أثر التفاعل بين معالم الفقرة ومستوى القدرة على دالة معلومات الفقرة في النموذج اللوجستي ثنائي المعلمة

د. أيمن عمر عبد الغني بني عامر
وحدة القياس والتقويم
كلية الشريعة والدراسات الإسلامية – جامعة القصيم
Aiman_banyamer@yahoo.com

أثر التفاعل بين معالم الفقرة ومستوى القدرة على دالة معلومات الفقرة في النموذج اللوجستي ثنائي المعلمة

د. أيمن عمر عبد الغني بني عامر
وحدة القياس والتقويم
كلية الشريعة والدراسات الإسلامية - جامعة القصيم

الملخص

هدفت هذه الدراسة التعرف إلى أثر التفاعل بين معالم الفقرة، والتفاعل بين معالم الفقرات والقدرة، على دالة معلومات الفقرة في النموذج اللوجستي ثنائي المعلمة. حيث تم استخدام برنامج WINGEN لتوليد استجابات (٦٠٠٠) مفحوص على (٦٠٠) فقرة، وتم استخدام برنامج BILOGM وبرنامج SPSS في التحليل. أظهرت النتائج وجود فروق دالة إحصائيا في متوسطات دالة معلومات الفقرة تعزى للتفاعل بين صعوبة الفقرة وتمييزها، ولوحظ أن دالة معلومات الفقرة تزداد بزيادة تمييز الفقرة وثبات الصعوبة. وقد تحققت أعلى قيمة لدالة معلومات الفقرة عند التفاعل بين المستوى المتوسط للصعوبة والمرتفع للتمييز. كما أظهرت النتائج وجود فروق دالة احصائيا في دالة معلومات الفقرة تعزى للتفاعل بين معالم الفقرة والقدرة، حيث لوحظ أن دالة معلومات الفقرة تزداد كلما كانت القدرة والصعوبة متقاربتين، وارتفع التمييز، وتنخفض كلما تباعدت القدرة عن الصعوبة وارتفع التمييز، وقد تحققت أعلى قيمة لدالة معلومات الفقرة عند تفاعل المستويات المرتفعة للقدرة والصعوبة والتمييز.

الكلمات المفتاحية: الصعوبة ، التمييز، القدرة، دالة المعلومات، نظرية استجابة الفقرة.

* تاريخ تسلم البحث: ٢٠١٧/١٠/٢٩م
* تاريخ قبوله للنشر: ٢٠١٨/٣/١٤م

The Effect of Interaction Between the Item Parameters and Ability Levels on the Item Information Function in the Two-Parameter Logistic model

Dr. Ayman O. Baniamer
Measurement and Evaluation Unit
Qassim University - Kingdom of Saudi Arabia

Abstract

This study aimed to identify the effect of interaction between item parameters, and interaction between item parameters and ability levels, on the Item information function in the two-parameter logistic model. WINGEN was used to generate (6000) responses tested on (600) items. BILOGM and SPSS were used in the analysis. The analyzed results showed that there were statistically significant differences in the item information function mean due to the interaction between difficulty and discrimination. More, the results indicated that the mean of the item information function increased by increasing the item discrimination. The highest value of the item information function mean was achieved when the intermediate level of difficulty and high level of discrimination. Also, the results showed that there were statistically significant differences in the item information function mean due to the interaction between item parameters and ability levels. It was observed that the item information function increases as the levels of ability and difficulty are close, and the level of discrimination increases. The highest value of the item information function was achieved when high levels of ability, difficulty and discrimination interacted.

Keywords: difficulty, discrimination, ability, information function, item response theory.

أثر التفاعل بين معالم الفقرة ومستوى القدرة على دالة معلومات الفقرة في النموذج اللوجستي ثنائي المعلمة

د. أيمن عمر عبد الغني بني عامر
وحدة القياس والتقويم
كلية الشريعة والدراسات الإسلامية – جامعة القصيم

المقدمة:

تعد نظرية استجابة الفقرة (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، وتقوم نظرية استجابة الفقرة على افتراض وجود متصل للسمة، بحيث يمكن تقدير احتمال الإجابة الصحيحة على هذا المتصل (Mislevy & Bock, 1990)، وأن العلاقة بين أداء الفرد على الفقرة وقدرته، يمكن أن تحدد من خلال ما يسمى بمنحنى خصائص الفقرة (Item Characteristic Curve)، كما تفترض أن مقدار الاحتمال يكون دالة متزايدة مطردة (Monotonically Increasing) مما يعني أن احتمال الإجابة الصحيحة يزداد بزيادة قدرة الفرد (Hambleton, 1994) .

وتعتمد نظرية استجابة الفقرة أكثر من نموذج في تصميم الاختبارات التربوية والنفسية وبنائها، وتتفق جميع هذه النماذج على قضية أساسية؛ هي تحديد العلاقة بين أداء المفحوص الملاحظ على الاختبار وبين السمة أو السمات الكامنة المسؤولة عن أداء المفحوص على الاختبار. وتصنف نماذج نظرية استجابة الفقرة بطرق مختلفة منها الاعتماد على استجابات المفحوصين، والتي تصنف على هذا الأساس إلى ثلاثة مستويات للاستجابة: استجابة ثنائية (dichotomous)، واستجابة متعددة (multi-chotomous)، واستجابة متصلة (continuous)، وتختلف هذه النماذج عن بعضها في الصيغة الرياضية لمنحنى خصائص الفقرة (Item Characteristic Curve)، وطرق التصحيح (Scoring Hambleton, Sawminthan & Rogers, 1991; Embretson & Reise, 2000)

ولكل نموذج رياضي مجموعة من الافتراضات حول البيانات التي يطبق عليها، ولأن نماذج نظرية استجابة الفقرة هي كغيرها من النماذج، فلها أيضاً مجموعة من الافتراضات

الواجب تحققها في البيانات التي تطبق عليها، وهذه الافتراضات هي: افتراض أحادية البعد (Unidimensionality)، وافتراض الاستقلال الموضعي (Local item independence)، وافتراض منحنى خصائص الفقرة (Item Characteristic Curve)، وافتراض التحرر من السرعة (Speededness)، وتتلخص هذه الافتراضات بما يلي:

• افتراض أحادية البعد (Unidimentionality) : أي وجود قدرة أو سمة وحيدة كامنة تفسر أداء الفرد على الاختبار، و تسمى عندها بالنماذج أحادية البعد. أما النماذج التي تفترض وجود أكثر من قدرة واحدة تكمن وراء أداء الفرد على الاختبار فتسمى النماذج متعددة الأبعاد.

• افتراض الاستقلال الموضعي (Local Item Independence) (LII): وينص هذا الافتراض على أن استجابات المفحوصين الذين لديهم نفس القدرة لفقرة ما مستقلة إحصائياً، أي أن استجابة المفحوص على فقرة ما يجب أن لا تؤثر في استجابته على أي فقرة أخرى، بمعنى إن إجابة الفقرة لا تعطي تلميحات أو معيقات لإجابة فقرة أخرى. وهذا الافتراض يوضح أن قدرة المفحوص وخصائص الفقرة؛ هي وحدها التي تؤثر في التحصيل أو الأداء على الاختبار.

• افتراض منحنى خصائص الفقرة (Item Characteristic Curve) (ICC): أي أن العلاقة بين تحصيل المفحوص على الفقرة وقدرته المقيسة بالاختبار، هي اقتران تراكمي صاعد، يعرف باسم منحنى خاصية الفقرة (ICC)، حيث يمثّل هذا المنحنى احتمالات إجابة المفحوصين للفقرة إجابة صحيحة في مستويات القدرة المختلفة، وكون المنحنى تراكمياً صاعداً؛ فإنه يشير بوضوح إلى أن احتمال إجابة الفقرة إجابة صحيحة يزداد بازدياد قدرة المفحوص، وفي العادة توصف هذه المنحنيات في نماذج الاختبارات المصممة لقياس سمة واحدة (أحادية البعد) بدلالة معلمة واحدة أو معلمتين أو ثلاث معالم، ويرجع الاختلاف الرئيسي بين نماذج السمات الكامنة المستخدمة إلى اختلاف صورة الدالة الرياضية التي ينتج عنها اختلاف في شكل منحنى الفقرة .

• افتراض التحرر من السرعة (Speededness): تفترض نماذج نظرية استجابة الفقرة أن عامل السرعة لا يلعب دوراً في الإجابة عن الفقرات، أي أن إخفاق المفحوصين في الإجابة عن فقرات الاختبار يرجع إلى انخفاض قدرته وليس إلى تأثير عامل السرعة على أجابته. بمعنى أن اختبارات السرعة لا توفر بيانات عن بعض الفقرات التي لا يصل إليها المفحوص بسبب ضيق الوقت المخصص للإجابة، ومن ثم لا تخضع مثل هذه الفقرات لأي نموذج. (Hambleton, Sawminthan & Rogers, 1991; Embretson & Reise, 2000).

ويلاحظ من الافتراضات السابقة أن هذه النظرية تقوم على افتراضات قوية قد يصعب تحققها، وأن ايجابيات استخدام هذه النظرية تتناسب طردياً مع مدى تحقق هذه الافتراضات،

و يمكن الكشف عن مدى تحققها من خلال الكشف عن مدى مطابقة البيانات للنموذج المعتمد.

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, 2003; Embreston & Rise, 2000).

ولم يقتصر البحث السيكومتري البحث 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 & Rogers, 1991)، وتعتبر فاعلية الاختبارات من الأمور المهمة؛ لما يترتب عليها من اتخاذ قرارات متعلقة بمستقبل الأفراد، فيما يتعلق بتعلمهم وعملهم وإلى غير ذلك من الأمور الأخرى (Thorndike, 1982). في النظرية التقليدية للقياس (Classical Test Theory) كان تباين أخطاء القياس مصطلحاً مهماً، كما كان معكوس تباين أخطاء القياس مؤشراً لدقة القرارات، التي تتخذ بشأن الأفراد، ثم ظهر تطور إيجابي في النظرية الحديثة للقياس (Item Response Theory)، وأكثر فائدة في دقة اتخاذ القرارات من خلال دالة معلومات الفقرة (Item Information Function)، والتي تختبر دقة كل فقرة من فقرات الاختبار عند كل مستوى من مستويات القدرة، ودالة معلومات الاختبار (Test Information Function) التي تحدد كفاءة الاختبار بشكل عام. (Hambelton & Cook, 1977)

واقترح مفهوم "المعلومات" في الإحصاء بداية من قبل فشر (Fischer,1974)، الذي عرّفها على أنها معكوس لمؤشر دقة تقدير معالم الفقرات، فإذا تم تقدير معالم الفقرات بشكل دقيق فسنحصل على مقدار أكبر من المعلومات، مقارنة فيما لو تم تقدير نفس المعالم بدقة أقل. (Baker, 2001).

وتقدم دالة معلومات الفقرة فائدة كبيرة في بناء، وتطوير، وتقييم الاختبارات، ولكن هذه الفائدة تعتمد على مطابقة منحنى خصائص الفقرة لبيانات الاختبار، فإذا كانت مطابقة البيانات لمنحنى خصائص الفقرة ضعيفة، فإن إحصاءات الفقرة ودالة معلومات الفقرة ستكون مضللة، حتى لو كانت مطابقة البيانات لمنحنى خصائص الفقرة جيدة، فيجب أن ننتبه لمعالم الفقرة، فالتمييز المنخفض والتخمين المرتفع للفقرة يقدم أيضاً نتائج غير دقيقة،

وكذلك فإن فائدة فقرات الاختبار تعتمد على الاحتياجات الخاصة لمطوريّ الاختبارات من اختبار ما، فيمكن لفقرة ما أن تقدم معلومات مرتفعة عند إحدى النقاط على إحدى نهايات متصل القدرة، ولكنها لن تقدم أي معلومات في أي نقطة أخرى على متصل القدرة.

وتتمتع دالة معلومات الاختبار (التي تمثل مجموع دوال معلومات الفقرات عند مستوى معين من القدرة) بميزة؛ وهي كون هذه الدالة مستقلة عن عينة المفحوصين، وبذلك تقدم النظرية الحديثة في القياس مميزات إضافية، فيما يتعلق بزيادة القدرة على تقدير أخطاء القياس (Baker, 2001).

ودالة المعلومات للفقرة تعطى من خلال العلاقة:

$$I_i(\theta)=\frac{p'_i(\theta)^2}{p_{i(\theta)}[1-p_{i(\theta)}]}$$

حيث: $Ii(\theta)$ دالة معلومات الفقرة (i)، واحتمال إجابة الفقرة (i) إجابة صحيحة من قبل مفحوص له القدرة θ تم اختياره عشوائياً من بين المفحوصين الذين لهم القدرة نفسها، و(Pi θ) هي مشتقة $\text{Pi}_{(\theta)}$ بالنسبة ل θ .

ومن خلال المعادلة السابقة يتضح أن دالة معلومات الفقرة تساوي مربع ميل منحنى خصائص الفقرة عند θ مقسوماً على تباين الأخطاء.

وتعطى دالة معلومات الفقرة في النموذج اللوجستي ثلاثي المعلمة من خلال العلاقة:

$$(\theta)=a_i^2 p_{i(\theta)}[1-p_{i(\theta)}]\left(1-\frac{c_i}{p_{i(\theta)}}\right)I_i$$

حيث: a_i معلمة تمييز الفقرة (i)، و c_i معلمة تخمين الفقرة (i).

وفي النموذج اللوجستي ثنائي المعلمة (عندما تكون c_i =0) تصبح دالة معلومات الفقرة عبارة عن مربع تمييز الفقرة مضروباً بتباين الفقرة عند θ كما يلي:

$$I_i(\theta)=a_i^2 p_{i(\theta)}[1-p_{i(\theta)}]I_i$$

و تصبح دالة معلومات الفقرة في النموذج اللوجستي أحادي المعلمة (نموذج راش) عبارة عن تباين الفقرة عند θ، كما في المعادلة:

$$(\theta)=p_{i(\theta)}[1-p_{i(\theta)}]I_i$$

وتمتلك دالة معلومات الاختبار والمعروفة ب $I(\theta)$ خاصية سهلة، حيث إنها عبارة عن مجموع دوال معلومات الفقرة عند مستوى محدد للقدرة، وذلك من خلال العلاقة:

ومن الواضح من خلال هذه المعادلة أن فقرات الاختبار تساهم بصورة مستقلة في دالة معلومات الاختبار، وهذا الأمر لم يكن موجوداً في النظرية التقليدية للقياس.

كما تلعب معالم الفقرة من صعوبة، وتمييز، وتخمين دوراً أساسياً في دالة معلومات الفقرة، وقد لخصها ديجرويتر وفاندركامب (De Gruijter & Van Der Kamp, 2003) بما يأتي:

دالة معلومات الفقرة تكون أكبر كلما كانت صعوبة الفقرة قريبة من قدرة الفرد، مقارنة فيما لو كانت صعوبة الفقرة بعيدة عن القدرة.

دالة معلومات الفقرة تكون أكبر عادة كلما كانت معلمة التمييز أكبر.

دالة معلومات الفقرة تزداد كلما اقتربت معلمة التخمين من الصفر.

كما تلعب قدرة الفرد أيضاً دوراً فاعلاً في دالة معلومات الفقرة، فقد أثبت بيرنبام (Birnbaum, 1968) أن الفقرة تقدم أعلى معلومات عند القيمة العظمى للقدرة (θ_{max})، وتعطى من خلال العلاقة:

$$\theta_{max} = b_i + \frac{1}{Db_i} ln\left[0.5(1 + \sqrt{1 + 8c_i})\right]$$

حيث: D: ثابت (Sampling Factor) مقداره ١,٧.

من خلال المعادلة السابقة يتضح أنه عندما تكون معلمة التخمين تساوي صفراً ($ci = 0$) تصبح القيمة العظمى للقدرة عند صعوبة الفقرة ($\theta_{max} = bi$)، وعندما تكون معلمة التخمين أكبر من صفر ($ci > 0$) تقدم الفقرة أعلى معلومات عندما يكون مستوى القدرة أعلى قليلاً من صعوبة الفقرة.

وبسبب التأثير الكبير لمعلمة التخمين على دالة معلومات الفقرة؛ فقد اتجه الباحثون لمطابقة بياناتهم للنموذج اللوجستي أحادي أو ثنائي المعلمة؛ بغرض الحصول على قيمة أعلى لدالة معلومات الفقرة. كما أن كمية المعلومات المقدمة من اختبار معين عند مستوى محدد للقدرة، ترتبط بعلاقة عكسية مع دقة تقدير القدرة عند تلك النقطة، وهي ما يسمى بالخطأ المعياري للتقدير $SE(\theta)$ حيث:

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

وهذه النتيجة تظهر عند تقدير القدرة باستخدام طريقة الأرجحية العظمى (Maximum Likelihood). والخطأ المعياري للتقدير يقوم بدور يشبه دور الخطأ المعياري للقياس في النظرية التقليدية، ولكن الخطأ المعياري للتقدير يتغير بحسب مستوى القدرة، في حين أن

الخطأ المعياري للقياس ثابت.

وقد اهتم الباحثون أيضاً بمقارنة دالة المعلومات لاختبارين أو أكثر يقيسان نفس السمة، لما لهذه المقارنة من أهمية في عملية تطوير الاختبارات.(Lord, 1977)؛ وذلك لتحديد مدى كفاءة وفاعلية الاختبار في قياس القدرة المحددة التي يقيسها، مما يساعد في تقويم تلك الاختبارات وانتقاء الأفضل والأكفأ منها، كما أن دوال معلومات الاختبار تمثل وسيلة قوية فعالة، تساعد مصمم الاختبار في تقويم خواص الاختبار وتعديلها بما يتناسب مع أفضل قياس يمكن أن يحققه الاختبار على متصل القدرة، كما أنها تساعد في تقييم مدى الدقة التي يقيس بها الاختبار عند المستويات المختلفة للسمة.(علام، ٢٠٠٥)

ومقارنة دوال المعلومات لاختبارات مختلفة يتم من خلال حساب الفاعلية النسبية (Relative Efficiency) $(RE(\theta)$ لاختبار ومقارنتها بالاختبار الآخر، ويتم حساب الفاعلية النسبية من خلال العلاقة الآتية:

$$RE(\theta) = \frac{I_A(\theta)}{I_B(\theta)}$$

حيث : دالة معلومات الاختبار A، دالة معلومات الاختبار B.

وقد حظي موضوع دالة المعلومات بالكثير من اهتمام الباحثين، ولكن من خلال مراجعة الأدب النظري لموضوع الدراسة، لم يقف الباحث على دراسة مباشرة لأثر التفاعل بين معالم الفقرة ومستوى القدرة على دالة معلومات الفقرة، حيث اتجهت الكثير من الدراسات لدراسة أثر متغيرات أخرى على دالة المعلومات.

ففي دراسة لهمبلتون وجونز (Hambleton & jones, 1994) والتي هدفت إلى التحقق من الآثار الناتجة عن اختيار الفقرات ذات القدرة التمييزية العالية على دقة دالة المعلومات؛ لأن تقدير معالم تلك الفقرات يتم بنسبة من الخطأ، مما يؤثر على تقدير دالة معلومات الفقرة ودالة معلومات الاختبار، بحيث نحصل على قيم أكبر من قيمها الحقيقية، وقد تّم أخذ متغيرين لهما أهمية في تحديد مقدار التأثير، وهما حجم عينة المفحوصين المستخدمة في معايرة الفقرات، ونسبة بنك الفقرات إلى طول الاختبار، واستخدم الباحثان البيانات المولدة، وقد توصلت الدراسة إلى أن دالة معلومات الاختبار تغالي في التقدير؛ وذلك بسبب الزيادة المفرطة في اختيار الفقرات ذات القدرة التمييزية العالية، كما بينت نتائج الدراسة أن حجم عينة المفحوصين أكثر تأثرا من نسبة بنك الفقرات إلى طول الاختبار على دقة المعلومات للاختبار.

وأجرى هياجنة (٢٠١١) دراسة هدفت للمقارنة بين ثلاث طرق لاختيار الفقرات من حيث تأثيرها على دالة معلومات الاختبار، حيث اعتمدت الطريقة الأولى على قيمة (θ_{max})، والثانية على قيمة ($I(\theta_{max})$

،بينما اعتمدت الطريقة الثالثة على الطريقة العشوائية. وأظهرت دوال الفاعلية النسبية أن تعظيمها لدوال معلومات الاختبار كان الأكثر عندما كان حجم العينة قليلا، وكذلك عندما كانت النسبة بين طول الاختبار إلى حجم بنك الفقرات كبيرة. كما بينت الدراسة أن حجم العينة الصغير يؤثر بشكل أكبر على تعظيم دالة معلومات الاختبار بغض النظر عن طول الاختبار. وكانت أخطاء تقديرات معالم فقرات الاختبارات التي تم تكوينها وفق طريقتي $I(\theta)_{max}$ والعشوائية أقل مقارنة بالطريقة المعتمدة على (θ_{max}).

كما أجرى مقدام زادة، وصالحي، وخودي (Moghadamzadeh, Salehi, Khodaie, 2011) دراسة لمقارنة دالة معلومات الفقرة والاختبار في النموذج الأحادي، والثنائي، والثلاثي المعلمة من نماذج نظرية استجابة الفقرة، حيث تم جمع البيانات من خلال استجابات المرشحين للقبول في جامعات الدولة الإيرانية على اختبار اختيار من متعدد، واختيرت البيانات من مجموعة الفيزياء والرياضيات لعام ٢٠٠٩، والذي تديره المنظمة الوطنية للاختبارات التربوية (NOET)، أظهرت النتائج أن قيم دالة معلومات الفقرة والاختبار في النموذج الثنائي المعلمة أعلى من قيم النموذج الأحادي والثلاثي المعلمة، كما أظهرت النتائج عدم وجود فروق ذات دلالة إحصائية بين قيم دالة معلومات الفقرة والاختبار في النموذج الأحادي المعلمة عن القيم المقابلة لها في النموذج الثلاثي المعلمة.

وفي دراسة بني عطا والشريفين (٢٠١٢) والتي هدفت إلى التحقق من أثر اختلاف شكل توزيع القدرة على معالم الفقرة ودالة المعلومات للاختبار، حيث استخرج الباحثان استجابات أربعة أشكال من توزيعات القدرة وفق النموذج اللوجستي ثلاثي المعلمة على ٦٠ فقرة ثنائية الاستجابة من خلال استخدام برنامج WINGEN. كشفت النتائج عن وجود فروق ذات دلالة إحصائية 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 أنّ دالة المعلومات للاختبار قد تباينت باختلاف شكل توزيع القدرة، إلا أنها جميعاً قد أعطت أعلى قيم لدالة المعلومات ضمن مدى القدرة (٠,٥) و(٠,٦٣) حيث كانت قيم الخطأ المعياري أقل ما يمكن.

وقام الزبون (٢٠١٣) بدراسة هدفت إلى معرفة أثر حجم العينة على تقدير دالة المعلومات للاختبار والخطأ المعياري في تقديرها، باستخدام النظرية الحديثة في القياس، واستخدم الباحث استجابات (٧٥٠٠) طالب وطالبة من الصف الثامن الأساسي على الاختبار الوطني لضبط نوعية التعليم لمادة الرياضيات، موزعين على خمس مجموعات، وقد أظهرت النتائج أن مقدار تقديرات دالة المعلومات يتغير بتغير حجم العينة، إذ يزداد بزيادة حجم العينة، كما أظهرت النتائج أن الخطأ المعياري في تقدير دالة المعلومات يتغير بتغيير حجم العينة؛ إذ يتناقص بزيادة حجم العينة.

وأجرى الشريفين وبني عطا (٢٠١٣) دراسة هدفت إلى تقصي أثر عدد خطوات أسئلة الاختبار، وشكل التوزيع لصعوبة الأسئلة في دقة التقديرات لقدرات الأفراد، ومعالم الصعوبة للأسئلة، ودالة المعلومات للاختبار، وفق نموذج التقدير الجزئي، حيث قام الباحثان بتوليد بيانات خاصة بثلاثة اختبارات مختلفة في عدد خطواتها من خلال برنامج WINGEN، كل منهما مكون من (٣٠) سؤالاً بخطوات صعوبة متساوية وفق ثلاثة أشكال توزيع مختلفة لصعوبة الأسئلة. وأشارت النتائج إلى أن الاختبار ذا الخمس خطوات أعطى أكبر كمية من المعلومات سواء كان شكل التوزيع لصعوبة الخطوات طبيعياً أو ملتوياً نحو اليمين، أو ملتوياً نحو اليسار، يليه الاختبار ذو الخطوات الأربع. وأخيرا جاءت كفاءة الاختبار ذي الخطوات الخمس أعلى من كفاءة الاختبار ذي الخطوات الثلاث؛ بمعنى أن الاختبار ذا الخطوات الخمس كان أكثر دقة لتقدير قدرات الأفراد.

ومن خلال ما ورد في الدراسات السابقة يلاحظ أن معظم الدراسات السابقة اتجهت إلى دراسة أثر عوامل تتعلق ببناء الاختبار على دالة معلومات الفقرة، مثل: طرق اختيار الفقرات، وطرق ترتيب الفقرات، وعدد البدائل، وموقع المموه القوي، وعدد خطوات الأسئلة. وبعضها ركز على النموذج المستخدم في التقدير، أو مقارنة أكثر من نموذج وأثرها على دالة المعلومات. وما تبقى منها ركز على دراسة أثر شكل توزيع القدرة أو شكل توزيع الصعوبة أو حجم العينة، ولذلك جاءت هذه الدراسة لتركز على عاملين مهمين من العوامل المهمة في تقدير دالة معلومات الفقرة، وهما معالم الفقرات، ومستوى القدرة، والتفاعل بينهما.

مشكلة الدراسة:

تلعب دالة معلومات الفقرة دوراً مهماً في تطوير وبناء وتقييم الاختبارات، وذلك لأن دالة المعلومات تعرض إسهام كل فقرة من فقرات الاختبار في تقدير معلمة القدرة عند كل نقطة

على متصل القدرة. وهذه الإسهامات تعتمد إلى حد كبير على معالم الفقرة من صعوبة وتمييز وتخمين، وعلى مستوى قدرة الفرد، وقد أثبت الأدب النظري والدراسات السابقة تأثر دالة المعلومات للفقرة بهذه المعالم، ولكن لم تتناول أي من هذه الدراسات أثر التفاعل بين معالم الفقرة، أو بين معالم الفقرة ومستوى القدرة على دالة معلومات الفقرة، فيما عدا التفاعل بين معلمة الصعوبة ومستوى القدرة، ولذلك كان الغرض من هذه الدراسة الكشف عن أثر التفاعل بين معالم الفقرة على دالة معلومات الفقرة، ودراسة تأثر دالة المعلومات للفقرة بالتفاعل بين معالم الفقرة وقدرة المفحوص في النموذج اللوجستي ثنائي المعلمة.

أهداف الدراسة:

تهدف هذه الدراسة إلى دراسة أثر التفاعل بين معالم الفقرة على دالة معلومات الفقرة، ودراسة أثر التفاعل بين معالم الفقرة وقدرة المفحوص أيضاً على دالة المعلومات الفقرة، وذلك بهدف الوصول إلى التفاعل اللازم للحصول على أعلى قيمة لدالة معلومات الفقرة، لاستخدامه في بناء الاختبارات وتطويرها، أو الوقوف على التفاعل الذي من شأنه أن يقدم قيمة منخفضة لدالة معلومات الفقرة، ومن ثم تجنبه. ويمكن اختصار هدف الدراسة بما يأتي:

- التعرف إلى أثر التفاعل بين معالم الفقرة (الصعوبة والتمييز) على دالة المعلومات للفقرة.
- التعرف إلى أثر التفاعل بين معالم الفقرة (الصعوبة والتمييز) ومستوى القدرة على دالة المعلومات للفقرة.

أسئلة الدراسة:

ستحاول هذه الدراسة الإجابة عن السؤالين الآتيين:

- ما أثر التفاعل بين معالم الفقرة (الصعوبة والتمييز) على دالة المعلومات للفقرة ؟
- ما أثر التفاعل بين معالم الفقرة (الصعوبة والتمييز) ومستوى القدرة على دالة المعلومات للفقرة ؟

أهمية الدراسة:

يوجد لدالة المعلومات استخدامات كثيرة في النظرية الحديثة للقياس، لم تكن متاحة من قبل، فيمكن استخدام دالة المعلومات في الحصول على معلومات حول مساهمة كل فقرة من فقرات الاختبار في تحديد مستوى قدرة المفحوصين، وفي مقارنة وتقييم الاختبارات المختلفة عند مستوى معين من مستويات القدرة، كما يمكن استخدامها أيضاً في بناء اختبارات متوازية، أو اختبارات لأغراض محددة.

ولذلك فمن المهم دراسة العوامل التي تؤثر في هذه الدالة، ومن أهم العوامل المؤثرة في دالة المعلومات معالم الفقرة ومستوى القدرة، والتفاعل بينهما. وقد أشار الأدب النظري – كما ذكرت سابقا- لأثر معالم الفقرة على دالة المعلومات، أو لأثر التفاعل بين مستوى صعوبة الفقرة ومستوى القدرة على دالة المعلومات، ولكن لم يشر الأدب النظري –حسب اطلاع الباحث- بشكل واضح لأثر التفاعل بين الصعوبة والتمييز، أو التفاعل بين الصعوبة والتمييز والقدرة وأثره على دالة المعلومات. ولذلك كان من المهم دراسة هذه التفاعلات وبيان أثرها على دالة معلومات الفقرة.

محددات الدراسة :

اقتصرت الدراسة على البيانات ثنائية الاستجابة، والمولدة من خلال برنامج WINGEN وفق النموذج اللوجستي ثنائي المعلمة. كما اعتمدت الدراسة على قيم معالم الفقرات ومعلمة القدرة، وقيم دالة المعلومات المستخرجة من خلال برنامج BILOG.

التعريفات الإجرائية :

معالم الفقرات (item characteristics): و هي خصائص الفقرات المقدرة من خلال برنامج BILOG باستخدام طريقة التقدير (Maximum Likelihood) للنموذج ثنائي المعلمة وهي:

الصعوبة (Item difficulty): وهي نقطة على متصل القدرة تقابل احتمال $\frac{1+c_i}{2}$ للإجابة على الفقرة i إجابة صحيحة، حيث c_i هي معلمة التخمين. و يرمز لها بالرمز b_i، ولها ثلاثة مستويات : صعوبة متدنية ، وصعوبة متوسطة $1 > b_i \geq -1$، وصعوبة مرتفعة $b_i \geq -1$.

التمييز (Item discrimination): وهي نسبة من ميل منحنى خصائص الفقرة الذي يقابل النقطة التي تكون فيها القدرة تساوي صعوبة الفقرة، و يرمز a_i لها بالرمز و لها ثلاثة مستويات: تمييز منخفض $a_i < 1$، وتمييز متوسط $2 > a_i \geq 1$، وتمييز مرتفع $a_i > 2$.

مستوى القدرة (Ability Level): وتعرف القدرة على أنها السمات الكامنة التي يمكن من خلالها التنبؤ بأداء مفحوص على فقرة ما، و يرمز لها بالرمز θ، وتم تقديرها باستخدام برنامج (BILOG) من خلال طريقة التقدير (Maximum Likelihood) وكان توزيعها

طبيعياً. وتم تقسيم القدرة إلى ثلاثة مستويات: قدرة منخفضة $\theta \leq -1$ ، وقدرة متوسطة $-1 < \theta < 1$ ، وقدرة مرتفعة $\theta \geq 1$.

النموذج اللوجستي ثنائي المعلمة (Two parameter logistic model 2PL): ويفترض هذا النموذج وجود معلمتين تمثل خصائص الفقرة وهما الصعوبة والتمييز، وبهذا فإن الفقرات تختلف في صعوبتها، وأيضاً تختلف في درجة تمييزها بين المفحوصين. ومنحنى خصائص الفقرة لهذا النموذج يعطى بالعلاقة:

$$Pi(\theta), = \frac{e^{Dai(\theta-bi)}}{1+e^{eDai(\theta-bi)}} \quad i=1, 2, \ldots, n$$

دالة معلومات الفقرة (item information function): وهو دالة رياضية تعبر عن كمية المعلومات التي تقدمها الفقرة، والمتمثلة في تمييز الفقرة بين مستويات القدرة المختلفة للمفحوصين، وفي هذه الدراسة تم الحصول على دالة معلومات كل فقرة من فقرات الاختبار وعند مستويات القدرة المختلفة من خلال برنامج (BILOG) (حيث يعطي البرنامج دالة معلومات كل فقرة عند ٦١ مستوى للقدرة مقسمة بين (٣- و ٣).

الطريقة والإجراءات:

توليد البيانات:

جرى توليد البيانات باستخدام برنامج (WINGEN) وهو من تطوير هان وهمبيلتون (Han & Hambleton, 2007). حيث يمكن من خلال هذا البرنامج توليد استجابات لاختبارات أحادية البعد سواء أكانت ثنائية التدريج أو متعددة التدريج، بالإضافة إلى اختبارات متعددة الأبعاد.

وقد تم توليد استجابات (٦٠٠٠) مفحوص على اختبار مكون من (٦٠٠) فقرة ضمن النموذج اللوجستي ثنائي المعلمة، وقد تم اختيار هذا العدد من المفحوصين والفقرات للحصول على عدد كاف من الفقرات في كل خلية تفاعل بين معالم الفقرات، أو بين معالم الفقرات ومستوى القدرة، للوصول إلى دقة أكبر في التقديرات كما اقترح لورد (Lord, 1980) بأن يكون طول الاختبار ٥٠ فقرة وعدد الأفراد ١٠٠٠ للحصول على أفضل التقديرات. وقد تم توليد البيانات بحيث يكون توزيع قدرات المفحوصين وفق التوزيع الطبيعي بمتوسط حسابي (٠)، وانحراف معياي (١). كما تم اختيار التوزيع المنتظم لمعلمتي الصعوبة والتمييز، بحيث يكون مدى الصعوبة بين (٣-) و(٣)، ومدى التمييز بين (٠) و (٣).

تحليل البيانات:

استخدم الباحث برنامج BILOG-MG3 و برنامج SPSS لتحليل البيانات المولدة كاملة، بداية تم التحقق من افتراضات النموذج الثنائي المعلمة، حيث تم التحقق من افتراض أحادية البعد لما له من أهمية في دقة التقديرات وذلك من خلال التحليل العاملي باستخدام طريقة المكونات الأساسية (Principle component analysis)، ومن ثم تدوير العوامل باستخدام طريقة (Varimax). حيث كانت نسبة التباين المفسر للعامل الأول (١٥,٢٨٤)، ونسبة التباين المفسر من العامل الثاني (٣,٦٦٧)، وأن ناتج قسمة الجذر الكامن للعامل الأول على العامل الثاني تساوي (٤,١٦٨) وهي أكبر من (٢)، مما يدل على وجود سمة واحدة سائدة للاختبار، وهذا يعني تحقق افتراض أحادية البعد. (Hatti, 1984; Reckase, 1997).

وبما أن افتراض أحادية البعد متحقق في البيانات، فهذا يعني تحقق افتراض الاستقلال الموضعي الذي يعد مؤشراً لأحادية البعد دي. (Emberston & Reise, 2000; Hambleton & Swaminthan, 1985).

بعد ذلك تم التحقق من مطابقة الأفراد والفقرات للنموذج اللوجستي ثنائي المعلمة، حيث تم تحليل البيانات المولدة باستخدام برنامج BILOG-MG3، وأظهرت نتائج التحليل أن جميع الفقرات جاءت مطابقة للنموذج، وكذلك جاءت جميع استجابات الأفراد.

المعالجة الإحصائية:

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

إيجاد معلمتي الصعوبة، والتمييز باستخدام النموذج اللوجستي ثنائي المعلمة باستخدام برنامج BILOG-MG3.

إيجاد قيم دالة المعلومات لكل فقرة من فقرات الاختبار باستخدام برنامج BILOG-MG3، حيث يعطي البرنامج قيم معلومات الفقرة عبر (٦١) قيمة على متصل القدرة، مقسمة على مدى من٣- إلى ٣.

تحليل التباين الثنائي (Two Way ANOVA) لدراسة أثر التفاعل بين معلمتي الصعوبة والتمييز على دالة معلومات الفقرة، حيث تم توزيع الفقرات حسب قيم الصعوبة (صعوبة منخفضة، صعوبة متوسطة، صعوبة مرتفعة) والتمييز (تمييز منخفض، تمييز متوسط، تمييز مرتفع) على تسع خلايا للتفاعل (ثلاثة مستويات للصعوبة تتقاطع مع ثلاثة مستويات للتمييز) وكما هو موضح في الجدول الآتي:

جدول (١)
توزيع الفقرات على مستويات الصعوبة والتمييز

		صعوبة منخفضة	صعوبة متوسطة	صعوبة مرتفعة
تمييز منخفض	المتوسط الحسابي للصعوبة	-١,٩٨	٠,٠٦	٢,٠٥
	المتوسط الحسابي للتمييز	٠,٤٩	٠,٥٠	٠,٥٨
	عدد الفقرات	٧٦	٨٩	٥٠
تمييز متوسط	المتوسط الحسابي للصعوبة	-٢,٠٧	-٠,٠٤	١,٩٠
	المتوسط الحسابي للتمييز	١,٤٤	١,٥٢	١,٤٦
	عدد الفقرات	٧٤	٦٦	٥٦
تمييز مرتفع	المتوسط الحسابي للصعوبة	-٢,٠٥	٠,٠١	٢,٠٥
	المتوسط الحسابي للتمييز	٢,٥٤	٢,٤٠	٢,٦٥
	عدد الفقرات	٦٢	٦٢	٧٠

تحليل التباين الثلاثي (Three Way ANOVA) وذلك لدراسة أثر التفاعل بين معالم الفقرات (الصعوبة والتمييز) مع مستوى القدرة على دالة معلومات الفقرة، حيث تم تقسيم القدرة إلى ثلاثة مستويات: قدرة منخفضة، قدرة متوسطة، قدرة مرتفعة ، بالإضافة لما تم في الخطوة السابقة من تقسيم لمعلمتي الصعوبة والتمييز، وتم عمل تقاطع بين مستويات القدرة الثلاث مع مستويات الصعوبة والتمييز لينتج (٢٧) خلية تفاعل.

النتائج ومناقشتها:

أولا: عرض نتائج السؤال الأول ومناقشتها

للإجابة عن السؤال الأول: "ما أثر التفاعل بين معالم الفقرة (الصعوبة والتمييز) على دالة المعلومات للفقرة ؟" تم إيجاد معلمتي الصعوبة والتمييز لكل فقرة من فقرات الاختبار في النموذج اللوجستي ثنائي المعلمة، كما تم استخراج قيم دالة المعلومات لكل فقرة من خلال برنامج BILOG-MG3، ومن ثم تم نقل هذه القيم (الصعوبة، والتمييز، ودالة المعلومات) لبرنامج SPSS لدراسة أثر التفاعل من خلال تحليل التباين الثنائي Two Way ANOVA، والجدول (٢) يبين المتوسط الحسابي والانحراف المعياري لدالة معلومات الفقرة عبر خلايا التفاعل المختلفة للصعوبة والتمييز.

جدول (٢)
المتوسط الحسابي والانحراف المعياري لدالة معلومات الفقرة عبر مستويات التفاعل المختلفة بين الصعوبة والتمييز

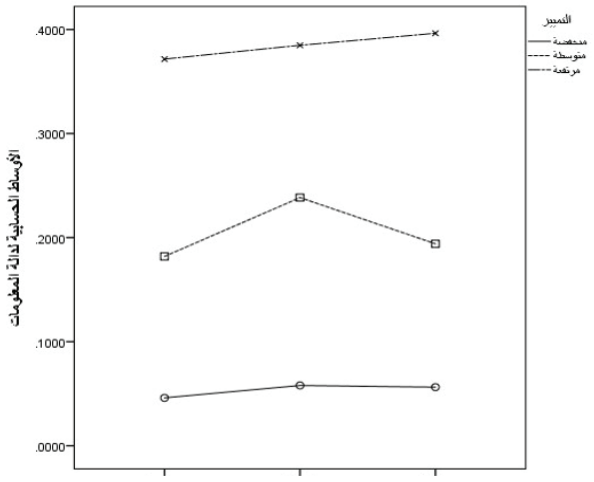
الصعوبة	التمييز	المتوسط الحسابي	الانحراف المعياري
منخفضة	منخفضة	٠,٠٤٦	٠,٠٣٥٢٤
	متوسطة	٠,١٨١٣	٠,٠٥١٧٤
	مرتفعة	٠,٣٦٧٩	٠,٠٦٥٩٢
متوسطة	منخفضة	٠,٠٥٨٢	٠,٠٤٦١٨
	متوسطة	٠,٢٤١٨	٠,٠٥٣٣٤
	مرتفعة	٠,٣٩٢٨	٠,٠٤٢٠٥
مرتفعة	منخفضة	٠,٠٥٦٣	٠,٠٣٥٧٥
	متوسطة	٠,١٩٣٤	٠,٠٥٥٢٧
	مرتفعة	٠,٣٩٢٣	٠,٠٩٧٤٤

كما يُظهر الجدول (٣) نتائج تحليل التباين الثنائي لأثر الصعوبة والتمييز والتفاعل بينهما على متوسط دالة معلومات الفقرة. يتضح من النتائج المبينة في الجدول (٣) وجود فروق ذات دلالة إحصائية عند مستوى الدلالة الإحصائية في الأوساط الحسابية لقيم دالة معلومات الفقرة تبعا لصعوبة وتمييز الفقرات، كما تظهر النتائج تأثر الأوساط الحسابية لقيم دالة معلومات الفقرة بالتفاعل بين معلمتي الصعوبة والتمييز عند مستوى الدلالة الاحصائية .

جدول (٣)
نتائج تحليل التباين الثنائي لأثر الصعوبة والتمييز على دالة معلومات الفقرة

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة F	الدلالة الاحصائية
الصعوبة	٠,١١١	٢	٠,٠٥٥	١٧,١٦١	٠
التمييز	١٠,٧٦٦	٢	٥,٣٨٣	١,٦٦٦,٣٠	٠
التفاعل بين الصعوبة والتمييز	٠,٠٦٥	٤	٠,٠١٦	٥,٠٠٦	٠,٠٠١
الخطأ	١,٩٠٩	٥٩١	٠,٠٠٣		
الكلي المصحح	١٣,١٥٥	٥٩٩			

كما يبين الشكل (١) التغير في متوسطات دالة معلومات الفقرة الناتج من التفاعل بين مستويات الصعوبة والتمييز.



شكل (١)
تأثير التفاعل بين الصعوبة والتمييز على دالة معلومات الفقرة

ويلاحظ من خلال الجدول (٢) أن متوسطات دالة معلومات الفقرة تأثرت تبعا للتغير في مستوى صعوبة الفقرات، حيث إنها تزداد بازدياد مستوى الصعوبة، ثم تعود للانخفاض عند مستويات الصعوبة المرتفعة، ويظهر ذلك بشكل واضح من خلال الشكل (١)، حيث إنه عند كل مستوى من مستويات التمييز، تبدأ متوسطات دالة معلومات الفقرة بالارتفاع تبعاً لزيادة مستوى الصعوبة، ثم تعود للانخفاض عند مستوى الصعوبة المرتفع، وأوضح ما يكون التغير في متوسطات دالة معلومات الفقرة عبر مستويات الصعوبة المختلفة عند المستوى المتوسط للتمييز، حيث كانت متوسطات دالة معلومات الفقرة عبر مستويات الصعوبة الثلاث (٠,١٨١٣، ٠,٢٤١٨، ٠,١٩٣٤) على الترتيب، وقد يعود السبب في ذلك إلى أن مستويات الصعوبة المنخفضة والمرتفعة تكون بعيدة عن مستوى القدرة الذي يتوزع طبيعيا عند كل مستوى من مستويات القدرة، أما في حالة الصعوبة المتوسطة فتكون قريبة من متوسط القدرة وبذلك تزداد قيمة معلومات الفقرة، وهذا يتوافق مع ما ذكره همبلتون وسواميناثن (Hambelton & Swaminathan,1985)، وما توصل إليه همبلتون وجونز (Hambleton & jones, 1994) من حيث إن قيم دالة معلومات الفقرة تكون مرتفعة كلما كان مستوى الصعوبة قريباً من مستوى القدرة، وكذلك تأثرت متوسطات دالة معلومات الفقرة تبعا للتغير في مستوى تمييز الفقرة، فلوحظ أن متوسطات دالة معلومات الفقرة تزداد بازدياد مستوى التمييز، وهذا بسبب العلاقة الطردية بين تمييز الفقرة ودالة معلومات الفقرة، والواضح في معادلة دالة المعلومات للفقرة في النموذج اللوجستي ثنائي المعلمة، وهذا يتفق مع ما ورد في الأدب النظري، والذي

أشار إلى أن زيادة مستوى تمييز الفقرة يؤدي إلى زيادة في قيم دالة معلومات الفقرة، ولكن يجب أن ننتبه إلى ما أشار إليه أيضاً همبلتون وجونز (Hambleton & jones, 1994) من حيث إن زيادة مستوى التمييز يؤدي إلى زيادة في دالة معلومات الفقرة، كما قد يؤدي إلى زيادة مبالغ فيها لدالة المعلومات. كما تأثرت متوسطات دالة معلومات الفقرة بالتفاعل بين مستويات الصعوبة والتمييز، فلوحظ أن قيم المتوسطات كانت ترتفع بارتفاع مستوى تمييز الفقرة وثبات مستوى الصعوبة، في حين أن هذا التأثر كان بسيطاً عند ارتفاع مستوى الصعوبة وثبات مستوى التمييز. وقد تحققت أعلى قيمة لمتوسطات دالة معلومات الفقرة عند مستوى الصعوبة المتوسط والتمييز المرتفع للفقرة حيث كان متوسط دالة معلومات الفقرة (٠,٣٩٢٨)، وقد كانت أيضاً هذه المتوسطات قريبة لقيمتها عند المستوى المرتفع لمعلمتي الصعوبة والتمييز (٠,٣٩٢٣). ويعود السبب في ذلك إلى أن مستوى الصعوبة المتوسطة يكون أقرب ما يكون لمستوى القدرة الذي يتوزع طبيعياً، ومن ثم فإن زيادة مستوى التمييز في هذه الحالة سيعطي أعلى قيمة لدالة المعلومات.

ثانيا: عرض نتائج السؤال الثاني ومناقشتها

وللإجابة عن السؤال الثاني: " ما أثر التفاعل بين معالم الفقرة (الصعوبة والتمييز) ومستوى القدرة على دالة المعلومات للفقرة ؟"

تم إيجاد معلمتي الصعوبة والتمييز لكل فقرة من فقرات الاختبار في النموذج اللوجستي ثنائي المعلمة، كما تم استخراج قيم دالة المعلومات لكل فقرة من خلال برنامج BILOG-MG3، وكانت (٦١) قيمة موزعة على متصل القدرة بين (٣-) و(٣)، ومن ثم تم نقل هذه القيم (الصعوبة، والتمييز، والقدرة، ودالة المعلومات) لبرنامج SPSS لدراسة أثر التفاعل من خلال تحليل التباين الثلاثي Three Way ANOVA، والجدول (٤) يبين المتوسط الحسابي والانحراف المعياري لدالة معلومات الفقرة عبر خلايا التفاعل المختلفة للصعوبة والتمييز والقدرة.

جدول(٤)
المتوسط الحسابي والانحراف المعياري لدالة معلومات الفقرة عبر خلايا التفاعل المختلفة للصعوبة والتمييز والقدرة

مستوى القدرة	مستوى الصعوبة	مستوى التمييز	المتوسط الحسابي	الانحراف المعياري
منخفضة	منخفضة	منخفض	٠,٠٧٢	٠,٠٦٥
		متوسط	٠,٣٩٥	٠,١٧٨
		مرتفع	٠,٩٦٤	٠,٥٨٢
	متوسطة	منخفض	٠,٠٤٦	٠,٠٣٨
		متوسط	٠,١٤٣	٠,١٤٠
		مرتفع	٠,١٣٦	٠,٢٣٠
	مرتفعة	منخفض	٠,٠٢١	٠,٠١٣
		متوسط	٠,٠١٤	٠,٠١٦
		مرتفع	٠,٠٢١	٠,٢٣٥
متوسطة	منخفضة	منخفض	٠,٠٤٧	٠,٠٤١
		متوسط	٠,١٣٩	٠,١٤٠
		مرتفع	٠,١٤٦	٠,٢٧٠
	متوسطة	منخفض	٠,٠٧٥	٠,٠٦٦
		متوسط	٠,٤٤١	٠,٢١٠
		مرتفع	٠,٨٧٣	٠,٤٧٨
	مرتفعة	منخفض	٠,٠٥٧	٠,٠٤٣
		متوسط	٠,١٥٦	٠,١٣٦
		مرتفع	٠,١٤٥	٠,٢٧٩
مرتفعة	منخفضة	منخفض	٠,٠١٩	٠,٠١٣
		متوسط	٠,٠١١	٠,٠١٣
		مرتفع	٠,٠٠٤	٠,٠٢٧
	متوسطة	منخفض	٠,٠٥٣	٠,٠٤٨
		متوسط	٠,١٣١	٠,١٣٦
		مرتفع	٠,١٤٥	٠,٢٥٠
	مرتفعة	منخفض	٠,٠٩٢	٠,٠٦٧
		متوسط	٠,٤١٢	٠,٢٠٣
		مرتفع	١,٠٢٢	٠,٦٥٤

كما يُظهر الجدول (٥) نتائج تحليل التباين الثلاثي لأثر الصعوبة والتمييز والقدرة والتفاعل بينهم على متوسط دالة معلومات الفقرة.

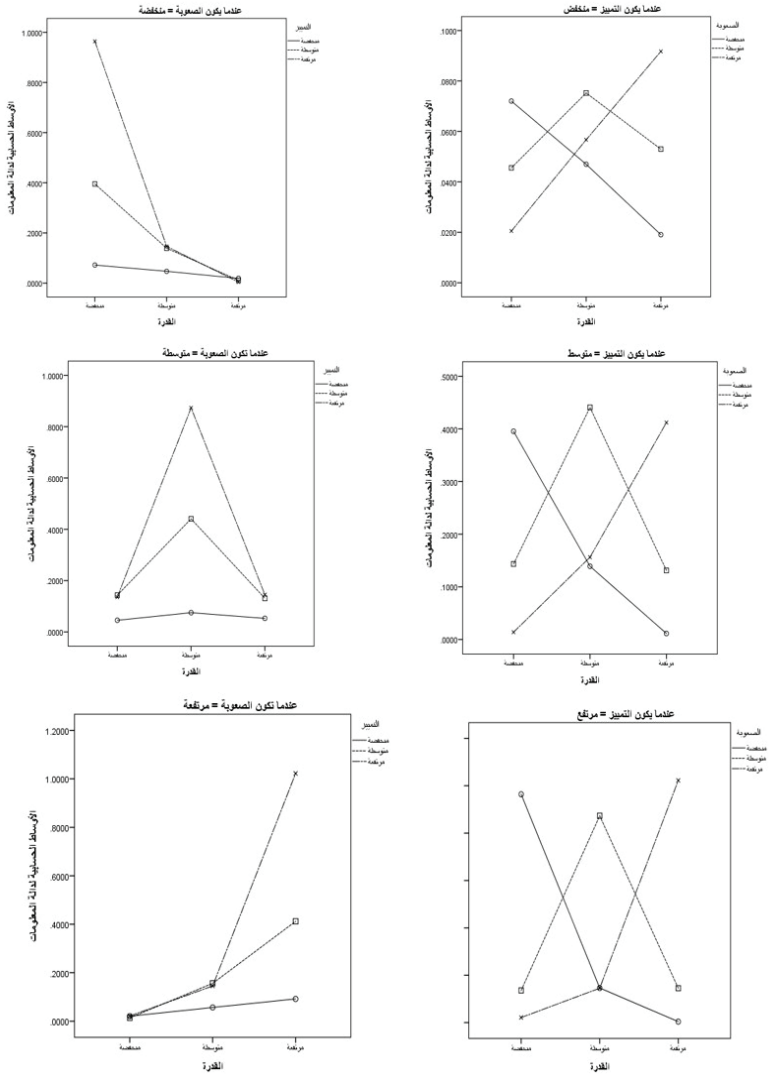
جدول (٥)
نتائج تحليل التباين الثلاثي لأثر الصعوبة والتمييز والقدرة على دالة معلومات الفقرة

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة F	الدلالة الاحصائية
القدرة	٥,٦٥٩	٢	٢,٨٣٠	٥٠,٩٢٩	٠,٠٠٠
الصعوبة	٤,٧٦١	٢	٢,٣٨٠	٤٢,٨٤٣	٠,٠٠٠
التمييز	٦٥٦,٦٢٢	٢	٣٢٨,٣١١	٥٩٠٩,١١٠	٠,٠٠٠
التفاعل بين القدرة والصعوبة	١٣١٨,٦٩٠	٤	٣٢٩,٦٧٣	٥٩٣٣,٦١٧	٠,٠٠٠
التفاعل بين القدرة والتمييز	٥,٣٠٣	٤	١,٣٢٦	٢٣,٨٦١	٠,٠٠٠
التفاعل بين الصعوبة والتمييز	٤,٢١٥	٤	١,٠٥٤	١٨,٩٦٦	٠,٠٠٠
التفاعل بين القدرة والصعوبة والتمييز	٩١٦,٦٩٦	٨	١١٤,٥٨٧	٢٠٦٢,٣٩٧	٠,٠٠٠
الخطأ	٢٠٣٢,٠٠٠	٣٦٥٧٣	٠,٠٥٦		
الكلي المصحح	٤٩٧٧,٣٢٨	٣٦٥٩٩			

تُظهر نتائج تحليل التباين الثلاثي الموضحة في الجدول (٥) وجود فروق ذات دلالة إحصائية عند مستوى الدلالة الإحصائية في الأوساط الحسابية لقيم دالة معلومات الفقرة تبعا لصعوبة وتمييز الفقرات، وكذلك تبعا لمستوى القدرة، كما تظهر النتائج تأثر الأوساط الحسابية لقيم دالة معلومات الفقرة بجميع حالات التفاعل بين معالم الصعوبة والتمييز والقدرة عند مستوى الدلالة الاحصائية .

ويتضح من خلال الجدول (٤) والشكل (٢) أن متوسطات دالة معلومات الفقرة تأثرت تبعا للتغير في مستوى صعوبة الفقرات ومستوى تمييز الفقرات، وهذا التغير في متوسطات دالة معلومات الفقرة مشابه لما نتج من السؤال الأول، كما تأثرت متوسطات دالة معلومات الفقرة تبعا للتغير في مستوى معلمة القدرة، فمن الملاحظ أن قيم دالة المعلومات تنخفض بزيادة مستوى القدرة عند المستوى المنخفض للصعوبة والتمييز، حيث كانت قيم متوسطات دالة المعلومات عبر مستويات القدرة الثلاث (٠,٠٧٢، ٠,٠٤٧، ٠,٠١٩) على الترتيب، كما يلاحظ أن متوسطات دالة المعلومات تزداد بزيادة مستوى القدرة عند المستوى المرتفع للصعوبة والتمييز، حيث كانت قيم متوسطات دالة المعلومات عبر مستويات القدرة الثلاث (٠,٠٢١، ٠,١٤٥، ١,٠٢٢) على الترتيب، أما في حالة المستوى المتوسط للصعوبة والتمييز، فقد ارتفعت متوسطات دالة المعلومات بارتفاع مستوى القدرة ثم عادت للانخفاض عند المستوى المرتفع للقدرة، ويعود السبب في ذلك إلى أن متوسطات دالة معلومات الفقرة تزداد كلما كان مستوى

القدرة ومستوى الصعوبة متقاربين وارتفع مستوى تمييز الفقرة، لتصل إلى أعلى قيمة لها (١,٠٢٢) عند المستوى المرتفع للقدرة والصعوبة والتمييز. في حين أن متوسطات دالة معلومات الفقرة تنخفض كلما تباعد مستوى القدرة عن مستوى الصعوبة وارتفع مستوى تمييز الفقرة؛ لتكون في أقل قيمة لها (٠,٠٠٤) عند المستوى المرتفع للقدرة والمنخفض للصعوبة والتمييز المرتفع، أما في حالة الصعوبة المرتفعة والقدرة المنخفضة، فإن ارتفاع مستوى التمييز أدى إلى انخفاض دالة معلومات الفقرة ثم عودتها للارتفاع عند المستوى المرتفع للتمييز، حيث كانت قيم دالة المعلومات (٠,٠٢١، ٠,٠١٤، ٠,٠٢١). ويعود السبب في ذلك إلى أن القيمة العظمى لدالة معلومات الفقرة تتحقق كلما اقترب مستوى الصعوبة من مستوى القدرة وارتفع مستوى التمييز، أما في حالة تباعد مستوى الصعوبة عن مستوى القدرة فهذا سيزيد قيمة الخطأ المعياري وستقدم الفقرة معلومات غير كافية عند مستويات القدرة هذه. لأن المعلومات هي دالة للقدرة، وهذا يعني أن الفقرة قد تكون غنية بالمعلومات عند مستويات القدرة القريبة من مستوى صعوبتها، وغير غنية بالمعلومات عند المستويات الأخرى للقدرة. مثلاً: الفقرات مرتفعة الصعوبة قد تقدم معلومات غنية للمستويات المرتفعة من القدرة، ولكنها لن تقدم معلومات كافية عن المستويات المنخفضة من القدرة، لأن إجاباتهم ستكون غالباً خاطئة، ومن ثم فإن احتمال إجابة الفقرة إجابة صحيحة سيكون (٠) تقريبا، وفي حالة الفقرات ذات الصعوبة المنخفضة فإن هذه الفقرات لن تقدم معلومات غنية عن المستويات المرتفعة من القدرة، لأن إجاباتهم عن هذه الفقرات ستكون صحيحة، ومن ثم فإن احتمال إجابة الفقرة إجابة صحيحة سيكون (١) تقريبا، وفي الحالتين فإن قيم دالة معلومات الفقرة ستساوي (٠) تقريباً وفق دالة معلومات الفقرة في النموذج اللوجستي ثنائي المعلمة. وهذا يتوافق مع ما أورده همبلتون وسواميناثن وروجرز (Hambelton, Swaminathan & Rogers, 1991) ومع ما أوردته امبرستون وريز (Embretson & Reise, 2000)، ومع ما توصل إليه همبلتون وجونز (Hambleton & jones, 1994) من حيث إن قيم دالة معلومات الفقرة تكون مرتفعة عندما يكون مستوى الصعوبة قريب من مستوى القدرة، وارتفع مستوى تمييز الفقرة. كما يتفق مع ما توصل إليه بيرنبام (Birnbaum, 1968) بالنسبة لمعلمة القدرة، حيث إن الفقرة تقدم أعلى معلومات عند القيمة العظمى للقدرة (θ_{max}).



شكل (٢)
تأثير التفاعل بين القدرة والصعوبة والتمييز على دالة معلومات الفقرة

وبناء على ما تقدم من نتائج تثبت دور معالم الفقرات ومستوى القدرة في تقدير دالة معلومات الفقرة، فيجب على القائمين على بناء وتطوير الاختبارات من مشرفين ومعلمين وغيرهم، أن يختاروا فقرات الاختبار باهتمام شديد، بهدف الحصول على أكبر قدر من المعلومات حول المتقدمين للاختبار، فاختيار فقرات ذات مستوى صعوبة يتناسب مع مستوى قدرات المفحوصين من شأنه أن يقدم معلومات غنية عن هؤلاء المفحوصين، فكلما كان مستوى

صعوبة الفقرات قريب من مستوى قدرات المفحوصين استطعنا الحصول على قيمة مرتفعة لدالة المعلومات بخاصة عند استخدام فقرات ذات تمييز مرتفع.

وفي حال لم تتوفر معلومات عن مستوى قدرات المفحوصين فيجب أن نختار فقرات ذات صعوبة متوسطة وتمييز مرتفع، لأنه على الأغلب ستتوزع قدرات المفحوصين توزيعاً طبيعياً. فاختيار فقرات متوسطة الصعوبة ومرتفعة التمييز من شأنه أن يعطي قيماً مرتفعة لدالة المعلومات. ويجب أن ننتبه لعدم المبالغة في مستوى تمييز الفقرات؛ لأن ذلك سيؤدي لقيم مبالغ فيها أيضاً لدالة المعلومات، وسنكون معلومات مضللة حول المفحوصين.

التوصيات:

في ضوء ما توصلت إليه الدراسة من نتائج يوصي الباحث عند بناء أو تطوير اختبارات، بضرورة الاعتماد على فقرات ذات مستوى مناسب من الصعوبة لمستوى قدرات المفحوصين، للحصول على قيم مرتفعة لدالة معلومات الفقرة، فاستخدام فقرات مستوى صعوبتها لا يتناسب مع مستوى قدرات المفحوصين يؤدي إلى انخفاض دالة المعلومات، كما أن استخدام فقرات ذات تمييز مرتفع يؤدي إلى رفع قيم دالة المعلومات، عندما تكون صعوبة الفقرات منسجمة مع قدرات المفحوصين، ولكن استخدام فقرات ذات تمييز مرتفع عندما لا تكون صعوبة الفقرات منسجمة مع قدرات المفحوصين يؤدي إلى انخفاض دالة المعلومات، كما يوصي بعدم المبالغة باستخدام الفقرات ذات التمييز المرتفع في بناء الاختبارات؛ لأن ذلك قد يعطي نتائج عكسية في قيم دالة معلومات الفقرة، بخاصة في حال كان مستوى صعوبة الفقرات منخفضاً ومستوى قدرات المفحوصين مرتفعاً .

المراجع:

بني عطا، زايد والشريفين، نضال (٢٠١٢). أثر اختلاف شكل توزيع القدرة على معالم الفقرة ودالة المعلومات للاختبار. المجلة الأردنية في العلوم التربوية، ٨(٢)، ١٥١-١٦٦.

الزبون، حابس (٢٠١٣). أثر حجم العينة على تقدير دالة المعلومات للاختبار والخطأ المعياري في تقديرها باستخدام النظرية الحديثة في القياس. مجلة جامعة النجاح للعلوم الإنسانية، فلسطين، ٢٧(٦)، ١٣١٣-١٣٣٤.

الشريفين، نضال وبني عطا، زايد (٢٠١٣). تقصي أثر عدد خطوات الأسئلة متعددة التدريج وشكل التوزيع لصعوبتها على تقديرات القدرة للأفراد، والصعوبة للأسئلة، ودالة المعلومات للاختبار وفق نموذج التقدير الجزئي. المجلة التربوية، ٢٨(١٠٩)، ٢١٣-٢٧٥.

علام، صلاح الدين (٢٠٠٥). نماذج الاستجابة للمفردات الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي (ط١). القاهرة: دار الفكر العربي.

هياجنة، علي (٢٠١١). أثر طريقة انتقاء الفقرات في دالة معلومات الاختبار. رسالة دكتوراه غير منشورة، جامعة اليرموك، إربد، الأردن.

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

Birnbaum, A. (1968). *Some latent trait models and their use in inferring an examinee's ability*. In F. M. Lord & M. R. Novick, Statistical theories of mental test scores. Reading, MA: Addison-Wesley.

De Gruijter, D., & Van der Kamp, L. (2003). *Statistical Test Theory for Education and Psychology*. Retrieved, from: www.leidenuniv,n1/ griuijterdnme.

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

Fischer, G. (1974). *Einfuhrung in die theorie psychologischer tests (introduction to mental test theory)*. Berne: Huber.

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

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

Hambleton, R., & Cook, L. (1977). Latent trait models and their use in the analysis of educational test data. *Journal of Educational Measurement*. *14*(2), 75-94.

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

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

Han, K., & Hambleton, R. (2007). *User's manual for wingen: windows software that generated art 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.

Lord, F. (1977). Practical applications of item characteristic curve theory. *Journal of Educational Measurement, 14*(2), 117-138.

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

Mislevy, R., & Bock, R. (1990). BILOG3: Item analysis and test scoring with binary logistic models, [Computer program]. Mooresville IN: Scientific software.

Moghadamzadeh, A., Salehi, K. & Khodaie, E. (2011). A Comparison the Information Functions of the Item and Test in One, Two and Three Parametric Model of the Item Response Theory (IRT). *Procedia - Social and Behavioral Sciences, 29,* 1359 – 1367.

Reckase, M. (1997). The past and future of multidimensional item response theory. *Applied Psychological Measurement, 21*(1), 25-36.

Thorndike, R. (1982). *Applied psychology Houghton Mifflin Company Boston Dalas Geneva,* Illinois Hopewell: New Jersey Palo Alto London.