المقارنة بين بعض مؤشرات المطابقة المتزايدة والمطابقة المطلقة وَفْق نظرية استجابة الفقرة باختلاف النموذج اللوغاريتمي وعدد الفقرات

إعـــــداد

د/ شادي يوسف الشواورة

أستاذ مساعد في القياس والتقويم
عميد شؤون الطلبة جامعة العقبة للتكنولوجيا
العقبة– الأردن

المقارنة بين بعض مؤشرات المطابقة المتزايدة والمطابقة المطلقة وَفْق نظرية استجابة الفقرة باختلاف النموذج اللوغاريتمي وعدد الفقرات

د/ شادي يوسف الشواورة*

المقدمة والإطار النظري:

يعد بناء الاختبارات ومعايرتها من أولويات حقل القياس والتقويم التربوي والنفسي؛ حيث يترتب على نتائج الاختبارات قرارات مهمة في إظهار الفروق في القدرات بين الأفراد؛ مما يتطلب من معدي الاختبارات توخي الموضوعية في القياس، والتي تعتمد بدورها على تحقيق الخصائص السيكومترية للاختبار. وأضافت النظرية الحديثة في القياس طرقا وأساليب جديدة في معايرة الفقرات والاختبارات؛ للوصول في القياس النفسي والتربوي إلى درجة عالية من الموضوعية.

تستند النظرية الحديثة في القياس إلى مجموعة من الافتراضات، والتي تُعَدّ ذات أهمية كبيرة، ويُعدّ انتهاكها تجاوزا خطيرا في مجالات القياس النفسي والتربوي، التي من شأنها تهديد صدق البناء للاختبارات المختلفة، ومن أهمها: أحادية البعد (Unidimensionality): ويعني هذا الافتراض أن أداء المفحوص على الاختبار يُعزى إلى قدرة واحدة ، أو سمة واحدة تفسر ذلك الأداء، ورغم أن هناك عوامل دخيلة على أداء المفحوص ،مثل: (قلق الاختبار، والدافعية ،وظروف تطبيق الاختبار...إلخ)، والتي قد تؤثر سلبا، إلا أن القدرة تبقى العامل الحاسم في نتيجة المفحوص. وأما الافتراض الثاني، فهو الاستقلال الموضعي) (local Inidependence، والذي يعني أن استجابات المفحوص للفقرات المختلفة في الاختبار مستقلة عن بعضها إحصائيا عند مستوى قدرة معين؛ أي أن أداء المفحوص على فقرة ما لا يتأثر سلبا ولا إيجابا بأدائه على الفقرات الأخرى والواردة في الاختبار.

* **د/ شادي يوسف الشواورة:** أستاذ مساعد في القياس والتقويم/ عميد شؤون الطلبة– جامعة العقبة للتكنولوجيا، العقبة، الأردن.

ويشير هامبلتون، وسومينثان (Hambleton & Swaminathan, 1991) إلى أن افتراض الاستقلال الموضعي يكافئ افتراض أحادية البعد، ويعني ذلك أنه إذا تحقق افتراض أحادية البعد في الاختبار، فإن الاختبار يحقق افتراض الاستقلال الموضعي. وبالإضافة إلى الافتراضين السابقين يجب هناك افتراضات أخرى، وهي: منحنى خصائص الفقرة (Item Characteristic Curve, ICC)، وهو اقتران رياضي يربط بين احتمال استجابة الفرد استجابة صحيحة على فقرة الاختبار، وبين القدرة التي تقيسها مجموعة الفقرات، أما افتراض التحرر من السرعة (Speediness) فيعني أن عامل السرعة في الأداء، ليس له تأثير على احتمالية الاستجابة بشكل صحيح عن فقرات الاختبار؛ أي أن إخفاق الفرد في الإجابة عن فقرات الاختبار يعود إلى انخفاض قدرته، وليس إلى تأثير السرعة في الإجابة.

وتُعدّ هذه الافتراضات غاية في الأهمية، وخاصة أحادية البعد، في دورها في الكشف عن السمة المقاسة، وأشار ستاوت (Stout, 1987) إلى أهمية أحادية البعد في التطبيقات الاختبارية المتعددة؛ للحفاظ على ترتيب الأفراد في المقياس الواحد، إضافة إلى تجنب التحيز الناشئ عن تقديرات معالم الفقرة وتقديرات قدرات الأفراد. وبين ووكر، وبرتراس (Walker & Beretras, 2000) أن النتائج غير الصحيحة لمعالم الفقرات والأفراد؛ بسبب زيادة خطأ القياس تُعزى إلى تطبيق بيانات متعددة الأبعاد في نموذج يفترض أحادية البعد ، وأشار زانج (Zhang, 2008)، إلى أن أحادية البعد تتأثر في عدد فقرات الاختبار، وأن الاختبارات القصيرة أكثر عرضة لانتهاك هذا الافتراض؛ مما يترتب على ذلك ضعف مؤشر صدق الاختبار، وقد بين الشواوره (٢٠١٣) أن طول الاختبار له الأثر في دقة تقدير معالم الفقرات باختلاف طريقة التقدير (طريقة الأرجحية العظمى الهامشية، وطريقة بييز)، ولقد شدد هامبلتون، وسواميناثان (Hambleton & Swaminathan, 1985) على وجوب تحقق افتراض أحادية البعد عند تطبيق نماذج نظرية استجابة الفقرة أحادية البعد؛ إذ إن انتهاك هذا الافتراض يؤثر في نتائج معادلة حساب أرجحية نمط استجابة المفحوصين.

أما هاتي (Hattie, 1985)، فقد أشار إلى ضرورة التحقق من فاعلية المؤشرات المستخدمة في الكشف عن أحادية البعد باختلاف ظروف أعداد

الاختبارات وتطبيقها، وأشار إلى خمسة مؤشرات تستخدم للكشف عن مدى توفر شرط أحادية البعد، وهي: نمط الاستجابة Answer Patterns، ومعاني الثبات Reliability، والمكونات الرئيسة Principal Components، والتحليل العاملي Factor Analysis، ومطابقة البيانات لأحد نماذج السمات الكامنة Latent Trait Models.

انبثقت نظرية استجابة الفقرة من افتراض إمكانية التنبؤ بأداء الأفراد أو تفسير أدائهم في اختبار ما، في ضوء خصائص مميزة لهذا الأداء تسمى القدرات، والتي لا يمكن ملاحظتها بشكل مباشر، وإنما من خلال أداء الأفراد في مجموعة من الفقرات؛ لذلك تهدف نظرية استجابة الفقرة إلى تحديد العلاقة الرياضية الاحتمالية بين أداء الفرد في اختبار ما، والسمة الكامنة وراء هذا الأداء، ويُعبر عن هذه العلاقة بالنماذج اللوغاريتمية (اللوجستية)، وتحاول هذه النماذج تقدير موقع الفرد (المفحوص) على متصل السمة أو القدرة باستخدام نمط استجاباته، وتقوم هذه النظرية على افتراض أساس مفاده أن احتمال استجابة الفرد على أي فقرة من فقرات الاختبار بشكل صحيح، يمثل اقتراناً لكل من قدرة الفرد التي يقيسها الاختبار من جهة وخصائص الفقرة التي يحاول ذلك الفرد الإجابة عنها من جهة ثانية؛ وهذا يتطلب الحصول على معلومات حول الفرد والفقرة؛ أي الحصول على تقديرات لقدرات الأفراد ومعالم الفقرات. (Crocker & Algina, 1986) وعلام (٢٠٠٥).

أشار بيكر (Baker, 2001) إلى النماذج الرياضية لنظرية استجابة الفقرة التي تختلف باختلاف معالم الفقرات المراد تقديرها،وتُعدّ هذه النماذج من أكثر نماذج نظرية استجابة الفقرة شيوعاً، وأكثرها ملاءمة مع الفقرات ثنائية التدريج، وتاليا توضيح للنماذج اللوغارتمية (اللوجستية) الثلاثة:

النموذج اللوغاريتمي أحادي المعلمة (One-Parameter Logistic Model): سعى واضع هذا النموذج، العالم الدنماركي جورج راش، إلى تحقيق الموضوعية في القياس النفسي التربوي، من خلال توصله إلى أدوات قياس اتسمت باستقلاليتها عن خصائص الأفراد المراد قياس قدرتهم، فضلاً عن استقلالية قدرة الأفراد المقاسة عن أدوات القياس ، ويوصف النموذج اللوجستي ،أحادي المعلمة بالمعادلة الآتية:

$$p_i(q) = \frac{e^{D\bar{a}_i(q-\mathbf{bi})}}{1+e^{D\bar{a}_i(q-\mathbf{bi})}}$$

Pi (θ) هي احتمال الإجابة الصحيحة للفرد j الذي قدرته jϴ على الفقرة i التي صعوبتها bi حيث إن (i=1 , 2,…… . ,n) ، و n هي عدد الفقرات. وأن D هو معامل التدريج اللوغاريتمي الذي يساوي واحد صحيح. وe هي الدالة الأسية ذات الأساس الطبيعي (2.71828)، وD ثابت التدريج (Sceling Fector)، ويساوي تقريباً (1.7).

النموذج اللوغاريتمي ثنائي المعلمة (Two-Parameter Logistic Model): يسمح النموذج ثنائي المعلمة باختلاف فقرات الاختبار؛ من حيث صعوبتها وتمييزها، ويمثل بالدالة الرياضية الآتية:

$$p_i(q) = \frac{e^{Dai(q-\mathbf{bi})}}{1+e^{Dai(q-\mathbf{bi})}}$$

النمـوذج اللوغـاريتمي ثلاثـي المعلمـة Three-Parameter Logistic (Model): ومن المشكلات التي واجهت النموذج اللوغاريتمي ثنائي المعلمة أن هنالك احتمالاً لإجابة الأفراد ذوي القدرة المنخفضة على الفقرة، وتكون نسبة إجابتهم أكبر من الصفر؛ وذلك بسبب عامل التخمين في الإجابات عن الفقرات الاختبارية. لذلك قام النموذج اللوغاريتمي الثلاثي على ثلاثة معالم وهي: معلمة الصعوبة (bi)، ومعلمة التمييز (ai)، ومعلمة التخمين (Ci)، ويحدد هذا المعلم على منحنى خصائص الفقرة في الجزء الأسفل منه، ويدعى خط التقارب الأدنى (Lower Asymptote) (Glas, Falcon 2003). ويصاغ النموذج الثلاثي على النحو التالي:

$$p_i(q) = Ci + (1-Ci)\frac{e^{Dai(q-bi)}}{1+e^{Dai(q-bi)}}$$

حيث إن $pi(q)$: احتمال إجابة المفحوص الذي اختير عشوائياً من مستوى القدرة (q)عن الفقرة (i) إجابةً صحيحة، bi: معلمة الصعوبة، ai: معلمة

التمييز، Ci: معلمة التخمين، q: معلمة القدرة D: عامل التدريج (Scaling Factor).

مؤشرات جودة المطابقة:

تأتي أهمية مؤشرات حسن المطابقة كمؤشر إحصائي؛ للكشف عن افتراض أحادية البعد وَفقا لنظرية استجابة الفقرة في القياس (Item Response Theory,(IRT))، وقد تم الإشارة إليها من قبل، وهو: (Mac Callum, Browne and Sugawara 1996, Kline, 2005) ، ويطلق عليها مؤشرات جودة المطابقة(لمطابقة البيانات مع النموذج المتوقع)، وتاليا عرضها مبينا أهمية كل مؤشر وكيفية قراءته.

أشار (Timothy, 2006) إلى مجموعة من مؤشرات حسن المطابقة منها النسبة بين χ^2 ودرجات الحرية المطلقة df ، ومؤشرات المطابقة المطلقة (Absolute Fit Indexes)، ومؤشرات المطابقة المتزايدة (Incremental Fit Indexes)، ويمكن تفصيلها كما يلي:

مؤشرات المطابقة المطلقة (Absolute Fit Indexes) :

تقوم هذه المؤشرات على مقارنة مصفوفة التغاير للعينة (Covariance Matrix) بالمصفوفة المحللة، ومؤشرات المطابق المطلقة عديدة، منها كما أشار إليها (Bentler & Bonett, 1980)، مؤشر حسن المطابقة الذي يدعى Goodness of Fit Index (GFI): ويقيس هذا المؤشر مقدار التباين في المصفوفة المحللة (البيانات الفعلية)، عن طريق النموذج المقترح موضوع الدراسة، وتتراوح قيمته بين (0، 1)، وتشير القيمة المرتفعة لهذا المؤشر إلى تطابق أفضل للنموذج مع بيانات العينة . ويتم الحصول على قيمة المؤشر من خلال المعادلة الآتية:

$$GFI = 1 - (V_{residual} / V_{total})$$

حيث إن:

$V_{residual}$: تباين البواقي في مصفوفة التغاير.

V_{total}: التباين الكلي في مصفوفة التغاير.

مؤشرات المطابقة المتزايدة Incremental Fit Indexes:

تتم عملية تقدير قيمها من خلال مقارنة النموذج المقترح مع النموذج الصفري، والذي يفترض وجود عامل عام واحد في مصفوفة الارتباطات المقاسة،

وتشمل المطابقة المتزايدة عدة مؤشرات (Bentler & Bonett, 1980)، منها مؤشر المطابقة المتزايد Incremental Fit Index (IFI): تتراوح قيمة هذا المؤشر بين (0، 1)، وعادة تشير القيمة المرتفعة منه (أكبر من 0.9) إلى تطابق أفضل للنموذج مع بيانات العينة. ونتوصل لقيمة المؤشر من خلال المعادلة الآتية:

$$IFI = \frac{c^2(\text{Null Model}) - c^2(\text{Proposed Model})}{c^2(\text{Null Model}) - df(\text{Proposed Model})}$$

وللحكم على حسن مطابقة النموذج لبيانات العينة؛ نلجأ لقيم مؤشرات المطابقة المطلقة والمتزايدة (GFI,IFI) ؛ ومن ثم قبول النموذج أو رفضه؛ بناءً على المحكات التي أشار لها بنترال وبونيت (Bentler & Bonett, 1980). انظر الجدول (١).

جدول١: علامة المحك لمؤشرات حسن المطابقة (GFI , IFI)

اسم المؤشر	الرمز	علامة المحك للمطابقة الجيدة
مؤشر حسن المطابقة Goodness of Fit Indexes	GFI	أكبر من 0.95
مؤشر المطابقة المتزايد Incremental Fit Index	IFI	أكبر من 0.95 أو 0.9

أجرى كارون، ومايكل، وكالن (Karon, Michael & Kallen, 2009) دراسة؛ لمعرفة أثر عدد الفقرات في تقييم أحادية البعد لنماذج استجابة الفقرة. اعتمد الباحثون في دراستهم على البيانات المولدة، وبواقع (47، 10) فقرة، مقابل البيانات الحقيقية، وبواقع (47، 10) فقرة. وكانت أعداد المفحوصين 745 لكل اختبار، وتم استخدم مؤشر جذر الخطأ التقريبي (RMSEA)، ومؤشر توكر لويس (TLI)،ومؤشر المطابقة المتزايد (IFI)، ومؤشر المطابقة المعياري (NFI)، ومؤشر المطابقة المقارن (CFI)؛ للكشف عن حسن المطابقة؛ ومن ثم تحقيق افتراض أحادية البعد، ولدى مقارنة فاعلية المؤشرات معا، أظهرت النتائج فاعلية مؤشر توكر لويس، ومؤشر المطابقة المعياري في حالة الحجوم الكبيرة للفقرات؛ في حين بينت أن مؤشر الجذر الخطأ التقريبي كان ضعيفا حيال الكشف

عن أحادية البعد. كما أظهر مؤشر المطابقة المقارن تحسنا يسيرا حيال أعداد الفقرات القليلة.

وقام مكجيل بدراسة (McGill, 2009) على مؤشرات التحقق من افتراض أحادية البعد: مؤشر كيزار the Kaiser (k>1) Criterion، والنسبة المئوية لنسبة التباين (greater than 50%, 40%, or 20%) ، ونسبة الجذور الكامنة، بالإضافة إلى طريقته المقترحة. وقد كانت متغيرات الدراسة عدد الفقرات (20, 40, 60) فقرة، وحجم العينة (100, 250, 500, 750, 1000) مختبرا، عندما توزعت القدرة بشكل طبيعي، وأثرها على افتراض أحادية البعد. توصل الباحث إلى أن مؤشر حسن المطابقة كيزر المعدل كان أكثر فاعلية بزيادة عدد المفحوصين في الكشف عن حسن المطابقة وافتراض أحادية البعد. وتلتها قوة طريقة الجذور الكامنة، ولم يكن لعدد الفقرات أي دلالة إحصائية حول حسن المطابقة وافتراض أحادية البعد.

قام ليو (Liu, 1992) بدراسة هدفت إلى معرفة أثر بعدية بيانات اختبار المولدة وَفق نماذج الاستجابة الفقرة ثنائية البعد التعويضية، وغير التعويضية، على مطابقة البيانات للنموذج. وقد استخدم نموذج استجابة الفقرة الثلاثي المعلمة، ونموذج استجابة الفقرة ثنائي البعد التعويضي، ونموذج استجابة الفقرة ثنائي البعد غير التعويضي. وتم استخدام بيانات اختبار متعددة الأبعاد في نموذج أحادي البعد، كمحك لفحص انتهاك افتراض أحادية البعد. وحسبت عدد الفقرات غير المطابقة للنموذج أحادي البعد، وقيم مربع كاي لبيانات الفقرة أحادية البعد، كمؤشر على مطابقة البيانات للنموذج. أظهرت النتائج أن الارتباط بين السمتين الكامنتين، لا يؤثر في افتراض أحادية البعد، وليس له أثر على مطابقة البيانات للنموذج.

قام كل من أورلاندو، و تيسب (Orlando & thissen,2000) بدراسة تعتمد على البيانات المولدة، و تهدف إلى تقييم دقة مؤشري مطابقة هما($S-X^2$ و $S-G^2$) باختلاف عدد الفقرات (80,40,10) فقرة ، والنموذج اللوغاريتمي (1PL,2PL,3PL) . وأظهرت النتائج أن الخطأ من النوع الأول لمؤشر المطابقة ($S-X^2$) باختلاف عدد الفقرات كان ثابتاً. أما الخطأ من النوع الأول لمؤشر المطابقة($S-G^2$) فقد زاد بزيادة عدد الفقرات . وأظهرت النتائج ضعف مؤشر ($S-X^2$) في الكشف عن حسن المطابقة للبيانات المولدة

وفق النموذج الثنائي والثلاثي ، وأفضلية مؤشر $(S-G^2)$ في الكشف عن مطابقة الفقرات، وخاصة عندما يكون عدد الفقرات قليلا.

قام تاي ودراسغو (Tay & Drasgow,2012) بدراسة محاكاة؛ للكشف عن عن فاعلية الوسط المعدل لمؤشر X^2 ؛ لتقييم جودة المطابقة من خلال مقارنة بأسلوب معدل مقترح ، وباستخدام النموذج اللوغارتمي (الأحادي والثنائي والثلاثي)، وباختلاف عدد الفقرات (45,30,5) فقرة ، وحجم عينة (5000,1500,1000,500) فرد. وأظهرت نتائج الدراسة أفضلية النموذج المعدل في الكشف عن حسن المطابقة.

قام الصرايرة (2008) بدراسة هدفت إلى المقارنة بين طرق فحص جودة مطابقة الفقرة، من خلال أثر تباين خصائص الفقرات، ونوع النموذج الرياضي، وطول الاختبار، وحجم العينة والتفاعل بينها، على معدل الخطأ من النوع الأول، وقوة الاختبار لإحصائيات جودة مطابقة الفقرة، بالاعتماد على أسلوب المحاكاة في توليد استجابات ثنائية. أظهرت أفضلية النموذج ثنائي المعلمة في جودة المطابقة مقارنة مع النموذج ثلاثي المعلمة، وتقل معدلات الخطأ من النوع الأول لإحصائيات جودة المطابقة $(G^2 \ ,X^2 \ B)$ بزيادة طول الاختبار وحجم العينة، مقارنة مع الإحصائي (RMS).

أجرت الحوري (2013) دراسة؛ بهدف الكشف عن افتراض أحادية البعد وَفق النظرية الحديثة في القياس، باختلاف عدد الفقرات (60,30,15) فقرة، وحجم عينة (1000) مفحوص، وأظهرت نتائج الدراسة أن مؤشرات حسن المطابقة (RMR,GFI,AGFI) ذات فاعلية عالية، باختلاف عدد الفقرات؛ في حين أظهرت الدارسة أن الموشرات أقل تأثرًا بعدد الفقرات والنموذج المستخدم، وخاصة مؤشر RMR وَفق النموذج اللوغارتمي الثلاثي.

مشكلة الدراسة وأسئلتها:

يُعدّ افتراض أحادية البعد من الافتراضات الأساسية التي يجب عدم تجاهلها في إعداد الاختبارات والمقاييس، ويؤدي تجاهل هذا الافتراض إلى أن قيمة المعلومات في الفقرة تقل، وكذلك تضعف مطابقة البيانات للنموذج؛ مما ينتج عن ذلك انخفاض دقة القياس الناتج عن تعظيم أخطاء التقدير والقياس. ومن هنا جاءت أهمية إجراء هذه الدراسة لبيان ضرورة توخي الدقة في اختيار مؤشر

حسن المطابقة للحكم على مطابقة النموذج لبيانات العينة في معايرة فقرات الاختبارات. وبالتحديد جاءت هذه الدراسة للإجابة عن الأسئلة الآتية:

السؤال الأول: هل تختلف دقة مؤشرات حسن المطابقة (المطلقة، والمتزايدة)، باختلاف عدد فقرات الاختبار (10، 40) فقرة؟

السؤال الثاني: هل تختلف دقة مؤشرات حسن المطابقة (المطلقة، والمتزايدة)، باختلاف النموذج اللوغاريتمي (أحادي المعلمة، ثنائي المعلمة)؟

أهداف الدراسة:

هدفت الدراسة إلى معرفة دقة بعض مؤشرات حسن المطابقة (المطابقة المطلقة، والمطابقة المتزايدة) في الكشف عن مطابقة النموذج لعينة الدراسة وَفق نظرية استجابة الفقرة، باختلاف عدد فقرات الاختبار والنموذج اللوغاريتمي المستخدم.

أهمية الدراسة:

تتمثل أهمية هذه الدراسة في جانبين: الأول الجانب النظري في مقارنة بعض مؤشرات حسن المطابقة في ظروف مختلفة (عدد الفقرات، والنموذج اللوغارتمي المستخدم)، وتأتي أهمية هذه الدراسة من أهمية التحقق من افتراض أحادية البعد ، خاصة في مجال بناء الاختبارات والمقاييس.

أما الجانب الثاني فهو الأهمية العملية لهذه الدراسة ،وتنبع من أهمية بناء الاختبارات والمقاييس، وضرورة أن تقيس قدرة واحدة. وهذا الأمر يخدم معدي الاختبارات، مثل: (المؤسسات التعليمية، والجامعات، والمدارس)، في إعداد الاختبارات وبنوك الأسئلة، والتي تتطلب توفر خصائص سيكومترية جيدة من شأنها استخدام صور متعددة للاختبار الواحد.

حدود الدراسة:

١. تتحدد نتائج هذه الدراسة بالأسلوب الذي يستخدمه الباحث، والقائم على توليد البيانات باستخدام برنامج (WinGen v1. 4).

٢. ستقتصر هذه الدراسة على مؤشر (IFI)، من مؤشرات المطابقة المطلقة (Absolute Fit Indexes)، ومؤشر (GIF)، من مؤشرات المطابقة المتزايدة .(Incremental Fit Indexes)

٣.ستقتصر هذه الدراسة على النموذج اللوغاريتمي أحادي المعلمة (1-PL)، والنموذج اللوغاريتمي ثنائي المعلمة (2-PL)، من نماذج نظرية استجابة الفقرة.

تعريف المصطلحات:

مؤشرات الكشف عن أحادية البعد:

هي إحصائيات تستخدم في الكشف عن افتراض أحادية البعد ضمن محكات متفق عليها في الأدب النظري لتلك المؤشرات، ويتم مقارنة المؤشرات معا، وكذلك مقارنة دقتها منفردة تبعا للنموذج اللوغارتمي وعدد الفقرات.

مؤشر حسن المطابقة Goodness of Fit Index (GFI):

مؤشر إحصائي يقيس مقدار التباين في المصفوفة المحللة (البيانات الفعلية)، عن طريق النموذج المقترح موضوع الدراسة، وتتراوح قيمته بين (1,0)، وتشير القيمة المرتفعة (GFI > 0,9) إلى تطابق أفضل للنموذج مع بيانات العينة.

مؤشر المطابقة المتزايد Incremental Fit Index (IFI):

مؤشر إحصائي تتراوح قيمته بين (1,0) ، وتشير قيمته إلى مدى تطابق النموذج مع بيانات العينة، وتدل القيمة المرتفعة منه (IFI>0.9) على تطابق أفضل للنموذج مع بيانات العينة.

منهجية الدراسة وإجراءاتها:

المتغيرات المستقلة: وهي عدد الفقرات (10,40) فقرة، والنموذج اللوغاريتمي (PL1,PL2).

المتغير المستقل: مؤشرات حسن المطابقة (IFI , GFI) .

مجتمع الدراسة:

تم الاعتماد على البيانات المولدة تبعا لمتغيرات الدراسة، وفق النموذج اللوغاريتمي أحادي المعلمة وثنائي المعلمة، إذ تم توليد (10) فقرات، و(40) فقرة، بواقع (1500) فردًا، وباستخدام برنامج توليد البيانات (WinGen V. 3).

وفي هذه الدراسة استخدمت طريقة المحاكاة؛ لتوليد البيانات اللازمة للبحث، والمسماة بطريقة مونت كارلو Monte Carlo Methods (MCM). وطبقت طريقة المونتي كارلو (MC) في مجال نظرية استجابة الفقرة IRT،

وذلك من خلال البرامج الحاسوبية المتاحة لتوليد استجابات الأفراد على الفقرة، وتقدير معالم الفقرة، ومن أشهر البرمجيات التي تستخدم هذه الطريقة برنامج (WinGen V. 3)، والذي تم استخدامه لتوليد البيانات في هذه الدراسة، وقد أكد العديد من الباحثين، مثل (Christion & George,2009)، أن التجربة من خلال (MCM) تقدم نتائج التجربة الحقيقية نفسها التي طبقت على الأفراد.

تم توليد بيانات وفقاً لمتغيرات الدراسة، باستخدام برمجية (WinGen V. 3) ، وكانت أعداد الأفراد (1500) فرداً، كما تم توليد تقديرات لقدرات الأفراد ومعالم الفقرات تبعا لمتغيرات الدراسة من حيث طول الاختبار (10،40) فقرة.

جدول (٢) الإحصاءات الوصفية للقدرة الحقيقية

القيمة الصغرى	القيمة العظمى	الوسط الحسابي	الانحراف المعياري
−3.529	3.556	−0.034	0.9678

يلاحظ من الجدول (٢) أن قيمة القدرة الحقيقية، قد تراوحت بين قيمة عظمى مقدارها 3.556، وقيمة صغرى مقدارها 3.529− ،بوسط حسابي مقداره 0.034 وانحراف معياري مقداره 0.9678

تم توليد مجموعة من الفقرات ، بالاعتماد على معلمة القدرة التي تم توليدها مسبقا، وفقا لعدد الفقرات (10، 40)، ونوع النموذج اللوغاريتمي (2PL,1PL)، ويمثل الجدول (٣) ملخصا للإحصاءات الوصفية لمعالم الفقرات.

جدول (٣) الإحصاءات الوصفية لمعلمة الصعوبة والتمييز
باختلاف النموذج اللوغاريتمي المستخدم

النموذج	عدد الفقرات	معلمة الفقرة	القيمة الصغرى	القيمة العظمى	الوسط الحسابي	الانحراف المعياري
النموذج الأحادى **1PL**	10	b	−1.953	2.031	0.269	1.120
	40	b	−1.601	1.691	0.196	0.809
النموذج الثنائى **2PL**	10	b	−1.574	2.097	0.896	1.004
		a	0.786	3.488	1.507	0.638
	40	b	−1.083	1.987	0.191	0.880
		a	0.433	3.275	1.584	0.835

يتبين من الجدول 3 أن قيم الوسط الحسابي لمعلمة الصعوبة وفقا لعدد الفقرات (10، 40)، ونوع النموذج الوغاريتمي (2PL,1PL) تراوحت بين

(0.191 و 0.896)، والانحراف المعياري (0.638 و1.120)؛ في حين أن قيم معلمة التمييز قد تراوحت ضمن المدى (5 ,0.5)، وبوسط حسابي (1.507 و 1.584)، وانحراف معياري (0.835 و 0.638).

وتم التوصل لمؤشر جودة المطابقة المطلقة(IFI) ،وجودة المطابقة المتزايدة (GFI) وفقا لمتغيرات الدراسة ، من خلال تمرير المعالم المولدة، باستخدام برنامج (WinGen V. 3) إلى برنامج (AMOS 21)، الذي يقدم عددا كبيرا من مؤشرات حسن المطابقة، حيث تم بناء المعادلة البنائية لكل متغير مستقل، وبعد ذلك تم إجراء بعض المعالجات الإحصائية بواسطة برمجية SPSS.

نتائج الدراسة:

تمحورت أسئلة هذه الدراسة حول المقارنة بين حسن المطابق المطلقة والمطابقة المتزايدة، من خلال استخدام مؤش(IFI) نموذجا للمطابقة المطلقة، ومؤشر (GFI) نموذجا للمطابقة المتزايدة في ضوء النموذج اللوغاريتمي المستخدم وعدد الفقرات، وللإجابة عن أسئلة الدراسة؛ تم استخدام البرمجيات سابقة الذكر للوصول إلى نتائج الدراسة.

وللإجابة عن السؤال الأول والذي ينص على: "هل تختلف دقة مؤشرات حسن المطابقة (المطلقة والمتزايدة)، باختلاف عدد فقرات الاختبار (10,40) فقرة؟" تم إجراء التحليلات المناسبة باستخدام برمجية AMOS 21، والجدول (٤)يبين نتائج هذه التحليلات.

الجدول (٤) نتائج قيم مؤشرات حسن المطابقة (GIF و IFI)، باختلاف النموذج اللوغاريتمي وعدد الفقرات

النموذج	عدد الفقرات	مؤشرات الكشف حسن المطابقة	
		IFI	GIF
النموذج الأحادى 1PL	10	0.975	0.976
	40	0.891	0.952
النموذج الثنائى 2PL	10	0.888	0.984
	40	0.753	0.934

يظهر الجدول (٤) أن مؤشر حسن المطابقة المطلقة (GFI) أظهر أفضلية في حسن مطابقة النموذج لبيانات العينة؛ حيث أعطت قيم هذا المؤشر مطابقة

عند عدد فقرات 40 و 10 وفق النموذجين (2PL و1PL)، فكانت قيم المؤشر أكبر من المحك المتفق عليه (value GIF >0.9).

في حين بين الجدول (٤) أن مؤشر حسن المطابقة المتزايد (IFI) يتأثر بشكل كبير بزيادة عدد فقرات الاختبار؛ حيث لم يحقق حسن مطابقة للنموذج مع بيانات العينة وفقا لمحك المؤشر لحسن المطابقة (value IFI >0.9)، فقد أظهر المؤشر عدم مطابقة النموذج للبيانات عند عدد فقرات (40) للبيانات المقدرة وفق النموذج اللوغاريتمي (2PL و1PL). وكانت الظرفية الوحيدة التي أظهر فيها مؤشر (IFI) حسن مطابقة هي عندما كان عدد فقرات الاختبار (10)، والمقدرة وفق النموذج الأحادي.

وللإجابة عن السؤال الثاني، والذي ينص على: "هل تختلف دقة مؤشرات حسن المطابقة (المطلقة، والمتزايدة)، باختلاف النموذج اللوغاريتمي (أحادي المعلمة، وثنائي المعلمة)؟" تم إجراء التحليلات المناسبة باستخدام برمجية AMOS 21، وقد أظهر الجدول (٤) قيم مؤشر حسن المطابقة المتزايد (IFI)، ومؤشرات حسن المطابقة (GFI) باختلاف النموذج اللوغاريتمي المستخدم (2PL و1PL)، وأشارت النتائج الواردة في الجدول (٤) إلى أن مؤشر (GFI) لحسن المطابقة المطلقة أعطى نتائج إيجابية لحسن مطابقة النموذج لبيانات العينة في معظم ظرفيات الدراسة. وقد أظهرت قيم المؤشر ارتفاعا ملحوظا عن المحك المتفق عليه (value GIF >0.95) باختلاف النموذج اللوغاريتمي (2PL و1PL)؛ في حين لم يعط مؤشر(IFI) مطابقة إيجابية للنموذج اللوغاريتمي الثنائي (2PL) باختلاف عدد الفقرات، لعدم اجتياز المحك (value IFI>0.90)، أما مطابقة النموذج اللوغاريتمي الأحادي (1PL) فقد تحقق حسن المطابقة عندما كان عدد فقرات الاختبار صغيرة (10 فقرات)، ولم يظهر حسن مطابقة عند حجم عينة كبيرة (40 فقرة).

ويتضح من نتائج هذه الدراسة أن مؤشر (GFI) لحسن المطابقة المطلقة أعطى أفضلية في مطابقة النماذج اللوغاريتمية المستخدمة (2PL و1PL) لبيانات العينة في معظم ظرفيات الدراسة، مقارنة بمؤشر(IFI) لحسن المطابقة المتزايد؛ فقد أشارت نتائج الجدول (٤) إلى اجتياز مؤشر(GFI) محك (value GIF >0.90) عند عدد فقرات 40 و10، والمولدة وفق النموذج الأحادي، وأظهر المؤشر أيضا مطابقة مرتفعة عند 10 فقرات وفق النموذج الثنائي ؛ في حين لم

يظهر مؤشر(IFI) مطابقة للنموذج اللوغاريتمي الثنائي عند عدد الفقرات الاختبارية 40 و 10 فقرة ؛ مما يشير إلى ضعف المؤشر في الكشف عن حسن المطابقة. واتفقت نتائج الدراسة مع ما توصلت إليه دراسة كل من: أورلاندو وتيسن (Orlando & thissen,2000)، وكارون، ومايكل، وكالن (Karon, Michael & Kallen, 2009)، والحوري (2013)، والصرايرة (2008)؛ من حيث تزايد حسن المطابقة مع زيادة عدد فقرات الاختبار ، وأفضلية مؤشر (GFI) في الكشف عن حسن مطابقة البيانات للنموذج؛ ويعزى ذلك إلى انخفاض تباين البواقي (v residual) في مصفوفة التغاير للبيانات، والتي تنعكس إيجابيا على ارتفاع قيم مؤشر حسن المطابقة (GFI) . وأظهرت النتائج الواردة في الجدول (٤) أن عدد فقرات الاختبار تؤثر نسبيا في اجتياز مؤشر حسن المطابقة للمحك المتفق عليه، وخاصة في النموذج اللوغاريتمي الثنائي (2PL) ،وهذا يتعارض مع دراسة مكجيل (McGill, 2009)، التي لم تكشف عن أثر عدد فقرات الاختبار على دقة المؤشرات المستخدمة في الكشف حسن المطابقة.

التوصيات:

يوصي الباحث بمزيد من الدراسات حول حسن المطابقة وفق بيانات حقيقية لاختبارات تحصيلية أو مقننة، وكذلك إجراء دراسات مقارنة على مؤشرات أخرى لحسن المطابقة،وإجراء دراسات على مقارنة حسن المطابقة باختلاف حجم العينة ومستويات القدرة.

المراجـــع

أولاً– المراجع العربية:

الحوري، أروى .(٢٠١٣).**التحقق من فاعلية مؤشرات الكشف عن افتراض أحادية البعد وفق نماذج نظرية استجابة الفقرة، في ضوء تغير طول الاختبار وشكل الفقرة،** رسالة دكتوراه غير منشورة، جامعة اليرموك،اربد، الأردن.

الشواورة، شادي.(٢٠١٣). **دقة تقدير معالم الفقرات بطريقتي الأرجحية العظمى الهامشية وبييز في ظروف مختلفة في عدد الفقرات وحجم العينة والنموذج اللوغاريتمي المستخدم،** رسالة دكتوراه غير منشورة، جامعة اليرموك، اربد، الأردن.

الصرايرة، راجي.(٢٠٠٨). **مقارنة ثلاث طرق إحصائية لفحص جودة مطابقة الفقرة تحت ظروف نوع النموذج الرياضي وطول الاختبار وحجم العينة والتفاعلات بينهما،** رسالة دكتوراه غير منشورة. الجامعة الأردنية. الأردن.

علام، صلاح الدين.(٢٠٠٥). **نماذج الاستجابة للفقرة الاختبارية أحدية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي**.دار الفكر العربي.القاهرة.

ثانيًا–المراجع الأجنبية:

Baker, F. (2001). The basics of item response theory. ***ERIC Clearinghouse on Assessment and Evaluation***, University of Maryland, College Park, MD.

Bentler, P.M., & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures**.** ***Psychological Bulletin***, *88*, 588-606.

Crocker, L. & Algina, J. (1986) ***INTRODUCTION TO CLASSICAL AND MODERN TEST THEORY*** .New York.

Christian, P. George, C. (2008). **Monte Carlo Statistical Methods.** Springer texts in Statistics, Second Edition.

Glas, C. A, W., & Falcon, J. C. S. (2003). A comparison of item – fit statistics for the three–parameter logistic model. **Applied Psychological Measurement**, 27 (2), 89- 106.

Hambleton, R. K., & Swaminathan, H. (1985) ***Item Response Theory. Principles and Applications*. kluwer**. Nijhoff publishing.

Hambleton, R. K, swaminathan, H. & Rogers, Jane. H. (1991).***Fundamental Of Item Response Theory***. Newbury park, CA: sage publications, Inc.

Hattie, J. (1985). Methodology Review: Assessing unidmensionality of tests and item**.** ***Applied Pschological Measurment, 9*** (2), 139-164.

Karon F. Cook Michael A. Kallen.(2009). ***Having a fit: impact of number of items and distribution of data on traditional criteria for assessing IRT's unidimensionality assumption***. Qual Life Res (2009) 18:447–460 DOI 10.1007/s11136-009-9464-4.

Kline, R.B. (2005), **Principles and Practice of Structural Equation Modeling (2nd Edition ed**). New York: The Guilford Press.

MacCallum RC, Browne MW, Sugawara HM (1996). Power analysis and determination of sample size for covariance structure modeling. **Psychological Methods**, 1(2):130-149.

McGill. M. T **(**2009**).** ***An Investigation of Unidimensional Testing Procedures under Latent Trait Theory using Principal Component Analysis*.**

Liu, X. (1992). ***The Dimensionality of test data generated by compensatory and non- compensatory two dimensional IRT models and its effects on model data fit*. Paper presented at the annual meeting of the Canadian**

Society for the study of Education (Charlottetown, Prince Edward Island, June, 1992).

Orlando, M., and Thissen, D. (2000). Likelihood-based item-fit indices for dichotomous item response theory models. ***Applied Psychological Measurement,* 24** (1), 50-64.

Stout, W.R. (1987). A nonparametric approach for assessing latent trait unidimensionality. ***Psychometrika***, *52*, 589-617.

Tay, L. & Drasgow, F. (2012). Adjusting the Adjusted X[superscript 2]/df Ratio Statistic for Dichotomous Item Response Theory Analyses:Does the Model Fit**?.** ***Educational and Psychological Measurement, 72*** (3), 510-528.

Timothy, A. (2006). ***Confirmatory Factor Analysis For Applied Research***. New York, NY, 10012.

Walker, C. & Betetras, S. (2000). **Using Multidimensional versus Unidimensional ability estimates to determine student proficiency in mathematics. Paper presented at the annual meeting of the American Educational Research Association**, New Orleans, LA, April24 – 28, 2000.

Fadlelmula ,F. K. (2011). Assessing Power of Structural Equation Modeling Studies: A Meta-Analysis. **Education Research Journal** Vol. 1(3), pp. 37-42.