

القدرة التنبؤية لعالم الفقرة بمؤشر مطابقتها باختلاف عدد فترات القدرة تبعا لنظرية استجابة الفقرة (IRT)

رنا ثاني المومني*

تاريخ قبوله 19/11/2015

تاريخ تسلم البحث 29/3/2015

The Ability of Item Parameters as Predictors of Item Fit Index According to Item Response Theory (IRT)

Rana Thani Momani, Measurement & Evaluation Department, Qassim University, KSA.

Abstract: This study aimed at inspecting the ability of item parameters difficulty, discrimination, and guessing as predictors of item fit index (χ^2) according to item response theory (IRT). To achieve the goals of study the researcher used an achievement test in statistics consist of (29) four multiple choice items applied to sample consists of (507) female students from Faculty of Education. SPSS package used to verify assumptions of IRT models, then BILOG-MG used to extract item parameters and item fit index for three parameters model, and identify fitting items, then extract correlation coefficients for item parameters and item fit index, finally estimate the regression line for them. Results showed that the number of fitted items varied according to number of ability intervals and was the best in case of six intervals, also item discrimination had a significant relationship with item fit index, the regression model predicted item fit index using transformed discrimination and difficulty parameters with superiority for six intervals case. Finally results discussed in light of literature review and it was recommended to study item fit for different sample sizes, fit indices, statistical models and item types.

Keywords: item response theory, item difficulty, item discrimination, item guess ability, item fit index.

ملخص: هدفت الدراسة إلى الكشف عن القدرة التنبؤية لمعالم الفقرة الثلاثة الصعوبة، والتمييز، والتخمين بمؤشر مطابقة الفقرة (χ^2) تبعا لنظرية استجابة الفقرة (IRT). ولتحقيق أهداف الدراسة تم تطبيق اختبار تحصيلي لمقرر المقدمة في الإحصاء النفسي تكون من (29) فقرة اختيار من متعدد بأربعة بدائل لعينة الدراسة المكونة من (507) طالبة من طالبات كلية التربية للعام الجامعي 2014/2015، ثم التحقق من الافتراضات الأساسية للنظرية باستخدام الحزمة (SPSS)، ثم تحليل البيانات باستخدام برنامج (Bilog-Mg) لاستخلاص معالم الفقرات ومؤشر مطابقة الفقرة باستخدام النموذج الثلاثي المعلمة وتحديد الفقرات المطابقة للنموذج، ثم استخدام الحزمة (SPSS) لاستخراج معاملات ارتباط مؤشر المطابقة ومعالم الفقرة ومن ثم استخراج معادلة انحدار مؤشر المطابقة على معالم الفقرة. أظهرت نتائج الدراسة اختلاف عدد الفقرات المطابقة للنموذج باختلاف عدد فترات القدرة وكانت الأفضل عند تقسيم متصل القدرة لست فترات، كما أظهرت وجود دلالة إحصائية لارتباط معلم التمييز مع (χ^2) وتنبؤ معلم الصعوبة والتمييز المحولين بمؤشر مطابقة الفقرة وكان الأفضل عند ست فترات. وقد تمت مناقشة النتائج التي توصلت إليها الدراسة في ضوء الأدب النظري، وأوصت الباحثة بدراسة المطابقة لعينات مختلفة الأحجام، ومؤشرات مطابقة ونماذج إحصائية مختلفة أنواع أخرى من الاختبارات.

الكلمات المفتاحية: نظرية استجابة الفقرة، صعوبة الفقرة، تمييز الفقرة، تخمين الفقرة، مؤشر مطابقة الفقرة.

مقدمة: احتلت الاختبارات وخصائصها مرتبة هامة على مر التاريخ؛ لما لها من أهمية بالغة في تصنيف الأفراد وانتقائهم، مما يحدد مصيرهم فيما بعد سواء أكان العلمي أم الوظيفي أو غير ذلك، مما أدى إلى ظهور نظريتي القياس المعروفتين بنظرية الاختبار الكلاسيكية (Classical Test Theory (CTT))، ومن ثم ما سُمي بالنظرية الحديثة، أو نظرية الاستجابة للفقرة (Item Response Theory (IRT))، حيث اهتمت كلتا النظريتين بتحديد قواعد وضع الاختبارات، ودراسة خصائصها المختلفة سعياً للوصول إلى ما يعرف بـ "الاختبار الجيد". وتعد كلتا النظريتين غاية في الأهمية في تقديم تقييمات مختلفة حول كل من فقرات الاختبار والاختبار ككل. (Capalleri, Lundy & Hays, 2014).

وقد أظهرت نظرية الاستجابة للفقرة نجاحاً ملحوظاً في تحسين وتطوير جودة الاختبار؛ مما أدى إلى شيوع استخدامها من قبل عدد من الباحثين والدارسين. وتقوم نظرية الاستجابة للفقرة على عدة افتراضات أساسية، ويشترط نجاح استخدام تطبيقاتها النظرية، والحصول على الفوائد المتوخاة، منها: حسن مطابقة النموذج المختار للبيانات، حيث يصبح من المهم فهم الأسباب التي تؤدي إلى عدم التطابق كجزء من فاعلية ومشروعية استخدام نماذج نظرية استجابة الفقرة (Maydeu, 2005).

نظرية استجابة الفقرة

تقدم نظرية استجابة الفقرة طرقاً عدة للوصول إلى ما يدعى بالاختبار المثالي (Optimal Test) من خلال تقديم مؤشرات إضافية عن نظيرتها الكلاسيكية (Tay, Meade & Cao, 2015). وقد اتسع نطاق استخدامها في الأونة الأخيرة من قبل الباحثين في غير المجال التربوي لتطبيق بشكل واسع في المجال الطبي تحديداً بشكل ملحوظ (Reise & Revicki, 2015). وتفترض هذه النظرية إمكانية التنبؤ بأداء المفحوصين أثناء الاختبارات من خلال تحديد خصائص المفحوصين، أو ما يعرف بسماتهم أو قدراتهم، وتقدير

* قسم علم النفس، جامعة القصيم، المملكة العربية السعودية.

© حقوق الطبع محفوظة لجامعة اليرموك، اربد، الأردن.

والنموذج ثلاثي المعالم من خلال المعادلة التالية:

$$P_i(\theta) = c_i + (1 - c_i) \frac{\exp[Da_i(\theta - b_i)]}{1 + \exp[Da_i(\theta - b_i)]} \quad (3)$$

حيث إن:

(θ): قدرة المفحوص.

(θ): $P_i(\theta)$: احتمال أن المفحوص الذي يمتلك مستوى قدرة (θ) أجاب عن الفقرة (i) إجابة صحيحة.

b_i : النقطة التي تقابل $P_i(\theta) = 0.5(c+1)$, وهو عبارة عن معلم صعوبة الفقرة.

a_i : ميل الدالة عند النقطة b على متصل السمة، وهو عبارة عن معلم تمييز الفقرة، ويساوي قيمة ثابتة لجميع الفقرات في النموذج أحادي المعلمة وتساوي (1,0).

c_i : المماس السفلي للدالة، وهو عبارة عن معلم التخمين ويفترض النموذج أحادي وثنائي المعلمة أن قيمة هذا المعلم تساوي صفراً).

مطابقة الفقرة

تعد دراسة مطابقة الفقرة للنموذج من أهم الموضوعات فيما يتعلق ببناء وتطوير الاختبارات، ويتم تقدير مطابقة الفقرة باستخدام طرق عدة. وتعد الطريقة الإحصائية التي تم تقديمها من قبل رايت وبانشابكيسان (Wright & Panchapakesan, 1969) الطريقة الأكثر شيوعاً لاختبار مطابقة الفقرة لنموذج أحادي المعلمة (راش)، الذي يمكن استخدامه للنموذجين ثنائي وثلاثي المعلمة والمعطى بالمعادلة:

$$\chi^2 = \sum_{i=1}^{n-1} \sum_{j=1}^n Y_{ij}^2 \quad (4)$$

والذي يمتلك توزيع مربع كاي (χ^2) بدرجات حرية تساوي ($n-2$)، حيث إن

$$Y_{ij} = \frac{[f_{ij} - E(f_{ij})]}{\sqrt{\text{var}(f_{ij})}} \quad (5)$$

f_{ij} : تكرار المفحوصين عند مستوى القدرة (i) والذين أجابوا إجابة صحيحة عن الفقرة (j). ويتوزع المتغير (Y_{ij}) طبيعياً بمتوسط حسابي مقداره (0)، وتباين مقداره (1) (Hambleton & Swaminathan, 1985).

ويتيح برنامج (BILOG-MG) إمكانية استخراج قيم (χ^2)، وما يقابلها من قيم احتمالية (Prob.) للاختبارات التي تتكون من (20) فقرة على الأقل، بحيث يكون حجم العينة كافياً، إذ يمكن

ما يمتلكونه من قدرات من خلال ما يعرف بدرجات القدرة (Ability Scores)، ثم استخدام هذه الدرجات لتفسير أداء كل من الفقرة والاختبار. وتقدم نظرية استجابة الفقرة إمكانية قياس المتغيرات التي لا يمكن قياسها مباشرة التي تسمى بالسّمات أو القدرات الكامنة بطريقة غاية في الدقة (De Ayala, 2009). وتقوم نماذج استجابة الفقرة على تحديد العلاقة بين أداء المفحوصين الملاحظ (Observable) على الاختبار والسّمات غير الملاحظة (Unobservable) المتوقعة للأداء على الاختبار، حيث توصف هذه العلاقة بدوال رياضية تقوم على افتراضات محددة فيما يخص بيانات الاختبار (Hambleton & Swaminathan, 1985). وقد تم تطوير هذه النماذج كطرق لتحليل الاختبارات التربوية، وطُبقت بشكل واسع في المجالين النفسي والاجتماعي، إضافة إلى توافرها بشكل جيد مع نتائج مقاييس الاتجاهات (Fontanella, Villano & Donato, 2015).

وتقوم نظرية استجابة الفقرة على عدد من الافتراضات الأساسية، هي:

- أحادية البعد (Uni-Dimensionality): وهو أن يقيس الاختبار سمة واحدة تفسر أداء المفحوصين على ذلك الاختبار (Hambleton, Swaminathan & Rogers, 1991)، إذ يحتوي الاختبار اتجاهها عاماً واسعاً يتضمن عوامل جزئية (Tay, Meade & Cao, 2015).
- الاستقلال الموضوعي (Local Independence): وهو الاستقلال الإحصائي لأداء المفحوصين على فقرات الاختبار، فأداء المفحوص على فقرة لا يتأثر بأدائه على أي فقرة أخرى (Hambleton & Swaminathan, 1985).
- منحنى استجابة الفقرة (Item Response Curve): ضمن إطار نظرية استجابة الفقرة يجب أن يمتلك منحنى استجابة الفقرة شكلاً لوجستياً. وقد عرف كل من هامبلتون وسواميناثان وروجرز (Hambleton, Swaminathan & Rogers, 1991) دالة استجابة الفقرة على أنه: تعبير رياضي يربط كل من احتمالية الإجابة الصحيحة على الفقرة والقدرة التي يقيسها الاختبار وخصائص الفقرة. ويشير بيكر (Baker, 2001) إلى أن نماذج استجابة الفقرة البارامترية الثلاثة هي النماذج الأكثر شهرة وانتشاراً، التي يمكن التعبير عنها من خلال المعادلة التالية للنموذج أحادي المعلمة:

$$P_i(\theta) = \frac{\exp[D(\theta - b_i)]}{1 + \exp[D(\theta - b_i)]} \quad (1)$$

والنموذج ثنائي المعالم من خلال المعادلة التالية:

$$P_i(\theta) = \frac{\exp[Da_i(\theta - b_i)]}{1 + \exp[Da_i(\theta - b_i)]} \quad (2)$$

وقام سيول (Seol, 1999) بدراسة قوة خمسة مؤشرات لمطابقة الفقرة للكشف عن الأداء التفاضلي للأفراد (DIF) باستخدام النموذج الأحادي، حيث استخدم المؤشر المعياري الموزون (W2) والمؤشر المعياري غير الموزون (UW2) والمؤشر المعياري المرجح ومؤشري (EC12, EC14)، باستخدام بيانات مولدة لـ (1000) مفحوص و (100) فقرة. أظهرت نتائج الدراسة عدم وجود علاقة خطية بين مؤشرات المطابقة ومعلم الصعوبة، حيث تراوحت قيم الارتباط بين (-0,04 و -0,02).

كما قام دودين (Dodeen, 2004) بدراسة خطية العلاقة بين معالم الفقرات، ومؤشر مطابقتها باستخدام النموذج ثلاثي المعلمة ومؤشري البواقي المعيارية (SRs)، ومربع كاي (χ^2)، وقد قام بتوليد بيانات ذات توزيع طبيعي بطرف مختلف لتسع مجموعات اختبارية تكون كل منها من (1000) مفحوص و (50) فقرة اختيار من متعدد بالاعتماد على النموذج المختار، وقد اختلفت المجموعات الاختبارية الثلاثة الأولى فقط في متوسط تمييزها، واختلفت المجموعات الاختبارية الثلاثة الثانية فقط في متوسط صعوبتها، بينما اختلفت الثلاثة الأخيرة فقط في متوسط تخمينها. بعد ذلك تم استخراج إحصاءات متوسط مطابقة الفقرة (Average item fit)، إضافة إلى كل من نسبة الفقرات المطابقة عند مستوى الدلالة (0,01) والارتباط بين معالم الفقرات المولدة ومؤشر مطابقة الفقرة، ثم تم استخدام تحليل التباين (ANOVA) لتحديد أثر اختلاف مستوى المعلمة على متوسط مؤشر المطابقة. وقد أظهرت النتائج أن كلا من متوسط ونسبة الفقرات المطابقة للنموذج اختلفت لكلا المؤشرين (χ^2 , SR_s)، وقد تبين وجود علاقة موجبة بين كلا المؤشرين ومعلمي التخمين والتمييز، فكلما ارتفع مستوى التخمين والتمييز ازدادت قيم مؤشر المطابقة؛ مما يشير إلى فقرات غير مطابقة، كما أظهرت النتائج عدم وجود علاقة بين مستوى الصعوبة ومؤشر المطابقة للفقرة.

ودرس كل من سينهاري ولو (Sinharay & Lu, 2008) العلاقة بين معالم الفقرة ومؤشرات مطابقة الفقرة (G^{2*} , χ^2 , $S - \chi^2$, $S - G^{2*}$) للتحقق من النتائج التي خرج بها دودين سابقاً، حيث قام الباحثان بتوليد بيانات لتسع مجموعات اختبارية تكون كل منها من (1000) مفحوص و (50) فقرة اختيار من متعدد باستخدام النموذج ثلاثي المعالم، وقد اختلفت المجموعات الاختبارية الثلاثة الأولى فقط في متوسط تمييزها، واختلفت المجموعات الاختبارية الثلاثة الثانية فقط في متوسط صعوبتها، بينما اختلفت الثلاثة الأخيرة فقط في متوسط تخمينها كما في دراسة دودين. وقد تم توليد البيانات للمجموعات التي اختلفت في تمييزها، حيث توزعت توزيعاً لوغاريتمياً طبيعياً، بينما تم استخدام توزيع بيتا لباقي المجموعات الاختبارية، بينما استخدم التوزيع الطبيعي في دراسة دودين. أظهرت نتائج الدراسة اختلاف نتائج الدراسة للباحثين مع نتائج دراسة دودين؛ إذ تبين وجود علاقة عكسية بين مؤشري المطابقة (χ^2) و (G^{2*})، ومعلم

استخراج ثلاث فئات على الأقل لقدرات المفحوصين. وأشارت الكثير من الدراسات إلى قوة مؤشر (χ^2) المستخرج ضمن برنامج (BILOG-MG)، حيث أفادت هذه الدراسات تمتعه بمقايير منخفضة من الخطأ من النوع الأول، كما تمت الإشارة لمشكلتين أساسيتين مرتبطتين بمؤشر (χ^2)، وهما: عدد فقرات متصل القدرة، وحجم العينة. ويعد القرار المتخذ بشأن عدد الفقرات قراراً شخصياً من قبل الباحثين إذا لم يتم الاتفاق حول العدد الأنسب لها، بينما قدم بعض الباحثين مقترحات حول عدد الفقرات المناسب كما هو الحال في مؤشر (Yen's Q1) الذي استخدم عشر فئات تطبيقية؛ إذ تشابه في نتائجه وشكل توزيعه مع مؤشر (χ^2) الذي تم تبنيه من قبل بعض البرامج الإحصائية كبرنامج (RUMM) (Reise, 1990). ولتحديد الفقرات المطابقة للنموذج ضمن نظرية الاستجابة للفقرة فإن القيم الاحتمالية التي تقل عن (0,05) تمثل الفقرات غير المطابقة لذلك النموذج (Haley et. al, 2009). وقد لوحظ ازدياد اهتمام الباحثين بدراسة المطابقة ضمن إطار نظرية استجابة الفقرة في السنوات الأخيرة، إلا أن البحث لا يزال جارياً لفهم ماهية المطابقة، وأسباب تعذرها لبعض الفقرات، الأمر الذي يدعو إلى أهمية دراسة المطابقة بجميع حيثياتها.

وتعد الدراسة التي قام بها روجرز وهاتي (Rogers & Hattie, 1987) من أوائل الدراسات ضمن هذا المجال، حيث درسا أداء ثلاثة من مؤشرات مطابقة المفحوص وثلاثة لمطابقة الفقرة، وهي (mean square residual, between t, total t) لنموذج راش أحادي المعلمة لبيانات مولدة لـ (500) مفحوص و (15) فقرة، وقد خلصت الدراسة إلى أن مؤشر (between t) أظهر حساسية نحو التخمين، وعدم تجانس معلم التمييز، وأوصى الباحثان بإمكانية استخدام هذا المؤشر لبحث مطابقة النموذج.

وقام ريس (Reise, 1990) بدراسة مقارنة بين مؤشر مطابقة الفقرة (χ_B^2)، ومؤشر مطابقة الفرد (Z_B) لبيانات مولدة لـ (1000) مفحوص و (50) فقرة من نوع الاختيار من متعدد، حيث تمت دراسة العلاقة بين مؤشري المطابقة ومعلم صعوبة الفقرة، حيث أظهرت النتائج عدم وجود علاقة بين أي منهما وصعوبة الفقرة.

ودرس كل من ماونت وشوماخر (Mount & Schumacker, 1998) تأثير التخمين على ثلاثة مؤشرات (الكلية الموزون، الكلية غير الموزون، المطابقة البيئية غير الموزونة)، حيث تم توليد بيانات لـ (100) مفحوص و (100) فقرة لمطابقة نموذج راش الأحادي، وقد تم استخدام ثلاثة مستويات من التخمين (0%, 25%, 50%)، ونوعين من أنواع التوزيع (normal & uniform). أظهرت النتائج عدم وجود تأثير لمستوى التخمين على متوسطات مؤشرات مطابقة الفقرات لنوعي التوزيع، كما أظهر المؤشر البيئي حساسية نحو مستوى التخمين؛ إذ ازداد عدد الفقرات غير المطابقة بازدياد التخمين، حيث أوصى الباحثان بالمؤشر البيئي كمؤشر جيد في بناء الاختبار.

الاختبار، وإمكانية التعامل مع البيانات بشكل أفضل، وبالتالي يقودنا إلى نتائج أكثر دقة وإمكانية الضبط والتحكم، وهو الهدف الأسمى فيما يخص مطابقة الفقرة. وتتحدد مشكلة الدراسة في الإجابة عن الأسئلة التالية:

1. ما الفقرات المطابقة للنموذج البارامترى ثلاثي المعلمة عند مستوى الدلالة ($\alpha = 0.05$)؟
2. هل توجد علاقة ارتباطية خطية بين مؤشر المطابقة (χ^2) لفقرات الاختبار للنموذج الثلاثي ومعالمها عند مستوى الدلالة ($\alpha = 0.05$) عند تغير عدد فقرات القدرة؟
3. هل يمكن التنبؤ بمؤشر مطابقة الفقرة (χ^2) من خلال معالم الفقرة الثلاثة (التمييز (a)، والصعوبة (b)، والتخمين (c)) عند مستوى الدلالة ($\alpha = 0.05$)؟

هدف الدراسة

تهدف هذه الدراسة إلى بحث طبيعة العلاقة وإمكانية التنبؤ بها بين مؤشر مطابقة الفقرة (χ^2)، ومعالم الفقرة الثلاث: الصعوبة، والتمييز، والتخمين باستخدام نظرية استجابة الفقرة بنموذجها الثلاثي المعلمة لبيانات تجريبية لاختبار تحصيلي، حيث تقدم هذه الدراسة طريقة لإمكانية فهم أسباب عدم المطابقة، وإمكانية التحكم بها؛ الأمر الذي يعود بالنفع في اختصار الوقت والجهد في عملية إنشاء الاختبارات، وبناء بنوك الفقرات، وبناء الاختبارات التكوينية المحوسبة وغيرها من التطبيقات التي تساهم في رفع سوية الاختبار.

أهمية الدراسة

الأهمية النظرية: تنبثق الأهمية النظرية للدراسة في محاولة التعرف إلى طبيعة العلاقة بين مؤشر مطابقة الفقرة ومعالمها، وإمكانية فهم أسباب عدم المطابقة، إضافة إلى إلقاء الضوء على طبيعة العلاقة الارتباطية بين مؤشر مطابقة الفقرة ومعالمها الثلاث لبيانات تجريبية؛ الأمر الذي قد يتيح التنبؤ بمطابقة الفقرة، أو عدمها من خلال خصائصها المختلفة، إضافة إلى ما توفره من معلومات قد تفيد باستخدام مؤشرات مطابقة أخرى أقل تأثراً بمعالم الفقرة، مما يوفر الوقت والجهد ويجعل من عملية تطوير وبناء الاختبارات وبنوك الفقرات عملية أكثر سهولة، إضافة إلى قلة الدراسات بشكل عام، وعدم وجود دراسات عربية بشكل خاص (في حدود اطلاع الباحثة)؛ مما قد يساهم في الإضافة للأدب النظري في هذا المجال عربياً.

الأهمية العملية: تنبثق الأهمية العملية للدراسة من خلال مساهمتها في بناء وتطوير الاختبارات المختلفة، وإنشاء بنوك الاختبارات، إضافة إلى إسهامها في بناء الاختبارات التكوينية المحوسبة (CAT)، فهي تتيح التنبؤ بالمطابقة للنموذج المختار من خلال خصائص الفقرة المختلفة.

التمييز للفقرة، إضافة لعدم وجود علاقة خطية بين مؤشر مطابقة الفقرة وتحميلها بعكس النتيجة التي خلص لها دودين، بينما توافقت الدراسات بعدم وجود علاقة لمؤشر المطابقة مع معلم الصعوبة؛ إذ إن النتائج لم تظهر نمطاً معيناً بهذا الخصوص. كما قام الباحثان ببحث العلاقة الخطية بين مؤشر مطابقة الفقرة ومعالمها على بيانات حقيقية مكونة من استجابات (8686) طالباً على اختبار للمهارات الأساسية مكون من (45) فقرة اختيار من متعدد. وقد أظهرت النتائج مطابقة (21) فقرة للنموذج، كما بينت النتائج وجود علاقة ارتباطية موجبة بين مؤشر المطابقة (χ^2) ومعلم التمييز كما في دودين، وعلاقة ارتباطية سالبة دالة إحصائياً بين معلم الصعوبة والتخمين ومؤشر المطابقة (χ^2) بعكس نتائج دراسة دودين، وعكس النتائج للدراسة نفسها على البيانات المولدة. وقد فسر الباحثان هذا التناقض بسبب شكل التوزيع المستخدم عند توليد البيانات، إضافة إلى أن البيانات المولدة تم وضعها بناء على النموذج المختار، وهو النموذج الثلاثي المعلم.

يتضح مما سبق وجود اختلاف وتضارب في نتائج بعض الدراسات السابقة، فبعضها أفاد بعدم وجود علاقة بين معالم الفقرة ومؤشر المطابقة كدراسة كل من ريس (Reise, 1990)، وماونت وشوماخر (Mount & Schumacker, 1998) وسيول (Seol, 1999)، كما أكد بعضها الآخر وجود علاقة دون تحديد نوعها وماهيتها كدراسة روجرز وهاتي (Rogers & Hattie, 1987)، بينما أفادت دراسات أخرى بوجود علاقة وتحديد شكلها على أنها خطية، إلا أن نتائجها تضاربت في اتجاه العلاقة عكسية أم طردية كدراستي دودين (Dodeen, 2004) وسينهاري ولو (Sinharay & Lu, 2008) على الرغم من أن كل منهما استخدم نفس عدد الفقرات والمفحوصين لكل مجموعة من مجموعات الدراسة، إضافة إلى تضارب نتائج دراسة سينهاري ولو (Sinharay & Lu, 2008) عند استخدام كل من البيانات المولدة والحقيقية.

مشكلة الدراسة

تعد دراسة المطابقة واحدة من أهم المشكلات التي يواجهها الباحثون ضمن إطار نظرية الاستجابة للفقرة (Item response theory (IRT))، حيث لا يزال الغموض قائماً حول أسباب عدم المطابقة، وكيفية التحكم والتنبؤ بها. وانطلاقاً من احتمالية وجود علاقة بين معالم الفقرة ومؤشرات مطابقتها، ونظراً لاختلاف وعدم اتفاق نتائج الدراسات السابقة حول طبيعة العلاقة بين مؤشرات مطابقة الفقرات ومعالمها، واعتمادها غالباً على بيانات غير تجريبية وبنى افتراضية (Simulation Data) تبلورت مشكلة هذه الدراسة كخطوة نحو اكتشاف طبيعة العلاقة بين معالم الفقرة ومؤشر مطابقتها، بل وإمكانية التنبؤ بمؤشر المطابقة للفقرة من خلال معالمها الثلاث: الصعوبة، والتمييز، والتخمين؛ الأمر الذي يساهم في إمكانية التنبؤ بمطابقة الفقرات لنماذج نظرية استجابة الفقرة، وتزويد الباحثين بمعلومات قد تفيد في استخدام مؤشرات مطابقة أخرى أقل تأثراً بمعالم الفقرات، مما يقود إلى تحسين نوعية

محددات الدراسة

محددات مكانية: جامعة القصيم في المملكة العربية السعودية.

محددات زمنية: الفصل الدراسي الأول للعام الدراسي 2015/2014م.

محددات بشرية: طالبات كليات التربية في جامعة القصيم.

محددات موضوعية: متغيرات الدراسة - مؤشرات المطابقة للنموذج الثلاثي المعلمة، ومعالمها الثلاث (الصعوبة، والتمييز، والتخمين).

مصطلحات الدراسة

- مطابقة الفقرة: وهي ملاءمة الفقرة للنموذج المختار من خلال محكات للحكم خاصة بكل نموذج، فيما يخص شكل دالة استجابة تلك الفقرة ومعالمها.
- النموذج ثلاثي المعلمة: وهو أحد نماذج استجابة الفقرة البارامترية، حيث تختلف الفقرات في قيم الصعوبة والتمييز والتخمين.
- معالم الفقرة: وهي خصائص الفقرة الثلاث تبعاً لنظرية استجابة الفقرة (الصعوبة، والتمييز، والتخمين).

الطريقة والإجراءات

منهج الدراسة: تم استخدام خطوات المنهج التجريبي للإجابة عن أسئلة الدراسة الذي يعد الأنسب لمثل هذه الدراسات.

مجتمع الدراسة: تكون مجتمع الدراسة من طالبات كليات التربية اللواتي يدرسن في جامعة القصيم للعام الجامعي 2015/2014م.

عينة الدراسة: تكونت عينة الدراسة من (507) طالبة من طالبات السنة الثانية في كليات التربية التابعة لجامعة القصيم للعام الجامعي 2015/2014م. وقد تم اختيار كلية التربية في الصفراء التي تضم قسمي علم النفس والتربية الخاصة باستخدام العينة العشوائية العنقودية، حيث كانت الكلية هي وحدة الاختيار، بعد ذلك تم اختيار أربع شعب من كل قسم بالطريقة العشوائية البسيطة، حيث بلغ عدد طالبات قسم علم النفس (259) طالبة، وبلغ عدد طالبات قسم التربية الخاصة (248) طالبة.

أداة الدراسة: تم استخدام اختبار تحصيلي لمقرر المقدمة في الإحصاء التربوي والنفسي مكون من (30) فقرة، وهو عدد مناسب للاختبارات التحصيلية وفقاً لما أشار إليه نانالي (Nunnally, 1978)، ويولي كل فقرة من فقرات الاختبار أربعة بدائل أحدها صحيح، حيث تعطى الإجابة الصحيحة الدرجة (1) والخطئة الدرجة (0)، وبذلك تكون أدنى درجة على الاختبار هي الدرجة (0)، وأعلى درجة هي الدرجة (30)،

وقد تم الالتزام بالشروط الفنية لكتابة الفقرة التي تمت الإشارة لها في عودة (2010) لتطوير الاختبارات التحصيلية.

صدق الأداة: قامت الباحثة بالتحقق من صدق الاختبار بطريقتين، هما: الصدق المنطقي، وصدق المحكمين.

أولاً: الصدق المنطقي: تم التحقق من الصدق المنطقي من خلال وضع فقرات الاختبار بالاستعانة بنتائج جدول المواصفات الذي يضمن أن تتوزع فقرات الاختبار توزيعاً عادلاً بين موضوعات المقرر والمستويات المعرفية المختلفة.

ثانياً: صدق المحكمين: تم عرض الاختبار بصورته الأولية على خمسة عشر محكماً ممن يحملون درجة الدكتوراه في تخصصات علم النفس والقياس والتقويم والرياضيات والإحصاء، حيث حُدثت نسبة الاتفاق لقبول الفقرة بما مقداره (80%). وبناءً على نتائج التحكيم تم التعديل في صياغة بعض الفقرات، كما تم حذف فقرة اتفق خمسة محكمين على ضرورة حذفها، وبذلك أصبح الاختبار بصورته المحكمة مكوناً من (29) فقرة.

ثبات الأداة: تم حساب ثبات الاتساق الداخلي للاختبار بطريقتي كرونباخ - ألفا، والتجزئة النصفية كمؤشر على ثبات نتائج الاختبار.

أولاً: التجزئة النصفية: تم تقدير ثبات الاختبار بطريقة التجزئة النصفية باستخدام معامل ثبات جتمان لنصفي الاختبار، وقد ساوت قيمة معامل الثبات (0,86).

ثانياً: كرونباخ - ألفا: تم تقدير ثبات الاختبار باستخدام معادلة كرونباخ - ألفا، وقد بلغت قيمته (0,9). وبناءً على ذلك يمكن اعتبار الاختبار مناسباً لجمع البيانات اللازمة للدراسة، حيث إن دلالات الثبات والصدق تفي لتلك الأغراض.

متغيرات الدراسة

أولاً: المتغيرات المستقلة

- صعوبة الفقرة (b_i).
- تمييز الفقرة (α_i).
- تخمين الفقرة (c_i).

ثانياً: المتغيرات التابعة

- مؤشر مطابقة الفقرة (χ^2)، الذي يُعامل كمتغير تابع فقط عند دراسة الانحدار.

- نسبة ما يفسره العامل الأول من التباين الكلي أكبر من (20%).
 - نسبة ما يفسره العامل الأول إلى العامل الثاني أكبر من (2).
- كما تم تمثيل الجذور الكامنة بيانياً من خلال ما يعرف بـ (Scree Plot)، حيث يمثل المحور الأفقي العوامل، ويمثل المحور العمودي مقدار الجذر الكامن الذي أكد أحادية البعد من خلال تمييز العامل الأول عن باقي العوامل. كما تم التحقق من افتراض الاستقلال الموضوعي عن طريق الإحصائي (Q3)، وذلك باستخدام الحزمة الإحصائية (R-Package)، إذ يمثل (Q3) معامل الارتباط للبواقي لأزواج الفقرات، وقد بلغت قيمة متوسط (Q3) للبيانات (0,035)، وهي قيمة واقعة ضمن فترة الثقة الخاصة به تبعاً لعدد فقرات الاختبار، وقد قل مستوى الانتهاك عن (0,0007)؛ مما يدل على تحقق فرض الاستقلال الموضوعي لبيانات الاختبار.

ثانياً: الخصائص السيكومترية للاختبار تبعاً للنظرية الكلاسيكية في الاختبار: للتعرف إلى الخصائص الكلاسيكية لاختبار الإحصاء كأداة لهذه الدراسة، تم استخراج عدد الاستجابات الصحيحة على كل فقرة وصعوبة الفقرة، إضافة إلى معاملات ارتباط الفقرة مع الاختبار، حيث إن ارتباط الفقرة بالسمة المقيسة يعد مؤشراً على تمييز الفقرة المستخرج باستخدام برنامج (bilog-mg). وقد أظهرت النتائج وجود بعض الفقرات السهلة كالفقرات (1، 2، 3، 4)، وبعض الفقرات متوسطة الصعوبة كالفقرات (16، 18، 21، 26)، كما يبين أن جميع الفقرات تمتعت بقيم مقبولة لمعامل ارتباط بيرسون، إذ لم تقل أي من قيم معامل ارتباط بيرسون عن (0,15) تبعاً لبراون (Brown, 1996)، كما يبين الجدول وجود بعض الفقرات عالية التمييز مثل الفقرات (10، 12، 16، 19)، والفقرات الأقل تمييزاً كالفقرات (1، 2، 3، 5).

ثالثاً: النتائج المتعلقة بالسؤال الأول الذي ينص على "ما الفقرات المطابقة للنموذج البارامتري ثلاثي المعلمة عند مستوى الدلالة ($\alpha = 0.05$)؟"

للإجابة عن هذا السؤال تم إيجاد معالم الفقرات والأخطاء المعيارية لها وإحصائيات جودة الملاءمة للتعرف على عدد فقرات الاختبار المطابقة للنموذج البارامتري ثلاثي المعلمة، وذلك كما هو موضح في الجدول (2).

- المعالجات الإحصائية:** تم استخدام المعالجات الإحصائية التالية:
1. التحقق من مطابقة البيانات للافتراضات الأساسية لنظرية استجابة الفقرة (أحادية البعد، والاستقلال الموضوعي) باستخدام الحزمة الإحصائية (SPSS).
 2. المعالجات الإحصائية لاستخراج معالم الفقرات (الصعوبة، التمييز، التخمين) للنموذج البارامتري ثلاثي المعلمة باستخدام البرنامج الإحصائي (BILOG-MG).
 3. المعالجات الإحصائية لاستخراج مؤشرات مطابقة الفقرات (χ^2) للنموذج البارامتري ثلاثي المعلمة باستخدام البرنامج الإحصائي (BILOG-MG).
 4. استخراج معاملات ارتباط مؤشر المطابقة للفقرة مع معالمها (الصعوبة، التمييز، التخمين) واستخراج معادلة انحدار معالم الفقرة على مؤشر المطابقة للفقرة باستخدام الحزمة الإحصائية (SPSS).

نتائج الدراسة

أولاً: التحقق من الافتراضات الأساسية لنماذج نظرية استجابة الفقرة، تم التحقق من الافتراضات الأساسية التي تسوغ استخدام نماذج نظرية استجابة الفقرة المختلفة، وهما: أحادية البعد، والاستقلال الموضوعي، وقد تم فحص أحادية البعد للبيانات الناتجة من تطبيق الاختبار باستخدام برنامج (SPSS)، حيث أجري التحليل بطريقة المكونات الأساسية (Principal Component Analysis)، وملاحظة نسب قيمة الجذر الكامن (Eigen Value) للعامل الأول إلى الجذر الكامن للعامل الثاني، إضافة إلى مقدار التباين الذي فسره العامل الأول والذي بلغ (28,78%) من التباين الكلي، وذلك كما هو مبين في الجدول (1).

جدول (1): العوامل الخمسة الأولى وجذورها الكامنة ونسب تباينها

| العامل | قيمة الجذر الكامن | نسبة التباين المفسر للعامل |
|--------|-------------------|----------------------------|
| 1 | 8.34 | 28.78% |
| 2 | 1.64 | 5.66% |
| 3 | 1.56 | 5.38% |
| 4 | 1.19 | 4.12% |
| 5 | 1.11 | 3.83% |

يلاحظ من الجدول (1) التأكيد على أحادية البعد من خلال:

جدول (2): قيم صعوبة وتمييز وتخمين فقرات الاختبار والأخطاء المعيارية وإحصائيات جودة الملاءمة ودرجات الحرية.

| ITEM | SLOPE S.E. | THRESHOLD S.E. | SYMPTOTE S.E. | CHISQ (PROB) | Df |
|--------|-----------------|-------------------|------------------|-----------------|----|
| ITEM01 | 0.657 0.131* | -1.678 0.364* | 0.298 0.105* | 8.7 0.2757 | 7 |
| ITEM02 | 0.66 0.123* | -1.85 0.359* | 0.274 0.101* | 11.4 0.1226 | 6 |
| ITEM03 | 0.76 0.130* | -1.182 0.261* | 0.246 0.091* | 4.3 0.6411 | 6 |
| ITEM04 | 1.087 0.163* | -1.156 0.171* | 0.181 0.072* | 6.8 0.3371 | 6 |
| ITEM05 | 0.762 0.193* | -0.056 0.247* | 0.311 0.086* | 32.1 0 | 7 |
| ITEM06 | 0.994 0.162* | -1.393 0.208* | 0.183 0.075* | 13.3 0.0389 | 6 |
| ITEM07 | 1.164 0.210* | -0.864 0.147* | 0.17 0.068* | 10.2 0.1173 | 6 |
| ITEM08 | 1.309 0.266* | -0.545 0.134* | 0.22 0.071* | 5 0.5396 | 6 |
| ITEM09 | 1.493 0.357* | -0.062 0.103* | 0.209 0.061* | 18.6 0.005 | 6 |
| ITEM10 | 3.094 0.870* | -0.019 0.060* | 0.242 0.042* | 5.8 0.4445 | 6 |
| ITEM11 | 1.232 0.293* | 0.296 0.125* | 0.231 0.058* | 15.3 0.018 | 6 |
| ITEM12 | 1.873 0.352* | -0.567 0.092* | 0.178 0.058* | 3.3 0.7748 | 6 |
| ITEM13 | 1.461 0.269* | -0.464 0.111* | 0.204 0.064* | 5.5 0.482 | 6 |
| ITEM14 | 1.322 0.252* | -0.32 0.123* | 0.226 0.065* | 4.9 0.6695 | 7 |
| ITEM15 | 0.866 0.145* | -0.877 0.194* | 0.194 0.076* | 9.4 0.1515 | 6 |
| ITEM16 | 1.754 0.304* | 0.097 0.065* | 0.108 0.034* | 11.3 0.0462 | 5 |
| ITEM17 | 1.376 0.239* | -0.177 0.105* | 0.189 0.056* | 6.6 0.3552 | 6 |
| ITEM18 | 2.154 0.599* | 0.382 0.079* | 0.248 0.039* | 9.8 0.1978 | 7 |
| ITEM19 | 1.57 0.243* | -0.394 0.089* | 0.143 0.052* | 8.3 0.2141 | 6 |
| ITEM20 | 1.181 0.171* | -0.8 0.126* | 0.14 0.057* | 9.3 0.1588 | 6 |
| ITEM21 | 1.672 0.267* | 0.248 0.060* | 0.06 0.025* | 11 0.0513 | 5 |
| ITEM22 | 1.156 0.183* | -0.939 0.140* | 0.146 0.060* | 15.8 0.0074 | 5 |
| ITEM23 | 1.867 0.357* | -0.249 0.086* | 0.212 0.054* | 2.2 0.824 | 5 |
| ITEM24 | 1.478 0.320* | 0.042 0.100* | 0.212 0.054* | 1.8 0.8754 | 5 |

| ITEM | SLOPE S.E. | THRESHOLD S.E. | SYMPTOTE S.E. | CHISQ (PROB) | Df |
|--------|-----------------|-------------------|------------------|-----------------|----|
| ITEM25 | 1.586 0.432* | 0.201 0.102* | 0.272 0.053* | 2.2 0.8987 | 6 |
| ITEM26 | 2.202 0.385* | 0.132 0.054* | 0.09 0.029* | 4.9 0.4298 | 5 |
| ITEM27 | 1.594 0.339* | -0.059 0.090* | 0.193 0.053* | 8 0.236 | 6 |
| ITEM28 | 1.281 0.294* | -0.054 0.132* | 0.27 0.065* | 7.5 0.2756 | 6 |
| ITEM29 | 1.146 0.184* | -0.267 0.113* | 0.151 0.057* | 7.1 0.4203 | 7 |

رابعاً: النتائج المتعلقة بالسؤال الثاني الذي ينص على "هل توجد علاقة ارتباطية خطية بين مؤشر المطابقة (χ^2) لفقرات الاختبار للنموذج الثلاثي ومعالمها عند مستوى الدلالة ($\alpha = 0.05$) في حال تغير عدد فترات القدرة؟"

للإجابة عن هذا السؤال تم استخراج معالم الفقرات الثلاث: التمييز (a)، والصعوبة (b)، والتخمين (c)، ومؤشر المطابقة (χ^2) للفقرة في ثلاثة حالات: الأولى عندما كان عدد فترات القدرة تسع فترات، وهو الخيار المقترح (default) من قبل برنامج (Bilog-Mg)، وهو ما تمت الإجابة عنه في السؤال الأول من أسئلة الدراسة. والحالة الثانية تم تحديد عدد فترات القدرة بست فترات، والحالة الثالثة تم تحديد عدد فترات القدرة باثنتي عشرة فترة؛ وذلك لدراسة أثر تغير عدد فترات القدرة على معالم الفقرة ومؤشر المطابقة والعلاقة بينها في حال كان عدد هذه الفترات أقل أو أكثر من العدد المقترح من قبل برنامج (Bilog-Mg). وهو تسع فترات. أظهرت النتائج أن نسب عدم المطابقة اختلفت باختلاف عدد فترات القدرة، حيث بلغت نسبة الفقرات غير المطابقة عندما تم تقسيم متصل القدرة إلى ست فترات (14%)، بينما بلغت هذه النسبة (21%) عندما كان عدد الفترات مساوياً لتسع فترات، وهو الخيار المقترح من قبل برنامج (Bilog-Mg)، في حين بلغت هذه النسبة عند تقسيم متصل القدرة اثنتي عشرة فترة (31%). ويتضح مما سبق التأثير الواضح لمؤشر المطابقة (χ^2) بعدد فترات متصل القدرة؛ فقد أظهرت الفقرات ميلاً للمطابقة يزداد مع قلة عدد الفترات لمتصل القدرة، وهذا يتوافق مع ما ذكره موراكي (Muraki, 1997) بأن ازدياد عدد المجموعات الفرعية الممثلة لمتصل القدرة يعمل على تجزئة التكرارات المتوقعة بشكل كبير، مما يؤدي إلى تضخم قيمة مؤشر المطابقة (χ^2)، وبالتالي فشل الفقرة في مطابقة النموذج. ومن ثم تم حساب معامل الارتباط بين معالم الفقرة الثلاث، ومؤشر مطابقة الفقرة (χ^2) تبعاً لتغير عدد الفترات (6، 9، 12) باستخدام الحزمة الإحصائية (SPSS).

ويوضح الجدول (3) العلاقة بين معالم الفقرات الثلاث: التمييز (a)، والصعوبة (b)، والتخمين (c)، ومؤشر المطابقة (χ^2) عندما كان عدد فترات القدرة (6) فترات.

يشير الجدول (2) إلى وجود بعض الفقرات السهلة كما هو الحال في الفقرات (1، 2، 3، 4، 6)، وبعض الفقرات متوسطة الصعوبة مثل الفقرات (9، 10، 16، 11، 18، 21، 25، 26، 24). كما يبين الجدول قيم تمييز فقرات الاختبار، إذ تمتعت بعض الفقرات بتمييز عالٍ مثل الفقرات (10، 18، 26). كما يقدم هذا الجدول إحصاءات مطابقة الفقرة إذ تشير القيم الاحتمالية التي تقل عن (0,05) إلى الفقرات غير المطابقة للنموذج ثلاثي المعلمة (Haley et al., 2009). وتبعاً لذلك تعد الفقرات (1، 2، 3، 4، 7، 8، 10، 12، 13، 14، 15، 17، 18، 19، 20، 21، 23، 24، 25، 26، 27، 28، 29) هي الفقرات المطابقة للنموذج الثلاثي المعلمة. وتعزى مطابقة هذا العدد من الفقرات (23) فقرة من أصل (29) فقرة؛ أي ما نسبته (79%) إلى النموذج المستخدم، وهو النموذج ثلاثي المعلمة الذي يأخذ اختلاف قيم صعوبة وتمييز وتخمين الفقرات بالاعتبار؛ مما يجعله نموذجاً ناجحاً من الناحية العملية والتطبيقية. كما يمكن عزو ذلك لمناسبة عدد الفترات المكونة لمتصل القدرة والبالغ عددها تسع فترات، وهي قيمة مقترحة من قبل برنامج (Bilog-Mg)؛ إذ إن مؤشر المطابقة يعتمد على عدة عوامل، أحدها عدد فترات متصل القدرة. ومن الجدير بالذكر إنه لا يوجد اتفاق حول عدد مثالي لعدد فترات متصل القدرة بمراجعة الأدبيات النظرية في هذا الخصوص إذ يتم اختيار عددها من قبل الباحثين اعتباطياً (Reise & Waller, 1990). كما يبين الجدول الفقرات التي لم تطابق النموذج ثلاثي المعلمة، وهي الفقرات (5، 6، 9، 11، 16، 22)، وتشكل الفقرات غير المطابقة للنموذج ما نسبته (21%) من فقرات الاختبار ككل.

وقد تم تقدير خصائص الاختبار بصورة عامة والمكون من (29) فقرة، حيث تمتع بصعوبة كلية مقدارها (-0,434)، وهو مستوى متوسط من الصعوبة، كما تمتع بتمييز كلي مقداره (1,405) وقد بلغ معامل تخمين الاختبار (0,2)، وهي قيم جيدة عملياً تبعاً لما تم ذكره من قبل بيكر (Baker, 2001).

يبين الجدول (5) وجود علاقة سالبة ضعيفة غير دالة إحصائياً بين مؤشر المطابقة (χ^2) ومعلم التمييز (a) على خلاف النتائج المستخرجة للعلاقة بينهما عندما كان عدد فترات القدرة (6) و(9) فترات، بينما كانت العلاقة بين مؤشر المطابقة (χ^2) وبين كل من معلمي الصعوبة (b) والتخمين (c) موجبة ولكنها ضعيفة، إلا أن جميعها غير دال إحصائياً، وهذا يتفق مع النتيجتين السابقتين عندما كان عدد فترات القدرة (6) و(9) فترات. يتبين مما سبق أن العدد الأنسب للفترات كان ست فترات، وذلك ما بينته النسب المئوية للفترات غير المطابقة، إضافة إلى قيم معامل الارتباط بين معالم الفقرة الثلاث ومؤشر المطابقة (χ^2). وتعزى مثل هذه النتيجة إلى ما أشار إليه مورافي (Muraki, 1997) بتضخم قيمة مؤشر المطابقة في حال ازدياد عدد الفترات للقدرة، وبالتالي فشل الفقرة في المطابقة، وهذا قد يقود إلى تساؤلات بحثية جديدة فيما يخص دراسة عدد الفترات لمتصل القدرة، وأثر ذلك على مؤشر المطابقة للفقرة، وارتباط ذلك بقيمة الخطأ من النوعين الأول والثاني. كما يمكن عزو ذلك لحجم عينة الدراسة الذي لا يعد كبيراً جداً؛ مما أدى لمثل هذه النتيجة خاصة عند تقسيم المتصل لاثنتي عشرة فترة، مما يدعو لدراسة أثر اختلاف حجم العينة وعدد فترات القدرة على مؤشر المطابقة وعلاقته بمعالم الفقرة من قبل الباحثين.

وتبعاً لما أشار إليه كثير من الباحثين، مثل جارسون واوسبوم وواترز (Garson, 2012; Osborne & Waters, 2002) أنه ولإثبات افتراض الخطية أو دحضه فإن ما يسمى بطرق التحويل (Method of Transformations) تعد من أهم وأنجح الطرق لاختبار افتراض الخطية. وتقوم هذه الطريقة بعمل تحويلات غير خطية للمتغيرين المستقل أو التابع، أو كليهما لإثبات خطية العلاقة بينهما، مما يتيح الحصول على معادلة انحدار خطية فيما بعد التي يمكن اعتبارها كمقدر مقبول للعلاقة الحقيقية بين المتغيرات والتي تتكون من المتغيرات المحولة والتابعة. والتحويلات الأكثر شهرة هي التحويل اللوغاريتمي الطبيعي، والجذر التربيعي، والقوة التربيعية. وتقوم عملية اختيار نوع التحويل المناسب على أسلوب المحاولة والخطأ، إضافة إلى الاعتماد على معايير هامة للحكم على نجاحها، إذ يتم الأخذ بالاعتبار معاملات الارتباط، ودالاتها الإحصائية، والتمثيل البياني عند دراسة العلاقة، وقيمة معامل التحديد (R^2) الذي يعد معياراً أساسياً لنجاح عملية التحويل عند دراسة الانحدار الخطي في حال الزيادة في مقداره، بينما يتم الحكم على هذه العملية بعدم النجاح في حال عدم التغير أو النقصان في قيمة (R^2).

وقد تم اختبار عدة أشكال من التحويلات للمتغيرات المستقلة، إلا أن التحويل اللوغاريتم الطبيعي للمتغيرات المستقلة أظهر علاقة خطية أكثر وضوحاً بين كل من المتغيرات المستقلة (معالم الفقرة)، والتابعة (مؤشر مطابقة الفقرة) عندما ساوى عدد فترات القدرة ست وتسع فترات، بينما لا يمكننا قول ذلك عندما كان عدد الفترات اثنتي عشرة فترة، وذلك باستخدام معامل ارتباط

جدول (3): مصفوفة معاملات ارتباط معالم الفقرة مع مؤشر المطابقة للفترات عند تقسيم القدرة لست فترات.

| | χ^2 | a | b | c |
|----------|----------|--------|---------|--------|
| χ^2 | - | 0.384* | 0.044 | 0.184 |
| a | | - | 0.638** | -0.284 |
| b | | | - | -0.202 |
| c | | | | - |

يبين الجدول (3) وجود علاقة سالبة ضعيفة دالة إحصائياً بين مؤشر المطابقة (χ^2)، ومعلم التمييز (a)، بينما كانت العلاقة بين مؤشر المطابقة (χ^2)، وبين كل من معلمي الصعوبة (b) والتخمين (c) موجبة ولكنها ضعيفة وغير دالة إحصائياً.

كما تم استخراج معاملات الارتباط بين معالم الفترات الثلاث: التمييز (a)، والصعوبة (b)، والتخمين (c)، ومؤشر المطابقة (χ^2) عندما كان عدد فترات القدرة (9) فترات، وهو العدد المقترح (default) من قبل برنامج (Bilog-Mg)، كما هو موضح في الجدول (4).

جدول (4): مصفوفة معاملات ارتباط معالم الفقرة مع مؤشر المطابقة للفقرة عند تقسيم القدرة لتسع فترات.

| | χ^2 | a | b | c |
|----------|----------|---------|---------|--------|
| χ^2 | - | -0.342* | 0.001 | 0.148 |
| a | | - | 0.637** | -0.269 |
| b | | | - | -0.179 |
| c | | | | - |

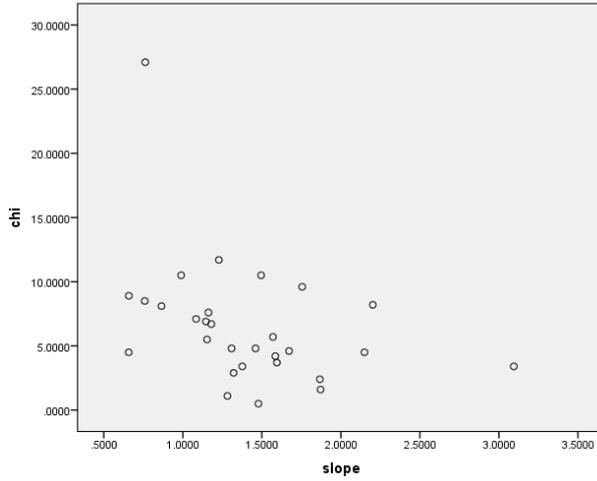
يبين الجدول (4) تقارب النتائج في الجدولين السابقين (3) و(4)، حيث يبين الجدول (4) وجود علاقة سالبة ضعيفة دالة إحصائياً بين مؤشر المطابقة (χ^2) ومعلم التمييز (a)، بينما كانت العلاقة بين مؤشر المطابقة (χ^2) وبين كل من معلمي الصعوبة (b) والتخمين (c) موجبة ولكنها ضعيفة وغير دالة إحصائياً، إلا أن العلاقة كانت أكبر بشكل طفيف عندما بلغ عدد فترات القدرة ست فترات.

كما تم استخراج معاملات الارتباط بين معالم الفترات الثلاث: التمييز (a)، والصعوبة (b)، والتخمين (c) ومؤشر المطابقة (χ^2) عندما كان عدد فترات القدرة (12) فترة، وذلك ما يوضحه الجدول (5).

جدول (5): مصفوفة معاملات ارتباط معالم الفقرة مع مؤشر المطابقة للفترات عند تقسيم القدرة لاثنتي عشرة فترة.

| | χ^2 | a | b | c |
|----------|----------|-------|---------|--------|
| χ^2 | - | -0.22 | 0.111 | 0.008 |
| a | | - | 0.638** | -0.284 |
| b | | | - | -0.202 |
| c | | | | - |

يتضح من الجدول (6) عدم وجود علاقة ارتباطية دالة إحصائياً بين مؤشر مطابقة الفقرة (χ^2)، وكل من معلمي الصعوبة والتخمين المحولين، بينما وجدت علاقة دالة إحصائياً بين مؤشر المطابقة (χ^2)، ومعلم التمييز المحول (loga) كما تبين سابقاً إلا أن هذه العلاقة ازدادت في قوتها، إضافة إلى أنها أصبحت دالة عند مستوى الدلالة الإحصائية (0,01)، كما أن التمثيل البياني التحويل اللوغاريتمي في توضيح العلاقة بين معلم التمييز الذي امتلك دلالة إحصائية للعلاقة بينه وبين مؤشر مطابقة الفقرة، وهذا ما يوضحه الشكلان (1) و(2).



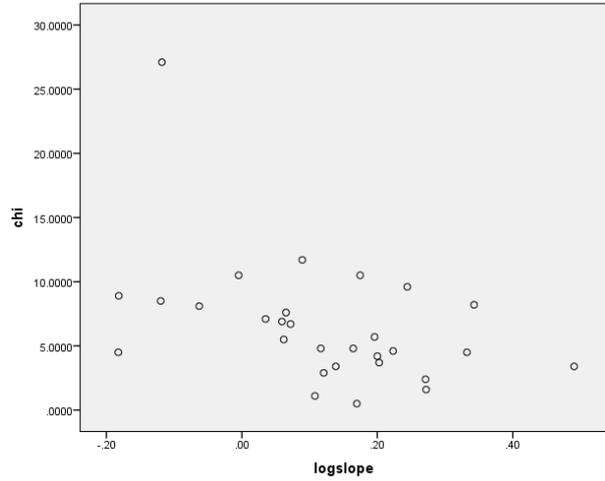
شكل (2): العلاقة بين مؤشر المطابقة والتمييز عند ست فترات

يتضح من الجدول (7) عدم وجود علاقة ارتباطية دالة إحصائياً بين مؤشر مطابقة الفقرة (χ^2) وكل من معلمي الصعوبة والتخمين المحولين، بينما وجدت علاقة دالة إحصائياً بازدياد طفيف بين مؤشر المطابقة (χ^2) ومعلم التمييز المحول (loga) كما تبين سابقاً، إلا أن التمثيل البياني (graphical method) أظهر فعالية التحويل اللوغاريتمي في توضيح العلاقة بين معلم التمييز الذي امتلك دلالة إحصائية للعلاقة بينه وبين مؤشر مطابقة الفقرة، وهذا ما يوضحه الشكلان (3) و(4).

بيرسون. ويبين الجدول (6) مصفوفة معاملات الارتباط لمؤشر مطابقة الفقرة (χ^2) ومعالمها المحولة عند تقسيم متصل القدرة لست فترات.

جدول (6): مصفوفة معاملات ارتباط معالم الفقرة المحولة مع مؤشر المطابقة للفقرات المطابقة لست فترات.

| | χ^2 | Loga | Logb | Logc |
|----------|----------|----------|---------|---------|
| χ^2 | - | -0.438** | 0.051- | 0.117 |
| Loga | | - | 0.730** | -0.364* |
| Logb | | | - | -0.257 |
| Logc | | | | - |

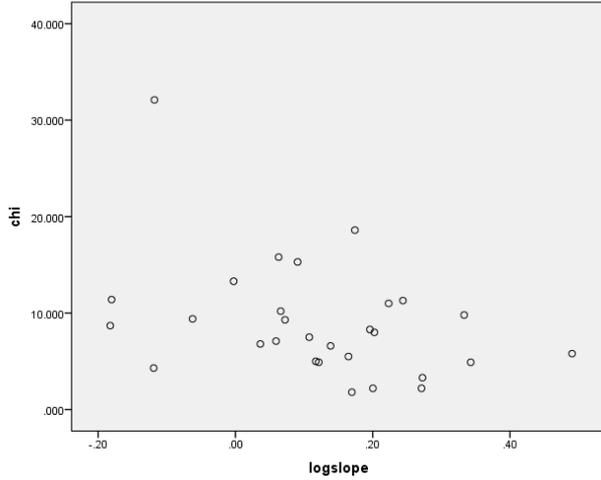


شكل (1): العلاقة بين مؤشر المطابقة ولوغاريتم التمييز عند ست فترات

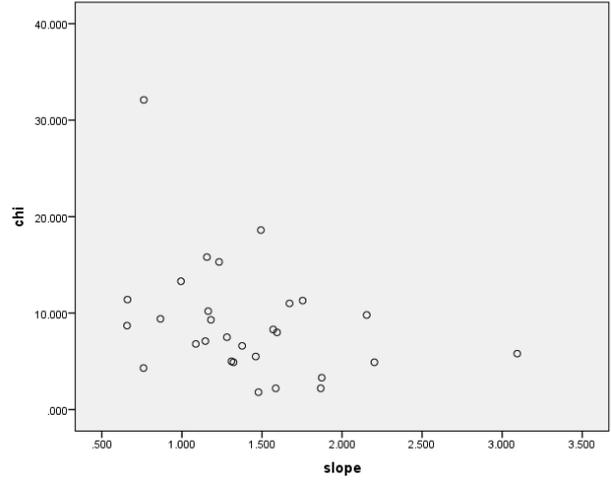
كما تم استخراج مصفوفة معاملات الارتباط لمؤشر مطابقة الفقرة (χ^2) ومعالمها المحولة عندما كان عدد فترات القدرة تسع فترات والموضحة في الجدول (7).

جدول (7): مصفوفة معاملات ارتباط معالم الفقرة المحولة مع مؤشر المطابقة للفقرات المطابقة لتسع فترات.

| | χ^2 | Loga | Logb | Logc |
|----------|----------|---------|---------|--------|
| χ^2 | - | -0.372* | 0.019- | 0.079 |
| Loga | | - | 0.671** | -0.201 |
| Logb | | | - | -0.351 |
| Logc | | | | - |



شكل (4): العلاقة بين مؤشر المطابقة ولوغاريتم التمييز لتسع فترات



شكل (3): العلاقة بين مؤشر المطابقة والتمييز لتسع فترات.

مطابقتها قد يرجع إلى أن نظرية استجابة الفقرة تهتم بتنوع صعوبة الفقرات وتظهر قوة بعكس التقليدية عند طرفي التوزيع؛ أي أن إحدى نقاط قوتها قدرتها على التعامل مع الفئتين العليا والدنيا من المفحوصين. وقد لوحظ أن الفقرات المطابقة كانت في الغالب تمتلك قوة تمييزية تزداد بازدياد الصعوبة نسبياً واقتربها من الصفر؛ أي مقدار متوسط من الصعوبة مع وجود قدر متدن ومقبول من التخمين. وتتفق هذه النتيجة مع نتائج دراسة كل من سينهاري ولو (Sinharay & Lu, 2008) للبيانات مولدة بوجود علاقة ارتباطية سالبة دالة إحصائياً بين مؤشر مطابقة الفقرة (χ^2) ومعلم التمييز المقدر، ووجود علاقة ارتباطية موجبة بين معلمي الصعوبة والتخمين، إلا أن الدلالة الإحصائية للعلاقة الأخيرة اقتضت على معلم الصعوبة، وهذا يخالف نتيجة الدراسة الحالية، ولعل السبب في مثل هذه النتيجة اختلاف طبيعة البيانات، حيث استخدم الباحثان بيانات مولدة بتوزيع طبيعي لوغاريتمي، إضافة إلى اختلاف حجم العينة، وعدد الفقرات. كما اتفقت نتيجة الدراسة الحالية جزئياً مع النتائج التي خلص لها كل من روجرز وهاتي (Rogers & Hattie, 1987) بعدم مطابقة الفقرات ذات التخمين المرتفع، كما اتفقت مع نتائج دراسة ريس (Reise, 1990) بعدم وجود علاقة بين مؤشر المطابقة ومعلم الصعوبة للفقرة، إلا أنها اختلفت مع النتائج التي خلصت لها دراسة دودين (Dodeen, 2004) بوجود علاقة طردية بين مؤشر المطابقة ومعلم التمييز، التي كانت موضع جدل؛ بسبب تعارضها مع أساسيات بناء الاختبار، حيث تفيد برفض الفقرات ذات التمييز المرتفع باعتبارها فقرات غير مطابقة. كما تعارضت نتيجة هذه الدراسة مع نتيجة كل من ماونت وشوماخر (Mount & Schumacker, 1998) اللذين اعتبرا أن لا علاقة بين التخمين وعدد الفقرات المطابقة، ونتيجة سيول (Seol, 1999) بعدم وجود علاقة خطية دالة إحصائياً بين معالم الفقرة ومؤشر مطابقتها، إضافة إلى تعارضها مع الجزء التجريبي من النتيجة التي خلص لها سينهاري ولو (Sinharay & Lu, 2008) بأن العلاقة بين مؤشر المطابقة وكل من معلمي الصعوبة والتخمين كانت سالبة، بينما

يوضح الشكلان (3) و(4) اتضاح العلاقة الخطية بين مؤشر المطابقة (χ^2) والتحويل لتمييز الفقرة (log a) وهذا ما تؤكدته الزيادة في قيمة معامل ارتباط بيرسون والدالة إحصائية.

كما تم استخراج مصفوفة معاملات الارتباط لمؤشر مطابقة الفقرة (χ^2) ومعالمها المحولة عندما كان عدد فترات القدرة اثنتي عشرة فترة والموضحة في الجدول (8).

جدول (8): مصفوفة معاملات ارتباط معالم الفقرة المحولة مع مؤشر المطابقة للفقرات لاثنتي عشرة فترة.

| | χ^2 | Loga | Logb | Logc |
|----------|----------|--------|---------|---------|
| χ^2 | - | -0.216 | 0.100 | -0.048 |
| Loga | | - | 0.730** | -0.364* |
| Logb | | | - | -0.257 |
| Logc | | | | - |

يبين الجدول (8) أن العلاقة بين مؤشر مطابقة الفقرة (χ^2) ومعالمها المحولة عند تقسيم متصل القدرة لاثنتي عشرة فترة قد قلت بشكل طفيف، وذلك يدل على عدم كفاءة التحويل في هذه الحالة. وقد تعزى النتائج السابقة لسبب منطقي، وهو أنه كلما ارتفعت القوة التمييزية للفقرة وكانت الفقرة ذات صعوبة متوسطة ازدادت نسبة قبولها ومطابقتها، وبذلك قلت قيمة مؤشر (χ^2)؛ الأمر الذي يرفع من قيمة الدلالة الإحصائية للفقرة فتكون بذلك الفقرة مطابقة ومقبولة. كما يمكن عزو ذلك أيضاً لمناسبة عدد فترات القدرة؛ إذ تقل احتمالية فقدان الفقرات وفشلها في المطابقة نتيجة التضخم غير المبرر في قيم مؤشر المطابقة؛ مما يجعل من العلاقة أكثر وضوحاً. بينما كانت العلاقة بين مؤشر المطابقة (χ^2) ومعلم التخمين (c) موجبة ولكنها ضعيفة وغير دالة. وتعزى هذه النتيجة لأن الفقرة ذات التخمين المرتفع هي فقرة غير مقبولة، وبذلك تمتلك قيمة مرتفعة لمؤشر (χ^2)؛ أي تقل قيمة الدلالة الإحصائية لها؛ مما يجعلها فقرة غير مطابقة للنموذج المختار. كما يبين وجود علاقة موجبة ضعيفة وغير دالة إحصائياً بين صعوبة الفقرة ومؤشر

إحصائياً، بينما فشل معلما التخمين (c) والصعوبة (b) في التنبؤ بمؤشر مطابقة الفقرة (χ^2)، ويعزى ذلك إلى ارتباط مؤشر المطابقة مع معلم التمييز، وهو الأمر الذي لم يتحقق لمعلمي الصعوبة والتخمين. وبذلك تكون معادلة الانحدار لتنبؤ المعلم بمؤشر المطابقة (χ^2) كالاتي:

$$\text{مؤشر مطابقة الفقرة } (\chi^2) = - 3,637 \text{ (تمييز الفقرة) } + 11,625$$

كما يمكن التعبير عنها باستخدام معادلة الانحدار المعيارية كالتالي:

$$\text{مؤشر مطابقة الفقرة } (\chi^2) = - 0,384 \text{ (تمييز الفقرة)}$$

كما تم استخراج نموذج الانحدار المكون من المعلم المحولة باستخدام الاقتران اللوغاريتمي في التنبؤ بمؤشر مطابقة الفقرة (χ^2) الذي أثبت نجاح معلم الفقرة المحولة والصعوبة والتمييز في التنبؤ بمؤشر مطابقتها، وذلك عند تقسيم متصل القدرة لست فترات، حيث بلغت قيمة معامل التحديد ($R^2=0.35$). ويعد ذلك دليلاً آخر على فعالية التحويل اللوغاريتمي في كشف العلاقة بين مؤشر مطابقة الفقرة ومعالمها، وهذا ما يبيئه الجدول (11).

جدول (11): تحليل التباين (ANOVA) لنموذج انحدار معلم الفقرة المحولة على مؤشر المطابقة (χ^2) عند ست فترات.

| Model | Sum of squares | DF | Mean Square | F |
|------------|----------------|----|-------------|--------|
| Regression | 235.262 | 2 | 117.631 | 6.872* |
| Residual | 445.04 | 26 | 17.117 | |
| Total | 680.301 | 28 | | |

كما يبين الجدول (12) معاملات الانحدار لنموذج الانحدار المكون من معلمي التمييز والصعوبة المحولين كمتنبئات بمؤشر مطابقة الفقرة (χ^2).

جدول (12): قيم معاملات الانحدار الخام والمعيارية ودالاتها لنموذج الانحدار لمعلم الفقرة (log b, log a) عند ست فترات

| Model | B | β | T |
|----------|---------|---------|----------|
| Log a | -26.786 | -0.858 | -3.693** |
| Log b | 24.075 | 0.575 | 2.476* |
| Constant | 0.2 | | 0.059 |

يبين الجدول (12) قدرة معلمي التمييز المحول (log a)، والصعوبة المحول (log b) في التنبؤ بمؤشر مطابقة الفقرة (χ^2)، حيث كانت معاملات الانحدار الخاصة به دالة إحصائياً، بينما فشل معلم التخمين المحول (log c) في التنبؤ بمؤشر مطابقة الفقرة (χ^2). مما سبق يلاحظ نجاح اللوغاريتم الطبيعي للصعوبة (log b) الذي لم ينجح كمتنبئ قبل عملية التحويل، كما أن معامل الارتباط بينه وبين مؤشر المطابقة لم يمتلك دلالة إحصائية. وقد تعزى مثل هذه النتيجة للارتباط القوي الدال إحصائياً بينه وبين معلم التمييز الذي بلغ (0.73)؛ مما أدى إلى حدوث تفاعل بين المتغيرين أدى

كانت موجبة لمعلم التمييز، وقد يعزى ذلك لاختلاف حجم العينة بشكل كبير بين الدراستين إضافة إلى اختلاف متوسطات معالم الاختبار وخصائص الفقرات.

خامساً: النتائج المتعلقة بالسؤال الثالث: الذي ينص على "هل يمكن استخدام معالم الفقرة الثلاث (الصعوبة، التمييز، التخمين) كمتنبئات لمطابقة الفقرة عند مستوى الدلالة $\alpha = 0.05$ "؟

للإجابة عن هذا السؤال تم استخراج معادلة الانحدار المتعدد (Multiple Regression Equation) باستخدام الحزمة الإحصائية (SPSS) لمعرفة قدرة معالم الفقرة على التنبؤ بمؤشر مطابقتها (χ^2). ولتحقيق هذا الغرض تم إجراء تحليل الانحدار مرتين الأولى عندما بلغ عدد فترات متصل القدرة ست فترات والثانية تسع فترات، حيث أظهر مؤشر المطابقة (χ^2) دلالة إحصائية لارتباطه بمعلم التمييز سابقاً، وقد تم إجراء تحليل تباين للنموذج الكلي للانحدار لاستخراج الدلالة لنموذج الانحدار المكون من مؤشر المطابقة (χ^2) كمتغير تابع ومعالم الفقرات كمتغير مستقل باستخدام طريقة (Backward method) للخروج بمعادلة انحدار تضم فقط المتغيرات الأكثر قدرة على التنبؤ، إذ تقوم هذه الطريقة على إدخال المتغيرات كاملة لمعادلة التنبؤ في البداية، ثم إقصاء المتغيرات التي تحمل أقل قيمة (F-to- Remove)، والانتهاؤ بمعادلة انحدار تضم المتغيرات المتبينة بالمتغير التابع بشكل مقبول (Flom & Cassell, 2007).

وقد أظهرت النتائج نجاح نموذج الانحدار المكون من معلم التمييز غير المحول في التنبؤ بمؤشر مطابقة الفقرة (χ^2)، وذلك عند تقسيم متصل القدرة لست فترات، حيث بلغت قيمة معامل التحديد ($R^2=0.15$) وهذا ما يبيئه الجدول (9).

جدول (9): تحليل التباين (ANOVA) لنموذج انحدار معلم الفقرة على مؤشر المطابقة (χ^2) عند ست فترات

| Model | Sum of squares | DF | Mean Square | F |
|------------|----------------|----|-------------|--------|
| Regression | 100.52 | 1 | 100.52 | 4.681* |
| Residual | 579.781 | 27 | 21.473 | |
| Total | 680.301 | 28 | | |

كما يبين الجدول (10) معاملات الانحدار لنموذج الانحدار المكون من معلم التمييز كمتنبئ بمؤشر مطابقة الفقرة (χ^2).

جدول (10): قيم معاملات الانحدار الخام والمعيارية ودالاتها لنموذج الانحدار لمعلم الفقرة (b, a)

| Model | B | β | T |
|----------|--------|---------|---------|
| A | -3.637 | -0.384 | -2.164* |
| Constant | 11.625 | | 4.627* |

يبين الجدول (10) قدرة معلم التمييز (a) في التنبؤ بمؤشر مطابقة الفقرة (χ^2)، حيث كانت معاملات الانحدار الخاصة به دالة

يبين الجدول (14) قدرة كل من اللوغاريتم الطبيعي لمعلم الصعوبة ($\log b$) واللوغاريتم الطبيعي لمعلم التمييز ($\log a$) في التنبؤ بمؤشر مطابقة الفقرة (χ^2)، حيث كانت معاملات الانحدار الخاصة بهما دالة إحصائياً، بينما فشل معلم التخمين المحول ($\log c$) في التنبؤ بمؤشر مطابقة الفقرة (χ^2). وكما لوحظ سابقاً فإن معلم الصعوبة المحول ($\log b$) لم يمتلك دلالة إحصائية للعلاقة مع مؤشر المطابقة للفقرة (χ^2)، بينما نجح كمتنبئ عندما كان عدد الفترات ست وتسع فقرات، وعزّي هذا لوجود تفاعل بين معلم الصعوبة والتمييز كما ذكر سابقاً نتج عن الارتباط القوي الدال إحصائياً فيما بينهما، كما يمكن إرجاعه إلى وجود ما يسمى بالمتغيرات المشوشة (Suppressor Variables) التي تؤثر في قيمة معامل الارتباط بالزيادة أو النقصان. وفي هذه الحالة ترجح الباحثة أن التأثير كان سالباً، إذ وجد أن الفقرات ذات الصعوبة المتوسطة كانت أكثر ميلاً للقبول كفقرات مطابقة. ولمعرفة مثل هذه المتغيرات توصي الباحثة بعمل دراسات تحليلية سببية (Causal Analysis) للوقوف على المتغيرات التي قد تؤثر في مؤشر المطابقة وطبيعة علاقتها مع بعضها بعضاً.

وبذلك فإنه يمكن التنبؤ بمؤشر مطابقة الفقرة (χ^2) من خلال معادلة الانحدار الخام كالتالي:

$$\text{مؤشر مطابقة الفقرة } (\chi^2) = 27,486 (\text{اللوغاريتم الطبيعي لصعوبة الفقرة}) - 29,544 (\text{اللوغاريتم الطبيعي لتمييز الفقرة}) + 1,673$$

كما يمكن التعبير عنها باستخدام معادلة الانحدار المعيارية كالتالي:

$$\text{مؤشر مطابقة الفقرة } (\chi^2) = 0,539 (\text{صعوبة الفقرة}) - 0,765 (\text{اللوغاريتم الطبيعي لتمييز الفقرة}).$$

يتضح من المعادلتين السابقتين أن معلم التمييز المحول ($\log a$) كان الأفضل في التنبؤ بمؤشر مطابقة الفقرة (χ^2)، حيث امتلك معاملات انحدار خام ومعيارية أكبر؛ مما يعني إسهامه بشكل أكبر في التنبؤ بمؤشر المطابقة بالمقارنة مع صعوبة الفقرة، إلا أن التنبؤ يتم بشكل عكسي؛ وذلك لأن معامل الانحدار يحمل إشارة سالبة، وهذا يتواءم مع قيمة معامل الارتباط السالبة بين معلم التمييز ومؤشر مطابقة الفقرة.

يتضح مما سبق أنه عند تقسيم متصل القدرة إلى ست فترات كانت معالم الفقرة المحولة هي الأفضل في التنبؤ بمؤشر مطابقة الفقرة (χ^2)، كما كان معلم التمييز المحول ($\log a$) الأكثر قدرة على التنبؤ، حيث بلغ معامل التحديد ($R^2=0.35$) وهي القيمة الأعلى بالمقارنة مع نماذج الانحدار الأخرى التي تمت دراستها.

لظهور مثل هذه النتيجة. وقد لوحظ ازدياد في قيم معامل الانحدار لمعلم التمييز المحول ($\log b$) كما أن دلالاته الإحصائية ازدادت قوة، ويعزى ذلك إلى ارتباط مؤشر المطابقة مع معلم التمييز المحول الذي ازداد وأصبح أكثر وضوحاً بعد إجراء التحويل اللوغاريتمي، وهو الأمر الذي لم يتحقق في معلمي الصعوبة والتخمين. وبذلك تكون معادلة الانحدار لتنبؤ المعالم بمؤشر المطابقة (χ^2) كالآتي:

$$\text{مؤشر مطابقة الفقرة } (\chi^2) = 24,075 (\text{اللوغاريتم الطبيعي لصعوبة الفقرة}) - 26,786 (\text{اللوغاريتم الطبيعي لتمييز الفقرة}) + 0,2$$

كما يمكن التعبير عنها باستخدام معادلة الانحدار المعيارية كالتالي:

$$\text{مؤشر مطابقة الفقرة } (\chi^2) = 0,575 (\text{صعوبة الفقرة}) - 0,858 (\text{اللوغاريتم الطبيعي لتمييز الفقرة})$$

ويتبين مما سبق أن القدرة التنبؤية لمعلم التمييز المحول أقوى منها لمعلم الصعوبة المحول.

كما أظهرت النتائج فشل نموذج الانحدار المكون من المعالم غير المحولة في التنبؤ بمؤشر مطابقة الفقرة (χ^2) عند تقسيم متصل القدرة إلى تسع فترات، حيث بلغت قيمة الاحصائي ($F=3.19$) التي لم تمتلك دلالة إحصائية لنموذج الانحدار الكلي. بينما نجحت المعالم المحولة لكل من معلمي الصعوبة والتمييز بالتنبؤ بمؤشر مطابقة الفقرة (χ^2) وهذا ما يبينه الجدول (13).

جدول (13): تحليل التباين (ANOVA) لنموذج انحدار معالم الفقرة المحولة على مؤشر المطابقة (χ^2) لتسع فترات.

| Model | Sum of squares | DF | Mean Square | F |
|------------|----------------|----|-------------|--------|
| Regression | 283.980 | 2 | 141.990 | 4.905* |
| Residual | 752.628 | 26 | 28.947 | |
| Total | 1036.608 | 28 | | |

يبين الجدول (13) نجاح نموذج الانحدار النهائي في التنبؤ بالمتغير التابع، وقد تكون هذا النموذج من اللوغاريتم الطبيعي لصعوبة الفقرة ($\log b$)، واللوغاريتم الطبيعي لتمييز الفقرة ($\log a$) كمتغيرات مستقلة مؤشر المطابقة (χ^2) كمتغير تابع، حيث بلغ مقدار التباين المفسر من قبلهما ($R^2=0.274$) ويبين الجدول (14) معاملات الانحدار للنموذج ودلالاتها الإحصائية.

جدول (14): قيم معاملات الانحدار الخام والمعيارية ودلالاتها لنموذج الانحدار لمعلم الفقرة ($\log b$, $\log a$).

| Model | B | β | T |
|----------|---------|---------|---------|
| Log b | 27.486 | 0.539 | 2.204* |
| Log a | -29.544 | -0.765 | -3.130* |
| Constant | 1.673 | | 0.390 |

Haley, S.; Pinkham, M.; Dumas, H.; Ni, P.; Gorton, G.; Watson, K.; Montpent, K.; Bilodeau, N.; Hambleton, R. & Tucker, C. (2009). Evaluation of Item Bank for a Computerized Adaptive Test of Activity in Children with Cerebral Palsy. *Physical Therapy*, 89, 6, 589-600.

Hambleton, R. K., Swaminathan, H. (1985). *Item Response Theory: Principles and Applications*. Norwell, MA: Kluwer-Nijhoff Publishing.

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

Maydeu, A. (2005). Further empirical results on parametric versus nonparametric IRT modeling of Likert-type personality data. *Multivariate Behavioral Research*, 40, 261-279.

Mount, R. E., & Schumacker, R. E. (1998). Identifying measurement disturbance effects using Rasch item fit statistic and the logit residual index. *Journal of Outcomes Measurement*, 2(4), 338-350.

Muraki, E. (1997). A generalized partial credit model. In W. van der Linden & R. K. Hambleton (Eds.), *Handbook of modern item response theory*. New York: Springer-Verlag

Nunally, J. (1978). *Psychometric Theory*, USA: McGraw-Hill Book Company. P 256-300.

Osborne, J. & Watera, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research & Evaluation*, 8(2).

Reise, S. (1990). A comparison of item and person fit method of assessing model data fit. *Applied Psychological Measurement*, 14 (2), 127-137.

Reise, S. & Rivicki, D. (2015). *Handbook of item response theory modeling*. Routledge: New York.

Reise, S. & Waller, N. (1990). Fitting the two parameter model to personality data. *Applied Psychological Measurement*, 14, 45-58.

Rogers, H. J., & Hattie, J. A. (1987). A Monte-Carlo investigation of several person and item person fit statistics for item response models. *Applied Psychological Measurement*, 11, 47-57.

Seol, H. (1999). Detecting differential item functioning with five standardized item-fit indices in the Rasch model. *Journal of Outcomes Measurement*, 3 (3), 233-247.

Sinharay, S. & Lu, Y. (2008). A further look at the correlation between item parameters and item fit statistics. *Journal of Educational Measurement*, 45 (1), 1-15.

Tay, L., Meade, M., & Smith, M. (2015). An overview and practical guide to IRT measurement equivalence analysis. *Organizational Research Methods*, 18, 3-46.

Wright, B. & Panchapakesan, N. (1969). A procedure for sample-free item analysis. *Educational & Psychological Measurement*, 29 (1), 23-48.

التوصيات

- في ضوء النتائج التي توصلت إليها الدراسة توصي الباحثة بما يلي:
1. دراسة أثر اختلاف حجم العينة وطول الاختبار في دراسة العلاقة بين مؤشر مطابقة الفقرة ومعالمها للنموذج ثلاثي المعلمة.
 2. دراسة العوامل المؤثرة في مطابقة الفقرة والبحث في مسببات عدم التطابق.
 3. دراسة العلاقة بين مؤشر مطابقة الفقرة ومعالمها باستخدام نماذج بارامترية ولابارامترية أخرى.
 4. دراسة حسن المطابقة البيانات ومسبباتها باستخدام أنواع أخرى من الاختبارات، مثل مقاييس الاتجاهات ومقاييس الشخصية.
 5. أهمية إجراء دراسات تحليلية سببية لمعرفة المتغيرات المؤثرة في مطابقة الفقرة إضافة إلى دراسة العلاقات البيئية بين هذه المتغيرات.
 6. دراسة أثر تغير عدد فترات متصل القدرة وأثر ذلك على مؤشرات مطابقة الفقرة، وأثر ذلك على الخطأ من النوعين الأول والثاني.

المراجع:

- عودة، أحمد. (2010) *القياس والتقويم في العملية التدريسية*، ط3، الأردن: دار الأمل للنشر والتوزيع.
- Baker, F. (2001). *The Basics of Item Response Theory*, 2nd ed USA: ERIC Clearinghouse on Assessment & Evaluation.
- Brown, J. D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall Regents.
- Capalleri, JC.; Lundy, Jason & Hays, RD. (2014). Overview of classical test theory and item response theory for the quantitative assessment of items in developing patient-reported outcomes measures, *Clin Ther*, 36(5), 648-662.
- De Ayala, R.J. (2009). *The Theory and Practice of Item Response Theory*. The Guilford Press, New York.
- Dodeen, H. (2004). The relationship between item parameter and item fit. *Journal of Educational Measurement*, 14 (3), 261-270.
- Flom, P. & Cassell, D. (2007). Stopping stepwise: Why stepwise and similar selection methods are bad, and what you should use. *The NorthEast SAS Users Group (NESUG)*, 2007.
- Fontanella, L.; Villano, P.; Donato, M. (2015). Attitudes towards Roma people and migrants: a comparison through a Bayesian multidimensional IRT model. *Quality & Quantity*, 49(1), 1-20.
- Garson, D. (2012). *Testing Statistical Assumptions*. Statistical Associates Publishing: NC.