Journal of Al-Quds Open University for Educational & Psychological Research & Studies No – (43) .December 2023
https://journals.qou.edu/index.php/nafsia
E-ISSN: 2307-4655 / P-ISSN: 2307-4647

The Factorial Structure of Scale in Light of Differential Items Functioning Type

Mr. Firas Mahmoud Ali Al-Mahasneh*[1], Dr. Amal Ahmad Mahmoud Al-Zoubi[2]

1Educational Psychology and Counsling, Yarmouk University, Irbid, Jordan.
Orcid No: 0009-0004-1537-2138
Email: firas390@hotmail.com

2Educational Psychology and Counsling, Yarmouk University, Irbid, Jordan.
Oricd No: 0000-0002-6993-3197
Email: Amal@yu.edu.jo

Received:
24/10/2022

Revised:
1/11/2022

Accepted:
10/09/2023

***Corresponding Author:**
firas390@hotmail.com

Citation: Al Mahasneh, F. M. A., & Al-Zoubi, A. A. M. The Factorial Structure of Scale in Light of Differential Items Functioning Type. Journal of Al-Quds Open University for Educational & Psychological Research & Studies, 14(43). https://doi.org/10.33977/1182-014-043-002

2023©jrresstudy. Graduate Studies & Scientific Research/Al-Quds Open University, Palestine, all rights reserved.

- **Open Access**

This work is licensed under a Creative Commons Attribution 4.0 International License.

Abstract
The current study aimed at verifying the factorial structure of scale in light of differential item functioning (DIF) type (uniform, non-uniform), in order to achieve the study objective. The experimental approach was followed by simulating the study conditions through generating the individual's responses of the sample study which consisted of 1000 individual for each of the focal group and the reference group on a scale of 50items in light of three parameters logistic model (3PLM) using R package. Also the data were replicated 300 times for each type of differential functioning (DIF). The results of the study indicated that the data fit the model of the study, and the indicators values of $\frac{\chi^2}{df}$, RMSEA was showing an acceptable goodness of fit values. In addition, the study revealed that the invariance levels of configure, metric, scalar, residuals property were confirmed using the (MGCFA) through all of the experimental conditions of the study, but it was differed according to the difference in the (DIF) type; due to the absolute goodness of fit indicators $\frac{\chi^2}{df}$, RMSEA. The study recommended the use of different percentages of (DIF) , and other types of goodness fit indices.
Keywords: Factorial structure, scales differential functioning type, invariant.

البناء العاملي للمقياس في ضوء نوع الأداء التفاضلي للفقرات

أ. فراس محمود علي المحاسنة*[1]، د. آمال أحمد محمود الزعبي[2]

[1]علم النفس التربوي والإرشادي، التربية، جامعة اليرموك، إربد، الأردن.

[2]علم النفس التربوي والإرشادي، التربية، جامعة اليرموك، إربد، الأردن.

الملخص
هدفت الدراسة إلى التعرف إلى البناء العاملي للمقياس في ضوء اختلاف نوع الأداء التفاضلي (منتظم، وغير منتظم) للفقرات، ولتحقيق ذلك اتبعت الدراسة المنهج التجريبي، إذ تم محاكاة ظروف الدراسة وتوليد استجابات لأفراد عينة الدراسة الإفتراضيين والبالغ عددهم (1000) فرد لكل من المجموعتين المرجعية والمستهدفة على مقياس افتراضي عدد فقراته (50) فقرة في ضوء النموذج ثلاثي المعلمة (3PLM) باستخدام برمجية (R)، وبواقع (300) مرة لنوعي الأداء التفاضلي، وقد خلصت الدراسة إلى تحقق خاصية اللاتغاير (الشكلي، والمتري، والقياس، والبواقي) بالاستناد إلى التحليل العاملي التوكيدي متعدد المجموعات (MGCFA)، عبر كل الظروف التجريبية للدراسة، إلا أنها اختلفت باختلاف نوع الأداء التفاضلي، ولصالح مستوى الأداء التفاضلي المنتظم(الشكلي، المتري)، ولصالح غير المنتظم (القياس، والبواقي)، وذلك من خلال مؤشرات جودة المطابقة المطلقة للنموذج ($\frac{\chi^2}{df}$, $RMSEA$). وقد أوصت الدراسة باستخدام نسب محددة من الأداء التفاضلي للفقرات، واستخدام أنواع إضافية من مؤشرات جودة المطابقة.

الكلمات المفتاحية: البناء العاملي، المقاييس، نوع الأداء التفاضلي، اللاتغاير.

المقدمة

نظراً لأهمية المقاييس في تصنيف الأفراد واختيارهم، وما ينتج عنها من قراراتٍ مهمةٍ تتعلق بمستقبلهم العلمي والوظيفي، وغير ذلك؛ فقد اتسعت دائرة استخدامها ودراسة خصائصها المختلفة، والمقصود في هذه الدراسة مقاييس القدرات والأداء العادي للأفراد، وقد أدى ذلك لظهور ما يعرف بنظريتي القياس التقليدية (CTT) والحديثة (IRT)، للتحقق من خصائصها (السيكومترية) الجيدة للوصول إلى تقييمات متعددة للفقرات والمقياس ككل(Capalleri et al., 2014).

ويُعدّ الحكم على مستوى عدالة المقياس خطوةً مهمةً للتحقق من قدرات الأفراد عليه، ويجري ذلك بالكشف عن الفقرات ذات الأداء التفاضلي باستخدام الأساليب الإحصائية المناسبة؛ لمعرفة الفقرات العادلة التي تناسب تقييم المقياس لمواضيع مادة دراسية معينة بالتطبيق على مجموعات متشابهة من الأفراد ذوي القدرات المتساوية، ولا بدّ من تساوي أدائهم على متغيرات الجنس أو العرق أو اللغة، وبذلك فإن وجود الفقرات ذات الأداء التفاضلي يعد دليلاً على وجود عوامل مرتبطة تعود إلى مجموعة معينة تؤثر في احتمالية الإجابة الصحيحة على الفقرة، وهذا يهدد عدالة المقياس بين الأفراد، وبناءً على ذلك جرى تطوير عدد من الطرق الإحصائية واستخدامها، وهدفت إلى الكشف عن الفقرات ذات الأداء التفاضلي للمقياس؛ لتحقيق المساواة بين الأفراد الذين لهم مستوى القدرة نفسه على السمة المقيسة التي تمثلها فقرات المقياس، ولكن الإجراءات والطرق المستحدثة لم تعط أدلة متسقة ومترابطة مما يقلل من تأثير الفقرات ذات الأداء التفاضلي على تطوير المقاييس، وهذا يتطلب مراجعة الدراسات المتعلقة بالأداء التفاضلي لما يترتب عليها من آثار تتعلق بعدالة نتائج المقاييس (Huang & Han, 2012; Tan & Jeral, 2005).

ويُعد الأداء التفاضلي للفقرات من أهم الأدلة التي يتم جمعها لدعم صدق بناء المقياس المستخدم؛ للتأكد من أن الفروق في استجابات العينة تُعزى إلى اختلافات حقيقية في السمة المقيسة، وليس بسبب انتماء الأفراد إلى مجموعات مختلفة حسب الجنس، أو العمر، أو العرق، أو اللغة وغيرها، ويمكن المقارنة بينها بعدالة، ويعتبر الأداء التفاضلي للفقرة من الأخطاء المنتظمة التي تهدد صدق درجات المقياس، وللكشف عنه هناك طرق متعددة منها ما يعتمد على النظرية الكلاسيكية للقياس مثل طريقة الإنحدار (اللوجستي)، حيث تستخدم الدرجة الكلية كمحكٍ للمطابقة، وأخرى تعتمد على نظرية استجابة الفقرة مثل طريقة نسبة (الأرجحية) التي تعتمد على السمة الكامنة كمحكٍ للمطابقة، وأحدث هذه الطرق تلك التي تعتمد على النمذجة بالمعادلة البنائية، وأن الفقرة التي تمثل الأداء التفاضلي لا يعني بالضرورة أنها متحيزة، وهنا يتم الرجوع إلى الأسس النظرية والأسباب المحتملة (عبد العزيز وكريش، 2016).

ويُعدّ البناء العاملي للمقياس مكوناً بالغ الأهمية؛ لمعرفة تجمع الفقرات التي تمثل السمة المقيسة، وتسمى بالعامل العام للمقياس، وما يتبعها من سماتٍ فرعيةٍ أو عوامل يتشبع عليها مجموعة من الفقرات أو المتغيرات ذات الصلة به؛ ولذلك يتم البحث في بناء المقياس، والتعرف إلى الخصائص (السيكومترية) لفقراته، وعندما يتم تكييف المقاييس إلى بيئاتٍ أخرى يجري التحقق من البناء العاملي لها، وذلك بمقارنة العوامل ذاتها للمقياس الأصلي مع النسخة المكيفة، وهنا يستخدم التحليل العاملي التوكيدي (CFA)؛ بهدف التأكد من تطابق البناء العاملي للمقياس، ويتم ذلك عند الحصول على العوامل نفسها، وتسمى العوامل تبعا للفقرات التي تمثلها، ويجب أن تكون ذات مغزى ومرتبطة بالمقياس الذي تم تحليله؛ وبالتالي يكون لها معنى ودلالة نظرية متصلة بالإطار النظري للدراسة (تيغزة، 2012).

ويُنظر للبناء العاملي للمقياس كشكل من أشكال الصدق البنائي الذي يتوصل إليه عن طريق التحليل العاملي، ويكثر الاهتمام بالصدق البنائي؛ لأن الخصائص النفسية التي يتم قياسها ما هي إلا بنى كامنة، ويعد هذا النوع من الصدق أحد أهم الأنواع التي يجري التحقق منها في المقاييس، فهو مفهوم شامل يتضمن أنواع الصدق جميعها، ويعد شرطا أساسيا للمقاييس بأنواعها التربوية والنفسية (عدس والكيلاني، 1993).

ويهدف التحليل العاملي إلى معرفة عدد العوامل الكامنة المشتركة التي تشكل أنماط الارتباطات بين أزواج المتغيرات المكونة للمقياس، والعلاقة بين العوامل والمتغيرات المشاهدة التي تمثل قيم التشبعات لهذه المتغيرات على تلك العوامل، ويستدل من التحليل العاملي على طبيعة العوامل الكامنة الممثلة للارتباطات الداخلية للمقياس، حيث تتشبع متغيرات المقياس على أكثر من عامل وبدرجات متفاوتة، وهنا تبرز مشكلة التفسير والتنقية لهذه العوامل؛ ويكمن الحل في إجراء عملية التدوير التي تقلل عدد العوامل إلى أقل ما يمكن، وأخيرا تستخرج نسبة التباين المفسّر من العوامل الكامنة لكل متغير مشاهد وتسمى عادة بالشيوع، ومتممتها تمثل نسبة التباين الفريد لهذا المتغير، وهي تمثل الحد الأدنى لثبات المتغير، ويمكن تلخيص ما سبق من خلال تحقيق

ثلاثة أهداف، هي: وصف العوامل الكامنة، والبرهنة على صحتها، واقتراح عوامل جديدة من البيانات الأصلية (Crocker & Algina, 1986; Eysenck, 2013).

ويُنظر للتحليل العاملي التوكيدي كأُسلوب من أساليب نمذجة المعادلة البنائية لاختبار صحة الفرضيات النظرية، والتحقق من البناء العاملي للمقاييس وتطويرها، ومطابقة البناء العاملي مع البيانات التي جرى جمعها، وأنه من أكثر الأساليب الإحصائية استخداما في الدراسات التطبيقية، حيث يُستخدم في التحقق من الخصائص (السيكومترية) لأدوات القياس، والصدق البنائي، ودراسة العلاقات والتأثيرات بين المتغيرات الكامنة والملاحظة، وتقييم ثبات التكافؤ العاملي، القدرة على التعامل مع مجموعات الدراسة في آنٍ واحدٍ، واستخدام المؤشرات الإحصائية في التحقق من مطابقة النماذج المقترحة (Brown, 2006; Frick et al., 2020).

ويُعد التحليل العاملي التوكيدي إحدى حالات النمذجة البنائية التي تتكون من نوعين: أحدهما نموذج القياس الذي يربط بين مجموعة الفقرات (المتغيرات المُقاسة) ومجموعة أصغر تُمثل العوامل (المتغيرات الكامنة)، والنوع الآخر هو نموذج البناء الذي يربط العوامل الكامنة بسلسلة من العلاقات المتكررة وغير المتكررة، ويُستخدم التحليل العاملي في تقييم نماذج القياس عبر مجموعات مختلفة (Brown, 2006).

وتهدف النمذجة البنائية إلى معرفة درجة مطابقة النموذج النظري لبيانات الدراسة الواقعية، وليس الوصول لأفضل نموذج يطابق البيانات؛ بمعنى التحقق من دعم بيانات العينة للنموذج النظري المقترح، وفي حال عدم تأييد البيانات للنموذج؛ فيتم تعديله، أو تطوير نماذج نظرية أخرى واختبارها (Byrne, 2010).

وأشارت (بايرن) (Byrne, 2001) إلى أربع خصائص تُميز نمذجة المعادلة البنائية (SEM) عن غيرها من الأساليب الإحصائية متعددة المتغيرات، وهي:

1. تأخذ نمذجة المعادلة البنائية مدخلا توكيديا (فروض يتم التحقق من صدقها) لتحليل البيانات؛ وذلك من خلال العلاقات المسبقة بين المتغيرات.
2. تُقدر نمذجة المعادلة البنائية تباين الخطأ، من خلال رسمة واحدة تُمثل نموذج القياس والعلاقات بين المتغيرات من البيانات مباشرة، في حين لا يتم هذا الأمر في الأساليب متعددة المتغيرات، حيث يتم تجاهل تباين الخطأ، واستخراج الخصائص (السيكومترية) للفقرات بأخذ العينة الاستطلاعية.
3. تتعامل نمذجة المعادلة البنائية مع المتغيرات الكامنة والمتغيرات الملاحظة، في حين تتعامل الأساليب متعددة المتغيرات مع المتغيرات الملاحظة فقط.
4. نمذجة المعادلة البنائية قادرة على نمذجة علاقات متعددة بين المتغيرات، وإظهار التأثيرات المباشرة وغير المباشرة، في حين الأساليب الإحصائية متعددة المتغيرات تُظهر فقط التأثيرات المباشرة.

وتُعبر النماذج بشكل عام عن مُحاكاةٍ أو تصورٍ للظواهر الواقعية والمشاكل البحثية، من خلال تمثيلها ورسمها؛ للوصول إلى فهم أعمق يُساعد في وضع الحلول وتصورها، وتُستخدم النماذج البنائية كطريقة لتحديد العلاقات الخطية المباشرة، وغير المباشرة وتقديرها وتقييمها بين العوامل الكامنة، والمتغيرات الملاحظة، وفيما يلي وصفا مختصرا لأهم النماذج البنائية المستخدمة في الدراسات التربوية والنفسية (Byrne, 2003; Schumacher & Lomax, 2004).

وتُقسم النماذج في النمذجة البنائية إلى قسمين: أحدهما نموذج القياس الذي يمثل العلاقات المتبادلة بين المتغيرات الكامنة، والفقرات التي تقيسها، وأخطاء القياس، والتشبعات على المتغيرات الكامنة، وفيه تظهر العلاقات المتبادلة بين المتغيرات الكامنة على شكل منحنى باتجاهين، وغالباً يُتحقق منه باستخدام التحليل العاملي التوكيدي (تقدير المعالم، وملاءمة مؤشرات المطابقة، ومؤشرات التعديل)، أما القسم الآخر فهو نموذج البناء، ويُبين هذا النوع العلاقات بين المتغيرات الكامنة وتأثيرها المباشر أو غير المباشر على المتغيرات الكامنة الأخرى في النموذج، وأخطاء القياس الناتجة منها، حيث تبدو هذه العلاقات بمسارات خطية باتجاه واحد، ويشمل نموذج البناء نموذج القياس حيث العلاقات المباشرة وغير المباشرة بين المتغيرات الكامنة والملاحظة (Albright & Park, 2009; Brown, 2006; Sarmento & Costa, 2019).

وتجري عملية تحليل نمذجة المعادلة البنائية، أو اختبار جودة النموذج المقترح التي تم الاجماع عليها من أغلبية مصادر نمذجة المعادلة البنائية مرتبة وفق الخطوات الخمس الآتية، هي (Albright & Park, 2009; Brown, 2006; DeCoster, 1998; Harrington, 2009):

- **أولاً: توصيف النموذج:** تعد أصعب خطوة، وتجري قبل جمع البيانات وتحليلها، ويتم توصيف النموذج العاملي المراد مطابقته من خلال وضع تصور لنموذج نظري باستخدام نظريات وبحوث مناسبة؛ لتحديد المتغيرات موضع الاهتمام، وعدد العوامل والتأثيرات المتبادلة فيما بينها والمتغيرات عليها وتوضيح القيود المفروضة على النموذج، ويجري ذلك بالاستناد إلى الأسس النظرية للدراسة، والتحليل العاملي التوكيدي، وتحليل المسار، والنمذجة البنائية كلها أمثلة تعتمد على منطق النظرية الموجهة (Theory Driven)، ومن الأهمية بمكان أن يكون النموذج النظري المقترح له جذور، ومشتق من الأدبيات القائمة.

وهناك عدة قواعد لا بدّ من اتباعها عند توصيف النموذج على صورة مسارات، وهذه القواعد أبرزها: وجود مبرر منطقي لكل تأثير مباشر أو غير مباشر، وتمثيل العوامل الكامنة والمتغيرات المشاهدة، وأخطاء القياس، وتوضيح العلاقات على النموذج، فهناك العلاقات أُحادية الاتجاه من السبب إلى النتيجة، والعلاقات ثنائية الاتجاه حيث إن كل متغير سبب ونتيجة في المتغير الآخر، والارتباط بين اثنين من المتغيرات المستقلة (علاقة منحنية)، ولا يعني أن أحدهما سبب أو نتيجة للآخر، ولا يظهر هذا الارتباط إلا في نماذج القياس في التحليل العاملي التوكيدي.

- **ثانياً: تعريف النموذج:** وتتمثل هذه الخطوة من خلال الإجابة على السؤال التالي:" هل يحتوي النموذج على قدرٍ كافٍ من المعلومات التي تسمح بالحصول على قيمٍ وحيدةٍ لتقديرات معالم النموذج؟ "أي لكي يكون النموذج معرفاً لا بدّ من الوصول لتقديرات لمعالمه، والتقديرات تكون للعلاقات بين المتغيرات المتضمنة بالنموذج.

وتُسمى عملية التوصل إلى حل وحيد ومحدد للمعالم الحرة في النموذج البنائي المقترح بالتعريف والتعيين للنموذج، وعدم التعيين يعني عدم الوصول لحل وحيد أو تقدير لكل معلم من المعالم الحرة للنموذج، وبالتالي يكون لكل معلم عدد كبير من الحلول أو التقديرات؛ مما يعني عدم الوصول لحل مناسب لكل معلم، ويكون النموذج غير قابل للتعريف أو التعيين عندما تُقدر المعالم بعدد أصغر من عدد المعلومات المتمايزة اللازمة؛ أي أن عدد المعلومات التي يحتاجها النموذج أكثر مما هو متوفر.

- **ثالثاً: تقدير النموذج:** وتهدف هذه الخطوة إلى تقدير معالم نموذج المعادلة البنائية (الأرقام التي تظهر على المسارات في النموذج)، والتي يتم من خلالها إيجاد القيم العددية للمعالم الحرة في النموذج، وذلك عندما تكون قيم تلك المعالم قادرة على إنتاج مصفوفة التباين والتباين المشترك للنموذج المقترح، وتكون أقرب ما يمكن من مصفوفة التغاير الناتجة من بيانات عينة الدراسة.

- **رابعاً: تقويم النموذج:** تعتمد عملية تقييم النموذج المقترح على أساس التبرير المنطقي لمعنى النموذج، وفائدته اعتمادا على النظريات ونتائج الدراسات السابقة، وبعد ذلك يأتي تقييم الحل الإحصائي الذي ينطوي على التحقق من مؤشرات حسن المطابقة، ويستدل على جودة النموذج المفترض لبيانات عينة الدراسة من خلال عدة مؤشرات (Brown, 2006; Joseph et al., 2017).

وتُعد مؤشرات حسن المطابقة الإحصائية دليلاً على جودة مطابقة النموذج المقترح لبيانات العينة، والتي تتحدد بعلامة قطع لكلٍ منها، وتفاوتت القيم المعتمدة لمؤشرات المطابقة حسب المراجع المتداولة في النمذجة البنائية، ومن خلالها يتم الحكم على مطابقة النموذج المقترح، ويتم التركيز عادة في أغلب الدراسات على استخدام المؤشرات المطلقة ومؤشرات المقارنة، وقد استخدم الباحثان في الدراسة الحالية المؤشرين ($\chi 2$ / df, RMSEA) كمتغيرات تابعة للدلالة على صدق البناء العاملي للمقياس، والحكم على مطابقة البيانات، والكشف عن خاصية التكافؤ العاملي للمقياس عبر متغير نوع الأداء التفاضلي، وفيما يلي عرض مفصل لهذين المؤشرين (Hair et al., 2006):

1. مؤشر مربع (كاي) المعياري($\chi 2$ / df):

لقد اقترح (جورسكوج) مؤشر مربع (كاي) النسبي أو المعياري كبديل للمؤشر (كاي) تربيع؛ لتأثره الشديد بحجم العينة، ويتم حسابه بإيجاد النسبة المئوية بين قيمة (كاي) تربيع إلى درجات الحرية، ويُنظر للقيمة (5) كحد أعلى للدلالة على مطابقة النموذج الحقيقي لنموذج التقدير، ونُقصان قيمته عن (5) يُدلل على مطابقة أفضل، وحدده بعضهم كمعيار بالمدى (2-1)، وزيادة قيمته عن (5) تدل على سوء المطابقة (Hair et al., 2006; Harrington, 2009; Ullman, 2001).

2. مؤشر جذر وسط خطأ التقريب (RMSEA):

ويعد من أفضل مؤشرات جودة المطابقة؛ لأنه يُبين مدى مطابقة النموذج المقترح مجهول المعالم (الذي يتضمن المعالم الحرة أو المجهولة التي جرى تقديرها بجودة عالية) لمصفوفة التباين والتباين المشترك لبيانات عينة الدراسة؛ وبالتالي يعبر هذا المؤشر عن الخطأ التقريبي لمجتمع الدراسة، فهو يقيس التباعد من خلال درجات الحرية؛ ولذلك يكون حساسا لعدد المعالم المُقدرة، وغير حساس لحجم العينة، وكلما زادت قيمته؛ دل ذلك على سوء المطابقة، حيث يصفه بعضهم كمؤشر لسوء المطابقة، وعندما تكون قيمة المؤشر مساوية للصفر؛ فتدل على أفضل مطابقة للنموذج، وإذا جاءت قيمته أقل أو يساوي (0.05)؛ دل ذلك على مطابقة جيدة، في حين قيمه المحصورة بين (0.05 و 0.1) دل على مطابقة مقبولة للنموذج، وفي حالة الزيادة عن القيمة (0.1)؛ يُوصف النموذج المقترح بأنه سيئ، ويُعاد مراجعته وتدقيقه (Brown, 2006; Chen, 2007; Hu & Bentler, 1999).

خامساً: تعديل النموذج: وتُسمى أيضاً بإعادة توصيف النموذج، ففي حالة وجود جوانب ضعف في النموذج المقترح، وذلك عند تقدير معالم النموذج واستخراج مؤشرات المطابقة، هناك طريقتان لتحديد المشكلة في النموذج، وهما: التحقق من مصفوفة البواقي، ومؤشرات التعديل، وهما متوفرتان في أغلب البرمجيات الإحصائية (Harrington, 2009).

ويُعد اللاتغاير في القياس، أو تكافؤ القياس، أو التكافؤ العاملي خاصية إحصائية تتمثل في ثبات البنية العاملية التي يتم قياسها عبر مجموعات متعددة؛ بمعنى عدم اختلاف البنية العاملية باختلاف مجموعات الدراسة، ويُستخدم للتحقق من تكافؤ تفسيرات الأفراد للإجراءات من الناحية النظرية على اختلاف لغتهم أو عرقهم (Vandenberg & Lance, 2000).

وتأتي أهمية التحقق من خاصية اللاتغاير في القياس في ظل المقارنة بين المجموعات المختلفة؛ من أجل المحافظة على صدق النتائج، حيث إن عدم تحقق اللاتغاير في القياس يُعد مصدراً من المصادر المحتملة لأخطاء القياس؛ وبالتالي وجود فرق بين ما أعد لقياسه وما تم تطبيقه بالفعل، وهذا يضعف نتائج المقاييس الإحصائية، ويؤثر في دقة التقدير، بمعنى الوصول لنتائج مضللة وتنبئات مُشكك فيها(Hult et al., 2008).

ويتم الكشف عن خاصية اللاتغاير(التكافؤ)، بعد الانتهاء من فحص البناء العاملي للمقياس، وغالبا ما يتم ذلك بعد تطبيق المقياس في بيئات مختلفةٍ ثقافياً وعرقياً عن المجتمع الأصلي للمقياس؛ وذلك للتحقق من تماثل خصائص الفقرات (السيكومترية)، ومدى مناسبة المقياس للأفراد بغض النظر عن جنسهم، أو ثقافتهم، أو أي متغيرات أخرى من الممكن أن تؤثر في أدائهم على المقياس، ومن أكثر الطرق المستخدمة للكشف عن تحقق خاصية اللاتغاير في القياس، طريقة التحليل العاملي متعدد المجموعات (Multiple-Group Confirmatory Factor Analysis: MG-CFA)، وبهذه الطريقة تتم المقارنة بين النماذج المتداخلة مع النموذج العام، وذلك للتحقق من أن المعالم المقيدة التي تجري إضافتها إلى النموذج الأبسط تكون متكافئة عبر مجموعات مختلفة، وتجري عملية التحقق من التكافؤ في القياس وفق أربع خطوات، أولها تكافؤ القياس الشكلي (Configurable)، ومن خلاله يتم التأكد من أن نموذج القياس يُحافظ على نسقٍ موحدٍ للعناصر المساوية وغير المساوية للصفر في مصفوفة تشبعات الفقرات والعوامل، بمعنى أن نمط المعالم الحرة والمقيدة يكون نفسه عبر المجموعات، فإذا تحقق هذا التكافؤ، تبدأ الخطوة الثانية المتمثلة بالتكافؤ المتري (Metric) للقياس، وفيها يتم التحقق من التكافؤ النسبي لتشبعات الفقرات على العوامل غير المعيارية عبر المجموعات المختلفة؛ بمعنى تساوي عدد الفقرات المتشبعة على العوامل في مجموعات الدراسة، وبعد التحقق من التكافؤ المتري، تبدأ عملية التحقق من التكافؤ العددي (Scalar)، وفيها يتم التحقق من تطابق قواطع (Intercept) الفقرات في مجموعات الدراسة أو التساوي النسبي لمتوسطات المؤشرات عبر المجموعات، وأخيراً في حال تحقق أنواع اللاتغاير الثلاثة السابقة، يتم التحقق من ثبات البواقي (Residuals)، ويستدل عليه من خلال تكافؤ بواقي الفقرات المتكافئة في المستويين المتري والعددي؛ أي أن مجموع التباين بين الفقرة والعامل غير المشترك وتباين الخطأ متماثل عبر مجموعات الدراسة، ويمثل كل مستوى منها جانبا من العلاقة بين البناء الكامن والمشاهدات الملاحظة، ويكون المستوى السابق شرطاً لتحقق المستوى اللاحق (Cheung & Rensvold, 2002; Millsap, 2011).

وظهرت معان عدة متقاربة للأداء التفاضلي للفقرة، حيث بين (براون) (Brawn, 2013) بأنه احتمالات مختلفة لإجابة الفقرة إجابة صحيحة من الأفراد الذين ينتمون إلى مجموعاتٍ مختلفةٍ، ولهم مستوى القدرة نفسه، وعادة يتم مقارنة توزيع استجابات الأفراد بين مجموعتين فرعيتين أو أكثر تسمى أحدهما المستهدفة وهي موضع الاهتمام مثل مجموعة النساء أو مجموعة الأقلية،

والتي يعتقد أنها تتأثر بالأداء التفاضلي للفقرة، ومجموعة أخرى هي المرجعية وتقارن مع المستهدفة مثل مجموعة البيض أو الذكور(Michel ides, 2008) .

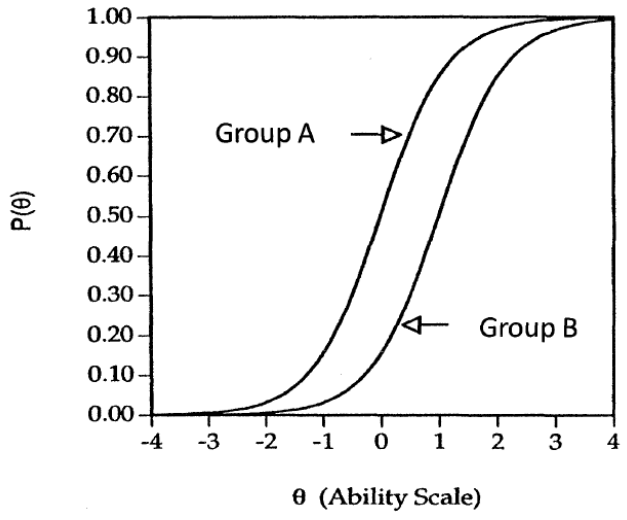
وذكر (هامبلتون) وآخرون(Hambleton et al., 2004) الأداء التفاضلي بأنـــه الفرق في احتمالات الاستجابة الصحيحة على الفقرة في المجموعات المختلفة والمتساوية في القدرة، في حين ذكر (مالر وأنجوف) (Angoff, 1993; Maller, 2003) بأن معنى الأداء التفاضلي للفقرة يدل على الاختلافات في الخصائص الإحصائية لفقرة ما بين مجموعتين من الأفراد لهم قدرة متساوية، وتكون المجموعتان هما مجموعة مرجعية وتشير إلى الأغلبية، ومجموعة مستهدفة وتشير إلى الأقلية.

وذكر (وولكر) (Walker, 2011) معنى الأداء التفاضلي للفقرة، بأنه حصول مجموعة من الأفراد على احتمالات مختلفة للإجابة الصحيحة على الفقرة عن مجموعة أخرى لهم مستوى متساوٍ من القدرة؛ وقد يرجع ذلك إلى أن الفقرة التي تظهر أداءً تفاضليا تقيس أكثر من سمةٍ أو جانب بسيط من السمة المقصودة، وبالتالي تكون غير عادلة لبعض الأفراد.

وأشار (زامبو) (Zumbo, 1999) إلى الطريقة (السيكومترية) للكشف عن وجود الفقرات ذات الأداء التفاضلي(DIF) الذي يرتبط باختلاف الموقع على متصل القدرة (الصعوبة أو العتبة)، وذلك من خلال مقارنة منحنيات خاصية الفقرة (ICCS) لمجموعاتٍ مختلفةٍ من الأفراد، حيث يمثل الأداء التفاضلي على الفقرة نفسها بمنحنى خاص لكل مجموعة، فإذا تطابقت المنحنيات أو كانت متقاربة جداً بعضها مع بعض؛ فإن الفقرة لا تُبدي أداء تفاضليا، أما إذا ابتعدت منحنيات الفقرة نفسها للمجموعات المختلفة بشكل واضح؛ فتُوصف الفقرة حينها بالأداء التفاضلي، وفي هذه الحالة تكون المساحة بين المنحنيات أقل ما يمكن، وقيم المعالم المتماثلة متكافئة تقريبا.

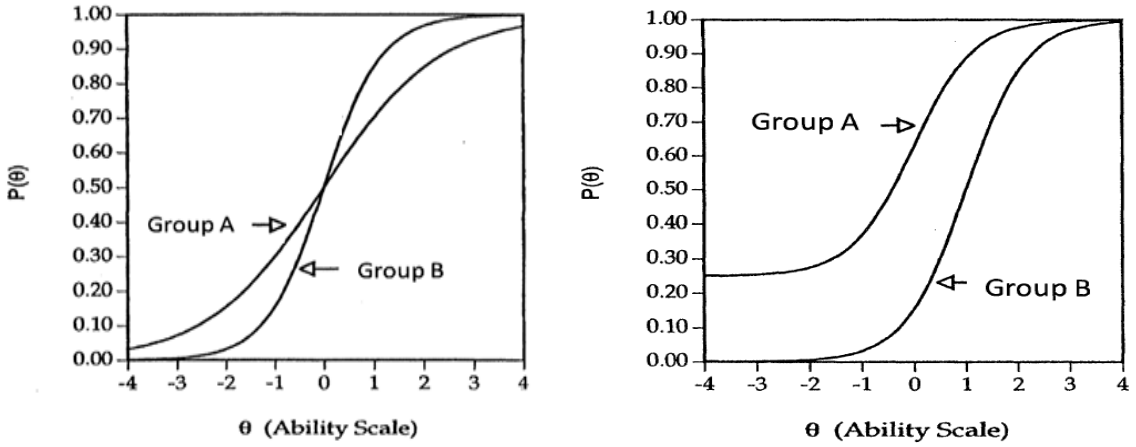
وينقسم الأداء التفاضلي للفقرة إلى نوعين؛ فإما أن يكون منتظما (Uniform)، أو غير منتظم (Non-Uniform)، ويعتمد ذلك على اختلاف مَعلمتي الصعوبة والتمييز لفقرات المقياس عبر المجموعتين (المرجعية، والمستهدفة)، ويعتمد ذلك على التفاعل بين مستوى القدرة وعضوية المجموعة.

الأداء التفاضلي المنتظم: يحدث الأداء التفاضلي المنتظم عندما تختلف مَعلمة الصعوبة (b) للفقرة، والتي تُعبر عن احتمال الإجابة الصحيحة على الفقرة عبر المجموعتين مع ثبات مَعلمتي التمييز(a)، والتخمين(c) للفقرة نفسها، ويتبين ذلك من خلال منحنى خصائص الفقرة (ICC) حيث تكون احتمالية الاستجابة الصحيحة على الفقرة أعلى وبشكلٍ منتظمٍ عند إحدى المجموعات، وعبر كل مستويات القدرة حيث تختلف المنحنيات، ولكنها لا تتقاطع؛ بمعنى عندما يتوازى منحنيا خصائص الفقرة (ICC) للمجموعتين المرجعية والمستهدفة بحيث تكون قيمة الفرق بين المجموعتين في احتمال الإجابة الصحيحة على الفقرة لها المقدار والاتجاه نفسه عبر مستويات القدرة المختلفة، وبالتالي لا يوجد تفاعل بين مستوى القدرة والمجموعة التي يتبعها الفرد، وفي هذه الحالة لا تمثل الفقرة قياسا متكافئا للسمة الكامنة نفسها عبر المجموعتين موضع الاهتمام كما هو مبين في الشكل رقم(1) (Gruijter & Kamp, 2008; Zumbo, 1999).



الشكل(1): الأداء التفاضلي المنتظم بين المجموعتين

1. **الأداء التفاضلي غير المنتظم:** يظهر الأداء التفاضلي غير المنتظم عندما تختلف مَعلمة التمييز(a) للفقرات عبر المجموعتين، وبصرف النظر عن اختلاف مَعلمة الصعوبة (b) والتخمين(c) للفقرة؛ بمعنى عند حصول التفاعل بين مستوى قدرة الفرد والمجموعة التي ينتمي لها، حيث إن الفرق في احتمالات الإجابة الصحيحة على الفقرة ليس متساويا عند كل مستويات القدرة، ويكون الأداء التفاضلي للفقرة لصالح المجموعة المستهدفة عند مستوى قدرة معين ولصالح المجموعة المرجعية عند مستوى قدرة آخر؛ بمعنى أن منحيي خصائص الفقرة (ICC) غير متوازيين، وبالتالي يكون الفرق في احتمال الإجابة الصحيحة على الفقرة بين المجموعتين مختلفاً في المقدار والاتجاه معاً (أو في أحدهما فقط)، وهنا قد تتقاطع منحنيات خصائص الفقرة (ICC) بين المجموعتين كما هو موضح في الشكل رقم(2) (Gruijter & Kamp, 2008; Zumbo, 1999).



الشكل(2): الأداء التفاضلي غير المنتظم بين المجموعتين.

وعند إجراء التحليل العاملي للمقياس، يتم التأكد من أن البناء العاملي متكافئ عبر المجموعات المختلفة، ويجري ذلك بطرق عدة أبرزها الكشف عن الأداء التفاضلي لفقرات المقياس المستخدم (Teresi & Fleishman, 2007)، ويعزو بعضهم ظهور الأداء التفاضلي إلى تعدد البعدية؛ ويحدث ذلك عندما تقيس الفقرة أبعادا عدة، وعندما تختلف المجموعات في مواقعها بالنسبة للمتغير أو المتغيرات الكامنة، وفي حال بقيت المجموعات متماثلة في مواقعها على المتغيرات فلا يحدث الأداء التفاضلي ولو تعددت الأبعاد (Ackerman, 1992).

إن معرفة الفقرات ذات الأداء التفاضلي والتخلص منها أو تعديلها؛ يزيد من دقة التنبؤ بالعوامل الكامنة، وتزداد أهمية الكشف عن الأداء التفاضلي في بناء وتكييف المقاييس التي ينتج عنها قرارات مهمة مثل قبول واختيار الأفراد للوظائف (Kalaycioglu & Berberoglu, 2011).

ويُعد الكشف عن الفقرات ذات الأداء التفاضلي في المقاييس النفسية أمراً مهماً جداً؛ لأنه قد يكون دليلاً على تحيز الفقرة، وهذا يعني عدم تحقق العدالة على المقياس، وهذا ينعكس على صدق المقاييس وثباتها، حيث يظهر الأداء التفاضلي الخصائص (السيكومترية) للفقرة، والفروق بين المجموعات المختلفة، في حين يهتم التحيز بالجانب الاجتماعي والعدالة والمساواة بين الفقرات (Braun, 2013; Salubayba, 2013).

ويحدث التحيز عندما تحتوي الفقرة على مصادر أخرى تحول دون إجابة الطلبة عليها إجابة صحيحة غير تلك المصادر التي ترتبط ببناء المقياس، وذلك يؤثر عكسياً على أداء الطلبة على المقياس (Huang & Han, 2012).

وقد أشار عودة (2014) إلى تعدد أشكال التحيز، فقد يظهر في عينة الفقرات من خلال زيادة تمثيل وحدة دون أخرى، أو تحيزاً في صياغة الفقرات عندما تكون الصياغة قوية في المفردات، والألفاظ لصالح المتفوقين لغوياً، أو تحيزاً في المحتوى لصالح جنس أو عرق معين.

ويمكن الكشف عن التحيز في محتوى الفقرات، والصياغة اللغوية لها من خلال إعداد قائمة للمراجعة والتحكيم، يقوم بها مجموعة من الخبراء في الموضوع الذي يقيسه المقياس، ومختصون في حقل القياس النفسي والتربوي، الذين لهم دراية، ومعرفة بخصائص

المفحوصين، حيث يقومون بمراجعة الفقرات تبعاً للمعايير الواردة بالقائمة؛ وبالتالي معرفة العوامل المؤثرة في استجابات الأفراد المرتبطة بالجنس، أو اللغة، أو العرق من أجل حذفها من الفقرات قبل تطبيق المقياس (Hambleton et al., 1991).

وقام السيد (2022) بدراسة هدفت إلى تقنين مقياس العوامل الستة المنقح للشخصية نسخة التقرير الذاتي للبيئة الأردنية (HEXACO-PI-R)، وإيجاد خصائصة (السيكومترية)، ولتحقيق أهداف الدراسة تم استخدام المنهج الوصفي، وترجم المقياس من اللغة الإنجليزية إلى اللغة العربية، وطبق على عينة مكونة من (1356) طالباً جامعياً، وتوزعت على ثلاث عينات (الاستطلاعية، والمكونات الأساسية، والتحليل العاملي التوكيدي)، واختيرت العينة بالطريقة المتاحة من ست جامعات حكومية، وأظهرت نتائج الدراسة وجود (25) عاملاً في البيئة الأردنية، وبينت النتائج وجود مطابقة جيدة للبيانات التي تم جمعها من عينة الدراسة مع النموذج النظري الذي تم بناء المقياس بالاستناد إليه، وتم استخدام خمسة مؤشرات للحكم على صحة المطابقة، هي: (χ2, CFI, TLI, RMR, RMSEA) حيث أظهرت تمتع المقياس بخاصية اللاتغاير تبعاً لمتغير الجنس، وأشارت النتائج تطابق المقياس مع نموذج الاستجابة المتدرجة، وعدم وجود أداء تفاضلي لفقراته باستثناء ثلاث فقرات أظهرت أداء تفاضليا تبعاً لمتغير الجنس، وقد تم حذفها من صورة المقياس النهائية.

وأجرت الشوبكي (2021) دراسة هدفت إلى التحقق من البناء العاملي لمقياس جودة الحياة لدى أفراد المجتمع الأردني، والكشف عن أثر خفض فقرات المقياس باستخدام مصفوفة المعاينة المتعددة، ثم الكشف عن اللاتغير في القياس لأفضل مقياس فرعي ناتج من المقياس الأصلي، ولتحقيق أهداف الدراسة؛ فقد طبق مقياس جودة الحياة المئوي، وبينت نتائج التحليل العاملي التوكيدي صدق البناء العاملي، حيث تطابق النموذج المقترح المعدّل مع البيانات المستخدمة، ولمعرفة أثر خفض فقرات مقياس جودة الحياة، استُخدمت مصفوفة المعاينة المتعددة، وقُسم المقياس الأصلي إلى تسعة مقاييس فرعية، وبلغت عينة الدراسة الكلية (3600) فرد جرى اختيارهم عشوائياً، وبينت نتائج التحليل العاملي التوكيدي أن المقياس الفرعي العاشر والرابع حققا أفضل تطابقٍ لمؤشرات حسن المطابقة، وعدم تحقق اللاتغير الشكلي في القياس للمقياس الفرعي العاشر والرابع، وتحقق اللاتغير التكويني فقط، وعدم تحقق اللاتغير المتري والعددي في القياس.

وقامت السعيد (2021) بدراسة هدفت إلى دراسة فاعلية طريقة المواءمة في الكشف عن اللاتغاير (ثبات القياس) للاختبار التقييمي للصف الثالث الأساسي لمهارات الرياضيات في ضوء متغير مديرية التربية والتعليم، ومتغير الجنس، حيث تكونت عينة الدراسة من (73448) طالباً وطالبةً، تم اختيارهم بطريقة عشوائية، ولتحقيق أهداف الدراسة، تم استخدام المنهج الوصفي التحليلي في تحليل نتائج الطلبة على الاختبار التقييمي لمهارات الرياضيات لضبط نوعية التعليم، وطبق على عينة الدراسة، وحللت استجابات الطلبة على (25) فقرة اختبارية من نوع الاختيار من متعدد، وتم استخدام أسلوبي التحليل العاملي التوكيدي من الدرجتين الأولى والثانية، والتحليل العاملي التوكيدي متعدد المجموعات، وللحكم على ملاءمة البناء العاملي للبيانات أستخدمت خمسة مؤشرات مطابقة (χ2, RMSEA, SRMR, CFI, TLI)، وأظهرت النتائج ملاءمة النموذج المقترح بمهاراته السبع لمحتوى الاختبار التقويمي، كما أظهرت مؤشرات الفرق (ΔRMSEA, ΔRMSR, ΔCFI, ΔTLI) حقق خاصية اللاتغاير في القياس تبعاً لمتغيري المديرية، وجنس الطالب.

وقامت الحربي والمدني (2021) بدراسةٍ هدفت للكشف عن أثر نسبة الفقرات ذات الأداء التفاضلي على قوة، وفاعلية اختبار منحنى خصائص الفقرة، حيث تم توليد البيانات لثلاثة اختبارات تكون كل منها من (100) فقرة، وتضمن كل واحد منها عدداً من الفقرات ذات الأداء التفاضلي بنسبٍ بلغت (10%)، (20%)، (30%)، وبلغ عدد أفراد الدراسة (3000) فرد توزعت قدراتهم توزيعا طبيعيا بمتوسط حسابي قيمته (0)، وانحرافٍ معياري قيمته(1)، وجاءت نتائج قوة الاختبار بطريقة راجو على النحو (0.619)، (0.786)،(0.703) للنسب (10%)، (20%)، (30%) على الترتيب، وجميعها حققت قوة أعلى من المتوسط حسب تصنيف اختبار كوهين.

وأجرى الشريفين (2018) دراسة هدفت إلى معرفة أثر نوع الأداء التفاضلي (منتظم، وغير منتظم) على الخصائص السيكومترية لفقرات الاختبار، باستخدام النموذج ثلاثي المعلمة، واللامعلمي لنظرية استجابة الفقرة. واستخدم برنامج (WINGEN) لتوليد نموذجي اختبار من نوع الاختيار من متعدد، وتكون كل منهما من (50) فقرة، تضمن الأول (10) فقرات ذات أداءٍ تفاضلي منتظم، والثاني (10) فقرات ذات أداءٍ تفاضلي غير منتظم، وتم توليد استجابات (1000) فرد، وحُللت البيانات وفق النموذج ثلاثي المعلمة، وأشارت النتائج إلى وجود فرق دال إحصائيا عند مستوى الدلالة (0.05) بين وسطي معالم التمييز، ومعالم التخمين لفقرات الاختبار، يعزى لنموذج نظرية استجابة الفقرة المستخدَم لصالح النموذج المعلمي مقارنة بالنموذج اللامعلمي،

وأيضا وجود فرق دال إحصائيا عند مستوى الدلالة (0.05) بين وسطي معالم الصعوبة، يعزى للنموذج المستخدم لصالح النموذج اللامعلمي مقارنة بالنموذج المعلمي، ووجود فروق دالة إحصائيا بين الأوساط الحسابية لمعالم كلٍ من: الصعوبة، والتخمين لفقرات الاختبار؛ تعزى للتفاعل بين النموذج المستخدم، ونوع الأداء التفاضلي، وبينت النتائج وجود فرق دال إحصائيا عند مستوى الدلالة (0.05) بين وسطي كمية معلومات الاختبار، يعزى لصالح النموذج اللامعلمي مقارنة بالنموذج المعلمي، حيث كانت كمية المعلومات التي يقدمها الاختبار أكبر عند مستويات القدرة جميعها وفق النموذج اللامعلمي، سواء أكان الأداء التفاضلي منتظما أم غير منتظم.

وأجرى أبو علام والمرابحة (2015) دراسة هدفت إلى الكشف عن أثر الفقرات ذات الأداء التفاضلي، وحجم العينة على قوة وفاعلية اختبار تحيز الفقرة المتزامن (SIBTEST) دراسة محاكاة، وتم توليد قدرات الأفراد على اختبار مكون من (50) فقرة حيث صممت (9) اختبارات لتحقيق أهداف المقارنة المتضمنة في الدراسة على العينات (400)، (800)، (1200) لكلا المجموعتين، وجرى اختبار شكل التوزيع الطبيعي، واستخدم النموذج ثلاثي المعلمة في توليد الاستجابات، واعتمدت قوة الاختبار كمعيار لتقييم فاعلية اختبار تحيز الفقرة المتزامن للكشف عن دالة الفقرة التفاضلية، وتم تحديد نسب فقرات الأداء التفاضلي (10%)، (20%)، (30%)، وحجم العينات للمجموعتين، وبينت نتائج الدراسة أن قوة اختبار تحيز الفقرة قد حققت مستوى مقبولا للحكم على قوة الاختبار الإحصائي عند النسب (20%)، (30%)، حيث بلغت قوة الاختبار(0.736) للنسبة (20%)، وبلغت (0.788) للنسبة (30%)، في حين بلغت (0.693) عند النسبة (10%). وقام اجباره (2015) بدراسةٍ لمعرفة أثر نسب الأداء التفاضلي على الخطأ من النوع الأول، وقوة الاختبار في الكشف عنه، مستخدما طريقتي الانحدار (اللوجستي)، ونسبة الأرجحية، وتم توليد البيانات بطريقة المحاكاة وفق نظرية استجابة الفقرة للنموذج أحادي المعلمة، حيث تم توليد اختبار بطول (40) فقرة واستجابات (1500) فرد لكل طريقة، واستخدمت نسب الأداء التفاضلي (10%)،(20%)، (40%)، وأشارت النتائج لوجود فروق دالة إحصائيا للخطأ من النوع الأول، تعزى لنسب الأداء التفاضلي، ولصالح النسب الأقل، ووجود فروق دالة إحصائيا تعزى لطريقة الكشف المستخدمة، ولصالح نسبة الأرجحية، وعدم وجود تفاعل بين نسب الأداء وطريقة الكشف المستخدمة.

وأجرى النوافلة (2013) دراسة هدفت لمعرفة أثر نسب فقرات الأداء التفاضلي، ومستوى الأداء التفاضلي في تقدير معالم الفقرات، وقدرة الأفراد وفق النموذج ثلاثي المعلمة، وتم توليد البيانات في ظروفٍ تجريبيةٍ مختلفةٍ، وبينت النتائج وجود دلالةٍ إحصائية بين متوسطات تقديرات معالم الصعوبة، تعزى إلى نسب فقرات الأداء التفاضلي لصالح المجموعة المستهدفة، ووجود فروقات متباينة بين متوسطات معالم التمييز والتخمين لصالح المجموعة المرجعية، وعدم وجود فروق ذات دلالةٍ إحصائية لتقديرات القدرة، وانخفاض كمية معلومات الاختبار بزيادة نسبة الأداء التفاضلي.

وقام جابريل (Gabriel, 2012) بدراسة هدفت للمقارنة بين فعالية ثلاث طرق للكشف عن الأداء التفاضلي مستخدماً النماذج المعلمية واللامعلمية، وهذه الطرق هي: اختبار تحيز الفقرة المتزامن (SIBTEST)، والأرجحية العظمى(MLE)، والانحدار اللوجستي(LOGREG)، واستخدمت مجموعة من الظروف التجريبية المختلفة (طول الاختبار، وحجم العينة، ونسبة الفقرات ذات الأداء التفاضلي، وتوزيعات مختلفة للقدرة للمجموعتين المرجعية والمستهدفة)، وذلك من خلال توليد استجابات الأفراد وفق النموذج اللوجستي ثلاثي المعلمة، وبينت النتائج تفوق طريق الانحدار اللوجستي في قوة الكشف عن الأداء التفاضلي، ومعدل الخطأ من النوع الأول، وعدم وجود طريقة واحدة ذات فاعلية تميز بين نوعي الأداء التفاضلي.

وأجرى زهان وزامبو(Zhan & Zumbo, 2009) دراسة للكشف عن أثر الأداء التفاضلي للفقرات على الخطأ من النوع الأول، وحجم الأثر، حيث استخدم النموذج اللوجستي ثلاثي المعلمة (3PLM) لتقدير معالم الفقرات بافتراض ثلاثة مستويات للأداء التفاضلي للفقرة، وبينت النتائج عدم تأثر مستوى الخطأ من النوع الأول بوجود فقرة واحدة ذات أداء تفاضلي من بين (38) فقرة، وأظهرت النتائج تضخماً لمستوى الخطأ من النوع الأول بوجود (16) فقرة ذات أداء تفاضلي، وكذلك فإن الزيادة في عدد الفقرات ذات الأداء التفاضلي تعمل على اتساع الفروق في حجم الأثر.

مشكلة الدراسة وأسئلتها

لقد توسعت استخدامات المقاييس النفسية، وجرى تطويرها في بلدان متعددة عبر لغات مختلفة؛ ونتيجة لذلك يتم التحقق من بنائها العاملي، وخصائصها (السيكومترية)، والكشف عما يسمى بالأداء التفاضلي لفقراتها، والذي يؤثر بشكلٍ كبيرٍ في الصدق البنائي لها؛ مما ينعكس على البناء العاملي، ودقة معاملات الصدق والثبات المستخرجة لها .

وتم التوصل إلى نتيجة مفادها، ندرة الدراسات التي تناولت موضوع البناء العاملي للمقياس في ضوء نوع الأداء التفاضلي للفقرات؛ وبالتالي تميزت هذه الدراسة عن غيرها، من خلال الربط ما بين موضوعين في غاية الأهمية في حقل القياس النفسي والتربوي وهما: الأداء التفاضلي والبناء العاملي للمقياس في ضوء ظروفٍ تجريبيةٍ متعددة.

وما تنفرد به هذه الدراسة، وتحاول الإجابة عليه يتمثل بمدى تأثر البناء العاملي للمقياس بنوع الأداء التفاضلي (منتظم، وغير المنتظم) للفقرات في ظل ظروفٍ تجريبيةٍ مختلفة، وذلك بتوليد بيانات الدراسة؛ وانطلاقاً مما سبق؛ فقد جاءت هذه الدراسة للبحث في البناء العاملي للمقياس في ضوء اختلاف نوع الأداء التفاضلي للفقرات (منتظم، وغير المنتظم)، وبالتحديد تحاول هذه الدراسة الإجابة عن السؤال الرئيس الآتي:

هل يتأثر البناء العاملي للمقياس في ضوء اختلاف نوع الأداء التفاضلي (منتظم، وغير المنتظم) للفقرات؟

ينبثق عن هذا السؤال الرئيس الأسئلة الفرعية الآتية:

السؤال الأول: هل يوجد أثر لنوع الأداء التفاضلي على التطابق الشكلي للبناء العاملي للمقياس؟
السؤال الثاني: هل يوجد أثر لنوع الأداء التفاضلي على التطابق المتري للبناء العاملي للمقياس؟
السؤال الثالث: هل يوجد أثر لنوع الأداء التفاضلي على تطابق قياس البناء العاملي للمقياس؟
السؤال الرابع: هل يوجد أثر لنوع الأداء التفاضلي على تطابق البواقي للبناء العاملي للمقياس؟

أهمية الدراسة

تبرز أهمية الأداء التفاضلي كمهدد لصدق البناء وثبات الاختبار، إضافة إلى أثره السلبي على دقة تقدير معالم الفقرات وقدرة الأفراد؛ وبالتالي يُسفر عن تقديرات ليست دقيقة وغير صادقة؛ وهذا بدوره يُثير مخاوف تتعلق بعدالة تصنيف الأفراد وسوء توزيعهم، والفهم الخطأ في الدراسات، ومرد ذلك كله هو اختلاف وظيفة الفقرة في التقييم عبر مجموعات الأفراد المختلفة (Croudace & Brown, 2012).

وتأتي هذه الدراسة لتكشف عن مدى تأثر البناء العاملي للمقياس في ضوء اختلاف نوع الأداء التفاضلي للفقرات، وتتمثل أهمية هذه الدراسة في الأهمية النظرية، والأهمية العملية.

الأهمية النظرية

تُعد المقاييس من أهم عناصر الدراسات والأبحاث التربوية والنفسية؛ لذلك وجب التأكد من صدقها في قياس ما تصبو إليه، ويُعد تحيز الفقرات من القضايا المهمة التي تواجه المقاييس في مدى قدرة الفقرات على قياس السمة التي أعدت من أجلها، والذي يمكن الاستدلال عليه إحصائيا من خلال ما يُعرف بالأداء التفاضلي للفقرات حيث يمكن الحكم على بقاء الفقرات أو إزالتها من المقياس؛ بسبب تحيزها لفئة دون أخرى، ولضمان صلاحية المقياس، وتتلخص الأهمية النظرية فيما يأتي:

- توفر هذه الدراسة مصدراً غنياً للمهتمين في بناء وتطوير المقاييس التربوية والنفسية، ومكملة لجهود الباحثين في مجالي البناء العاملي والأداء التفاضلي، حيث تم تناول ومناقشة الإطار النظري، والدراسات السابقة ذات الصلة، وربطها بنتائج الدراسة الحالية.
- أهمية دراسة العوامل التي تؤثر في البناء العاملي للمقياس ومنها اختلاف نوع الأداء التفاضلي للفقرات (منتظم، وغير منتظم).
- تُعد واحدةً من الدراسات القليلة التي تناولت البناء العاملي للمقياس في ضوء اختلاف نوع (منتظم، وغير المنتظم) الأداء التفاضلي للفقرات.
- أهمية البناء العاملي للمقياس، والتحقق منه من خلال التحليل العاملي لأدوات القياس المستخدمة.

- توجيه نظر الباحثين إلى أهمية التحقق من الأداء التفاضلي للفقرات، ونوعه (منتظم، وغير المنتظم) في الخصائص (السيكومترية) للفقرات، وأثر ذلك على البناء العاملي للمقياس.
- ونظراً لأهمية خاصية اللاتغاير فى القياس النفسي والتربوي؛ قدمت الدراسة الحالية إطاراً إثرائياً يشتمل على الجوانب النظرية، والإحصائية المتعلقة بخاصية اللاتغاير في القياس (MI)، ودورها في تحسين البناء العاملي للمقاييس، ويُشكل فرصة للدارسين للمقاييس والاختبارات من الرجوع إليه عند دراسة البناء العاملي للمقاييس في ضوء اختلاف نوع الأداء التفاضلي (منتظم، وغير المنتظم) للفقرات.

الأهمية العملية

تُعد المقاييس على اختلاف أنواعها أحد أهم أنواع الأدوات التربوية والنفسية استخداماً وتطبيقاً؛ ولذلك تتجلى الأهمية العملية لهذه الدراسة فيما يأتي:

- ضرورة القيام بفحص الأداء التفاضلي للفقرات كخاصية (سيكومترية) مهمة كغيرها من الخصائص، مثل: الصعوبة والتمييز، وتقديم الأدلة العملية على تأثر البناء العاملي للمقياس بنوع الأداء التفاضلي للفقرات (منتظم، وغير المنتظم).
- الإفادة من نتائج هذه الدراسة، من خلال تزويد المهتمين بالمجالين التربوي والنفسي بنوع الأداء التفاضلي للفقرات (منتظم، وغير المنتظم) المسموح به ضمن مقاييسهم واختباراتهم، بحيث تكون قادرة على تحقيق أهدافها والمحافظة على البناء العاملي.
- الاهتمام بتطبيق التحليل العاملي كأُسلوب إحصائي في الدراسات النفسية والتربوية.
- يُتوقع استخدام نتائج هذه الدراسة من المختصين ببناء المقاييس، وتكييفها في بيئاتٍ وبلدان مختلفة.
- تشجيع الباحثين على استخدام البرمجيات الحديثة المختصة في توليد البيانات وتحليلها.
- استخدام البيانات المولدة في بناء وتطوير المقياس، ووضع آليةٍ لتقييم نتائجهم عند وجود فقراتٍ ذات أداءٍ تفاضلي على اختلاف نوع الأداء التفاضلي للفقرات (منتظم، وغير المنتظم).

التعريفات النظرية والإجرائية

البناء العاملي: مجموعة من المتغيرات غير المشاهدة أو العوامل الكامنة المرتبطة بفقرات المقياس، ويمثل نوعاً من صدق البناء، ويتم التوصل إليه من خلال التحليل العاملي (أبو حطب وصادق، 1991؛ تيغزة، 2012).

اللاتغاير في القياس: خاصية إحصائية تُشير إلى قياس البناء العاملي نفسه عبر مجموعاتٍ محددة (Vanderberg & Lance, 2000)، ويُشير إلى ثبات العمليات أو الوظائف التي من المفترض أن يقيسها المقياس عبر المجموعات المختلفة؛ أي تكافؤ المحتوى والتفسير للمقياس عند التحقق منه عبر المجموعات المختلفة من الأفراد (Byrne & Watkins, 2003).

الأداء التفاضلي للفقرات: دالة مشتقة إحصائيا، تعبر عن الفرق في احتمالات الاستجابة الصحيحة على الفقرات بين المجموعات المختلفة عند مستوى قدرة واحد للمفحوصين (Hambleton, 1999).

نوع الأداء التفاضلي للفقرات: نوع الأداء الذي تُبدي فيه الفقرة فروقاً في احتمالات الإجابة عليها إجابةً صحيحةً بين مجموعتيّ المقارنة عند مستوى القدرة نفسه، ويكون منتظماً (Uniform) إذا كان احتمال الإجابة على الفقرة لمجموعة ما أكبر أو أقل من مجموعة أخرى عند مستويات القدرة، وغير المنتظم (Non – Uniform) إذا كان احتمال الإجابة الصحيحة على الفقرة مختلفاً بين المجموعتين، ويكون لصالح مجموعة دون أخرى عند مستوى معين من القدرة، ولصالح مجموعة أخرى عند مستوى قدرة مختلف (Walker, 2011; Zumbo, 1999).

محددات الدراسة

1. تتحدد نسب الفقرات ذات الأداء التفاضلي بنسبٍ عشوائية من الأداء التفاضلي ضمن المدى (10%–30%) من فقرات المقياس.
2. بيانات هذه الدراسة افتراضية، تم توليدها من خلال برمجية (R) لتحاكي الظروف التجريبية.
3. استخدام النموذج ثلاثي المعلمة (3PLM) من نماذج نظرية استجابة الفقرة، وشكلٍ واحدٍ من الفقرات ثنائية التدريج (0،1).
4. اقتصرت معايير جودة مطابقة النموذج للبيانات على مؤشري جودة المطابقة المطلقة (χ^2 / df ,RMSEA).

5. تتحدد تعميم نتائج الدراسة في ضوء التعريفات الإجرائية التي تمثل المفاهيم والمصطلحات الواردة فيها.
6. دقة تقدير نتائج البرامج الحاسوبية المستخدمة في توليد وتحليل بيانات الدراسة الدراسات السابقة.

منهجية الدراسة

لتحقيق أهداف الدراسة المتمثلة في الكشف عن تأثير نوع الأداء التفاضلي(منتظم، وغير المنتظم) على البناء العاملي للمقياس، اتبعت هذه الدراسة المنهج التجريبي بالاعتماد على أسلوب المحاكاة في توليد البيانات في ظل تحديدٍ مسبق لأحجام العينات، مستوى القدرة ومعالم الفقرات.

عينة الدراسة

تكونت عينة الدراسة من مجموعة الأفراد الافتراضيين الذين تم محاكاة توزيع قدراتهم باستخدام برمجية(R) لتتبع التوزيع الطبيعي وفق الظروف التجريبية وبحجم بلغ (2000) فرد موزعة بالتساوي على المجموعتين المرجعية والمستهدفة ولكل من مرات التوليد.

إجراءات الدراسة

تم محاكاة الظروف التجريبية من خلال توليد بيانات بمعدل (300) مرة لكل ظرف تجريبي بحيث يكون حجم العينة لكل من مرات التوليد (2000) فرد موزعين على مجموعتين (مستهدفة، ومرجعية) بالتساوي لكل من الظروف التجريبية، وبلغ طول المقياس (50) فقرة حسب ما استخدمه وأوصى به كل من (دراسجو ولورد) (Drasgow, 1982; Lord, 1980) بأن يكون طول الاختبار (50) فقرة، وعدد الأفراد (1000) فرد للحصول على أفضل دقة لتقديرات معالم النموذج ثلاثي المعلمة لاستجابة الفقرة؛ إذ إنه بزيادة حجم العينة وطول الاختبار يتناقص الخطأ المعياري لتقدير المعالم في النموذج ثلاثي المعلمة، وكذلك للإيفاء بمتطلب التحليل العاملي التوكيدي (Confirmatory Factor Analysis) من حجم عينة يتناسب مع عدد الفقرات (تيغزة، 2012). ولتحقيق أهداف الدراسة تم اتّباع الإجراءات الآتية:

- **أولاً:** محاكاة الأفراد حيث تم استخدام برمجية (R) لتوليد القدرة والتي تتبع التوزيع الطبيعي لكل من المجموعتين المرجعية (Reference) والمستهدفة (Focal) وذلك تماشياً مع المبدأ الذي يستند إليه الأداء التفاضلي والذي يقضي بأن المجموعات مختلفة الاستجابة على الرغم من تساوي القدرة، وقد كانت القدرة لكل من المجموعتين تتبع التوزيع الطبيعي بمتوسط حسابي (0) وانحراف معياري 1 (0.1)N~.
- **ثانياً:** محاكاة المقاييس من خلال توليد معالم فقرات المقياس حيث تم استخدام برمجية (R) في توليد توزيع كل من معالم الفقرات وفقا للنموذج ثلاثي المعلمة (3PLM)، ويبين الجدول (1) توزيع معالم فقرات المقياس.

جدول(1): توزيع معالم فقرات المقياس المولدة.

المعلمة	بدون أداء تفاضلي	ذات الأداء التفاضلي
الصعوبة	$\sim N(0,1)$	$\sim N(0.2,1)$
التمييز	$\sim Unif(0.5,1.5)$	$\sim Unif(1,2)$
التخمين	$\sim Unif(0,0.5)$	$\sim Unif(0,0.5)$

- **ثالثاً:** محاكاة استجابات الأفراد على فقرات المقاييس، حيث استخدمت برمجية (R) لتوليد استجابات الأفراد المولدة قدراتهم في الخطوة (أولاً) على الفقرات المولدة معالمها في الخطوة (ثانياً) في ضوء الظروف التجريبية والمتمثلة في نوع الأداء التفاضلي، حيث تم توليد استجابات الأفراد على فقرات المقياس بواقع (600) مرة موزعة على (2) من الظروف التجريبية بالتساوي (300) مرة لكل ظرف، ويحدد هذا العدد عند استقرار التقديرات للحصول على نتائج أكثر موثوقية كما هو مبين في الجدول (2).

الجدول (2): وصف الظروف التجريبية لعينة الدراسة.

الظرف	نوع الأداء التفاضلي	عدد مرات التوليد	حجم العينة المرجعية	حجم العينة المستهدفة
1	منتظم	300	1000	1000
2	غير منتظم	300	1000	1000

- **رابعاً:** تم التحقق من افتراضات نظرية الاستجابة للفقرة للبيانات المتمثلة في استجابات الأفراد، ولصعوبة التحقق من استجابات الأفراد جميعها والبالغة (600) مرة، تم التحقق من عينة عشوائية اختيرت من الاستجابات، وأظهرت النتائج تحقق افتراضات أحادية البعد، الاستقلال الموضعي، ومنحنى خصائص الفقرة، في حين لم يتم التحقق من افتراض التحرر من السرعة بسبب أن البيانات افتراضية، ولا يمكن قياس زمن الاستجابة.
- **خامساً:** تم إجراء تحليل عاملي توكيدي متعدد المجموعات على الحالة الافتراضية التي تتمثل في استجابات أفراد افتراضيين استجابوا على فقرات تخلو من الأداء التفاضلي؛ للحصول على النموذج الأصلي للمقياس والذي يمثل عاملا واحدا مع (50) فقرة باستخدام برمجية (AMOS).

وتم استخراج مؤشرات المطابقة للنموذج للتحقق من المطابقة كما يبينها الجدول (3).

الجدول (3) :مؤشرات المطابقة للنموذج الأصلي.

المؤشر	قيمة للنموذج الحالي	محك المطابقة
χ^2	1196.963	الأصغر
Df	1175	–
χ^2/df	1.019	أقل من 2
RMSEA	0.006	أقل من 0.08

من الجدول (3) لمؤشرات المطابقة يتبين أن النموذج مطابق للبيانات؛ إذ بلغت قيمة مؤشر نسبة مربع كاي إلى درجة الحرية (χ^2/df) (1.019) وهي قيمة أقل من (2) حسب المعايير التي أوردها أولمان (Ullman, 2001)، وبلغت قيمة جذر متوسط مربعات الخطأ (RMSEA) (0.006)، وهي قيمة أقل من (0.08) حسب المعايير التي أوردها كلاين (Kline, 2015)، وبذلك تم اعتماد هذا النموذج ليمثل البناء العاملي للمقاييس التي يتم البحث في أثر الأداء التفاضلي على جودة مطابقته ومدى الوثوق في مؤشراته.

متغيرات الدراسة

المتغير المستقل: ويمثل نوع الأداء التفاضلي للفقرات: ولهُ مستويان (منتظم، وغير منتظم)، والمتغير التابع :وهو البناء العاملي للمقاييس، ويتمثل في مؤشرات المطابقة عبر مستويات اللاتغاير في البناء العاملي.

المعالجات الإحصائية

تم استخدام البرمجيات (AMOS, WINGEN, R, SPSS, LDID, BILOG) لتوليد البيانات التي تحاكي الظروف التجريبية، وللتحقق من افتراضات نظرية استجابة الفقرة، والتحقق من ملاءمة النموذج التوكيدي، واستخراج مؤشرات المطابقة، وللإجابة عن أسئلة الدراسة تم استخدام التحليل العاملي التوكيدي للتحقق من المطابقة (الشكلية، والمترية، والقياس، والبواقي) للبناء العاملي للمقياس في ضوء الظروف التجريبية المختلفة المتمثلة بنوع الأداء التفاضلي (منتظم، وغير المنتظم) للفقرات، واستخراج مؤشرات المطابقة للنموذج التوكيدي (RMSEA, $\frac{\chi^2}{df}$)، وللكشف عن دلالة تلك الفروق تم استخدام اختبار (t) للعينات المستقلة.

نتائج الدراسة ومناقشتها

النتائج المتعلقة بسؤال الدراسة الأول ومناقشتها: "هل يوجد أثر لنوع الأداء التفاضلي للفقرات على التطابق الشكلي للبناء العاملي للمقياس؟ "وللإجابة عن هذا السؤال، تم إجراء تحليل عاملي توكيدي متعدد المجموعات (MGCFA) على استجابات الأفراد التي تم توليدها على فقرات المقاييس الافتراضية بوجود نوعين من الأداء التفاضلي (منتظم، وغير المنتظم)، وقد تم فحص أثر نوع الأداء التفاضلي على مؤشرات المطابقة الشكلية من خلال استخراج الأوساط الحسابية لكل من تلك المؤشرات باختلاف نوع الأداء التفاضلي كما هو مبين في الجدول رقم (4).

الجدول(4): الأوساط الحسابية والانحرافات المعيارية لمؤشرات المطابقة الشكلية حسب نوع الأداء التفاضلي.

غير منتظم		منتظم		
الإنحراف المعياري	الوسط الحسابي	الانحراف المعياري	الوسط الحسابي	المؤشر
0.035	1.081	0.033	1.072	$\frac{\chi^2}{df}$
0.002	0.009	0.002	0.008	RMSEA

وبملاحظة قيم الأوساط الحسابية في الجدول (4) يتبين:

- تحقق المطابقة الشكلية للنموذج بين المجموعتين المرجعية والمستهدفة باختلاف نوع الأداء التفاضلي (منتظم، وغير المنتظم) حيث بلغت قيمة مؤشر مربع (كاي) المعياري (1.072)، (1.081) على الترتيب لنوعي الأداء التفاضلي (منتظم، وغير المنتظم)، وهي أقل من (2)، وأن قيم مؤشر جذر وسط مربعات الخطأ (RMSEA) لم تتجاوز القيمة (0.08) حيث بلغت قيمته (0.008)، (0.009) على الترتيب لنوعي الأداء التفاضلي (منتظم، وغير المنتظم).
- ويلاحظ من قيم الأوساط الحسابية وجود فروق ظاهرية في قيم مؤشري المطابقة الشكلية مربع (كاي) المعياري $\frac{\chi^2}{df}$، وجذر وسط مربعات الخطأ التقريبي(RMSEA) باختلاف نوع الأداء التفاضلي (منتظم، وغير المنتظم)، وللكشف عن دلالة تلك الفروق تم استخدام اختبار (t) للعينات المستقلة في الجدول(5).

الجدول (5): اختبار(t) للعينات المستقلة لفحص الفروق في مؤشري مربع كاي المعياري ووسط الخطأ التقريبي للمطابقة الشكلية.

المؤشر	قيمة t	درجة الحرية	القيمة الاحتمالية للخطأ
$\frac{\chi^2}{df}$	-3.374	598	.001
RMSEA	-3.351	598	.001

وأظهرت نتائج اختبار (t) في الجدول (5) وجود دلالة للفروق في مؤشر مربع (كاي) المعياري العائدة لنوع الأداء التفاضلي، حيث بلغت قيمة اختبار(t) (3.374−)، والقيمة الاحتمالية للخطأ للاختبار (P-value) (0.001)، وهي قيمة أقل من مستوى الدلالة (α=0.05)؛ مما يشير إلى وجود اختلاف في مؤشر مربع (كاي) المعياري للمطابقة الشكلية باختلاف نوع الأداء التفاضلي (منتظم، وغير المنتظم)، ولصالح الأداء التفاضلي غير المنتظم، وأظهرت النتائج وجود دلالة للفروق في مؤشر وسط مربعات الخطأ التقريبي العائدة لنوع الأداء التفاضلي (منتظم، وغير المنتظم)، حيث بلغت قيمة اختبار(t) (3.351 −)، والقيمة الاحتمالية للخطأ للاختبار (P-value) (.001)، وهي قيمة أقل من مستوى الدلالة (α=.05)؛ مما يشير لوجود اختلاف في مؤشر وسط مربعات الخطأ التقريبي للمطابقة الشكلية باختلاف نوع الأداء التفاضلي (منتظم، وغير المنتظم)، ولصالح الأداء التفاضلي غير المنتظم.

وقد تبين من نتائج السؤال الأول وجود مطابقة شكلية عبر كل الظروف التجريبية للدراسة، ولكنها تختلف باختلاف نوع الأداء التفاضلي، ولصالح مستوى الأداء التفاضلي المنتظم، ويمكن تفسير ذلك من خلال اعتماد المطابقة الشكلية على مدى الارتباط بين فقرات المقياس، وذلك للتعبير عن السمة الكامنة والتي تتمثل في عامل واحد هنا، ولذلك ظهرت الفروق بين مستويي الأداء التفاضلي، حيث إن مستوى الأداء التفاضلي المنتظم يتم محاكاته من خلال زيادة صعوبة الفقرة، والذي ينعكس على شكل إزاحة في احتمالية الاستجابة على الفقرة عبر مستويات القدرة، ولكنها تكون إيجابية عبر متصل القدرة، وفي حين يعمل مستوى الأداء التفاضلي غير المنتظم على تغيير شكل منحنى خصائص الفقرة لتتشوه الفروق في احتمالية الاستجابة واتجاهاتها لتكون موجبة عند بعض مستويات القدرة وسالبة عند بعضها الآخر، وهذا يؤثر على ارتباط الفقرة بباقي فقرات المقياس بشكل سلبي وأكبر من تأثير الأداء التفاضلي المنتظم، ومن هنا وجدت الفروق بين نوعي الأداء التفاضلي في المطابقة الشكلية، وهذا ينطبق على مؤشرات المطابقة الشكلية المطلقة مثل ($\frac{\chi^2}{\mathrm{df}}$ ،$RMSEA$).

- **النتائج المتعلقة بسؤال الدراسة الثاني ومناقشتها: "هل يوجد أثر لنوع الأداء التفاضلي على التطابق المتري للبناء العاملي للمقياس؟**

وللإجابة عن هذا السؤال، تم إجراء التحليل العاملي التوكيدي متعدد المجموعات (MGCFA) على استجابات الأفراد التي جرى توليدها على المقاييس الافتراضية بوجود نوعين من الأداء التفاضلي (منتظم، وغير المنتظم)، وقد تم فحص أثر نوع الأداء التفاضلي (منتظم، وغير المنتظم) على مؤشرات المطابقة المترية من خلال استخراج الأوساط الحسابية لكل من تلك المؤشرات باختلاف نوع الأداء التفاضلي كما هو مبين في الجدول (6).

الجدول(6): الأوساط الحسابية والانحرافات المعيارية لمؤشرات المطابقة المترية حسب نوع الأداء *التفاضلي*.

	منتظم		غير منتظم	
المؤشر	**الوسط الحسابي**	**الانحراف المعياري**	**الوسط الحسابي**	**الانحراف المعياري**
$\frac{\chi^2}{df}$	1.120	.038	1.185	.043
RMSEA	.011	.002	014.	.002

وبملاحظة قيم الأوساط الحسابية في الجدول (6) يتبين:

- تحقق المطابقة المترية للنموذج بين المجموعتين المرجعية والمستهدفة باختلاف نوع الأداء التفاضلي (منتظم، وغير المنتظم)، حيث بلغت قيم مؤشر مربع (كاي) المعياري (1.120)، (1.185) على الترتيب لنوعي الأداء التفاضلي (منتظم، وغير المنتظم)، وكانت أقل من (2) لجميع الظروف التجريبية، وأن قيم مؤشر جذر متوسط مربعات الخطأ لم تتجاوز القيمة (.08)، حيث بلغت قيمته (.011)، (.014) على الترتيب لنوعي الأداء التفاضلي (منتظم، وغير المنتظم).
- ويلاحظ من قيم الأوساط الحسابية وجود فروق ظاهرية في قيم مؤشري المطابقة المترية مربع (كاي) المعياري ($\frac{\chi^2}{df}$)، وجذر وسط مربعات الخطأ التقريبي (RMSEA) باختلاف نوع الأداء التفاضلي (منتظم، وغير المنتظم)، وللكشف عن دلالة تلك الفروق تم استخدام اختبار(t) للعينات المستقلة المبين في الجدول رقم(7).

الجدول (7) اختبار(t) للعينات المستقلة الفروق في مؤشري مربع كاي المعياري ووسط التقريب للمطابقة المترية العائدة لنوع الأداء التفاضلي.

المؤشر	**قيمة t**	**درجة الحرية**	**القيمة الاحتمالية للخطأ**
$\frac{\chi^2}{df}$	−19.679	586.620	.000
(RMSEA)	−19.021	598	.000

وأظهرت نتائج اختبار (t) الواردة في الجدول رقم (7) وجود دلالة للفروق في مؤشر مربع (كاي) المعياري العائدة لنوع الأداء التفاضلي (منتظم، وغير المنتظم)، فقد بلغت قيمة اختبار(t) (19.679−)، والقيمة الاحتمالية للخطأ للاختبار (P-value) (0.000) وهي قيمة أقل من مستوى الدلالة (α=0.05)؛ مما يشير لوجود اختلاف في مؤشر مربع (كاي) المعياري للمطابقة المترية باختلاف نوع الأداء التفاضلي (منتظم، وغير المنتظم)، ولصالح الأداء التفاضلي غير المنتظم، وأظهرت النتائج وجود دلالة للفروق في مؤشر وسط مربعات الخطأ التقريبي العائدة لنوع الأداء التفاضلي (منتظم، وغير المنتظم؛ إذ بلغت قيمة اختبار (t) (19.021−)، والقيمة الاحتمالية للخطأ للاختبار (P-value) (.000) وهي قيمة أقل من مستوى الدلالة (α= .05)؛ مما يشير لوجود اختلاف في مؤشر وسط مربعات الخطأ التقريبي للمطابقة المترية باختلاف نوع الأداء التفاضلي (منتظم، وغير المنتظم)، ولصالح الأداء التفاضلي غير المنتظم.

ويتبين من نتائج السؤال الثاني وجود مطابقة مترية عبر كل الظروف التجريبية، ولكنها تختلف باختلاف نوع الأداء التفاضلي، ولصالح مستوى الأداء التفاضلي المنتظم هذا بالنسبة لمؤشرات المطابقة المطلقة ($RMSEA$ ،$\frac{\chi^2}{df}$).

- **النتائج المتعلقة بسؤال الدراسة الثالث ومناقشتها:" هل يوجد أثر لنوع الأداء التفاضلي على تطابق قياس البناء العاملي للمقياس؟"**

وللإجابة عن هذا السؤال، تم إجراء التحليل العاملي التوكيدي متعدد المجموعات (MGCFA) على استجابات الأفراد التي تم توليدها على المقاييس الافتراضية بوجود نوعين من الأداء التفاضلي (منتظم، غير المنتظم)، وقد تم فحص أثر نوع الأداء التفاضلي (منتظم، وغير المنتظم) على مؤشرات مطابقة القياس باستخراج الأوساط الحسابية لكل من تلك المؤشرات باختلاف الأداء التفاضلي (منتظم، وغير المنتظم) في الجدول رقم (8):

وبملاحظة قيم الأوساط الحسابية في الجدول رقم(8) يتبين:

الجدول (8) الأوساط الحسابية والإنحرافات المعيارية لمؤشرات مطابقة القياس حسب نوع الأداء التفاضلي.

المؤشر	المنتظم		غير المنتظم	
	الوسط الحسابي	الإنحراف المعياري	الوسط الحسابي	الإنحراف المعياري
$\frac{\chi^2}{df}$	3.035	.361	1.195	.046
RMSEA	.045	.004	.014	.002

تحقق مطابقة القياس للنموذج بين المجموعتين المرجعية والمستهدفة باختلاف نوع الأداء التفاضلي (منتظم، وغير المنتظم)، حيث بلغت قيم مؤشر مربع (كاي) المعياري (3.035)، (1.195) على الترتيب لنوعي الأداء التغاضلي (منتظم، وغير المنتظم)، وكانت قريبة من (2) لجميع الظروف التجريبية، وأن قيم مؤشر جذر متوسط مربعات الخطأ التقريبي لم تتجاوز القيمة (0.08)، حيث بلغت قيمته (.045)، (.014) على التوالي لنوعي الأداء التفاضلي (منتظم، وغير المنتظم).

- ويلاحظ من قيم الأوساط الحسابية وجود فروق ظاهرية في قيم مؤشري مطابقة القياس مربع (كاي) المعياري($\frac{\chi^2}{df}$)، وجذر وسط مربعات الخطأ التقريبي(RMSEA) باختلاف نوع الأداء التفاضلي (منتظم، وغير المنتظم)، وللكشف عن دلالة تلك الفروق تم استخدام اختبار(t) للعينات المستقلة في الجدول رقم (9).

الجدول (9):اختبار(t) للعينات المستقلة لفحص الفروق في مؤشري مربع كاي المعياري ووسط الخطأ التقريبي العائدة لنوع الأداء التفاضلي.

المؤشر	قيمة t	درجة الحرية	القيمة الاحتمالية للخطأ
$\frac{\chi^2}{df}$	87.662	308.536	.000
(RMSEA)	124.348	394.370	.000

وأظهرت نتائج اختبار (t) الواردة في الجدول رقم (9) وجود دلالة للفروق في مؤشر مربع (كاي) المعياري العائدة لنوع الأداء التفاضلي (منتظم، وغير المنتظم)، فقد بلغت قيمة اختبار(t) (87.662)، والقيمة الاحتمالية للخطأ للاختبار(P-value) (.000)، وهي قيمة أقل من مستوى الدلالة (α=.05)؛ مما يشير لوجود اختلاف في مؤشر مربع (كاي) المعياري للمطابقة المترية باختلاف نوع الأداء التفاضلي (منتظم، وغير المنتظم)، ولصالح الأداء التفاضلي المنتظم، وأظهرت النتائج وجود دلالة للفروق في مؤشر وسط مربعات الخطأ التقريبي العائدة لنوع الأداء التفاضلي (منتظم، وغير المنتظم)، حيث بلغت قيمة اختبار(t) (124.346)، والقيمة الاحتمالية للخطأ للاختبار (P-value) (.000) وهي قيمة أقل من مستوى الدلالة (α=0.05)؛ مما يشير إلى وجود اختلاف في مؤشر وسط مربعات الخطأ التقريبي للمطابقة المترية باختلاف نوع الأداء التفاضلي (منتظم، وغير المنتظم) ولصالح الأداء التفاضلي المنتظم.

ويتضح من نتائج السؤال الثالث وجود مطابقة قياس عبر كل الظروف التجريبية، ولكنها تختلف باختلاف نوع الأداء التفاضلي، ولصالح مستوى الأداء التفاضلي غير المنتظم هذا بالنسبة لمؤشرات المطابقة المطلقة ($RMSEA$ ،$\frac{\chi^2}{df}$).

- **النتائج المتعلقة بسؤال الدراسة الرابع ومناقشتها:"هل يوجد أثر لنوع الأداء التفاضلي على تطابق البواقي للبناء العاملي للمقياس؟"**

وللإجابة عن هذا السؤال، تم إجراء التحليل العاملي التوكيدي متعدد المجموعات (MGCFA) على استجابات الأفراد التي تم توليدها على المقاييس الافتراضية بوجود نوع الأداء التفاضلي (منتظم، غير المنتظم)، وقد تم فحص أثر نوع الأداء التفاضلي على مؤشرات مطابقة البواقي من خلال استخراج الأوساط الحسابية لكل من تلك المؤشرات باختلاف نوع الأداء التفاضلي كما هو مبين في الجدول رقم (10).

الجدول(10): الأوساط الحسابية والإنحرافات المعيارية لمؤشرات مطابقة القياس حسب نوع الأداء التفاضلي.

	المنتظم		غير المنتظم	
المؤشر	الوسط الحسابي	الانحراف المعياري	الوسط الحسابي	الانحراف المعياري
$\frac{\chi^2}{df}$	3.620	.490	1.120	.050
RMSEA	.051	.005	.014	.002

وبملاحظة قيم الأوساط الحسابية في الجدول (10) يتبين:

- تحقق مطابقة البواقي للنموذج بين المجموعتين المرجعية والمستهدفة باختلاف نوع الأداء التفاضلي (منتظم، وغير المنتظم)، حيث بلغت قيمة مؤشر مربع (كاي) المعياري (3.620)، (1.120) لنوعي الأداء التفاضلي (منتظم، وغير المنتظم) على الترتيب، وهي قريبة من (2)، وأن قيم مؤشر جذر وسط مربعات الخطأ (RMSEA) لم تتجاوز القيمة (0.08) حيث بلغت قيمته (0.051)، (0.014) على الترتيب لنوعي الأداء التفاضلي (منتظم، وغير المنتظم).
- ويلاحظ من قيم الأوساط الحسابية وجود فروق ظاهرية في قيم مؤشري مطابقة البواقي مربع (كاي) المعياري ($\frac{\chi^2}{df}$)، وجذر وسط مربعات الخطأ التقريبي (RMSEA) باختلاف نوع الأداء التفاضلي (منتظم، وغير المنتظم)، وللكشف عن دلالة تلك الفروق تم استخدام اختبار(t) للعينات المستقلة في الجدول رقم (11).

الجدول (11) :اختبار(t) للعينات المستقلة للفروق في مؤشري مربع كاي المعياري ووسط الخطأ التقريبي لمطابقة البواقي.

المؤشر	قيمة t	درجة الحرية	القيمة الاحتمالية للخطأ
$\frac{\chi^2}{df}$	85.085	305.216	.000
(RMSEA)	123.826	383.593	.000

وأظهرت نتائج اختبار (t) الواردة في الجدول (11) وجود دلالة للفروق في مؤشر مربع (كاي) المعياري العائدة لنوع الأداء التفاضلي (منتظم، وغير المنتظم)، فقد بلغت قيمة اختبار(t) (85.085)، والقيمة الاحتمالية للخطأ للاختبار (P-value) (.000)، وهي قيمة أقل من مستوى الدلالة (α=.05)؛ مما يشير لوجود اختلاف في مؤشر مربع (كاي) المعياري لمطابقة البواقي باختلاف نوع الأداء التفاضلي (منتظم، وغير المنتظم)، ولصالح الأداء التفاضلي المنتظم، وأظهرت النتائج وجود دلالة للفروق في مؤشر وسط مربعات الخطأ التقريبي العائدة لنوع الأداء التفاضلي، حيث بلغت قيمة اختبار(t) (123.826)، والقيمة الاحتمالية للخطأ للاختبار (P-value) (.000)، وهي قيمة أقل من مستوى الدلالة (α=.05)؛ مما يشير لوجود اختلاف في مؤشر وسط مربعات الخطأ التقريبي لمطابقة البواقي باختلاف نوع الأداء التفاضلي، ولصالح الأداء التفاضلي المنتظم.

ويتبين من نتائج السؤال الرابع وجود مطابقة البواقي عبر كل الظروف التجريبية، ولكنها تختلف باختلاف نوع الأداء التفاضلي، ولصالح مستوى الأداء التفاضلي غير المنتظم هذا بالنسبة لمؤشرات المطابقة المطلقة ($RMSEA$ ،$\frac{\chi^2}{df}$).

التوصيات

هدفت هذه الدراسة إلى التحقق من البناء العاملي للمقياس في ضوء مستويات مختلفة لنوع الأداء اللتفاضلي، وتم التوصُل إلى التوصيات الآتية:

توصيات نظرية:

- إجراء الدراسة في ضوء نوع النموذج (اللوجستي) بوجود نسب مختلفة من الأداء التفاضلي المنتظم.
- إجراء الدراسة في ضوء أحجام مختلفة من العينات، وأحجام مختلفة من الأداء التفاضلي.
- إجراء الدراسة على مقاييس متعددة الأبعاد.

توصيات عملية:

- التخلص من الفقرات ذات الأداء التفاضلي عند بناء المقاييس في المراحل الأولية، وقبل التطبيق الرئيس.
- تطبيق المقاييس على عينات مختلفة، واعتبار التحقق من الأداء التفاضلي للفقرات كخاصية (سيكومترية) يضاف إلى ذلك الصدق، والثبات، ومعالم الفقرات.
- إجراء دراسات واقعية على مقاييس عالمية للتحقق من البناء العاملي لها في ضوء نسبة الفقرات ذات الأداء التفاضلي ونوعه.
- التنويع في مؤشرات المطابقة المنوي التحقق منها، ويُنصح بالمؤشرات المطلقة عند وجود أداء تفاضلي منتظم، ومؤشرات المقارنة عند وجود أداء تفاضلي غير المنتظم.
- اعتبار مؤشرات المطابقة مؤشرات لاحتمالية وجود أداء تفاضلي في فقرات المقياس.

المصادر والمراجع العربية:

- اجباره، محمد. (2015). *أثر نسبة الفقرات ذات الأداء التفاضلي على الخطأ من النوع الأول وقوة الاختبار في الكشف عنه باستخدام طريقتي الإنحدار اللوجستي ونسبة الأرجحية العظمى*. أطروحة دكتوراة غير منشورة، جامعة اليرموك.
- تيغزة، محمد.(2012). *التحليل العاملي الاستكشافي والتوكيدي عمان:* دار المسيرة.
- الحربي، آلاء والمدني، فاطمة. (2021). الكشف عن أثر نسبة المفردات ذات الأداء التفاضلي على قوة وفاعلية اختبار منحنى خصائص المفردة في النموذج ثلاثي المعلم. *مجلة العلوم التربوية والنفسية، 5*(14)، 46-62.
- أبو حطب، فؤاد وصادق، آمال. (1991). *مناهج البحث وطرق التحليل الإحصائي في العلوم النفسية والاجتماعية والتربوية*. مكتبة الأنجلو المصرية.
- السعيد، أريج. (2021). *فعالية طريقة المواءمة في الكشف عن اللاتغير في أداء طلبة الصف الثالث الأساسي على الاختبار التقييمي لمهارات الرياضيات عبر مديريات التربية والتعليم الأردنية*. أطروحة دكتوراه غير منشورة، جامعة اليرموك، الأردن.
- السيد، سعيد. (2022). *مقياس العوامل الستة المنقح للشخصية (HEXAC0-PI-R) للبيئة الأردنية باستخدام نظرية الاستجابة للفقرة والتحقق من بنيته العاملية*. .أطروحة دكتوراة غير منشورة، جامعة اليرموك.
- الشريفين، نضال .(2018). أثر نوع الأداء التفاضلي للفقرات على الخصائص السيكومترية للفقرات والاختبار وفق النماذج المعلمية والنماذج اللامعلمية لنظرية الاستجابة للفقرة *دراسات، العلوم التربوية، 45* (4)، 605-632.
- الشوبكي، فاطمة. (2021). *البناء العاملي واللاتغير في القياس لمقياس جودة الحياة بدلالة طول المقياس*. أطروحة دكتوراة غير منشورة، جامعة اليرموك، الأردن.
- عبد العزيز، بوسالم وأحمد، كريش. (2016). الأداء التفاضلي للبند في الاختبارات النفسية، مصادره وتفسيره. *مجلة العلوم الاجتماعي ة، 10* (2)، 127-142.
- عدس، عبد الرحمن والكيلاني، عبد الله(1993). *برنامج التربية للقياس والتقويم في التعلم والتعليم*.
- أبو علام، رجاء والمرابحة، عامر.(2015). أثر نسبة المفردات ذات الأداء التفاضلي وحجم العينة على قوة وفاعلية اختبار تحيز المفردة المتزامن (دراسة محاكاة). *المجلة العربية للعلوم الاجتماعية، 2*(7) ، 139-165.
- عودة، أحمد(2014)، *القياس والتقويم في العملية التدريسية*. دار الأمل.
- النوافلة، علي. (2013). *أثر نسبة الفقرات ذات الأداء التفاضلي ومستواه في تقدير معالم فقرات الاختبار وقدرات الأفراد وفق نموذج الاستجابة ثلاثي المعلمة*. أطروحة دكتوراة غير منشوره، جامعة اليرموك، الأردن.

References

- Abdel-Alaziz, B., &Ahmed, K. (2016). Differential item functioning its resources and interpretations. (in Arabic), Journal of Social Sciences, 10(2),127-142.
- Abo-Alam, R., & Al-Marabhah, A. (2015). The effect differential item functioning ratio & sample size on the strength & power of SIBTEST. (in Arabic), Arabian Journal for Social Sciences, 2(7), 139-165.
- Abo-Hatab, F., & Sadeq, A. (1991). Research Methodology &Statistical analysis Methods for Psychological and Educational Sciences. (in Arabic), Englo-Eygptian for Publishing and Distribution.
- *Ackerman, T. (1992). A didactic explanation of item bias impact, and item validity from a multidimensional perspective. Journal of Educational Measurement, 29, 67-91.*
- Addas, A., &Al-Kalany, A. (1993). Measurement and evaluation program for teaching and learning. (in Arabic)
- *Albright, J., & Park, H. (2009). Confirmatory factor analysis using AMOS, LISREL, Mpluse, SAS/STAT CALIS. Working Paper. The University Information Technology Services (UITS) Center for Statistical and Mathematical Computing, Indiana University.*
- Al-Harbi, A., & Madany, F. (2021). Detecting the effect of differential item functioning on ICC test strength and power for third logistic IRT model. (in Arabic), Journal of Psychological & Educational Sciences,5(14), 45-62.
- Al-nawafleh, A. (2013). The effect of differential item functioning percentage and its level on the estimations of items and individuals parameters (Un published doctoral dissertation), (in Arabic),Yarmouk University, Jordan.
- Al – Saeed, A. (2020). The Effectiveness of the Alignment Method in Detecting the Invariance in the Performance of the Third Grade Students on the Evaluation Test of Mathematics Skills across the Directorates of Jordanian Education, (Un published doctoral dissertation), (in Arabic), Yarmouk university, Jordan.
- Alseid, S. (2022). Standardizing the Revised Scale of the Personality Six Factors (HEXACO-PI-R) for Jordanian Environment using Item Response Theory (IRT) and Verifying its Factorial Structure, (Un published doctoral dissertation), (in Arabic), Yarmouk University, Jordan.
- Al-Sharefeen, N. (2018). The effect of differential item functioning type on Psychological characteristics for item and test in light of IRT parametric and non-parametric models. (in Arabic), Journal of Educational Sciences,45(4), 605-632.
- Al – Shobaki, F. (2021). Factor Structure and invariance of Measurement of the Quality-of-Life Scale According to Scale Length, (Un-published doctoral dissertation), (in Arabic), Yarmouk University, Jordan.
- *Angoff, W. (1993).Perspectives on differential item functioning methodology. In P.W.*
- *Braun, H. (2013). Differential item performance and the procedure in Mantel-Haenszel test Validity. Routledge.*
- *Brown, T. (2006). Confirmatory factor analysis for applied research. Guilford Press.*
- *Byrne, M. (2001). Structural equation modeling with AMOS, EQS, and LISREL: comparative approach for the factor validity of a measuring instrument. International. Journal of Testing, 1, 55-86.*
- *Byrne, B., & Watkins, D. (2003). The issue of measurement invariance revisited. Journal of Cross- Cultural Psychology,34(2),155-157.*
- *Camilli, G. & Shepard, L. (1994). Methods for identifying bias test item. Sage publication.*
- *Capalleri, JC., Lundy, Jason & Hays, RD. (2014). Overview of classical test theory and item response theory for the quantitative assessment of items in developing patient-reported outcomes measures. Clin Ther, 36(5), 648-662.*
- *Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. Structural equation modeling: a multidisciplinary journal, 14(3), 464-504.*
- *Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit Index's for testing measurement invariance. Structural equation modeling, 9(2), 233-255.*
- *Crocker, L. & Algina, J. (1986). Introduction to Classical and Modern Test Theory. University of Florida.*
- *DeCoster, J. (1998). Overview of Factor Analysis. Tuscaloosa. Retrieved from https://www.researchgate.net/file.PostFileLoader.html?id=54464aefd2fd6405388b45bb&assetKey=AS%3A273627033866242%401442249176796.*
- *Drasgow, F. (1982). Biased test items and differential validity. Psychological Bulletin,92,526-531.*
- Ejbara, M. (2015). The Effect of Percentage of items with DIF on Type I Error & Test Power in detecting, Using Logistic Regression & Likelihood Ratio Method, (Un published doctoral dissertation), (in Arabic), Yarmouk university, Jordan.
- *Eysenck, H. J. (2013). The structure of human personality (psychology Revivals). Routledge.*
- *Frick, P., Barry, C., & Kamphaus, R. (2020). Clinical Assessment of Child and Adolescernt Personality and Behavior. Springer.*

- *Gabriel E. L. (2011). Detection and Classification of DIF Types Using Parametric and Nonparametric Methods: A comparison of the IRT- Likelihood Ratio Test, Crossing-SIBTEST, and Logistic Regression Procedures Unpublished dissertation. University of South Florida.*
- *Gruijter.J & Kamp.V. (2008). Statistical Test Theory for the Behavioral Sciences. Tylor and Frances Group L.L.E.*
- *Hair, F., Black, C., Babin, J., Anderson, E. & Tatham, L. (2006). Multivariate data analysis (6 ed.): Prentice Hall.*
- *Hambleton, R, Swaminathan, H, & Rogerrs, H. (1991). Fundamentals of item response theory, Sage.*
- *Hambleton, R., Merenda, P., & Spielberger, C. (Eds.). (2004). Adapting educational and psychological tests for cross-cultural assessment. Psychology Press.*
- *Harrington, D. (2009). Confirmatory factor analysis. University Press.*
- *Huang, J., & Han, T. (2012). Revisiting differential item functioning; Implications for fairness investigation. International Journal of Education, 4(2), 74-86.*
- *Hu, L. & Bentler, P. (1999). Evaluating model fit .In R.h. Hoyle (Ed.), Structural equation modeling: Concepts, issues, and application (pp. 76- 99). Thous and Oaks, CA: Sage.*
- *Hult, G. T. M., D. J. KetchenJr, D. A. Griffith, B. R. Chabowski, M. K. Hamman, B. J. Dykes, W. A. Pollitte and S. T. Cavusgil (2008). 'An assessment of the measurement of performance in international business research', Journal of International Business Studies, 39, 1064– 1080.*
- *Joseph, F., Hair, J., & Marco, J. (2017). Advanced Issues in Partial Structural Equation Modling: Conceptual and methodological Norwell, MA: Kluwer Academic.*
- *Kalaycioglu, D. & Berberoglu, G. (2011). Differential item functioning analysis of the science and mathematics items in the university entrance examinations in turkey. Journal of Psycho Educational Assessment, 29 (5), 467-478.*
- *Kline, R. B. (2015). Principles and practice of structural equation modeling. Guilford Press.*
- *Lord, F. (1980). Application of item response theory to practical testing problems. Lawrence Erlbaum Associates, Publishers.*
- *Maller, S. (2003). Best practices in detecting bias in nonverbal tests. Springer.*
- *Michel ides, P. (2008). An illustration of a Mantel-Haenszel procedure to flag misbehaving common itemsin test equating. Practical Assessment Research and Evaluation, 13(7), 1-13.*
- *Millsap, R. E. (2011). Statistical approaches to measurement invariance. Routledge.*
- Odeh, A. (2014). Measurement and evaluation in teaching process, (in Arabic), Dar Al- for Publishing and Distribution, Irbid, Jordan.
- *Salubayba, T. (2013). Differential item functioning detection in reading comprehension test using Mantel-Haenszel, Item response theory, and logical data Analysis. The International Journal of Social Sciences, 14(1), 76-82.*
- *Sarmento, R. P., Costa, V. (2019). Confirmatory Factor Analysis--A Case study. Ar Xiv preprint arXiv: 1905 .05598.*
- *Schumacher, E., & Lomax, G. (2004). A beginner's guide to structural equation modeling. Psychology Press.*
- *Tan, X., & Jeral, M. (2005). Using local DIF analyses to assess group differences on multilingual examination. Applied Measurement in Education, 2,313-334.*
- Tegza, M. (2012). The exploratory and confirmatory factor analysis, Dar Al-Maserah for Publishing and Distribution, (in Arabic), Amman, Jordan.
- *Teresi, J. & Fleishman, J. (2007). Differential item functioning and healt h assessment. Quality of Life Research, 16 (1), 33–42.*
- *Ulman, J. B. (2001). Structural equation modeling. In B. G. Tabachnick & L. S. Fidell (2001). Using Multivariate Statistics (4th ed & pp 653- 771). Needham Heights, MA: Allyn & Bacon.*
- *Vandenberg, R. and Lance, C. (2000). A review and synthesis of the measurement invariance Literature: Suggestions, practices, and recommendations for organizational research. Organizational Research Methods 3(1), 4-69.*
- *Walker, C. (2011). What's the DIF? Why differential item functioning analyses are an important part of instrument development and validation. Journal of Psychoeducational Assessment, 29 (4), 364-376.*
- *Zhan, L. and Zumbo, B. (2009). Impact of Differential Item Functioning on Subsequent Statistical concluons Based on observed test Score Data, Psychological, 30(2), 243-370.*
- *Zumbo, D. (1999). A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and likert type item scores: Directorate of Human Resources Research and Evaluation, Department of National Defense.* Abo-Alam, R., & Al-Marabhah, A. (2015). The effect differential item functioning ratio & sample size on the strength & power of SIBTEST. (in Arabic), Arabian Journal for Social Sciences, 2(7), 139-165.