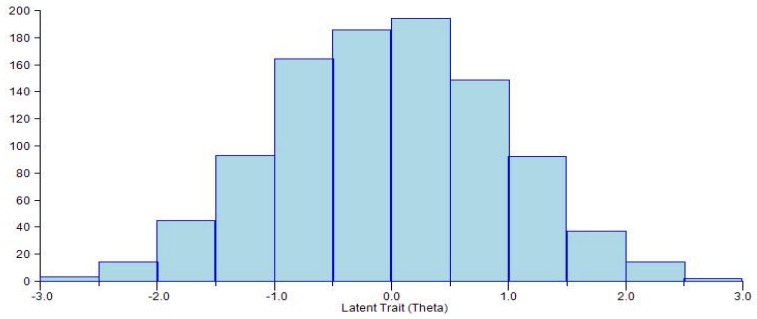
وفي دراسة أجراها سونج (Seong, 1990) هدفت للتحقق من أثر شكل توزيع القدرة وحجم العينة وعدد نقاط التربيع باستخدام النموذج اللوجستي ثنائي المعلمة على حساسية تقديرات معالم الفقرة باستخدام برمجية (BILOG)، حيث تم توليد بيانات إلى (45) فقرة تحت ثلاثة ظروف لتوزيع القبلي (التوزيع الطبيعي، توزيع موجب الالتواء وتوزيع سالب الالتواء) وثلاثة ظروف لتوزيع القدرة (التوزيع الطبيعي، توزيع موجب الالتواء وتوزيع سالب الالتواء) وعينتان بحجم (100) و(1000) فرد ومستويين من عدد نقاط التربيع (10) و(20) نقطة. وأظهرت النتائج أن تقديرات معالم الفقرة أكثر دقة عندما يكون التوزيع القبلي للقدرة يتشابه مع توزيع قدرة الذي تم توليد البيانات تحت افتراضاته وكذلك عند حجم العينة الكبير، كما أظهرت تحسن دقة تقديرات معالم الفقرة بزيادة عدد نقاط التربيع.

ومن خلال استعراضنا للدراسات السابقة نجد أنها ركزت على أثر شكل توزيع القدرة أو تباين حجم العينة على دالة معلومات الفقرة والاختبار ودراسة واحدة بحثت أثر عدد نقاط التربيع في تقدير معالم الفقرة – في حدود علم الباحث – في حين لم يتم التطرق إلى أثر عدد نقاط التربيع في دالة معلوماتية الفقرة وفي الثبات النظري والامبريقي. وجاءت هذه الدراسة لسد هذه الحاجة من خلال فحص أثر عدد نقاط التربيع على دالة معلوماتية الفقرة وعلى الثبات النظري والامبريقي تحت افتراضات النموذج اللوجستي ثلاثي المعلمة والذي لم تتناوله دراسة سابقة – في حدود علم الباحث. وكما تأتي أيضا هذه الدراسة استكمالا لتلك الجهود المبذولة في هذا المجال. ومما يجدر الاشارة إليه أن عرض نتائج الدراسات السابقة ليس من أجل تأكيد النتائج أو نفيها أو توافقها مقارنة بنتائج الدراسة الحالية بل من أجل العرض التراكمي لبناء المعرفة حول موضوع دالة معلومات الاختبار والثبات الامبريقي والنظري لأنها لم تتطرق إلى أثر عدد نقاط التربيع في دالة معلوماتية الفقرة وفي الثبات النظري والامبريقي.

منهجية الدراسة وإجراءاتها

تم أستخدم إسلوب المحاكاة بهدف الإجابة على أسئلة الدراسة حيث تم محاكاة معالم الفقرات لاختبار مكون من (50) فقرة تبعا للنموذج اللوجستي ثلاثي المعلمة باستخدام برمجية (Wingen3) إذ تم محاكاة قيم معلمة التمييز للفقرات وفقا للتوزيع اللوغريتم الطبيعي (Lognormal) بوسط حسابي صفر وانحراف معياري مقداره (0.276)، حيث كان الوسط الحسابي والانحراف المعياري لمعلمة التمييز للفقرات المحاكاة (0.998) و(0.267) على التوالي، وكذلك تم محاكاة قيم معلمة الصعوبة للفقرات وفقا للتوزيع الطبيعي ((Normal Distribution (0,1)) حيث كان الوسط حسابي والانحراف المعياري لمعلمة الصعوبة للفقرات المحاكاة (0.054 -) و(0.94) على التوالي، وتم محاكاة

قيم معلمة التخمين للفقرات وفقا لتوزيع بيتا Beta Distribution(a = 10, b=30)، وهذا التوزيع يحاكي قيم لمعلمة التخمين لاختبار اختيار من متعدد بأربعة بدائل، حيث كان الوسط الحسابي والانحراف المعياري لقيم معلمة التخمين للفقرات المحاكاة (0.252) و(0.072)على التوالي، ويبين الملحق (1) قيم معالم الفقرات التي تم توليدها، وتم محاكاة استجابات الأفراد تحت افتراض التوزيع الطبيعي للقدرة بوسط صفر وانحراف معياري واحد وبحجم عينة (1000) فرد، حيث كان الوسط الحسابي والانحراف المعياري لقدرات الافراد المحاكاة (0.029) و(0.994) على التوالي، حيث تم محاكاه استجابات أفراد العينة المولدة بالاعتماد على معالم الفقرات المولدة سابقا، والشكل (1) يبين توزيع قدرة أفراد الدراسة.



شكل 1. توزيع قيم القدرة المحاكاة لأفراد الدراسة

تحليل البيانات

للتحقق من مدى مناسبة الاستجابات التي تمت محاكاتها لافتراضات النموذج اللوجستي ثلاثي المعلمة إذ تم فحص افتراض أحادية البعد (Unidimensionality) للاستجابات المحاكاة باستخدام التحليل العاملي اعتمادا على طريقة المكونات الأساسية (principle component analysis). ومن ثم تدوير العوامل باستخدام طريقة الفاريماكس (varimax) باستخدام برمجية (SPSS)، والجدول (1) يبين نتائج التحليل العاملي.

الجدول (1)

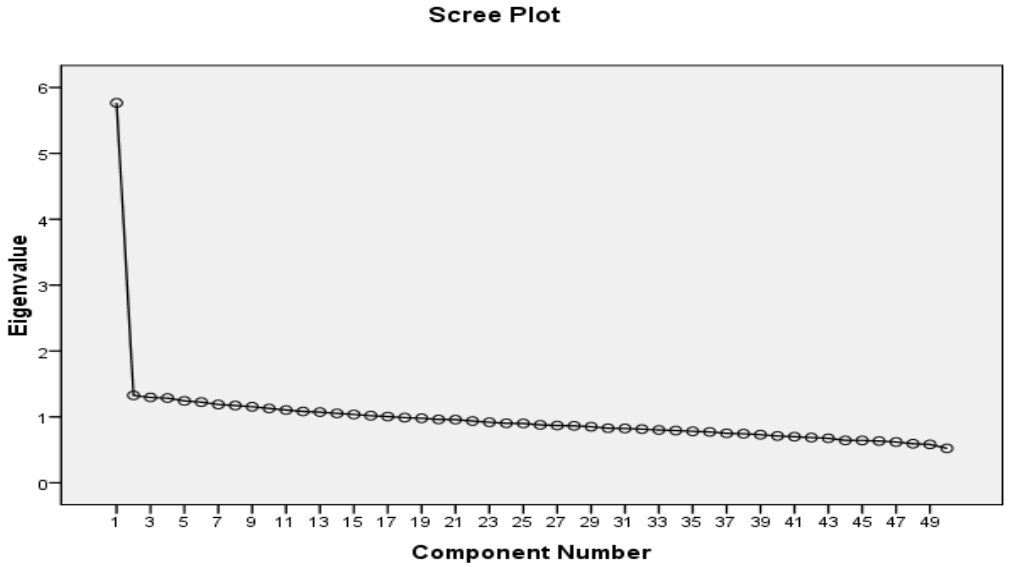
نتائج التحليل العاملي

العامل الأول		العامل الثاني		ناتج القسمة (الجذر الكامن الأول)/ (الجذر الكامن الثاني)
الجذر الكامن	نسبة التباين المفسر	الجذر الكامن	نسبة التباين المفسر	
5.766	11.532	1.327	2.653	4.35

تبين نتائج التحليل العاملي الواردة في الجدول (1) واعتمادا على ما أشار إليه هاتي (Hatti, 1985) تحقق أحادية البعد لاستجابات الأفراد، إذ كانت نسبة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني أكبر من (2)، مما يشير إلى أن الاستجابات المولدة تعكس عاملاً واحداً تقيسه فقرات الاختبار المحاكى، وكما تطابقت نتائج اختبار فرز العوامل بالرسم البياني (Scree plot) بنتائج التحليل العاملي، حيث أشار التحليل البياني للجذور الكامنة بوجود عامل واحد كما يبينها الشكل (2).

وفيما يتعلق بفحص افتراض الاستقلال المحلّي وأفترض 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 إجراءًا من غير ذي معنى. واكتفى الباحث بالتحقق فقط من افتراض أحادية البعد، وأخيرا تم التحقق من جودة مطابقة البيانات للنموذج باستخدام برمجية (BILOG-MG3) وأظهرت نتائج التحليل أن جميع الفقرات كانت متطابقة مع النموذج اللوجستي ثلاثي المعلمة ما عدا ثلاث فقرات هي (33) و(38) و(39) حيث تم رفض الفرضية الصفرية، إذ كانت قيمة مربع كاي ذات دلالة إحصائية، وكذلك أظهرت تطابق

توقعات النموذج مع جميع استجابات الأفراد، باستثناء الأفراد ذوي الأرقام (879) و(421) و(212) و(763) حيث تم رفض الفرضية الصفرية إذ كانت قيمة مربع كاي ذات دلالة إحصائية، وبناءً على هذه النتائج تم حذف الفقرات من إجراءات التحليل اللاحقة والإبقاء على الأفراد غير المطابقين لقلة عددهم مقارنة بالعدد الكلي.



شكل 2. التمثيل البياني للجذور الكامنة

المعالجة الاحصائية

للإجابة عن أسئلة الدراسة تم تحليل البيانات باستخدام برمجية (BILOG-MG3) لتقدير القيمة القصوى لدالة معلوماتية الفقرة، والثبات الامبريقي والنظري للاختبار تبعا لاختلاف عدد نقاط التربيع، حيث إن هذه البرمجية تتيح تقدير القيمة القصوى لدالة معلومات الفقرة عند أعداد مختلفة من نقاط التربيع ضمن أمر يسمى (CALIB NQPT) ويتم تغيير العدد وفق عدد النقاط المراد تحليل الفقرات عندها والملحق (2) يبين أوامر التحليل لتقدير القيمة القصوى لدالة معلوماتية الفقرة والثبات الامبريقي والنظري للاختبار في حالة (10) نقاط تربيع وبعد ذلك تم استخدام إسلوب تحليل التباين الثنائي للمجموعات المعشاة One Way ANOVA Randomize Block Design) واستخدام اختبار أدنى فرق (LSD) (Least Significant Difference) لفحص الفروق في متوسط القيمة القصوى لدالة معلوماتية الفقرة باختلاف عدد قيم التربيع وأخيرا تم استخدام اختبار فيليدت للعينات المستقلة (W) للمقارنات بين قيم معاملات الثبات المحسوبية.

نتائج الدراسة ومناقشتها

السؤال الأول: هل تختلف القيمة القصوى لدالة معلوماتية الفقرة باختلاف عدد قيم التربيع؟

لإجابة هذا السؤال استخدم إسلوب تحليل التباين الأحادي للمجموعات المعشاة للقيم القصوى لدالة معلوماتية الفقرة تبعا لاختلاف عدد قيم التربيع والجدول (2) يبين نتائج التحليل.

الجدول (2)

تحليل التباين الأحادي للمجموعات المعشاة للقيم القصوى لدالة معلوماتية الفقرة تبعا لاختلاف عدد قيم التربيع

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة (ف)	الدلالة الاحصائية
الفقرة	2.131	46	.046	3.785E3	.000
عدد القيم التربيع	0.005	4	0.001	92.180	.000
الخطأ	.002	183	1.224E-5		
المجموع	2.137	233			

يبين الجدول (2) وجود أثر دال لعدد قيم التربيع في متوسط القيم القصوى لدالة معلوماتية الفقرة، حيث كانت قيمة (ف) تساوي (92.180) وهي دالة إحصائيا عند مستوى دلالة ($\alpha \leq 0.05$) ولتحديد إلى أي من مستويات متغير عدد قيم التربيع تعود دلالة الفروق، تم استخراج الوسط الحسابي والخطأ المعياري وفترة الثقة للقيم القصوى لدالة معلوماتية الفقرة تبعا لمستويات متغير عدد قيم التربيع، كما يبين الجدول (3).

الجدول (3)

الوسط الحسابي للقيم القصوى لدالة معلوماتية الفقرة والخطأ المعياري وفترة الثقة

عدد الفقرات	عدد قيم التربيع	المتوسط الحسابي	الخطأ المعياري	فترة الثقة عند مستوى ثقة (95%)	
				الحد الأدنى	الحد الأعلى
47	10	0.157	0.001	0.156	0.158
47	15	0.165	0.001	0.164	0.166
47	25	0.168	0.001	0.167	0.169
47	35	0.169	0.001	0.168	0.170
47	50	0.169	0.001	0.168	0.170

بالرجوع إلى الجدول (3) يتبين أن هنالك فروقا ملاحظة بين الأوساط الهامشية للقيم القصوى لدالة معلومات الفقرة تبعا لمستويات عدد قيم التربيع ولتحديد مواقع الفروق بين المتوسطات تم إجراء المقارنات البعدية باستخدام اختبار أدنى فرق (LSD) (Least Significant Difference) والجدول (4) يبين نتائج التحليل.

الجدول (4)

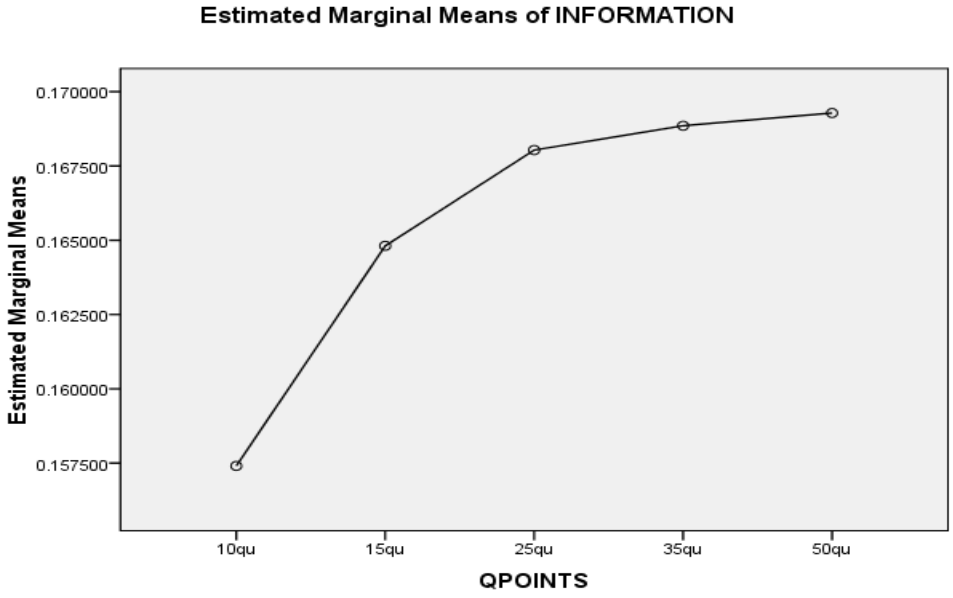
نتائج المقارنات البعدية باستخدام اختبار أدنى فرق (LSD)

عدد قيم التربيع	15	25	35	50
10	0.008*	0.011*	0.012*	0.012*
15	*	0.003*	0.004*	0.004*
25	*	*	0.001	0.001
35	*	*	*	0.00

وبالرجوع للجدول (4) يتبين أن هناك فروقاً دالة إحصائيا عند مستوى دلالة ($\alpha \leq 0.05$) بين متوسطات الحسابية الهامشية للقيم القصوى لدالة معلوماتية الفقرة عند عدد نقاط التربيع (10) من جهة وعدد نقاط التربيع (50,35,25,15) من جهة أخرى ولصالح عدد قيم التربيع الأكثر (50,35,25,15) نقطة. وبينت النتائج أيضا أن هناك فروقاً دالة إحصائيا عند مستوى دلالة ($\alpha \leq 0.05$) بين متوسطات الحسابية للقيم القصوى لدالة معلوماتية الفقرة عند عدد نقاط التربيع (15) من جهة وعدد نقاط التربيع (50,35,25) من جهة أخرى ولصالح عدد قيم التربيع الأكثر (50,35,25) نقطة في حين لم يكن هناك فروق دالة إحصائيا عند مستوى دلالة ($\alpha \leq 0.05$) بين متوسطات الحسابية للقيم القصوى لدالة معلوماتية الفقرة عند عدد نقاط التربيع (25) من جهة وعدد نقاط التربيع (50,35) من جهة أخرى. وكذلك لم يكن هناك فروق دالة إحصائيا عند مستوى دلالة ($\alpha \leq 0.05$) بين متوسطات الحسابية للقيم القصوى لدالة معلوماتية الفقرة عند عدد نقاط التربيع (35) من جهة وعدد نقاط التربيع (50) من جهة أخرى، إذ أشارت النتائج الواردة في الجدول (3) إلى تساوي متوسط القيم القصوى لدالة معلومات الفقرة عندهما وكما يوضح الشكل (3) العلاقة بين الأوساط الهامشية للقيم القصوى لدالة معلوماتية الفقرة وعدد نقاط التربيع، حيث لخص الشكل (3) النتائج التي تم التوصل إليها في الجداول السابقه، اذ ازدادت المتوسطات الهامشية للقيم القصوى لدالة معلوماتية الفقرة بزيادة عدد نقاط التربيع عبر النقاط (25,15,10) حيث كان ميل المنحنى كبيراً جداً في حين قل ميل المنحنى عبر النقاط (25,35,50)، مما يدل على أن الزيادة في قيمة المتوسطات الهامشية عبر هذه النقاط ليس ذا

دلالة وهذا ما تؤيده نتائج التحليل الاحصائي المشار إليها سابقا. وعليه فإن القيم القصوى لدالة معلوماتية الفقرة تزداد بزيادة عدد نقاط التربيع ولغاية (25) نقطة، وبعد ذلك تصبح الزيادة غير دالة احصائيا إذ تثبت الزيادة بعد (35) نقطة تربيع.

ويمكن تفسر ذلك بناء على الصيغة الرياضية لدالة معلوماتية الفقرة، فكلما زادت قيمة البسط الذي يمثل ميل المماس لمنحنى خصائص الفقرة تزداد قيمة دالة معلوماتية الفقرة، وأيضا زيادة عدد نقاط التربيع لمتصل القدرة يزيد من الدقة في تحديد القيمة القصوى للمشتقة الأولى لاقتران منحنى خصائص الفقرة التي يكون عندها أكبر ميل لمماس منحنى خصائص الفقرة، وبالتالي تزداد قيمة البسط والذي بدوره يؤدي إلى زيادة قيمة دالة معلوماتية الفقرة.



شكل 3. العلاقة بين الاوساط الهامشية للقيم القصوى لدالة معلوماتية الفقرة وعدد نقاط التربيع

السؤال الثاني: هل تختلف قيم معامل الثبات الامبريقي والثبات النظري باختلاف عدد قيم التربيع؟

لإجابة السؤال تم تقدير قيم معاملات الثبات الامبريقي والثبات النظري من خلال استخدام برمجية (BILOG-MG3) تبعا لاختلاف عدد قيم التربيع كما يبينها الجدول (5).

الجدول (5)

قيم معاملات الثبات الامبريقي والنظري

نوع معامل الثبات	عدد نقاط التربيع				
	10	15	25	35	50
الامبريقي	0.835	0.843	0.847	0.848	0.848
النظري	0.845	0.850	0.852	0.853	0.853

يتبين من الجدول (5) أن كلا من قيم معاملي الثبات الامبريقي والنظري يزداد بزيادة عدد قيم التربيع ولغاية (35) نقطة تربيع، بحيث تصبح الزيادة بعد هذا العدد من نقاط التربيع ليس له أي أثر ملموس على زيادة قيم معاملات الارتباط، إذ تثبت قيمة معامل الثبات بعد هذا الحد من قيم التربيع وأيضا تبين النتائج أن قيم معاملات الثبات النظرية أعلى من قيم معاملات الثبات الامبريقي كذلك عبر جميع نقاط التربيع ولمعرفة دلالة الفروق بين معاملات الثبات المحسوبة عند كل مستوى من مستويات نقاط التربيع لكل نوع على حده من أنواع الثبات المحسوب وكذلك للمقارنة قيم معاملات نوعي الثبات عند كل نقطة تربيع تم استخدام اختبار فيليدت للعينات المستقلة (W) للمقارنات بين قيم معاملات الثبات المحسوبية

وبدرجات حرية (N2-1)(N1-1) والصيغة الرياضية للاختبار كالآتي:

$$W = \frac{1-\hat{\alpha}_2}{1-\hat{\alpha}_1}$$

α_1: معامل الثبات الأصغر α_2 : معامل الثبات الأكبر. والجدول (6) يبين قيمة الإحصائي (W) للمقارنات بين قيم معاملات الثبات ودلالاتها الإحصائية.

الجدول (6)

قيمة الإحصائي (W) للمقارنات بين قيم معاملات الثبات ودلالاتها الإحصائية

عدد نقاط التربيع	المقارنة بين قيم معاملات الثبات الامبريقي				المقارنة بين قيم معاملات الثبات النظري			نقطة التربيع	المقارنة بين معاملات الثبات الامبريقي والنظري عند كل نقطة تربيع
	15	25	(35,50)		15	25	(35,50)		
10	0.95 (*0.21*)	0.93 (*0.11*)	0.92 (*0.09*)	0.92 (*0.09*)	0.97 (*0.30*)	0.96 (*0.30*)	0.95 (*0.20*)	10_10	0.93(*0.16*)
								15_15	0.95(*0.23*)
15	****	0.97 (*0.34*)	0.96 (*0.30*)	0.96 (*0.30*)	****	0.95 (*0.23*)	0.98 (*0.37*)	25_25	0.96(*0.29*)
								35_35	0.96(*0.29*)
25	****	****	0.99 (*0.45*)	0.99 (*0.45*)	****	****	0.99 (*0.45*)	50_50	0.96(*0.29*)
								**********	

**حيث إن القيمة بين القوسين تمثل الدلالة الإحصائية (p) والقيمة الأخرى تمثل المقدار الاحصائي (w).

تبين النتائج الواردة في الجدول (6) أنه لا يوجد فروق دالة احصائيا للمقارنات المختلفة بين قيم معاملات الثبات. ويمكن تفسير ذلك بأن تأثير عدد نقاط التربيع في تقديرات معامل الثبات النظري والامبريقي يكون غير جوهري عند ظرف الدراسة الحالية وهي الاختبارات الطويلة وحجم العينة الكبير.

التوصيات

تبيَّن من النتائج التي توصلت إليها الدراسة أن هناك أثراً لعدد قيم التربيع قي القيمة القصوى لدالة معلوماتية الفقرة، حيث إن القيم القصوى لدالة معلوماتية الفقرة يزداد بزيادة عدد نقاط التربيع ولغاية (25) نقطة وبعد ذلك تصبح الزيادة غير دالة احصائياً، إذ تثبت الزيادة بعد (35) نقطة تربيع. وكما أشارت النتائج إلى أنه لا يوجد فروق دالة إحصائيا للمقارنات المختلفة بين قيم معاملات الثبات. وتوصي الدراسة باستخدام عدد نقاط تربيع (25) نقطة وتوصي الدراسة الحالية أيضا بإجراء دراسات تتعلق في أثر عدد نقاط التربيع في حالة الاحجام الصغيرة للعينات وكذلك في الاختبارات القصيرة وعند التوزيعات المختلفة للقدرة على دالة معلوماتية الفقرة والاختبار ومعاملات الثبات النظري والامبريقي.

المصادر والمراجع

بني عطا، زايد ونضال الشريفين، 2012، أثر اختلاف شكل توزيع القدرة على معالم الفقرة ودالة المعلومات للاختبار، المجلة الأردنية في العلوم التربوية، مجلد 8 العدد (2)، (155-166).

التقي، أحمد، 2009، النظرية الحديثة في القياس، دار المسيرة، عمان، الأردن.

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

الزبون، حابس 2013، أثر حجم العينة على تقدير دالة المعلومات للاختبار والخطأ المعياري في تقديرها باستخدام النظرية الحديثة في القياس، مجلة جامعة النجاح للأبحاث (العلوم الإنسانية)، المجلد 27 (6).

علام، صلاح الدين، 2005، نماذج الاستجابة للمفردات الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي، الطبعة الأولى، القاهرة، دار الفكر العربي.

Baker, F. B and Kim, S.H. 2004. *Item response theory: parameter estimation techniques*, (2nd ed.) New York: M arcel Dekker, Inc.

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

Du Toit, M. (Ed.). 2003. *IRT from SSI. BILOG-MG, MULTILOG, PARSCALE, TESTFACT*, Lincolnwood, IL: Scientific Software International.

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

Hambleton, R. K. and Jones, R. W. 1994. Item parameter estimation and their influence on test information function, *Applied Measurement in Education*, 7 (3): 171-186.

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

Hambleton, R.K. and Swaminathan, H. and Rogers, H.J. 1991. *Fundamentals of Item Response Theory: International Educational and Professional*, Publisher Newbury park.

Hambleton, R.K. and Swaminathan, H. 1985. *Item Response Theory: Principles and applications*, Boston MA: Kluwer-Nyjhoff.

Han, K. T. and 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 No642, 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.

Seong, T. J. 1990. Sensitivity of marginal maximum likelihood estimation of item and ability parameters to the characteristics of the prior ability distributions, *Applied Psychological Measurement*, 14: 299-311.

The Effect of Quadrature Points Number on the Maximum Value of Item Information Function, Empirical and Theoretical Reliability

*Raji Awad Sarierh**

ABSTRACT

This study aimed at verifying the effect number of quadrature points on the maximum value of item information function, empirical and theoretical reliability under three-parameter logistic assumptions. Moreover, to answer questions of the study was to generate responses according to three parameter logistic model using (WINGEN3) program on (50) dichotomous items. The responses of the Individuals were analyzed by (BILOG– MG 3) programs according to number of quadrature points under the three parameter logistic model assumptions, the results of the One Way ANOVA Randomize Block Design revealed that there were statistical significant differences at the ($\alpha \leq 0.05$) level between means of the maximum value of item Information function. In addition, the results showed that the maximum value of item Information function increases of when the number of quadrature points. Moreover, no statistical significant differences at the ($\alpha \leq 0.05$) level between empirical and theoretical reliability values according difference of the quadrature point's number.

Keywords: Number of Quadrature Points, Item Information Function, Empirical Reliability, Theoretical Reliability, Three-Parameter Logistic.

* Faculty of Educational Sciences, Mutah University, Jordan. Received on 14/1/2014 and Accepted for Publication on 6/4/2014.