مجلـــة العلـــوم التربـــوية والدراســـات الإنســـانيـــة
ISSN: 2709-0302 (online)

Humanities and Educational Sciences Journal
ISSN: 2617-5908 (print)

أثر جودة مطابقة الفقرات للنموذج ثلاثي المعلم على البنية العاملية للمقياس باستخدام التحليل العاملي التوكيدي (دراسة محاكاة)(*)

الباحثة/ بشائر عبدالله الغدير
باحثة دكتوراه بقسم علم النفس التربوي
كلية التربية - جامعة القصيم السعودية
421200438@qu.edu.sa

أ.د/ ربيع عبده أحمد رشوان
أستاذ علم النفس التربوي
كلية التربية - جامعة القصيم السعودية
rashwan_75@yahoo.com

*) تاريخ تسليم البحث 8/11/2023 تاريخ قبوله للنشر 27/12/2023
*) موقع المجلة: http://hesj.org/ojs/index.php/hesj/index

أثر جودة مطابقة الفقرات للنموذج ثلاثي المعلم على البنية العاملية للمقياس باستخدام التحليل العاملي التوكيدي (دراسة محاكاة)

الباحثة/ بشائر عبدالله الغدير
باحثة دكتوراه بقسم علم النفس التربوي
كلية التربية - جامعة القصيم السعودية

أ.د/ ربيع عبده أحمد رشوان
أستاذ علم النفس التربوي
كلية التربية - جامعة القصيم السعودية

الملخص

هدفت الدراسة الحالية للتعرف على تأثير ملائمة الفقرات للنموذج ثلاثي المعلم في نظرية الاستجابة للمفردة الاختبارية على الأوزان الانحدارية المعيارية والبواقي كمؤشرات لجودة البنية العاملية للمقياس باستخدام التحليل العاملي التوكيدي في ضوء حجم العينة وعدد الفقرات، وتكونت عينة الدراسة من أحجام مختلفة تم توليدها باستخدام المحاكاة، وباتباع المنهج الوصفي المقارن واستخدام الأساليب الإحصائية المناسبة من خلال برمجيات R. SPSS. AMOS توصلت الدراسة للعديد من النتائج أهمها أنه توجد فروق دالة إحصائيًا في الأوزان الانحدارية المعيارية ترجع لاختلاف نسبة الفقرات غير المطابقة للنموذج ثلاثي المعلم والفروق لصالح النماذج الأقل في نسبة الفقرات غير المطابقة، والأعلى في عدد الفقرات وحجم العينة الأقل، وبالنسبة للبواقي فكانت الفروق لصالح النموذج الأقل في نسبة الفقرات غير المطابقة، والأعلى في عدد الفقرات حجم العينة الأقل، كذلك توجد تأثيرات دالة للتفاعلات المشتركة بين متغيرات الدراسة على الأوزان الانحدارية المعيارية والبواقي، ووفقًا للنتائج تم تقديم مجموعة من التوصيات والمقترحات.

الكلمات المفتاحية: نظرية استجابة الفقرة، النموذج ثلاثي المعلم، مطابقة الفقرات، الصدق البنائي، دراسة محاكاة.

Items Fit to the Three Parameter Logistic Model and Their Effect on the Factorial Structural of Scale Using Confirmatory Factor Analysis: Simulation Study

Bashaer Abdullah Alghadeer
Doctoral researcher. Department of Psychology College of Education. Qassim University. Kingdom of Saudi Arabia

Pro. Dr. Rabie Abdo Ahmed Rashwan
Professor of Educational Psychology College of Education - Qassim University. Kingdom of Saudi Arabia

Abstract

The current study aimed to identify the effect of fitting the items to the three-parameter model in item response theory on the standard regression weights and residuals as indicators of the quality of the factor structure of the scale using confirmatory factor analysis in light of the sample size and number of items.

The study sample consisted of different sizes that were generated using simulation. following the comparative descriptive approach and using appropriate statistical methods through R. SPSS. AMOS. The study reached many results. the most important of it: there are statistically significant differences in the standard regression weights due to the difference in the percentage of items that do not fit the three- parameter model. and the differences are in favor of models with the lowest percentage of non- fit items. and the highest in The number of items and the smaller sample size. and for the residuals. the differences were in favor of the model with the lower percentage of non- fit items. and the higher number of items and smaller sample size. There are also significant effects of the joint interactions between the study variables on the standard regression weights and the residuals. and according to the results. a set of recommendations and proposals were presented.

Keywords: Item Response Theory. Three- Parameters Model. Item Fit. Structural Validity. Simulation.

مقدمة:

لاحظ المختصون في مجال البحوث النفسية تزايد الاهتمام بتطوير وإعداد الاختبارات مع التركيز على سلامتها وتطبيق النظريات وتنميتها، ويظهر هذا التركيز بشكل خاص في مجال القياس النفسي، حيث أصبحت النظرية الحديثة في القياس والمعروفة بنظرية الاستجابة للفقرة (IRT) نقطة تحول رئيسة، وقد شهدت نماذج هذه النظرية تنوعًا كبيرًا، ناتجًا عن تفاوت خصائص الفقرات والمعالم المقدرة، وهذا التنوع أدى إلى ظهور مجموعة متنوعة من الاختبارات التي تُستخدم لتقييم قابلية استخدام نماذج النظرية الحديثة في مجال القياس، ويعد بناء الاختبارات ومعايرتها من أبرز الأولويات في ميدان القياس؛ حيث تترتب على نتائج الاختبارات قرارات حيوية؛ ولذلك يتعين على معدي الاختبارات أن يسعوا جاهدين لضمان الموضوعية في القياس، وتحقيق هذه الموضوعية يعتمد بشكل كبير على تحقق الخصائص السيكومترية للاختبار، وقد أدت النظرية الحديثة في ميدان القياس إلى إضافة طرق وأساليب جديدة في معايرة الفقرات والتأكد من مطابقتها للنموذج بهدف تحقيق درجة عالية من الموضوعية في القياس (الشواورة، 2017، 101).

وتعنى مطابقة الفقرة في نظرية الاستجابة للمفردة الاختبارية "قياس درجة التناسب بين النموذج الرياضي الذي تم اعتماده وبين نمط استجابة الأفراد للفقرة بواسطة الإحصائيات المختلفة للمطابقة" (الشريفين، 2003، 29)؛ وتعرف أيضًا بأنها مؤشر إحصائي يحدد الفارق بين البيانات الفعلية والبيانات المتوقعة من خلال النموذج المستخدم، ويتم ذلك عن طريق مقارنة قيم هذه المؤشرات بقيم حرجة محددة بهدف تحديد أنماط الاستجابة التي تتناسب أو لا تتناسب مع النموذج، وتعمل هذه المؤشرات كوسيلة لتحديد مدى صلاحية الفقرة في التنبؤ بالسلوك المستهدف أو الخصائص المقصودة (D'Agostino & Stephens, 1986, 1).

ويقدم مفهوم مطابقة الفقرة العديد من الطرق التي يمكن استخدامها لتحقيق ما يُعرف بالاختبار المثالي، حيث اتسع استخدام هذا المفهوم في السنوات الأخيرة وانتشر بشكل كبير في مجالات البحث المتنوعة، ويفترض هذا المفهوم إمكانية التنبؤ بأداء الفرد في الاختبار عن طريق تحديد قدراته أو خصائصه الكامنة التي لا يمكن قياسها مباشرة بشكل دقيق، وقد تعتمد نماذج مطابقة الفقرة على تحديد العلاقة المتوقعة بين أداء الفرد في الاختبار والسمات الكامنة، كما يمكن وصف هذه العلاقة بواسطة دوال رياضية مع مراعاة بعض الافتراضات الخاصة المتعلقة ببيانات الاختبار، ويتيح ذلك فهم العلاقة بين الأداء الملاحظ والسمات الكامنة مما يساعد في تفسير أوجه القوة والضعف لدى الأفراد بشكل أكثر دقة (De Ayala, 2009, 5).

ومنذ ظهور نظرية الاستجابة للفقرة أصبحت مركز اهتمام الباحثين في تخصص القياس والتقويم، حيث تجاوزت العديد من التحديات التي واجهت نظرية القياس التقليدية مثل مشكلة بنوك الأسئلة واللاتغاير (Invariant)، وتقوم النظرية الحديثة بتحليل استجابات الأفراد على الفقرة، وهذا ما جعلها تسمى بنظرية الاستجابة للفقرة، ويعني ذلك أنها تأخذ في اعتبارها كيفية استجابة الأفراد لكل فقرة على حدة، وتستفيد من هذه البيانات لتقدير مستوى القدرة أو السمة الكامنة بشكل دقيق، وتقدم النظرية أسلوبًا أكثر دقة في تحليل الأداء مما يجعلها أداة هامة في مجالات متعددة من القياس (Hambleton & Swaminathan, 1985).

وعندما يتم التأكد من مطابقة الفقرات يعني ذلك أن هناك توافقًا أو اتفاقًا بين أداء الأفراد على هذه الفقرات وما يتوقعه النموذج اللوجستي الذي تم اختياره، ويتم تحديد هذا التوافق من خلال معايير خاصة تُستخدم لتقييم معالم الفقرة في كل نموذج، ويظهر هذا التوافق في شكل ومعالم دالة استجابة الفقرة المحددة وفقًا للنموذج المستخدم، وبتحليل هذه الدالة يمكن للباحثين فهم كيفية تأثير تلك الفقرة على التوقعات والأداء، وبالتالي تحسين جودة وفعالية الاختبار (المومني، 2016، 72- 73).

وتعتبر نمذجة المعادلة البنائية من الأساليب المتبعة للتحقق من مقبولية أو منطقية نموذج القياس، وهي عملية تبدأ عادةً من وضع نموذج يستند إلى نظرية توضح العلاقات بين مجموعة من المتغيرات، ويتم استخدام هذا النموذج لتوضيح طبيعة العلاقات وفهم الظواهر المعينة (عامر، 2018، 16)؛ ويرى Crisci (2012) أنها أسلوب إحصائي متقدم يتيح فحص النظريات والنماذج والبُنى الكامنة للتعرف على مفاهيم أو ظواهر معينة، ويمكن أن تشمل على تحليل العلاقات بين متغيرات متعددة في سياق واحد، ويسمح بتقدير التأثيرات المباشرة وغير المباشرة بين هذه المتغيرات، وتُستخدم في مجالات متنوعة مثل علم النفس، وعلوم الاجتماع، والتعليم، والأعمال، حيث تُمثل أداة قوية لاختبار وتطوير النظريات والنماذج العلمية.

وأكد Kline (2002) على إمكانية استخدام أحد أساليب نمذجة المعادلة البنائية، وهو التحليل العاملي التوكيدي للتحقق من الصدق البنائي عندما يكون للباحث معرفة نظرية أو تجريبية أو كليهما في بناء المتغير الكامن، ويقوم الباحث بعد ذلك باختبار هذا البناء الافتراضي إحصائيًا، ويمكن الحكم على جودة النموذج النظري من خلال مجموعة من المؤشرات المعروفة باسم مؤشرات جودة المطابقة (Goodness Fit Indices)، ويتطلع الهدف من نمذجة المعادلة البنائية إلى تحديد مدى تطابق النموذج النظري مع البيانات الميدانية، وفي حال تأييد بيانات العينة للنموذج النظري يصبح من الممكن افتراض نماذج نظرية أكثر تعقيدًا، وأما إذا لم تدعم البيانات النموذج النظري فيمكن تعديله واختباره مرة أخرى أو تطوير نماذج نظرية جديدة لفهم الظواهر المدروسة؛ وتعد مؤشرات جودة المطابقة من أحد أهم المعايير التي تُستخدم لتقييم جودة نموذج القياس كنموذج للنمذجة البنائية، حيث تُعتبر هذه المؤشرات دليلًا على صدق البناء (المهدي، 2007؛ محمد، 2020).

والدراسة الحالية تحاول بحث تأثير جودة مطابقة الفقرات في النموذج ثلاثي المعلم على الأوزان الانحدارية المعيارية والبواقي كمؤشرات لجودة البنية العاملية للمقياس باستخدام التحليل العاملي التوكيدي، وفقًا لحم العينة وعدد فقرات المقياس.

مشكلة الدراسة:

يتضح من خلال ما تم عرضه سابقًا أن التحليل العاملي التوكيدي يتم فيه التحقق من التصور النظري لأدوات القياس ومدى المطابقة بينه وبين البيانات التي يتم الحصول عليها من خلال أداة القياس؛ وبالطبع لا يكون التطابق تامًا بين النموذج المفترض والبيانات الفعلية، والنموذج الجيد هو الذي تتوافر له أفضل مؤشرات لجودة المطابقة أما نظرية الاستجابة للفقرة فهي تعتمد على دراسة العلاقة بين السمات الكامنة موضوع الاهتمام واحتمالية الإجابة الصحيحة للفقرة، والفقرة الجيدة هي التي يتسق فيها أداء الأفراد على الفقرات مع مستوى السمة الكامنة.

ويتضح بذلك أن تركيز نظرية الاستجابة للفقرة يكون على فقرات أداة القياس ومعالمها، بينما في النمذجة البنائية يكون التركيز على بنية المقياس وما تتمتع به من صدق في تمثيل الظاهرة موضوعالقياس (Kline, 2002, 9-10؛ علام، 2005، 107؛ Abedalaziz & Leng, 2003, 65؛ بوصلب وصحراوي، 2016، 62-67).

وقد حظي موضوع مطابقة الفقرات باهتمام بالغ من قبل الباحثين في الآونة الأخيرة، وقد ظهرت مجموعة من الاختبارات التي تستخدم في التحقق من مدى مطابقة الفقرة للنموذج التي تعتبر محور النظرية الحديثة في القياس مثل اختبار مربع كاي بأشكاله المختلفة (Kang & Troy, 2008, 395)، أو من خلال اختبار البواقي (Haberman et al., 2013, 420-421)؛ كما بحثت بعض الدراسات في المطابقة بين الفقرات حسب النظرية الكلاسيكية ونماذج النظرية الحديثة مثل دراسة المرابحة وبن صالح (2021)، كما كان لأثر الخطية على مطابقة الفقرة للنموذج اهتمام في دراسة ماكدونالد (McDonald, 1982)؛ ونلاحظ اهتمام دراسة الحاج محمود (2018) بأثر التحيز في الفقرات على مطابقتها لنموذج نظرية الاستجابة للفقرة في حين قارنت بعض الدراسات بين طرق التحقق من مطابقة الفقرة للنموذج في ضوء طول الاختبار وصعوبة الفقرات مثل دراسة محمود (2012)، ودراسة الشواورة (2017)، ودراسة Cook et al. (2009)، ودراسة McGill (2009)، ودراسة Orlando and Thissen (2000)، ودراسة Tay and Drasgow (2012)، ودراسة الصرايرة (2008)، ودراسة الحوري (2013).

كما كانت هناك جهود أخرى تبذل للتحقق من صدق المقاييس والاختبارات وذلك من خلال تأكيد سلامة البنية العاملية لتلك المقاييس والاختبارات باستخدام التحليل العاملي التوكيدي والذي قد يكون جزئيًا في بعض الأحيان بسبب المطابقة الجزئية للنموذج المفترض للبيانات الفعلية وكلما زاد التوافق بين النموذج المفترض والبيانات الفعلية تزداد جودة البنية العاملية للمقياس والتي يستدل عليها من خلال العديد من المؤشرات لذلك تم اللجوء لما يسمى بمطابقة الفقرة للنموذج التوكيدي وهو النموذج الذي يؤكد أو ينفي ارتباط الفقرة بعاملها المفترض تشبعها عليه، وقد تم الحرص على مطابقة الفقرة للنموذج التوكيدي في كثير من الأبحاث والدراسات للتحقق من البنية العاملية لاختبارات ومقاييس مشهورة مثل دراسة أجريت في جامعة العلوم الإسلامية الماليزية بالاشتراك مع كلية القيادة والإدارة (2017)، ودراسة الخالدي (2018).

واعتمدت بعض الدراسات على التحليل العاملي التوكيدي في الكشف عن البنية العاملية للمقياس مثل دراسة بلعيد وقماري (2018) ودراسة بلعربي (2018)، وبحثت بعض الدراسات في ثبات البنية العاملية للمقاييس والاختبارات وعدم اختلافها باختلاف الفئات العرقية أو النوع أو التخصص على سبيل المثال وذلك باستخدام التحليل العاملي التوكيدي متعدد المجموعات كما في دراسة السمكي والسوالمة (2017)؛ ودراسة Young et al. (2019)؛ ودراسة رشوان (2021)؛ ودراسة حسب الله (2023)؛ ودراسة Brown et al. (2017)؛ ودراسة Cheung and Rensvold (2002)؛ ودراسة Wu et al. (2007).

ومما سبق يتضح أن هناك انفصال واضح بين الاتجاهين (نظرية الاستجابة للفقرة، النمذجة البنائية واستخداماتها في التأكد من الصدق البنائي لأدوات القياس) وفي حدود إطلاع الباحثان هناك ندرة في الدراسات التي حاولت الاستفادة من ميزات الاتجاهين في بناء أدوات القياس حيث تناولت دراسة Rodriguez et al. (2021) تقييم الخصائص السيكومترية لمقياس الشعور بالوحدة النفسية بناءً على نموذج المعادلة البنائية ونظرية استجابة الفقرة وأظهرت النتائج أن نموذج المعادلة البنائية يكشف عن وجود ثلاثة عوامل تتوافق مع الافتراضات النظرية للمقياس، كما تم استخدام معادلة خصائص الفقرة من خلال تطبيق نموذج IRT متعدد الأبعاد فدلت على أن جميع الفقرات الملائمة تؤدي أداءً جيدًا في النموذج البنائي؛ وفي دراسة Lin and Tsai (2021) والتي هدفت لتقييم الخصائص السيكومترية لمقياس القراءة في الرياضيات بناًء على منظور تحليلي مزدوج: نمذجة المعادلات البنائية SEM ونظرية الاستجابة للفقرة، وكان من بين النتائج دعم أحادية البعد باستخدام النمذجة البنائية مع وجود تأثير واضح لمطابقة الفقرة لنظرية الاستجابة على الفقرة.

ويذكر Baker (2001, 51-63) أن الفقرات المطابقة تحمل مؤشرات مطابقة مرتفعة كما تدل على مناسبة النموذج المستخدم، وقد توجد فقرات لا تتطابق مع النموذج، وبالتالي تحتاج إلى فحص وإعادة صياغة وإذا لم تقبل التعديل فيجب أن تحذف أو أن النموذج غير مناسب، وقد أجريت العديد من الدراسات التي كان من ضمن ما تدرسه مطابقة الفقرات، حيث اتفقت معظم الدراسات في أن الفقرات الجيدة في المطابقة تزيد من دالة معلومات الاختبار والثبات وجودة الخصائص السيكومترية، والفقرات الضعيفة تحذف لتأثيرها السلبي كما في دراسة أبو فودة (2013) ودراسة الصيخان (2019).

ومن ثم فإن الدراسة الحالية تحاول تتبع العلاقة بين ملائمة الفقرات في النموذج ثلاثي المعلم وبين مؤشرات جودة الصدق البنائي للمقياس والتي يتم تقديرها باستخدام التحليل العاملي التوكيدي متمثلة في (تشبعات الفقرات "الوزن الانحداري المعياري"، البواقي) باختلاف حجم العينة وعدد فقرات المقياس، ويمكن أن تتحدد مشكلة الدراسة الحالية في التساؤلات الأتية:

1- هل تختلف الأوزان الانحدارية المعيارية "التشبعات" كمؤشرات لجودة الصدق البنائي للمقياس باختلاف مطابقة الفقرات للنموذج ثلاثي المعلم في نظرية الاستجابة للفقرة وفقًا لحجم العينة وعدد الفقرات؟
2- هل تختلف البواقي كمؤشرات لجودة الصدق البنائي للمقياس باختلاف مطابقة الفقرات للنموذج ثلاثي المعلم في نظرية الاستجابة للفقرة وفقًا لحجم العينة وعدد الفقرات؟

أهداف الدراسة:

تهدف الدراسة الحالية إلى التعرف على:

1- درجة اختلاف الأوزان الانحدارية المعيارية (التشبعات) كمؤشر على جودة البنية العاملية للمقياس باستخدام التحليل العاملي التوكيدي باختلاف ملائمة الفقرات للنموذج الثلاثي في نظرية الاستجابة للفقرة وفقاً لحجم العينة وعدد الفقرات.
2- درجة اختلاف البواقي كمؤشر على جودة البنية العاملية للمقياس باستخدام التحليل العاملي التوكيدي باختلاف ملائمة الفقرة للنموذج الثلاثي في نظرية الاستجابة للفقرة وفقًا لحجم العينة وعدد الفقرات.

أهمية الدراسة:

الأهمية النظرية: تبرز الأهمية النظرية للدراسة الحالية من خلال التالي:

- دمجها لاثنين من أهم الاتجاهات في بناء أدوات القياس في مجال علم النفس التربوي بصفة عامة والقياس والتقويم بصفة خاصة.
- الربط بين مطابقة فقرات أدوات القياس باستخدام نماذج الاستجابة للفقرة ومؤشرات الصدق البنائي لهذه الأدوات من الدراسات النادرة في مجال القياس والتقويم- على حد علم الباحثان- وبالتالي تسهم النتائج المتوقعة للدراسة الحالية لسد هذه الفجوة البحثية.
- قد تسهم النتائج التي سيتم التوصل إليها في تزويد الباحثين بمؤشرات خاصه بصدق البناء العاملي من خلال البحث في مطابقة الفقرات لنموذج استجابة الفقرة، أو العكس إمكانية أن يزودنا البناء العاملي بمؤشرات عن مطابقة البيانات لنماذج استجابة الفقرة، وهو ما يعد مطلب وأمر ضروري لزيادة الثقة في أدوات القياس والنتائج المترتبة عليها.

الأهمية التطبيقية: تبرز الأهمية التطبيقية للدراسة الحالية من خلال التالي:

- قد تسهم النتائج التي سيتم التوصل إليها في التعرف على حجم العينة وعدد الفقرات المناسبة للحصول على توافق بين نماذج الاستجابة للفقرة والصدق البنائي باستخدام التحليل العاملي التوكيدي، ولهذا الأمر العديد من الجوانب التطبيقية في بناء أدوات القياس، وهو ما يمكن أن يفيد الباحثين في الدراسات المستقبلية بصفة عامة.
- تسهم النتائج الحالية في تحديد أهم مؤشرات جودة المطابقة في النمذجة البنائية الأكثر تأثرًا بجودة مطابقة الفقرات في نماذج الاستجابة للفقرة، وهو ما يلفت انتباه الباحثين للاهتمام بمثل هذه المؤشرات والتي يتم إغفال الكثير منها من قبل الباحثين ممن يستخدمون التحليل العاملي التوكيدي.
- المتفحص للدراسات التي اهتمت بتوظيف نظرية الاستجابة للفقرة أو التحليل العاملي التوكيدي يلاحظ انفصال تام بين الاتجاهين مع زيادة الاهتمام بالتحليل العاملي التوكيدي نظرًا لسهولة إجراءاته ووضوح نتائجه وتوافر البرامج الإحصائية المستخدمة، والربط بين الاتجاهين قد ييسر كثيرًا على الباحثين، ففي حالة توافر مؤشرات معينة بشأن معالم الفقرات من خلال التحليل العاملي التوكيدي يمكن ضمان تحقق جودة مطابقة الفقرات لنماذج نظرية الاستجابة للفقرة التي مازالت إلى الآن لم توظف في المجال التطبيقي بالقدر المتوقع لها نظرًا لصعوبة إجراءاتها وصعوبة الوصول للبرامج الإحصائية المستخدمة.

مصطلحات الدراسة:

1- نظرية الاستجابة للفقرة (Item Response Theory):

تقوم على افتراض أساسي يشير إلى وجود صفات أو سمات كامنة تكمن وراء أداء المفحوصين على أداة القياس (Harris, 1989, 157)؛ فهي عبارة عن تنظيم لمجموعة من النماذج يتم استخدامها لتوضيح وتفسير كيفية حدوث التطابق بين المتغيرات الكامنة والملاحظة، وكل نموذج من هذه النماذج يتمتع بمعالم أو بارامترات فريدة تميزه عن الآخرين، وتُستخدم هذه المعالم لتفسير العلاقات بين المتغيرات الكامنة والمشاهدة (De Ayala, 2009, 4).

النموذج اللوجستي ثلاثي البارامتر (النموذج ثلاثي المعلم) Three Parameter Logistic Model:

وهو نموذج معقد يتميز بخاصية خط التقارب السفلي الذي يكون أكبر من الصفر، وهذه الخاصية تعني أن احتمالية الاستجابة الصحيحة عندما تكون قدرات المفحوص منخفضة لم يعد يقترب من الصفر، فيمكن للمفحوص أن ينجح في الاستجابة للفقرة حتى عندما تكون قدراته منخفضة نظرًا لظاهرة التخمين (De Ayala, 2009, 123)؛ وتتباين فيه الفقرات في بارامتراتها الثلاثة (الصعوبة، التمييز، التخمين) (Baker, 2001, 28)؛ ويمكن توضيح معنى هذه البارامترات كما يلي:

بارامتر الصعوبة: Difficulty Parameter

يرمز لهذا البارامتر عادة بالرمز b، وهو يُوضّح مدى صعوبة الفقرة، وهو نقطة على محور القدرة أو السمة حيث تتقاطع النقطة مع نقطة الانعطاف في المنحنى، ويُعرف أحيانًا بمعلم الوضع في النموذج الأحادي والثنائي، وتقابل هذه النقطة على محور القدرة احتمال 50% للإجابة بشكل صحيح على الفقرة، أما في النموذج الثلاثي تكون النقطة حيث يكون احتمال الإجابة الصحيحة على الفقرة مساويًا لـ 2/ (1+c)، حيث يُعبر c عن معلم التخمين (Harris,1989, 157).

ويمكن القول بأنها المعلمة التي تحدد الطريقة التي تتصرف بها الفقرة على طول مقياس القدرة، كما أنه يعتبر عامل يؤثر على احتمال استجابة الفرد بطريقة معينة على فقرة، حيث يستجيب الفرد الذي يتمتع بمستوى عالي من القدرة بشكل صحيح على الفقرة باحتمال أكبر من الفرد الذي لديه مستوى منخفض من القدرة (Bajwa et al., 2016, 5).

بارامتر التمييز: Discrimination Parameter

يرمز لهذا البارامتر عادة بالرمز a، وهو عبارة عن انحدار الخط المماس للمنحنى المميز للفقرة عند نقطة الانقلاب، ويُستخدم لقياس قدرة الفقرة على التمييز بين القدرات المختلفة للممتحنين، وعندما يكون الميل أعلى يكون التمييز أكبر في حين يشير انخفاض الميل إلى أن الفقرة غير مميزة بشكل كبير (Baker, 2001, 9).

بارامتر التخمين: Pseudo - Guessing Parameter

يرمز لهذا البارامتر عادة بالرمز c، يعبر عن احتمال الاستجابة الصحيحة عند مستويات القدرة المنخفضة، ويُمثل هذا المعلم نقطة على محور احتمالية الاستجابة الصحيحة، حيث تتراوح القيم فيه بين صفر وواحد، ويمكن للأفراد ذوي القدرة المنخفضة أو الذين يعتمدون على التخمين أن يحصلوا على الإجابة الصحيحة بشكل أكبر في هذه الحالة (Harris, 1989, 157)؛ ويعبر عن أقل قيمة لخط التقارب السفلي للمنحنى عندما تقترب قدرة المستجيب من (−∞) (De Ayala, 2009, 124).

مطابقة الفقرة للنموذج (Item Fit):

تُعبّر عن مدى تناسب أو توافق الاستجابات على الفقرة مع التوقعات المشتقة من النموذج (المرابحة، بن صالح، 2021، 19)، وتُشير إلى مدى تناسب الفقرة مع النموذج المحدد، ويتم تحديد ذلك من خلال مجموعة من المحكات المعينة لكل نموذج، والتي تُستخدم للتقييم الشامل للدالة ومعالم الفقرة (المومني، 2016، 75).

نمذجة المعادلات البنائية (Structural Equation Modeling):

تُعتبر إحدى الأساليب الإحصائية الشاملة لاختبار الفروض التي تفسر العلاقات بين المتغيرات الملاحظة والمتغيرات الكامنة. (Hoyle, 1995, 1)، ومن أهم تطبيقاتها التحليل العاملي التوكيدي.

– التحليل العاملي التوكيدي (CFA: Confirmatory Factor Analysis):

يُعد إحدى تطبيقات نمذجة المعادلات البنائية، ويهدف إلى التحقق من صدق البنية العاملية للمقياس من خلال اختبار العلاقات بين المتغيرات المقاسة والمتغير أو المتغيرات الكامنة (Ismail et al., 2018, 73).

– البنية العاملية للمقياس (Factorial Structure of the Scale):

تُمثل العوامل أو الأبعاد أو المتغيرات الكامنة التي تُظهر العلاقات المشتركة بين شتى المتغيرات المقاسة، والتي تُشكل المساحات المشتركة للدلالة على مفهوم معين، ويتم استنتاج هذه العوامل من خلال التحليل العاملي الاستكشافي، ويتم تقييم جودتها عبر التحليل العاملي التأكيدي (Ismail et al., 2018; Pop & Frent, 2008)، (تيغزة، 2012، 21-22).

دراسة محاكاة (Simulation Study):

تُستخدم في الغالب للحصول على نتائج تجريبية حول أداء الطرق الإحصائية في سيناريوهات معينة، بدلًا من الاعتماد على النتائج المستمدة من دراسات فعلية، ويُشير ذلك إلى التحديات التي تواجهها بعض الدراسات الواقعية في تحقيق شروط محددة في البيانات المختصة بالتحليل الإحصائي (Morris et al., 2018, 2074).

حدود الدراسة:

الحدود الموضوعية:

1- تناولت الدراسة نموذجين مهمين في بناء أدوات القياس وهما التحليل العاملي التوكيدي باستخدام النمذجة البنائية وأحد نماذج نظرية استجابة الفقرة وهو النموذج ثلاثي المعلم.

2- تم التركيز في جودة المطابقة بواسطة النموذج ثلاثي المعلم على ملائمة الفقرة للنموذج، والتركيز على الأوزان الانحدارية المعيارية "التشبعات"، والبواقي في النمذجة البنائية.

3- اعتمدت الدراسة الحالية على بيانات مولده باستخدام حزمة mirt في برنامج (R) باختلاف حجم العينة وعدد الفقرات.

الحدود الزمانية: تم توليد البيانات في الفصل الدراسي الثاني من العام الجامعي 1444 هجري 2023 ميلادي.

الإطار النظري:

أولًا: نظرية الاستجابة للفقرة الاختبارية:

تعتبر من النظريات الحديثة المعاصرة التي اجتهد فيها العلماء لمعالجة قصور النظرية التقليدية (الكلاسيكية) وتقوم على قياس البُنى الكامنة لدى المفحوص وذلك بعد استجابته على مجموعة من أدوات القياس المختلفة، كما تتميز باعتمادها على افتراضات قوية يشترط توفرها في البيانات وهذا الذي أكسبها دقة في القياس (Sijtsma, 1998, 3)؛ حيث توصف افتراضاتها بالقوية حيث تقوم النظرية على افتراض أحادية البعد وهو ما يعني وجود سمة واحدة تسيطر على أداء الفرد في الاختبار، وأن استجابات الأفراد من نفس مستوى القدرة على فقرات

الاختبار المختلفة تكون مستقلة عن بعضها البعض، كذلك تستند النظرية لافتراض التحرر من السرعة والذي يعنى أن الفشل في الاستجابة على الفقرة لا يرجع لعامل الوقت، ويعد المنحنى المميز للفقرة أو منحنى خصائص الفقرة والذي يمثل العلاقة بين السمة السامنة واحتمال الاستجابة الصحيحة أحد أهم افتراضات النظرية واللبنة الأساسية لها حيث يلخص هذا المنحنى جميع معلومات الفقرة (علام، 2005، 63-64؛ Hijazi & Alkateeb, 2014; Yu, 2012; Maldonado & Loreto, 2021؛ Hu et al., 2021؛ ويشير Magno (2009) أن لنظرية الاستجابة للفقرة العديد من الميزات منها:

1- حل المشكلات المتعلقة بالتحليل المتكرر لمجموعات البيانات في كل مرة يتم فيها إجراء اختبار بسبب الجهود المتسقة لإنشاء نتائج أفضل للاختبار.

2- اعتماد النتائج على تقديرات أكثر ثباتًا لموثوقية الاختبارات والمقاييس وصعوبات الفقرات التي يتم نشرها، مقارنة بالنظرية الكلاسيكية.

3- الحاجة إلى إتاحة طرق لاستخدام نظرية IRT، كما يجب إتاحة حزم برامج IRT للمستخدمين بسهولة ويجب جعلها أكثر سهولة في الاستخدام.

نماذج نظرية الاستجابة للفقرة:

هناك العديد من النماذج الرياضية لنظرية الاستجابة للفقرة والتي تختلف في مناسبتها لعدد الأبعاد (أحادية البعد، متعددة الأبعاد) وفي شكل التدريج (ثنائية الاستجابة، متعددة الاستجابات). كما أن هناك بعض النماذج تفترض توفر قدرات أو سمات كثيره تكمن وراء أداء الفرد على الاختبار ويمكن تسميتها بالنماذج متعددة الأبعاد، ويجدر الإشارة هنا إلى وجود تنوع في نماذج السمات الكامنة الأحادية البُعد الثنائية التدريج، وذلك تبعًا لاختلاف عدد (بارامترات) أو معالم الفقرات (علام، 2005، 67)، يمكن توضيحه في الآتي:

- **النموذج أحادي المعلم** 1PLM:

ويعتمد هذا النموذج على معلم صعوبة الفقرة (bi) فقط، حيث يفترض أن معلم التمييز (a) متساوي لكل الفقرات، كذلك يفترض هذا النموذج عدم وجود أثر للتخمين في الاستجابة على الفقرة (الصيخان، 2019، 31) وعندما تساوي قيمة معلم التخمين (1) صحيح يعرف النموذج بنموذج راش Rasch (Baker, 2001) والدالة الرياضية لهذا النموذج تتمثل في:

$$Pi(\Theta)=1/(1+e^{-(\Theta-bi)})$$

حيث $Pi(\Theta)$ تعني احتمال إجابة المفحوص الذي له قدرة (Θ) على الفقرة (i) إجابة صحيحه؛ e = قيمة ثابته = 2.718 تقريبًا.

- **النموذج ثنائي المعلم** "2PL":

يسمى أيضا بنموذج بيرنبوم "Birnbaum Model"، وهذا النموذج يستخدم إذا كانت الفقرات غير متساوية في معلم التمييز، أي هناك تغيير في ميل المنحنى لجميع الفقرات، ويصف معلم التمييز قدرة الفقرة على التمييز بين الأفراد الذين لهم قدرة مختلفة، ويعتبر البعض النموذج الثنائي أكثر منطقية من النموذج الأحادي لأنه من الصعب إيجاد فقرات لها نفس القوة التمييزية، وتتراوح قيم معلم التمييز نظريًا بين (-∞) و (+∞) (Al-Qahtani. 2022؛ القضاة، De Ayala, 2009; 2020)؛ ويمكن التعبير عن هذا النموذج بالدالة الرياضية التالية:

$$Pi(\Theta)=1/(1+e^{-ai(\Theta-bi)})) \text{ (Baker. 2001)}$$

– النموذج ثلاثي المعلم "3PL":

ويســـمى بنموذج لورد "Lord Model"، وتظهر فيه معلمة التخمين والذي لا يوجد طريقة لتحديده في النظرية الكلاسيكية، ويعتبر هذا النموذج هو الحالة العامة للنموذجين الأسبقين، وكنتيجة حقيقية موجودة فإن جزء من احتمالية الاستجابة الصحيحة ترجع لمعلم التخمين، ويمكن تمثيله في الدالة اللوجستية بخط تقاربي سفلي لا يساوي الصفر (Al-Qahtani, 2022)؛ ويمكن التعبير عن هذا النموذج رياضيا كالتالي:

Pi(Θ)=Ci+(1-Ci) (1/(1+e^(-ai(Θ-bi) (Baker. 2001)

مطابقة الفقرات في نظرية الاستجابة للفقرة:

يعد اختبار مطابقة الفقرة خطوة مهمة عند معايرة وتحليل اختبار يستند إلى نظرية استجابة الفقرة (IRT)، حيث يعد مطابقة النموذج شرطًا أساسيًا ضروريًا لاستخلاص استنتاجات صحيحة من المعلمات المقدرة، وتشير المطابقة الضعيفة للفقرة إلى أن معلمات الفقرة لها صلاحية مشكوك فيها (أي أنها لا تمثل بدقة كيفية استجابة المفحوصين لفقرات الاختبار)، والحصول على دليل يشير إلى توافق بيانات النموذج عند استخدام نموذج IRT مع عمل استنتاجات من مجموعة بيانات، وهذا الدليل شرط يوصى به كمعايير للاختبار التربوي والنفسي من قبل الجمعية الأمريكية للبحوث التربوية، والجمعية الأمريكية لعلم النفس، والمجلس الوطني للقياس في التعليم، ويؤدي عدم تحقيق هذا الشرط إلى فشل تطبيق IRT في تقييم مجموعة البيانات الحقيقية قبل استخلاص أي استنتاجات من النموذج المناسب، ويمكن أن يؤدي ذلك إلى أي استنتاجات قد تكون مضللة مستمدة من نماذج غير جيدة (Maydeu-Olivares, 2015; Liu & Maydeu-Olivares, 2014)؛ وتوجد العديد من طرق تقييم مطابقة الفقرة منها (طريقة فحص النموذج التنبؤية اللاحقة، طريقة تحليل البواقي، مربع كاي (X2) لجودة المطابقة) (Oribhabor & Altuariariagbon, 2020)؛ وتوجد العديد من العوامل التي تؤثر على مطابقة الفقرة في نماذج نظرية الاستجابة للفقرة منها:

أ– حجم العينة:

يتطلب التقدير الفعال لمعلمات الفقرة في نماذج نظرية الاستجابة للفقرة IRT استجابات من عينات كبيرة من الأشخاص على الرغم من عدم وجود قواعد عملية لتحديد أحجام العينات، فمن المقبول عمومًا أن هناك حاجة لأحجام أكبر للعينات مع زيادة عدد المعلمات في النموذج على سبيل المثال إذا كان ما لا يقل عن (150) شخصًا كافيًا لتقدير فعال لمعلمة الصعوبة في النموذج أحادي المعلم PL1، فقد تكون هناك حاجة إلى عينات أكبر (10) مرات لتقدير دقيق للمعلمات في نموذج IRT متعدد الاستجابة، ويلزم وجود عينات كبيرة بشكل خاص عند معايرة مجموعات أكبر من الفقرات لإدراجها في بنوك الأسئلة، ومن هذا يتضح تأثير حجم العينة على مطابقة الفقرات أثناء حساب مؤشرات الجودة (Van Der Linden, 2018).

ب–تحديد النموذج:

أجري العديد من الدراسات التي تختبر نماذج مختلفة من نظرية الاستجابة للفقرة لمعرفة تأثير اختلاف النماذج على التحقق من افتراضات هذه النظرية وكانت النتائج تشير إلى أهمية استخدام نماذج مختلفة للتأكد من النتائج

التي يتم الوصـــل لها (Cao et al., 2014)؛ وقد حدد علام (2005) المحكات التي يجب مراعاتها عند اختيار النموذج المناسـب على النحو التالي: (تحقق افتراضـات النموذج- أنواع البيانات الاختبارية المراد تحليليها- تحمل النموذج لمخالفة افتراضـاته- خبرة الباحث بالنموذج المسـتخدم، توافر برامج الحاسـوب لتنفيذ ما يتطلبه النموذج من عمليات إحصائية وحسابية).

ج- طريقة التقدير:

مما لا شـك فيه أن صـدق وموضـوعية الاختبارات أو المقاييس يعتمد على مجموعة من الأمور ومنها دقة الأسـاليب التي اسـتخدمت في بناءها واختيار فقراتها وتفسـير نتائجها، وكذلك في وصـفها للقدرة التي يقيسـها الاختبار، ومن طرق التقدير التي شـاع انتشـارها طريقة الأرجحية العظمى، طرق التقدير البيزية، وطريقة بروكس وغيرها من الطرق والتي تختلف دقة نواتجها باختلاف تحقق الافتراضات القائمة عليها (ضعضع وآخرون، 2020).

ثانيًا: التحليل العاملي التوكيدي:

ترجع بدايات التطور التاريخي للنمذجة البنائية SEM إلى القرن العشـرين، حيث ظهرت النمذجة بعد التعرض لتطور النماذج الثلاثة الأساسية (تحليل الانحدار - التحليل العاملي- تحليل المسار)، وكان لكارل بيرسون Karl Pearson جهود في تحليل الانحدار بعد أن أوجد معادلة الارتباط، ثم بعد ذلك اسـتخدم تشـارلز سـبيرمان Charles Spearman معامل الارتباط لتحديد الفقرات التي تتجمع معًا وذلك لإيجاد النموذج العاملي، وفي عام (1940) أضاف كل ثرستون ولولي Thurston and Lolly تطبيقات إضافية أيضا في التحليل العاملي ترجع الجهود للعالم (هو) Ho عام 1922 وكلا من واندرسـون وروين Anderson and Rowen (1956)، وقد أعطى كارل جورسـكوج Carl Gorskog نظرة متكاملة للتحليل العاملي التوكيدي، وأخيرًا العالم البيولوجي سـيول راتب Salary Seoul الذي قام بتطوير ما يعرف بتحليل المسـار والذي يستخدم في معاملات الارتباط وتحليل الانحدار من أجل نمذجة علاقات أكثر تعقيدًا وقد استخدمت التطبيقات الأولية لتحليل المسار مع نماذج السلوك الحيواني وكانت بعيدة عن الأنظار في الخمسينيات (Byrne, 2010).

وقد تضـافرت جهود العلماء في تحديد تطبيقات نمذجة المعادلة البنائية وتم التوصـل إلى أربعة أنواع من هذه التطبيقـات تتمثـل في نموذج تحليـل المسـار "Path analysis model" ونموذج الانحـدار البنـائي "Structural regression models" ونموذج التغيرات الكامنة "Latent change Model" ونموذج التحليـل العـاملي التوكيـدي (Byrne, Confirmatory factor analysis model 2010; Raykov & Marcoulides, 2006; Fan, et al., 2016)

وبالنسـبة للتحليل العاملي التوكيدي كأحد تطبيقات النمذجة بالمعادلات البنائية فيذكر (De Coster 1998) أن هناك بعض الاستخدامات الشائعة له والتي من أهمها: اثبات صدق نموذج العامل العام مقارنة جودة مطابقة نماذج نظرية مختلفة لنفس مجموعة البيانات، اختبار أهمية تشـبعات محددة، اختبار العلاقة بين عاملين أو أكثر، اختبار ما إذا كانت مجموعة من العوامل مترابطة أم غير مرتبطة، تقييم الصـدق التقاربي والصـدق التمييزي لأداة القياس؛ ولتحقيق ذلك يتم الاعتماد على العديد من مؤشرات حسن المطابقة والتي من أهمها كما يذكر تيغزة (2012) مؤشر مربع كاي، مؤشر حسن المطابقة (GFI)، مؤشر توكر– لويس (TLI)، مؤشر المطابقة المقارن (CFI)، مؤشـر المطابقة المعياري (NFI)، الجذر التربيع لمتوسـط مربعات البواقي (RMR)، الجذر التربيعي لمتوسط مربعات خطأ التقريب (RMSEA).

العوامل المؤثرة في مؤشرات جودة البنية العاملية للمقياس:

أ- حجم العينة:

تعتبر مسألة حجم العينة مهمة في نمذجة المعادلة البنائية، ولا توجد قاعدة عامة قاطعة لتحديد عدد المشاركين المطلوبين حتى يكون نموذج التحليل العاملي التوكيدي موثوقًا به، ومع ذلك فإن معظم الباحثين يتفقون على أن هناك حاجة إلى 200 مشاركًا على الأقل إن لم يكن 300 مشاركًا، ويقترح كابلان Kaplan (1995) أن تقييم النماذج يجب أن يعتمد في المقام الأول على اعتبارات القوة، وليس على حجم العينة، وهناك حاجة إلى عينات أكبر على التوالي عند استخدام العديد من المتغيرات المقاسة في نموذج معقد في التحليل العاملي التوكيدي لأن عدد المعلمات التي سيتم تقييمها يزداد، ويجب توخي الحذر عند إجراء دراسة بحجم عينة أقل من (200) (Phakiti et al., 2018).

ب- عدد الفقرات واتساقها:

تشير العديد من الدراسات إلى تأثير عدد الفقرات على موثوقية النتائج ومؤشرات جودة المطابقة، ويجب أن يحتوي كل مقياس فرعي على عدد فقرات كافية مع مراعاة أن 30% من الفقرات قد يتم حذفها أثناء التحليل، ويصبح من الصعب تحديد ما إذا كانت هناك "فقرات كافية" أم لا لكن المقياس الذي يحتوي على أقل من 10 فقرات يكون الوضع فيه حرج (Ellis, 2017)؛ ويرى العمري (2011) أن خصائص الفقرات ومدى اتساقها ومضمونها وطريقة تركيبها من المشكلات المهمة والحديثة التي تتعلق ببناء المقياس النفسي.

ج- ثبات وصدق المتغيرات الملاحظة:

يحتاج الباحثون إلى تقديم دليل على ثبات وصحة بياناتهم، والثبات من أهم الاعتبارات التي يجب أن تؤخذ في الاعتبار عند استخدام التحليل العاملي التوكيدي لأنه إذا كان لا يمكن ملاحظة بعض السلوكيات أو الأفكار أو المهارات بشكل متسق، فلا يمكن للبيانات المرتبطة بها أن تكون صادقة، فيجب ألا يقل معامل ثبات ألفا كرونباخ على سبيل المثال عن 0.7 ومعامل ثبات كيودر وريتشاردسون KR20 يجب ألا يقل عن 0.8 إن لم يكن 0.9 (Phakiti et al., 2018).

د- توزيع الدرجات:

تتأثر مؤشرات جودة البنية العاملية للمقياس بطبيعة توزيع الدرجات، وهنا يعد شرط التوزيع الطبيعي متعدد المتغيرات للمتغيرات الملاحظة شرط ضروري لاستخدام التحليل العاملي التوكيدي، وتتأثر بعض طرق التقدير بعدم تحقق هذا الشرط، والذي يشير إلى أن الملاحظات مستقلة وموزعة بشكل متماثل، وكذلك يجب أن يكون التفرطح متعدد المتغيرات صفرًا أو قريبًا من الصفر، ويمكن التغلب على هذه المشكلة إحيائنا بإزالة القيم المتطرفة متعددة المتغيرات قبل إجراء التحليل (Phakiti. et. al., 2018).

ثالثًا: أسلوب المحاكاة:

المحاكاة هي إجراء حاسوبي مكثف لتقييم الأساليب الإحصائية بالنظر إلى نماذج ومعلمات معروفة، وفي أغلب الأحيان يتم ذلك باستخدام أرقام عشوائية لتوليد البيانات وتحليلها باستخدام النموذج وتمكننا العينات العشوائية التي تم توليدها من إنشاء توزيع عينات تجريبي للمعلمات في ظل الظروف التي نحتاج لها وتتكون خطوات المحاكاة من صياغة مشكلة، وتصميم دراسة محاكاة، وتوليد استجابات الفقرات وتقدير معلمات النموذج، وتحليل نتائج المحاكاة (Lee et al., 2018).

ويذكر أن هذا الأسلوب شائع جدًا في الأدبيات البحثية المتعلقة بنظرية الاستجابة للفقرة (IRT)، ويتطلب هذا الأسلوب مراعاة خمسة أنواع من البيانات وهي كما يلي: (بيانات الاستجابة المشاهدة (المولدة)، توليد المعلمات المستندة إلى النموذج، تقديرات المعلمات، وظائف الاستجابة بناًء على المعلمات الحقيقية، وظائف الاستجابة بناًء على المعلمات المقدرة) ويقصد بالبيانات المرصودة الدرجات التي يتم توليدها بموجب نموذج IRT، ثم يتم التعامل مع المعلمات المستندة إلى نموذج التوليد كما هو معروف في دراسة المحاكاة ويتم استخدامها لإنشاء البيانات وربما تقييم أي تقديرات ناتجة، كما تُستخدم البيانات المرصودة أيضًا لتقدير المعلمات (أي حساب الإحصائيات) ضمن نموذج نظرية الاستجابة للفقرة IRT معين (Luecht & Ackerman, 2018).

إجراءات الدراسة:

مجتمع الدراسة:

اعتمد الباحثان في الدراسة الحالية على المحاكاة في تحقيق أهداف الدراسة والإجابة عن أسئلتها وبالتالي سيتم الاعتماد على بيانات افتراضية تتوافر بها خصائص معينة، وبلغ إجمالي عدد المعلومات التي تم توليدها (900000) معلومة، وإجمالي عدد النماذج موضوع المقارنة في الدراسة بلغ (100) نموذجًا.

عينة الدراسة:

تكونت عينة الدراسة من عدد مختلف من أحجام العينة (200- 400- 600- 800-1000) مفحوصًا افتراضيًا تم توليدهم كقدرات باستخدام برمجيات المحاكاة حيث تتبع قدراتهم التوزيع الطبيعي بمتوسط حسابي (0) وانحراف معياري (1)، وتم تقديرها من خلال: طريقة المربعات الصغرى الموزونة WLS، حيث تستخدم هذه الطريقة عندما تكون البيانات التي نعمل عليها ذات توزيع غير ثابت، أي أن تباين البيانات ليس متساويًا في جميع النقاط لأنها تقوم على معالجة اختلافات التوازن في البيانات عن طريق تخصيص أوزان مختلفة لكل نقطة في عملية المعايرة لضمان أن النموذج المعاير يكون دقيقًا وموثوقًا حتى عندما تكون البيانات مختلفة في توزيعها (Raposo, 2016).

أداة الدراسة:

تم الاعتماد على استجابات افتراضية على مقاييس افتراضية بطول (20- 40- 60- 80-100) فقرة والتي تم توليدها من خلال البرمجيات المتعلقة بتوليد بيانات وتحديدًا برنامج (R) والتي تحاكي ظروف اختبارية تتعلق بالدراسة.

الأساليب الإحصائية التي تم استخدامها:

لتحقيق أهداف الدراسة والاجابة على الأسئلة تم بالمجمل استخدام مجموعة من الأساليب الإحصائية المناسبة لإجراء المعالجات الإحصائية اللازمة حيث تم الاستعانة ببرمجية AMOS. SPSS. R. Excel.

نتائج الدراسة وتفسيراتها:

أولًا: الاجابة على السؤال الأول:

نص السؤال الأول للدراسة الحالية على "هل تختلف الأوزان الانحدارية المعيارية "التشبعات" كمؤشرات لجودة الصدق البنائي للمقياس باختلاف مطابقة الفقرات للنموذج ثلاثي المعلم في نظرية الاستجابة للفقرة وفقًا لحجم العينة وعدد الفقرات؟".

وللإجابة عن هذا السؤال تم حساب المتوسطات الحسابية والانحرافات المعيارية للأوزان الانحدارية المعيارية "التشبعات" باختلاف مطابقة الفقرات للنموذج ثلاثي المعلم في نظرية الاستجابة للفقرة وفقًا لحجم العينة وعدد الفقرات، كما هو موضح بجدول (1):

جدول (1) المتوسطات الحسابية والانحرافات المعيارية للأوزان الانحدارية المعيارية باختلاف مطابقة الفقرات للنموذج ثلاثي المعلم وفقاً لحجم العينة وعدد الفقرات

نسبة الفقرات غير المطابقة	عدد الفقرات	حجم العينة											
		200		400		600		800		1000		الكلي	
		م	ع	م	ع	م	ع	م	ع	م	ع	م	ع
صفر%	20	0.511	0.077	0.515	0.062	0.475	0.058	0.487	0.051	0.490	0.049	0.496	0.061
	40	0.530	0.531	0.480	0.065	0.482	0.058	0.472	0.054	0.472	0.048	0.488	0.063
	60	0.594	0.084	0.519	0.072	0.504	0.069	0.502	0.058	0.490	0.057	0.490	0.057
	80	0.651	0.099	0.537	0.061	0.514	0.063	0.503	0.050	0.500	0.052	0.541	0.052
	100	0.716	0.100	0.556	0.066	0.529	0.058	0.523	0.065	0.502	0.044	0.565	0.104
%5	20	0.477	0.136	0.439	0.126	0.455	0.110	0.454	0.115	0.471	0.109	0.459	0.182
	40	0.508	0.125	0.459	0.115	0.456	0.115	0.451	0.101	0.450	0.095	0.465	0.112
	60	0.572	0.126	0.496	0.115	0.482	0.117	0.480	0.115	0.469	0.105	0.500	0.121
	80	0.628	0.157	0.509	0.138	0.490	0.129	0.479	0.110	0.477	0.114	0.517	0.142
	1005	0.682	0.167	0.529	0.133	0.507	0.123	0.490	0.116	0.480	0.113	0.538	0.151
%10	20	0.458	0.163	0.413	0.158	0.433	0.150	0.432	0.151	0.446	0.157	0.436	0.153
	40	0.484	0.174	0.435	0.148	0.432	0.146	0.432	0.132	0.430	0.128	0.442	0.146
	60	0.543	0.188	0.472	0.151	0.459	0.145	0.456	0.148	0.444	0.142	0.475	0.159
	80	0.607	0.197	0.483	0.172	0.466	0.164	0.456	0.148	0.455	0.149	0.493	0.176
	100	0.412	0.145	0.499	0.178	0.482	0.163	0.466	0.154	0.458	0.149	0.463	0.160
%20	20	0.400	0.217	0.368	0.198	0.385	0.190	0.393	0.191	0.396	0.200	0.389	0.195
	40	0.428	0.230	0.408	0.207	0.390	0.186	0.384	0.184	0.385	0.172	0.399	0.196
	60	0.485	0.255	0.424	0.203	0.414	0.199	0.412	0.199	0.402	0.192	0.402	0.212
	80	0.545	0.269	0.439	0.216	0.419	0.210	0.408	0.201	0.410	0.199	0.444	0.225
	100	0.579	0.310	0.451	0.218	0.432	0.214	0.419	0.197	0.411	0.194	0.458	0.238

يلاحظ من نتائج جدول (1) وجود فروق ظاهرية بين المتوسطات الحسابية لقيم الأوزان الانحدارية المعيارية (التشبعات)، وحسب المتغيرات حجم العينة وعدد الفقرات ونسبة الفقرات غير المطابقة، ولمعرفة الدلالة الإحصائية لتلك الفروق الظاهرية تبعًا للمتغيرات حجم العينة وعدد الفقرات ونسبة الفقرات غير المطابقة والتفاعلات المشتركة بينهما، فقد تم استخدام تحليل التباين الثلاثي كما هو موضح بجدول (2):

جدول (2) دلالة الفروق في الأوزان الانحدارية المعيارية باختلاف مطابقة الفقرات للنموذج ثلاثي المعلم وفقًا لحجم العينة وعدد الفقرات

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة "ف"	مستوى الدلالة
نسبة الفقرات غير المطابقة	6.011	3	2.004	86.512	0.01
عدد الفقرات	2.883	4	0.721	31.121	0.01
حجم العينة	4.770	4	1.192	51.487	0.01
نسبة الفقرات غير المطابقة*عدد الفقرات	0.638	12	0.053	2.296	0.01
نسبة الفقرات غير المطابقة * حجم العينة	0.304	12	0.025	1.095	0.359
عدد الفقرات*حجم العينة	1.379	16	0.086	3.721	0.01
نسبة الفقرات غير المطابقة*عدد الفقرات*حجم العينة	2.414	48	0.050	2.172	0.01
الخطأ	136.639	5900	0.023		
الكلي	1588.894	6000			
الكلي المصحح	**163.739**	**5999**			

يتضح من جدول (2) أنه:

1- توجد فروق ذات دلالة إحصائية عند مستوى (0.01) بين متوسطات قيم الأوزان الانحدارية المعيارية (التشبعات) ترجع لاختلاف نسبة الفقرات غير المطابقة، ولعدد الفقرات، ولحجم العينة.

2- توجد فروق ذات دلالة إحصائية عند مستوى (0.01) في متوسطات قيم مؤشر الأوزان الانحدارية المعيارية (التشبعات) ترجع لتأثير التفاعل بين نسبة الفقرات غير المطابقة وعدد الفقرات، ولتأثير التفاعل بين عدد الفقرات وحجم العينة، ولتأثير التفاعل بين عدد الفقرات وحجم العينة ونسبة الفقرات غير المطابقة.

3- لا توجد فروق ذات دلالة إحصائية في متوسطات قيم مؤشر الأوزان الانحدارية المعيارية (التشبعات) ترجع لتأثير التفاعل بين حجم العينة ونسبة الفقرات غير المطابقة.

ونلاحظ بشكل عام وجود تأثير جميع المتغيرات حجم العينة وعدد الفقرات ونسبة الفقرات غير المطابقة والتفاعلات المشتركة بينهما باستثناء التفاعل بين نسبة الفقرات غير المطابقة وحجم العينة على التشبعات "الأوزان الانحدارية المعيارية"، ولمعرفة اتجاه الفروق في حالة التأثيرات الرئيسية لنسبة الفقرات غير المطابقة وحجم العينة وعدد الفقرات على الأوزان الانحدارية المعيارية تم الاعتماد على اختبار مدى توكي للفرق الدال الصادق (Tukey HSD) كاختبار للمقارنات البعدية، فكانت النتائج كما هي موضحة في التالي:

1- نسبة الفقرات غير المطابقة:

جدول (3) المقارنات البعدية للأوزان الانحدارية المعيارية باختلاف نسبة الفقرات غير المطابقة للنموذج ثلاثي المعلم

نسبة الفقرات غير المطابقة	**0% (م=0.535)**	**5% (م= 0.510)**	**10% (م= 0.469)**
5% (م= 0.510)	*0.025		
10% (م= 0.469)	*0.066	*0.041	
20% (م= 0.436)	*0.099	*0.074	*0.033

* الفرق بين المتوسطين دال عند مستوى (**0.05**) ** الفرق بين المتوسطين دال عند مستوى (**0.01**).

يتضح من جدول (3) أنه كلما زادت نسبة الفقرات غير المطابقة تقل الأوزان الانحدارية المعيارية (التشبعات)، حيث كانت الفروق دالة إحصائيًا عند مستوى دلالة 0.05 في الأوزان الانحدارية المعيارية بين كل نموذج والتالي له لصالح النموذج الأقل في نسبة المفردات غير المطابقة؛ وهذا يؤكد تأثر الأوزان الانحدارية المعيارية سلبًا بنسبة الفقرات غير المطابقة، وبالتالي يمكن استنتاج أن ارتفاع الأوزان الانحدارية المعيارية (التشبعات) في التحليل العاملي التوكيدي يعد دليلًا على جودة مطابقة الفقرات لنموذج نظرية الاستجابة للمفردة، فعندما تحتوي البيانات على فقرات غير مطابقة يمكن أن تكون هناك عوامل مشتركة خفية قد تؤثر على العلاقة بين المتغيرات، وهذه العوامل يصعب قياسها أو التحكم فيها، وبالتالي يمكن أن يؤثر على قيم الأوزان الانحدارية المعيارية ويجعلها أقل دقة.

ويتأكد مما سبق أن حذف الفقرات غير المطابقة للنموذج لثلاثي المعلم يعد إجراء مهم لضمان ارتفاع تشبعات الفقرات على عواملها الكامنة، وعندما تحتوي البيانات على فقرات غير مطابقة لنموذج نظرية الاستجابة للفقرة، فقد يرجع ذلك لتأثر الأداء على هذه الفقرات بعوامل أخرى تختلف عن السمة التي يتم قياسها، وهو ما يفسر انخفاض قيم الأوزان الانحدارية المعيارية (التشبعات)، أو أن الفقرة تتشبع على أكثر من متغير كامن، ومجمل النتائج هنا تؤكد على ضرورة التعرف على الفقرات غير المطابقة وحذفها.

2- عدد الفقرات:

جدول (4) المقارنات البعدية للأوزان الانحدارية المعيارية باختلاف عدد فقرات المقياس

عدد الفقرات	20 (م= 0.445)	40 (م= 0.448)	60 (م= 0.480)	80 (م= 0.498)
40 (م= 0.448)	0.003			
60 (م= 0.480)	*0.035	*0.032		
80 (م= 0.498)	*0.053	*0.050	*0.018	
100 (م= 0.506)	*0.061	*0.058	*0.026	0.008

***الفرق بين المتوسطين دال عند مستوى (0.05)، ** الفرق بين المتوسطين دال عند مستوى (0.01).**

يتضح من جدول (4) وبصفة عامة أن الأوزان الانحدارية المعيارية (التشبعات) تزداد بزيادة عدد فقرات المقياس، حيث كانت هناك فروق دالة إحصائيًا عند مستوى دلالة 0.05 بين متوسطي الأوزان الانحدارية المعيارية لكل نموذجين متتاليين لصالح النموذج الأعلى في عدد الفقرات ماعدا الفروق بين النموذجين (20 فقرة – 40 فقرة)، والنموذجين (80فقرة، 100 فقرة).

ومن النتائج السابقة يتضح تأثر الأوزان الانحدارية المعيارية بزيادة عدد الفقرات وكلما زاد عدد الفقرات كلما زادت قيم التشبعات، والذي قد يرجع لأن زيادة عدد الفقرات يجعل الفقرات أكثر اتساقًا فيما بينها وهو ما يسهم في زيادة ثباتها، والذي ينعكس إيجابيا على التشبع، كذلك التشبع هو مقدار إسهام الفقرة في تكوين المتغير الكامن (التباين المشترك) فكلما زاد عدد الفقرات يزداد التباين المشترك بينها وهو ما يسهم في زيادة التشبع؛ ويجب ملاحظة أن تأثير عدد الفقرات على التشبعات كان تأثيرًا ضعيفًا، وعليه فيجب ألا يكون عدد الفقرات مبالغ فيه بشكل كبير، ويكفي أن يكون عدد الفقرات كافي لتمثيل المتغير الكامن تمثيلا جيدًا، والنتائج السابقة يمكن

تفسيرها وفقًا لأن الزيادة في عدد الفقرات يسهم في زيادة قيم معاملات الثبات ويسهم في قلة أخطاء القياس الأمر الذي يؤثر إيجابيًا في قيم معاملات الارتباط المكون الأساسي لحساب التشبعات المعيارية وغير المعيارية (Goodwin & Leech, 2006).

3- حجم العينة:

جدول (5) المقارنات البعدية للأوزان الانحدارية المعيارية باختلاف حجم العينة

حجم العينة	200 (م= 0.566)	400 (م= 0.484)	600 (م= 0.469)	800 (م=0.461)
400 (م= 0.484)	*0.082			
600 (م= 0.469)	*0.097	0.015		
800 (م= 0.461)	*0.105	*0.023	0.008	
1000 (م= 0.455)	*0.111	*0.029	0.014	0.006

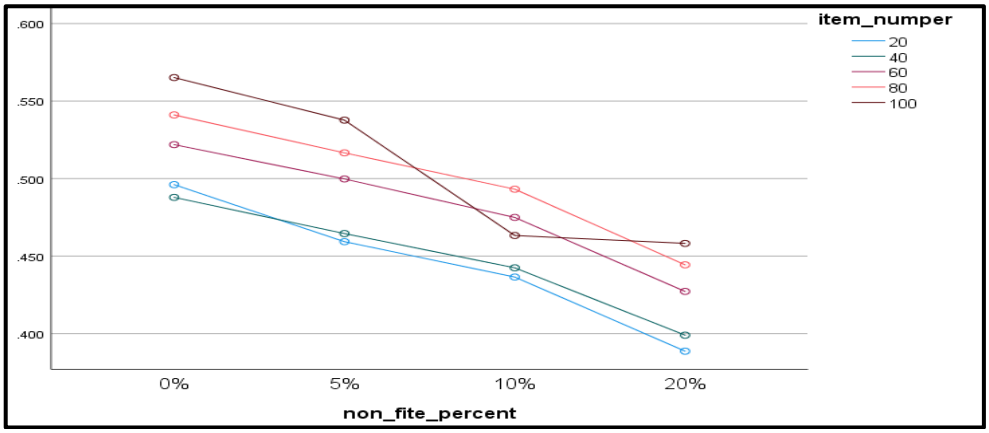
***الفرق بين المتوسطين دال عند مستوى 0.05، ** الفرق بين المتوسطين دال عند مستوى (0.01).**

يتضح من جدول (5) أنه بزيادة حجم العينة تقل الأوزان الانحدارية المعيارية (التشبعات)، حيث كانت هناك فروق دالة إحصائيًا عند مستوى دلالة 0.05 بين متوسط التشبعات عند حجم عينة 200 ومتوسط التشبعات عند باقي حجوم العينات لصالح متوسط التشبعات عند حجم عينة 200، كذلك كانت هناك فروق دالة إحصائيًا عند مستوى دلالة 0.05 بين متوسط التشبعات عند حجم عينة 400 وعند حجم عينة 800 وكذلك عند حجم عينة 1000 لصالح متوسط التشبعات عند حجم العينة 400، بينما كانت الفروق غير دالة إحصائيًا بين متوسطات التشبعات عند حجوم العينات 600، 800، و1000.

ومما سبق يمكن استنتاج أنه مع الزيادة الكبيرة لحجم العينة (زيادة حجم العينة عن 600) فإنه تصبح الفروق بين الأوزان الانحدارية المعيارية (التشبعات) غير دالة إحصائيًا، وهذه النتيجة على قدر كبير من الأهمية في مجال النمذجة البنائية، والتي كثر فيها الجدل حول حجم العينة المناسبة للتحليل العاملي التوكيدي، والدراسة الحالية تؤكد أن حجم العينة بين 400 إلى 600 كافي لتقدير تشبعات مستقرة للفقرات على العوامل الكامنة؛ ويمكن تفسير زيادة قيم التشبعات في حالة العينات الصغيرة (200) والمتوسطة 400 إلى 600 وعدم تغيرها في حالة العينات الكبيرة في ضوء أن حجم العينة لا تؤثر بصورة مباشرة على قيم التشبعات في التحليل العاملي التوكيدي إلا أنها تؤثر بصورة غير مباشرة من خلال تأثرها في قيم معاملات الارتباط المستخدمة في حساب التشبعات، حيثأشار كل من جودون وليبتش Goodwin & Leech (2006) إلى وجود ست عوامل تؤثر على قيمة معامل الارتباط في البيانات أهمها التباين الذي له علاقة مباشرة بحجم العينة، فلكما زاد التباين في الدرجات أثر ذلك إيجابيًا علي قيم معاملات الارتباط، كما أن التشبعات المعيارية يتم فيها حساب تباين درجات المتغيرات المشاهدة والكامنة، وكلما قل حجم العينة كلما زاد التباين في درجات الأفراد سبب ما يطلق عليه تقلب العينة Sample Fluctuation؛ أما بالنسبة لتأثير التفاعلات فيمكن توضحه من خلال الرسوم البيانية كما هو موضح في التالي:

1-التفاعل الثنائي بين نسبة الفقرات غير المطابقة وعدد الفقرات

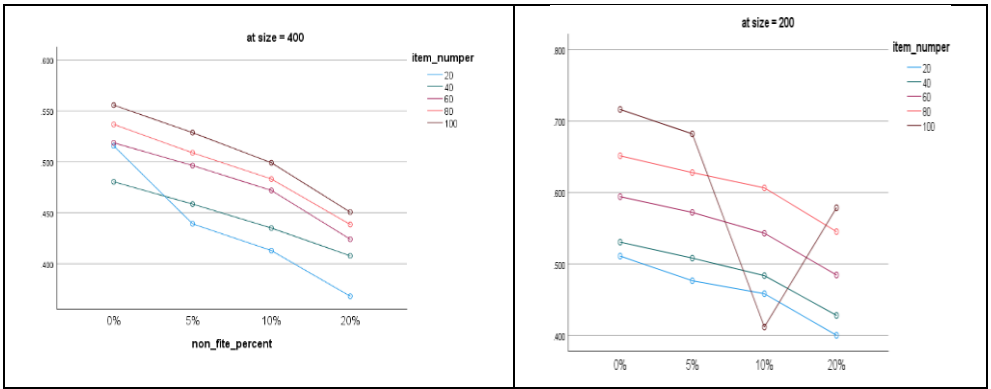
شكل (1) تأثير التفاعل الثنائي بين عدد الفقرات ونسبة الفقرات غير المطابقة على الأوزان الانحدارية المعيارية



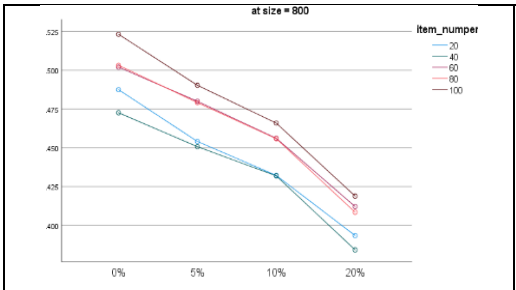
يلاحظ من شكل (1) أن الفروق الناتجة من التفاعل الثنائي بين نسبة الفقرات غير المطابقة وعدد الفقرات تعزى لنسبة الفقرات غير المطابقة (0%) وعدد فقرات (100)؛ وهذا يؤكد أن الأوزان الانحدارية المعيارية تكون أفضل إذا كانت نسبة الفقرات غير المطابقة صفر وكان عدد الفقرات كبير، والسبب في ذلك أن نسبة الفقرات غير المطابقة الأقل وحجم العينة الكبير يساهمان في تحسين دقة وثبات الأوزان الانحدارية المعيارية مما يزيد من موثوقية النموذج الإحصائي، ويعني أن النموذج يمكن أن يقدم نتائج أكثر دقة واستقرارًا في تمثيل العلاقات بين المتغيرات والتنبؤ بالنتائج بشكل أفضل.

2-التفاعل الثلاثي بين (حجم العينة وعدد الفقرات ونسبة عدم المطابقة):

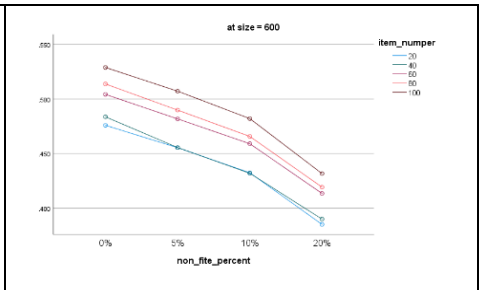
شكل (2) تأثير التفاعل الثنائي بين عدد الفقرات ونسبة الفقرات غير المطابقة عند أحجام العينة المختلفة على الأوزان الانحدارية المعيارية



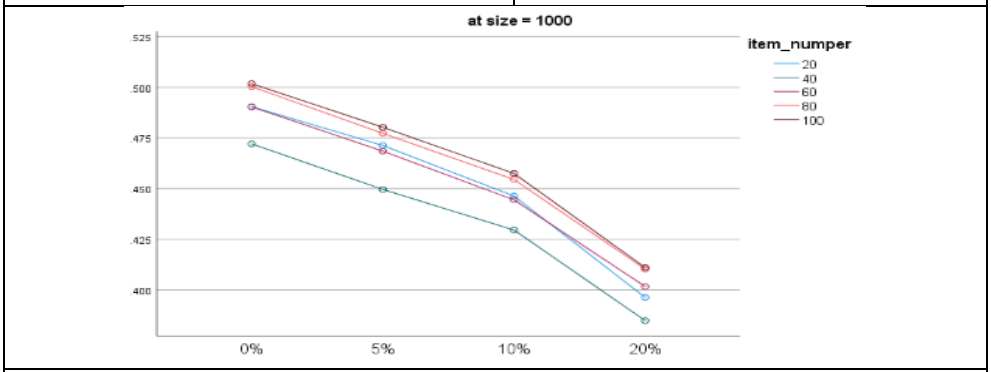
أ- عند حجم عينة (200)	أ- عند حجم عينة (400)



د- عند حجم عينة (800)



ح- عند حجم عينة (600)



د- عند حجم عينة (1000)

يلاحظ من شكل (2) أن زيادة حجم العينة يسهم في استقرار تقديرات الأوزان الانحدارية المعيارية حيث تزداد التشبعات بزيادة عدد الفقرات ونقص نسبة الفقرات غير المطابقة للنموذج ثلاثي المعلم، حيث تقترب التقديرات في حالة عدد الفقرات (80) منها في حالة عدد الفقرات (100) كذلك تقترب التقديرات في حالة عدد الفقرات (60) منها في حالة عدد الفقرات (40) ومن الواضح نقص قيم التشبعات في حالة انخفاض عدد الفقرات (20) عنها في باقي الحالات، وربما كان هذا الاستقرار في النتائج راجع لدقة تقدير مطابقة الفقرات للنموذج الثلاثي بزيادة حجم العينة، فهناك توصيات في بعض الأدبيات بأن يكون حجم العينة في حالة النموذج الثلاثي حجم كبير ولا يقل عن (1000)، ومن الواضح أن زيادة نسبة الفقرات غير المطابقة في حالة عدد الفقرات المنخفض (20 فقرة) يتسبب في ضعف في قيم التشبعات والتي يتضح من الشكل أنها أقل من (0.2)، ومن المهم هنا التأكيد على جزئية أن زيادة حجم العينة يسهم في دقة التقديرات حتى لو استخدمنا أدوات عدد فقراتها صغير (20 فقرة) بشرط جودة هذه الفقرات أما في حالة زيادة عدد الفقرات غير المطابقة فمن الضروري أن يكون عدد الفقرات كبير وأكبر من (20) فقرة، وهنا يمكن أن توضح النتائج السابقة لماذا يفضل استخدام عينات كبيرة في حالة النموذج ثلاثي المعلم هذا من جانب ومن جانب أخر في حالة المقاييس الطويلة والتي تتضمن عدد كبير من الفقرات، والذي يمكن تفسيره بأن زيادة حجم العينة يسهم في دقة تقدير المطابقة للنموذج الثلاثي والذي ينعكس إيجابيًا على جودة البنية العاملية للمقياس وخاصة ما يتعلق بتشبعات الفقرات على المتغير الكامن.

وبشـــكل عام تؤكد النتائج الســـابقة أن تقدير قيم الأوزان الانحدارية المعيارية (التشـــبعات) يتأثر كثيرًا بعدد الفقرات حيث تزداد قيم التشـــبعات بزيادة عدد الفقرات، وتقل التشـــبعات بزيادة نســـبة الفقرات غير المطابقة للنموذج الثلاثي، والذي يتطلب في هذه الحالة – زيادة نســـبة الفقرات غير المطابقة- عينات كبيرة للوصـــول لتقديرات دقيقة للتشبعات، ومن جهة أخرى يحتاج لزيادة عدد فقرات المقياس.

وتؤكد النتائج السابقة أنه مع زيادة حجم العينة تستقر تقديرات الأوزان الانحدارية المعيارية في حالة زيادة عدد الفقرات عن (20) ويكاد يكون هناك تشـابه بين التشـبعات في حالة عدد الفقرات (40) وعدد الفقرات (60) وكذلك في حالة عدد الفقرات (80) وعدد الفقرات (100)، وهو ما يعني أن دقة تقديرات التشـبعات لا يحتاج لعدد كبير من الفقرات في حالة جودة مطابقة هذه الفقرات للنموذج الثلاثي أما في حالة زيادة نسـبة الفقرات غير المطابقة فنحتاج لزيادة عدد الفقرات وزيادة حجم العينة للوصـول لتقديرات مسـتقرة للتشـبعات، ويمكن أن تفسـر النتائج الســـابقة وفقا لأن زيادة عدد الفقرات وزيادة حجم العينة يســـهم في زيادة تقدير العلاقات المشـــتركة بين الفقرات، وهو ما يســـهم كذلك في دقة تمثيل المتغير الكامن، وبالتالي يكون النموذج البنائي أكثر دقة ومطابقة للبيانات الفعلية موضـوع المعالجة، وبالنسـبة للفقرات غير المطابقة والتي تعد فقرات غير جيدة وفقًا للنموذج ثلاثي المعلم فإن انخفاض نسبتها يسهم في جعل الفقرات أكثر تجانسًا وأعلى في تشبعها على المتغير الكامن، والناتج من زيادة الارتباطات بين الفقرات وهو ما يجعل تقدير الأوزان الانحدارية المعيارية يتسم بالدقة والاستقرار وهو ما يتفق مع ما يشير إليه (ضعضع وآخرَيْن، 2020؛ والثوابية، 2010).

ثانيًا: نتائج الاجابة على السؤال الثاني:

نص الســـؤال الثاني للدراســـة الحالية على "هل تختلف البواقي كمؤشـــرات لجودة الصـــدق البنائي للمقياس باختلاف مطابقة الفقرات للنموذج ثلاثي المعلم في نظرية الاستجابة للفقرة وفقًا لحجم العينة وعدد الفقرات؟".

وللإجابة عن هذا الســـؤال تم حســـاب المتوســـطات الحســـابية والانحرافات المعيارية للبواقي باختلاف جودة مطابقة الفقرات للنموذج ثلاثي المعلم في نظرية الاســـتجابة للمفردة الاختبارية وفقًا لحجم العينة وعدد الفقرات، فكانت كما هي موضحة بجدول (6):

جدول (6) المتوسطات الحسابية والانحرافات المعيارية للبواقي باختلاف مطابقة الفقرات للنموذج ثلاثي المعلم وفقًا لحجم العينة وعدد الفقرات

نسبة الفقرات غير المطابقة	عدد الفقرات	حجم العينة											
		200		400		600		800		1000		الكلي	
		م	ع	م	ع	م	ع	م	ع	م	ع	م	ع
صفر %	20	0.012	0.005	0.011	0.004	0.008	0.003	0.007	0.002	0.012	0.005	0.011	0.004
	40	0.029	0.008	0.024	0.006	0.018	0.004	0.015	0.003	0.029	0.008	0.024	0.006
	60	0.046	0.009	0.033	0.007	0.026	0.005	0.021	0.004	0.046	0.009	0.033	0.007
	80	0.064	0.011	0.045	0.008	0.034	0.007	0.028	0.005	0.064	0.011	0.045	0.008
	100	0.075	0.013	0.055	0.009	0.042	0.007	0.033	0.007	0.075	0.013	0.055	0.009
%5	20	0.056	0.022	0.041	0.015	0.031	0.012	0.025	0.010	0.056	0.022	0.041	0.015
	40	0.014	0.005	0.010	0.004	0.008	0.003	0.007	0.003	0.014	0.005	0.010	0.004
	60	0.029	0.008	0.022	0.006	0.018	0.004	0.014	0.004	0.029	0.008	0.022	0.006
	80	0.045	0.011	0.033	0.008	0.025	0.006	0.020	0.005	0.045	0.011	0.033	0.008

تابع جدول (6) المتوسطات الحسابية والانحرافات المعيارية للبواقي باختلاف مطابقة الفقرات للنموذج ثلاثي المعلم وفقًا لحجم العينة وعدد الفقرات

نسبة الفقرات غير المطابقة	عدد الفقرات	حجم العينة											
		200		400		600		800		1000		الكلي	
		م	ع	م	ع	م	ع	م	ع	م	ع	م	ع
	100	0.061	0.014	0.043	0.010	0.033	0.008	0.027	0.006	0.061	0.014	0.043	0.010
%10	20	0.075	0.012	0.054	0.010	0.041	0.007	0.034	0.006	0.075	0.012	0.054	0.010
	40	0.055	0.022	0.040	0.016	0.030	0.012	0.025	0.010	0.055	0.022	0.040	0.016
	60	0.013	0.006	0.010	0.004	0.007	0.003	0.006	0.003	0.013	0.006	0.010	0.004
	80	0.028	0.009	0.021	0.007	0.170	0.005	0.013	0.004	0.028	0.009	0.021	0.007
	100	0.045	0.102	0.033	0.009	0.025	0.006	0.020	0.006	0.045	0.102	0.033	0.009
%20	20	0.059	0.016	0.042	0.011	0.032	0.008	0.026	0.006	0.059	0.016	0.042	0.011
	40	0.070	0.018	0.052	0.013	0.040	0.010	0.325	0.008	0.070	0.018	0.052	0.013
	60	0.052	0.023	0.038	0.017	0.029	0.012	0.024	0.010	0.052	0.023	0.038	0.017
	80	0.012	0.006	0.008	0.005	0.007	0.004	0.006	0.003	0.012	0.006	0.008	0.005
	100	0.030	0.011	0.020	0.007	0.016	0.006	0.013	0.004	0.030	0.011	0.020	0.007

يلاحظ من نتائج جدول (6) وجود فروق ظاهرية بين المتوسطات الحسابية للبواقي، وفقًا لمتغيرات حجم العينة وعدد الفقرات ونسبة الفقرات غير المطابقة، ولمعرفة الدلالة الإحصائية لتلك الفروق الظاهرية تبعًا لهذه المتغيرات والتفاعلات المشتركة بينهما، فقد تم استخدام تحليل التباين الثلاثي فكانت النتائج كما هي موضحة بجدول (7):

جدول (7) دلالة الفروق في البواقي باختلاف مطابقة الفقرات للنموذج ثلاثي المعلم وفقًا لحجم العينة وعدد الفقرات

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	ف	مستوى الدلالة
نسبة الفقرات غير المطابقة	1.528	4	0.382	2101.833	0.01
عدد الفقرات	0.035	3	0.012	63.545	0.01
حجم العينة	1.139	4	0.285	1566.992	0.01
نسبة الفقرات غير المطابقة*عدد الفقرات	0.083	12	0.007	38.140	0.01
نسبة الفقرات غير المطابقة * حجم العينة	0.248	16	0.016	85.309	0.01
عدد الفقرات*حجم العينة	0.159	12	0.013	72.929	0.01
نسبة الفقرات غير المطابقة*عدد الفقرات*حجم العينة	0.607	48	0.013	69.548	0.01
الخطأ	1.072	5900	0.000		
الكلي	19.202	6000			
الكلي المصحح	**6.451**	**5999**			

ومن جدول (7) يتضح أنه:

1- توجد فروق ذات دلالة إحصائية عند مستوى (0.01) بين متوسطات قيم البواقي ترجع لاختلاف نسبة الفقرات غير المطابقة، ولاختلاف عدد الفقرات، ولاختلاف حجم العينة.

2- توجد فروق ذات دلالة إحصائية عند مستوى (0.01) في متوسطات قيم البواقي ترجع لتأثير التفاعل بين نسبة الفقرات غير المطابقة وعدد الفقرات، ولتأثير التفاعل بين عدد الفقرات وحجم العينة، ولتأثير التفاعل بين عدد الفقرات وحجم العينة ونسبة الفقرات غير المطابقة.

3- لا توجد فروق ذات دلالة إحصائية في متوسطات قيم البواقي ترجع لتأثير التفاعل بين حجم العينة ونسبة الفقرات غير المطابقة.

وللكشف عن تأثير نسبة الفقرات غير المطابقة وحجم العينة وعدد الفقرات على البواقي تم الاعتماد على اختبار مدى توكي للفرق الدال الصادق (Tukey HSD) كاختبار للمقارنات البعدية، فكانت النتائج كما هي موضحة في التالي:

1-تأثير نسبة الفقرات غير المطابقة:

جدول (8) المقارنات البعدية للفروق في البواقي باختلاف نسبة الفقرات غير المطابقة

نسبة الفقرات غير المطابقة	0% (م=0.050)	5% (م= 0.049)	10% (م= 0.039)
5% (م= 0.049)	0.001		
10% (م= 0.039)	*0.011	*0.010	
20% (م= 0.045)	*0.005	*0.004	*0.006

*الفرق بين المتوسطين دال عند مستوى (0.05)، ** الفرق بين المتوسطين دال عند مستوى (0.01).

يتضح من جدول (8) أنه توجد فروق دالة إحصائيًا عند مستوى دلالة 0.05 بين متوسطي البواقي لكل نموذجين لصالح النموذج الأقل في نسبة الفقرات غير المطابقة، والنتائج هنا تؤكد في مجملها أن البواقي تقل بزيادة نسبة الفقرات غير المطابقة.

2-عدد الفقرات:

جدول (9) المقارنات البعدية للفروق بين البواقي باختلاف عدد الفقرات

عدد الفقرات	20 (م=0.012)	40 (م= 0.028)	60 (م= 0.041)	80 (م=0.053)
40 (م= 0.028)	*0.016			
60 (م= 0.041)	*0.029	*0.013		
80 (م= 0.053)	*0.041	*0.025	*0.012	
100 (م= 0.058)	*0.046	*0.030	*0.017	*0.005

*الفرق بين المتوسطين دال عند مستوى (0.05)، ** الفرق بين المتوسطين دال عند مستوى (0.01).

يتضح من جدول (9) أنه توجد فروق دالة إحصائيًا عند مستوى 0.05 بين متوسطي البواقي بين كل نموذجين لصالح النموذج الأعلى في عدد الفقرات، والنتائج هنا تؤكد في مجملها أن البواقي تزداد بزيادة عدد الفقرات، والذي ربما يرجع لزيادة تعقيد النموذج، فبزيادة عدد الفقرات يزداد تعقيد النموذج والذي يترتب عليه زيادة معالم النموذج التي يتم تقديرها، وهو ما يؤثر على التباين المشترك بين الفقرات والذي ينعكس على البواقي فيؤدي إلى زيادتها نوعًا ما، فكلما كان عدد المتغيرات المتشبعة على المتغير الكامن أقل كلما زاد التباين المشترك المقدر بينها خاصة وأن زيادة عدد الفقرات في الدراسة الحالية يترتب عليه زيادة عدد الفقرات غير المطابقة للنموذج ثلاثي المعلم.

3-حجم العينة:

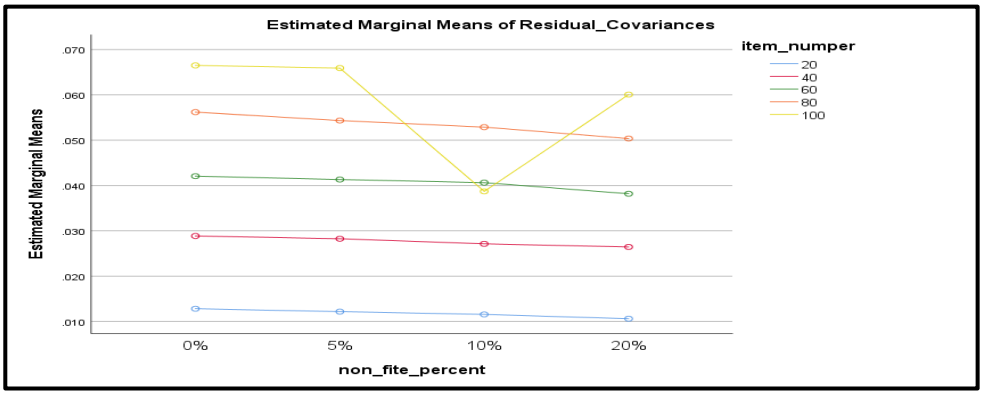
جدول (10) المقارنات البعدية للفروق في البواقي باختلاف حجم العينة

حجم العينة	200(م=0.084)	400 (م=0.053)	600 (م=0.039)	800 (م=0.030)
400 (م= 0.053)	*0.031			
600 (م=0.039)	*0.045	*0.014		
800 (م=0.030)	*0.054	*0.023	*0.009	
1000 (م= 0.024)	*0.060	*0.029	*0.015	*0.006

يتضح من جدول (10) أنه توجد فروق دالة إحصائيًا عند مستوى دلالة 0.05 بين متوسطي البواقي في كل نموذجين لصالح النموذج الأعلى في حجم العينة؛ والنتائج هنا في مجملها تؤكد أن البواقي تقل بزيادة حجم العينة، فبزيادة حجم العينة تصبح العلاقات بين المتغيرات أكثر استقرار، وهو ما ينعكس إيجابيًا على التباين المشترك وبالتالي تقل البواقي والتي تعبر عن التباين غير المفسر في الفقرة؛ **ولدراسة تأثير التفاعلات بين المتغيرات المستقلة على البواقي تم الاعتماد على الرسم البياني لهذه التفاعلات وهو ما يتضح في التالي:**

1-تأثير التفاعل الثنائي بين نسبة الفقرات غير المطابقة وعدد الفقرات

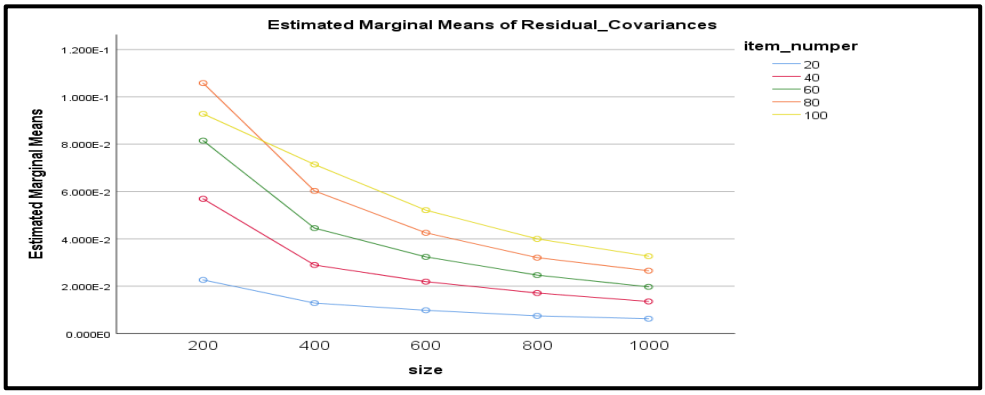
شكل (3)تأثير التفاعل الثنائي بين عدد الفقرات ونسبة الفقرات غير المطابقة على البواقي



يلاحظ من الشكل (3) أن قيم البواقي تزيد بزيادة عدد الفقرات في النموذج وتقل نوعًا ما بزيادة نسبة الفقرات غير المطابقة، وهذه النتيجة تفسيرها المنطقي قد يكمن في زيادة تعقيد النموذج، فكلما زاد تعقيد النموذج يزداد عدد الفقرات المتشبعة على المتغير الكامن وبالتالي تزداد البواقي والتي تعبر عن التباينات غير المفسرة في الفقرات، وهذا يتأكد من أن البواقي في حالة بساطة النموذج بتضمنه لأقل عدد من الفقرات (20) لم تتغير كثيرًا في حالة زيادة نسبة الفقرات غير المطابقة، بينما زيادة عدد الفقرات بصفة عامة ينتج عنه زيادة في البواقي وهو ما يتفق مع ما يشير إليه (عطا، 2014؛ Sahin & Anil, 2017).

2-ثانيًا: تأثير التفاعل الثنائي بين عدد الفقرات وحجم العينة

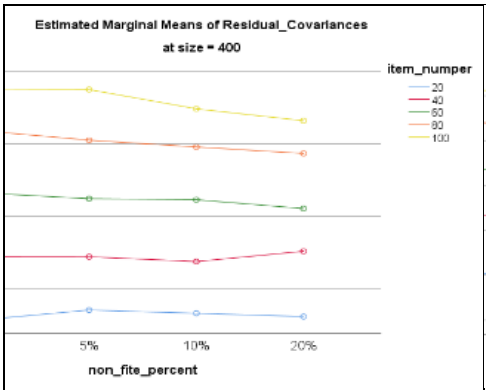
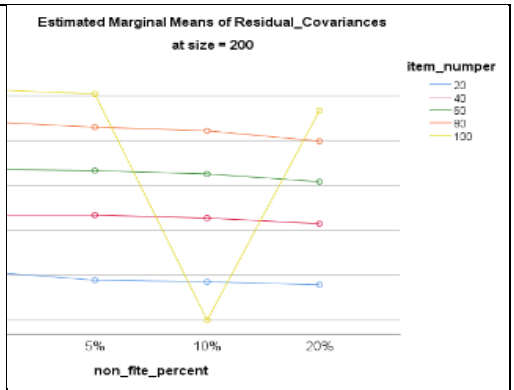
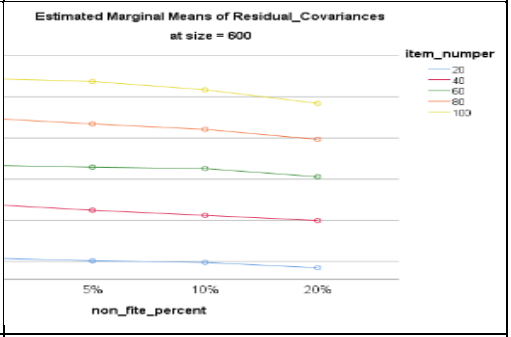
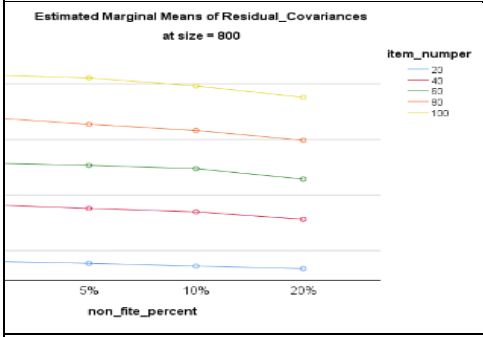
شكل (4) تأثير التفاعل الثنائي بين حجم العينة وعدد الفقرات على البواقي

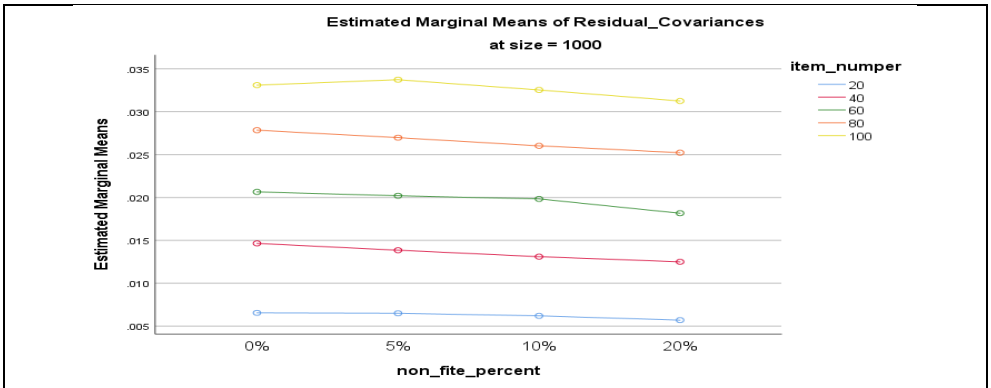


يلاحظ من الشـــكل (4) أن زيادة عدد الفقرات وحجم العينة يؤدي إلى نقص في قيم البواقي في النموذج البنائي، وكلما زاد حجم العينة تقل البواقي حتى في عدد الفقرات الكبير (100)، وهو ما يؤكد أن زيادة عدد الفقرات يتطلب حجم عينة كبير حتى يمكن تقدير التباين المشـــترك بدقة، بينما كلما زاد عدد الفقرات وقل حجم العينة فستكون البواقي عند أعلى قيمة لها، ومن الواضح أن نقص حجم العينة (200) وزيادة عدد الفقرات يؤدي إلى عدم استقرار في التقديرات التي يمكن الوصول إليها في النموذج البنائي وخاصة البواقي، حيث نلاحظ أن هناك زيادة في البواقي في حالة عدد الفقرات (80) عنها في حالة عدد الفقرات (100)، والنتيجة العامة هنا لتأثير التفاعل بين حجم العينة وعدد الفقرات يؤكد أن زيادة عدد الفقرات يتطلب حجوم كبيرة للعينات حتى نضمن دقة تقديرات معالم النموذج، والعكس في حالة عدد الفقرات الأقل (20) فقرة فممكن الاعتماد على عينات أقل في الحجم حيث كانت أقل قيمة للبواقي عد حجم عينة (200) وعدد الفقرات (20)، والذي يمكن تفسيره أيضًا وفقًا لفكرة تعقيد النموذج وبساطته، وهو ما يتفق بدرجة كبيرة مع ما يشير إليه (نصراوين، 2018 وعطا، 2014).

رابعًا: تأثير التفاعل الثلاثي بين (حجم العينة وعدد الفقرات ونسبة عدم المطابقة):

شكل (5) تأثير التفاعل الثنائي بين عدد الفقرات ونسبة الفقرات غير المطابقة عند أحجام العينة المختلفة على البواقي

<table>
<tr>
<td> 
</td>
<td> 
</td>
</tr>
<tr>
<td>أ- عند حجم عينة (200)</td>
<td>أ- عند حجم عينة (400)</td>
</tr>
<tr>
<td> 
</td>
<td> 
</td>
</tr>
<tr>
<td>ح- عند حجم عينة (600)</td>
<td>د- عند حجم عينة (800)</td>
</tr>
</table>



د- عند حجم عينة (1000)

يلاحظ من شـــكل (5) الاتفاق الكبير في تأثير حجم العينة وعدد الفقرات ونســـبة الفقرات غير المطابقة، حيث تزداد قيم البواقي بزيادة عدد الفقرات وتقل بزيادة نســـبة الفقرات غير المطابقة للنموذج الثلاثي في نظرية الاستجابة للفقرة، وبزيادة حجم العينة يصبح تأثير عدد الفقرات ونسبة الفقرات غير المطابقة أكثر استقرارًا، فبزيادة عدد الفقرات يصـبح النموذج أكثر تعقيدًا، وتزداد قيم البواقي، وتأثير نسـبة الفقرات غير المطابقة على البواقي في الدراســة الحالية نتيجة غير متوقعة، حيث أن بزيادة عدد الفقرات غير المطابقة يقل التجانس بين الفقرات وبالتالي يقل التباين المشـترك بينها وكان المتوقع أن ذلك يسـهم في زيادة البواقي لكن النتيجة هنا تؤكد على عكس ذلك، حيث تقل البواقي بزيادة نسبة الفقرات غير المطابقة للنموذج ثلاثي المعلم في نظرية الاستجابة للمفردة.

بناًء على النتائج التي تم الوصـــول إليها في الدراســـة الحالية يمكن القول بأن هناك توافق في مطابقة الفقرات للنموذج التوكيدي والنموذج الثلاثي لنظرية الاستجابة للفقرة، وبمعنى آخر النتائج تشير إلى أن هذين النموذجين قد أظهرا أداًء مماثلًا أو مشابهًا عند تطبيقهما في ضوء حجم العينة وعدد الفقرات.

توصيات ومقترحات الدراسة:

بناء على النتائج التي تم التوصل إليها فإن يمكن تقديم التوصيات الآتية:

- على المراكز المتخصـــصـــة بالبحث وإعداد المقاييس لفت نظر الباحثين لإمكانية توجيه الدراســـة الحالية نحو دراســـات ميدانية أو تجارب على أرض الواقع لاختبار النتائج والمفاهيم التي تم تطويرها في بيئة حقيقية، فقد يساهم في توثيق تطبيقات عملية لنتائج الدراسة الحالية.
- على الباحثين دراسة كيف يمكن تحسين تصميم الاختبارات والفقرات بالاعتماد على نظرية الاستجابة للمفردة الاختبارية والتحليل العاملي التوكيدي في محالة الاستفادة من ميزات الاتجاهان.
- من أهم نتائج الدراســة الحالية التأكيد على التأثير الســلبي لنســبة عدد الفقرات غير المطابقة للنموذج ثلاثي المعلم في نظرية الاســـتجابة للمفردة، وأن هذا التأثير يقل بزيادة عدد الفقرات وحجم العينة، ويجب الأخذ بعين الاعتبار هذه النتيجة في البحوث المستقبلية التي تهدف لتقنين أدوات القياس المختلفة في المجال النفسي والتربوي.

- تكرار الدراسة الحالية أو جزء منها بالاعتماد على بيانات حقيقية لاختبار صدق النتائج التي تم التوصل إليها في الدراسة الحالية.
- ضرورة توجيه الباحثين لدراسة تأثير عوامل إضافية على مطابقة النموذج التوكيدي ونماذج نظرية الاستجابة للفقرة، كعدد بارامترات النموذج وطرق التقدير، والاعتدالية.

المراجع:

أبو فودة، باسل. (2013). استخدام النموذج اللوجستي ثلاثي المعلمة في تحليل فقرات اختبار مقرر التقييم والتقويم لدى عينة من طالبات جامعة الأميرة نورة بنت عبد الرحمن. *المجلة العلمية لجامعة الملك فيصل (العلوم الإنسانية والإدارية)*، 13(2)، 287-327.

بلعربي، عادل عبد الرحمان. (2018). البناء العاملي الاستكشافي والتوكيدي لاستبانة فرق العمل-دراسة سيكومترية. *مجلة العلوم الاجتماعية*، 7(29)، 32-42.

بلعيد، سارة وقماري، محمد. (2018). الكشف عن البنية العاملية لنموذج الميول لجون هولاند (استخدام التحليل العاملي التوكيدي لنموذج هرمي من الدرجة الثانية والثالثة). *الرواق*، 4(1)، 85-105.

بوصلب، عبد الحكيم وصحراوي، عبد الله. (2016). النمذجة البنائية (SEM) ومعالجة صدق المقاييس في البحوث النفسية والتربوية نموذج البناء العاملي لعلاقات كفاءات التسيير الإداري بالمؤسسة التعليمية. *مجلة العلوم النفسية والتربوية*، 3(2) 61-91.

تيغزة، أمحمد. (2012). *التحليل العاملي الاستكشافي والتوكيدي مفاهيمها ومنهجيتها بتوظيف حزمة SPSS وليزرول LISREL*. الأردن، عمان: دار المسيرة للنشر والتوزيع.

الثوابية، أحمد. (2010). أثر حجم العينة على تقدير صعوبة الفقرة والخطأ المعياري في تقديرها باستخدام نظرية الاستجابة للفقرة. *مجلة جامعة دمشق*، 26 (1+2)، 225- 256.

جامعة العلوم الإسلامية الماليزية، وكلية القيادة والإدارة. (2017). اختبار صلاحية النموذج المقترح لقياس فاعلية الإدارة باستخدام التحليل العاملي التوكيدي دراسة تطبيقية على موظفي وزارة العمل والتأهيل بليبيا. *International Journal of Management and Human Science*. 1(1). 31-43.

الحاج محمود، فريال محمود محمد. (2018). الأداء التفاضلي لفقرات وموهات اختبار محكي المرجع في العلوم للصف السادس الابتدائي مصمم وفق نظرية استجابة الفقرة. *مجلة جامعة القدس المفتوحة للأبحاث والدراسات*، 2(44)، 123-135.

حسب الله، عبد العزيز محمد. (2023). تأثير بعض المتغيرات المرتبطة بعينة البحث على تقديرات الخصائص السيكومترية للمقياس النفسي ومفرداته "مقياس تقدير الذات لروزنبرج نموذجًا". *مجلة كلية التربية جامعة بنها*، 133(3)، 482-572.

الحوري، أروى. (2013). *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، إربد، الأردن.

الخالدي، هاني. (2018). البنية العاملية لمقياس قلق المستقبل لدى طلبة الجامعة: دراسة استكشافية توكيدية مقارنة. *المجلة العربية للدراسات الأمنية*، 34(3)، 372-385.

رشوان، ربيع عبده أحمد. (2021). الأداء التفاضلي للبنود وتأثيره على الأداء التفاضلي للاختبار باستخدام نماذج نظرية الاستجابة للمفردة والتحليل العاملي التوكيدي متعدد المجموعات. *مجلة العلوم التربوية والدراسات الإنسانية*، 6(15)، 44-93.

السمكي، لطفي والسوالمة، يوسف محمد. (2017). اللاتغير في القياس لاختبار تمس (TIMSS) في العلوم للصف الثامن عبر موقع المدرسة وجنس الطالب في الأردن. *المنارة للبحوث والدراسات*، 23(4)، 437-460.

الشريفين، نضال. (2003). بناء مقياس اتجاهات معلمي العلوم نحو العمل المخبري. *المجلة الاردنية في العلوم التربوية*، 3(2)،169-187.

الشواورة، شادي. (2017). المقارنة بين بعض مؤشرات المطابقة المتزايدة والمطابقة المطلقة وَفْق نظرية استجابة الفقرة باختلاف النموذج اللوغاريتمي وعدد الفقرات. *العلوم التربوية: مجلة علمية محكمة ربع سنوية*، *25*(2-الجزء 1)، 99-117.

الصرايرة، راجي. (2008). *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.

الصيخان، رحاب. (2019). *مقارنة بين النظرية الكلاسيكية والنموذج ثلاثي المعلم في انتقاء فقرات اختبار تحصيلي للغة الإنجليزية*. رسالة ماجستير، كلية التربية، جامعة القصيم، المملكة العربية السعودية.

ضعضع، هبه وطومان، منار وطيفور، مصطفى. (2020). أثر حجم العينة وطرائق التقدير في دقة تقدير معالم نموذج راش. *مجلة جرش للبحوث والدراسات*، 21 (1)، 131-170.

عامر، عبد الناصر. (2018). *نمذجة المعادلة البنائية للعلوم النفسية والاجتماعي الأسس والتطبيقات والقضايا*. الرياض: دار جامعة نايف للنشر.

عطا، زايد. (2014). تقصي دقة تقدير النموذج اللوجستي ثلاثي المعلمة لمعالم الفقرة وقدرة الأفراد في ضوء تغير طول الاختبار وحجم العينة. *مجلة جامعة الشارقة للعلوم الإنسانية والاجتماعية*، 11(7)، 1-37

علام، صـــلاح الدين. (2005). *نماذج الاســـتجابة للفقرة الاختبارية أحادية البعد ومتعددة الابعاد وتطبيقاتها في القياس النفسي والتربوي*. القاهرة: دار الفكر العربي.

العمري، حســـان. (2011). أثر ترتيب الفقرات في المقاييس النفســـية على التجانس الداخلي للمقياس وخصائص فقراته. *مجلة جامعة دمشق*، 27(3، 4)، 111-129.

القضـــاة، محمد. (2020). *أثر حجم العينة وطول الاختبار في دراســـة العلاقة بين مؤشـــر مطابقة المفردة ومعالمها للنموذج ثلاثي المعلمة*. رسالة دكتوراه غير منشورة، كلية التربية، جامعة المنصورة، مصر.

محمد، محمد إبراهيم. (2020). مؤشرات المطابقة: دراسة مقارنة بين نماذج التحليل العاملي التوكيدي من الدرجة الأولى والهرمية والثنائية في مقاييس التقرير الذاتي. *مجلة البحث في التربية وعلم النفس*، 35(4)، 413-480

المرابحة، عامر جبريل وبن صالح، عفيف. (2021). تقييم مدى مطابقة فقرات اختبار علم النفس للنظرية الكلاســـيكية في القياس ونظرية الاســـتجابة للفقرة وفقا للنموذج اللوجســـتي ثنائي المعلم (دراســـة مقارنة). *مجلة العلوم النفسية والتربوية*، 7(2)، 14-34.

المهدي، ياسـر فتحي الهنداوي. (2007). منهجية النمذجة بالمعادلة البنائية وتطبيقاتها في بحوث الإدارة. *مجلة التربية والتنمية*، 15(40)، 9-41.

المومني، رنا. (2016). القدرة التنبؤية لمعالم الفقرة بمؤشـــر مطابقتها باختلاف عدد فترات القدرة تبعا لنظرية استجابة الفقرة (IRT). *المجلة الأردنية في العلوم التربوية*، 12(1)، 71-84.

نصراوين، معين سلمان سليم. (2018). دقة تقدير معالم الفقرات عند استخدام اربعة نماذج لوجستية في إطار نظرية الاستجابة للفقرة. *دراسات "العلوم التربوية"* ، 45 (4)، 179-205.

Abedalaziz, A. & Leng, C. H. (2013). The Relationship between CTT and IRT Approaches in Analyzing Item Characteristics. *The Malaysian Online Journal of Educational Science*, 1(1), 64-70.

Al-Qahtani, M. (2022). The Effect of Test Length on the Accuracy of Estimating Ability Parameter in the Two- and Three-Parameter Logistic Models: Comparison by Using the Bayesian Method of Expected Prior Mode and Maximum Likelihood Estimation. *Journal of Educational and Social Research*, 12(1), 168-187.

Bajwa, D.; Koeszegi, S. & Vetschera, R. (2016). *Group Decision and Negotiation: Theory, Empirical Evidence, and Application*. USA: Springer.

Baker, F. (2001). *The Basics of Item Response Theory*. 2nd Ed., USA: ERIC Clearinghouse on Assessment & Evaluation.

Bean. G. J. & Bowen. N. K. (2021). Item Response Theory and Confirmatory Factor Analysis: Complementary Approaches for Scale Development. *Journal of Evidence-Based Social Work*, 18(6), 597–618 https://doi.org/10.1080/26408066.2021.1906813

Brown, G., Harris, L., O'Quin, C., & Lane, K. (2017) Using multi-group confirmatory factor analysis to evaluate cross-cultural research: identifying and understanding non-invariance. *International Journal of Research & Method in Education*, 40(1), 66-90

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. 2nd ed., New York: Routledge.

Cao. Y, Lu. R & Tao. W. (2014). Effect of Item Response Theory (IRT) Model Selection on Testlet-Based Test Equating. *Research Reports*, 14-19. doi:10.1002/ets2.12017.

Cheung, G. & Rensvold, R. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling*, 9(2), 233–255

Cook, K. & Kallen, M. & Amtmann, D. (2009). Having a Fit: Impact of Number of Items and Distribution of Data on Traditional Criteria for Assessing IRT's Unidimensionality Assumption. *Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation*, 18, 447-460.

Crisci, A. (2012). Estimation Methods for the Structural Equation Models: Maximum Likelihood, Partial Least Squares and Generalized Maximum Entropy. *Journal of Applied Quantitative Methods Inquires*, 7(2), 3-17.

D'Agostino. R. B. & Stephens. M. A. (1986). *Goodness-of-Fit Techniques*. New York: Marcel Dekker, INC

De Ayala, R. J. (2009). *The Theory and practice of item response theory*. New York: The Guilford Press

DeCoster, J. (1998). Overview of Factor Analysis. Researc-gate. 2014, *https://www.researchgate.net/publication/255620387*

Ellis. J. (2017). Factor analysis and item analysis. Applying Statistics in Behavioral. *https://www.applyingstatisticsinbehaviouralresearch.com/documenten/factor_analysis_and_item_analysis_version_11_.pdf*

Fan. Y., Chen. J.; Shirkey. G.; John. R.; Wu, R. S; Park, H. & Shao, C. (2016). Applications of structural equation modeling (SEM) in ecological studies: an updated review. *Ecological Processes*, 5, 19. DOI 10.1186/s13717-016-0063-3

Goodwin, L. D., & Leech, N. L. (2006). Understanding correlation: factors that affect the size of r. *The Journal of Experimental Education*, 74(3), 249–266.

Haberman, S. J.; Sinharay, S. & Chon, K. H. (2013). Assessing Item Fit for Unidimensional Item Response Theory Models Using Residuals from Estimated Item Response Functions. *Psychometrika*, 78(3), 417-440. DOI: 10.1007/s11336-012-9305-1

Hambleton, R. & Swaminathan, H. (1985). *Item Response Theory: Principles and Applications*. Boston, MA: Kluwer Academic Publishers.

Harris, D. (1989). Comparison of 1-, 2-, and 3-parameter IRT models. *Instructional Topics in Educational Measurement*, 8(1), 157-163.

Hijazi, T. & Alkateeb, A. (2014). The Compatibility between the Classical Test Theory and the Two Parameter Logistic Models in the Item Stratification of Criterion Referenced Test in Provisions of Recitation and Tajweed. *AnNajah University Journal for Research, Humanities*, 28(10), 2240-2270.

Hoyle, H (1995). *Structural Equation Modeling: Concepts, Issues and Applications*. New York: Sage Publications.

Hu, Z., Lin, L., Wang, Y. & Li, J. (2021). The Integration of Classical Testing Theory and Item Response Theory. *Psychology*, 12, 1397-1409. doi: 10.4236/psych.2021.129088.

Ismail, T. H.; Ameen, A., & Habtoor, N. (2018). The Confirmatory Factor Analysis (CFA) of Motivation ،Job Satisfaction and Employee Performance. *International Journal of Management and Human Science*, 2 (4)73-83.

Kang, T. & Troy, T. (2008). Performance of the Generalized S-X2 Item Fit Index for Polytomous IRT Models. *Journal of Educational Measurement*, 45 (4), 391-406.

Lee, S. Choi, Y. & Cohen, A. S. (2018). Automating Simulation Research for Item Response Theory using R. *International Journal of Assessment Tools in Education*, 5(4), 682–700.

Lin, H, Y. & Tsai, T. L. (2021). Development of Mathematics Reading Assessment: Psychometric Evaluation Based on SEM and IRT. *International Journal of Education, Psychology and Counselling*, 6(38), 46-56. DOI: 10.35631/IJEPC.638005

Liu, Y. & Maydeu-Olivares, A. (2014). Identifying the source of misfit in item response theory models. *Multivariate Behavioral Research*, 49, 354-371.

Luecht, R. & Ackerman, T. R. (2018). A Technical Note on IRT Simulation Studies: Dealing with Truth. *Educational Measurement: Issues and Practice. Estimates, Observed Data, and Residuals*, 37(3), 65-76.

Magno. C. (2009). Demonstrating the Difference between Classical Test Theory and Item Response Theory Using Derived Test Data. *The International Journal of Educational and Psychological Assessment*, 1(1), 1-11.

Maldonado, C. & Loreto, F. (2021). *Statistical Methods for the Analysis of Polytomous Response Data in non-Cognitive Tests*. [Ph.D. in Statistics], Pontificia Universidad Católica de Chile.

Maydeu-Olivares, A. (2015). Evaluating fit in IRT models. In Steve P. Reise & Dennis A. Revicki (Eds.). *Handbook of item response theory modeling: Application to typical performance assessment*, (pp. 111-127), New York: Routledge.

McDonald, R. (1982). Linear Versus Models in Item Response Theory. *Applied Psychological Measurement*, 6(4), 379-396.

McGill, M. (2009). *An Investigation of Unidimensional Testing Procedures under Latent Trait Theory using Principal Component Analysis*. [Ph. D. dissertation], faculty of the Virginia Polytechnic Institute and State University.

Morris, T., P., White, R & Crowther, J., M. (2018). Using Simulation Studies to Evaluate Statistical Methods. *Wiley Statistics in Medicine*, 38, 2074–2102.

Oribhabor. C. B. & Aituariagbon. K. E. (2020). Trend Analysis of the Item Fit of the WAEC Mathematics Multiple Choice Test Items using Three Parameter Logistic Model. *ASSEREN Journal of Education*, 5(1), 1-12.

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

Phakiti, A., De Costa, P, Plonsky, L. & Starfield, S. (2018). *The Palgrave Handbook of Applied Linguistics Research, Methodology*. London: Palgravr.

Pop, H. F., & Frentiu, M. (2008). *Applications of Principal Components Methods*. First International Conference on Complexity and Intelligence of the Artificial and Natural Complex Systems. Medical Applications of the IEEE Complex Systems. Biomedical Computing, Pp. 103-109.

Raposo, F. (2016). Evaluation of analytical calibration based on least-squares linear regression for instrumental techniques: A tutorial review. *Postprint of TrAC Trends in Analytical Chemistry*,77, 167-185. DOI: https://doi.org/10.1016/j.trac.2015.12.006

Raykov, T. & Marcoulides, G. A. (2006). *A first course in structural equation modeling*. 2nd ed. N.J: Lawrence Erlbaum Associates.

Rodrıguez, T.; Tomas, J.; Hontangas, M.; Leon, V.; Leon, A.; Palomino, B.; Bossio, R.; Pena-Calero, B. & White, M. (2021). Validation of the De Jong Gierveld Loneliness Scale in Peruvian Old adults: A Study Based on SEM and IRT Multidimensional Models. *The Journal of General Psychology*. DOI: 10.1080/00221309.2021.1913398. Online Ahead of Print.

Sahin, A. & Anıl, D. (2017). The Effects of Test Length and Sample Size on Item Parameters in Item Response Theory. *Educational Sciences: Theory & Practice*, 17(1), 321- 335. DOI:10.12738/estp.2017.1.0270

Sijtsma, K. (1998). Methodology review: Nonparametric IRT approaches to the analysis of dichotomous item scores. *Applied Psychological Measurement*, 22(1), 3-31.

Tay, L., & Drasgow, F. (2012). Adjusting the adjusted χ^2/df ratio statistic for dichotomous item response theory analyses: Does the model fit? *Educational and Psychological Measurement, 72(3), 510–528*

Van Der Linden, W. J. (2018). *Handbook of Item Response Theory, Volume Three: Applications*. 1st Edition. New York, (Chapman & Hall/CRC Statistics in the Social and Behavioral Sciences). PP 17.

Wu, A.D., Li, Z. & Zumbo, B. D. (2007). Decoding the Meaning of Factorial Invariance and Updating the Practice of Multi-Group Confirmatory Factor Analysis: A Demonstration with TIMSS Data. *Practical Assessment, Research & Evaluation*, 12, 1-26.

Young, S. R.; Keith, T. Z., & Bond, M. A. (2019). Age and Sex Invariance of the International Cognitive Ability Resource (ICAR). *Intelligence*, 77, 101399.

Yu, C. H. (2012). A Simple Guide to the Item Response Theory (IRT) and Rasch Modeling. *https://www.fisica.net/enem/A-Simple-Guide-to-the-Item-Response-Theory-(IRT)-and-Rasch-Modeling.pdf*