فاعلية طريقة المواءمة في الكشف عن اللاتغير في أداء طلبة الصف الثالث الأساسي على الاختبار التقييمي لمهارات الرياضيات عبر مديريات التربية والتعليم الأردنية

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

إعداد

أريج حسن عبدالله السعيد

إشراف

الأستاذ الدكتور نضال كمال الشريفين

التخصص – القياس والتقويم

للعام الدراسي 2021/2020

2021/2020 م

فاعلية طريقة المواءمة في الكشف عن اللاتغير في أداء طلبة الصف الثالث الأساسي على الاختبار التقييمي لمهارات الرياضيات عبر مديريات التربية والتعليم الأردنية

إعداد

أريج حسن عبدالله السعيد

ماجستير قياس وتقويم، جامعة اليرموك 2010م

قدمت هذه الرسالة استكمالاً لمتطلبات الحصول على درجة الدكتوراه في تخصص القياس والتقويم في جامعة اليرموك، اربد، الأردن

وافق عليها

الأستاذ الدكتور نضال كمال الشريفين ..... رئيسًا

أستاذ قياس وتقويم، جامعة اليرموك.

الأستاذ الدكتور أحمد عودة ..... عضوًا

أستاذ قياس وتقويم، جامعة اليرموك.

الأستاذ الدكتور محمود فيصل القرعان ..... عضوًا

أستاذ قياس وتقويم، جامعة اليرموك.

الدكتور عمر صالح بني ياسين ..... عضوًا

أستاذ مشارك قياس وتقويم، جامعة البلقاء.

2021/2020 م

إهداء

إلى من يعيش في ذاكرتي، إلى نبع الحنان، وميض الحب، إلى رفقاء دربي أبي الحبيب، وأمي الغالية، وإلى من وقف إلى جانبي يعضدني ويشاركني كفاح الحياة المرير بصبر وجلد لا ينقطع إلى أب أولادي، وإلى أولادي وأخوتي وأخواتي الذين يهتفون لكل خطوة أخطوها إلى الأمام، وإلى أستاذي الكبير الدكتور نضال كمال الشريفين، وإلى كل من رأى في مسيرتي العلمية بصيص أملٍ، وحافزًا لإكمال مسيرته دون تردد إلى كل هؤلاء بكل الحب أهدي رسالتي هذه

شكر وتقدير

الحمد لله الذي تتمُّ بنعمتِه الصالِحات؛ الذي هدانا لهذا، وما كنا لنهتدي، لولا أن هدانا الله، الذي منَّ عليَّ بإتمام هذه الدراسة، وصلوات اللهِ وتسليماتُهُ على رحمتِهِ المُهداةِ للعالمينَ؛ سيدنا وإمامنا وأُسوتِنا وحبيبنا محمَّدٍ، وعلى آله وصحبه الطاهرين ومن اتبعهم بإحسانٍ إلى يوم الدينِ. وبعد،

فأتقدم بداية لأستاذي الدكتور نضال كمال الشريفين بعظيم الشكر والتقدير؛ لِما اكتنفني به من جزيل حلمه وصبره، وإحاطته إيّاي بموفور علمه ورعايته، حتى غدت ملحوظاته وتوجيهاته النبراس الذي أنار لي الطريق، وذلّل لي الصعاب، وأضحت إرشاداه خير الأنيس، وخير الرفيق أثناء إنجازي هذا العمل، فجزاه الله عني خير الجزاء، وجعل ذلك كله في ميزان حسناته.

ويسعدني في هذا المقام أن أتوجه بالشكر والعرفان إلى السادة الأفاضل أعضاء لجنة المناقشة وهم: الأستاذ الدكتور نضال الشريفين مشرفًا ورئيسًا، والدكتور عمر بني ياسين عضوًا خارجيًا، والأستاذ الدكتور أحمد عودة عضوًا، والدكتور محمود القراعان عضوًا، لتفضلهم بمناقشة الرسالة وإبداء الملحوظات القيمة، التي سيكون لها أكبر الأثر في إثراء هذه الرسالة، وإغنائها بما هو مفيد ونافع.

والحمد لله أولًا وأخيرًا على ما يسَّر لي من سُبل البحث، وهيأ لي من ظروف مكنتني من الوصول إلى ما أصبو إليه، كما أرجوه سبحانه أن يكون هذا العمل خالصًا لوجهه الكريم، وأن يمُن على أمتنا بالنصر والسؤدد، إنه سميع مجيب.

الباحثة

جدول المحتويات

الموضوع	الصفحة
العنوان	أ
قرار لجنة المناقشة	ب
إهداء	ج
شكر وتقدير	د
جدول المحتويات	و
قائمة الجداول	ز
قائمة الأشكال	ح
قائمة الملاحق	ط
الملخص باللغة العربية	ي
الفصل الأول: خلفية الدراسة وأهميتها	1
المقدمة	1
اللاتغير في القياس (ثبات القياس) (Measurement Invariance) (MI)	15
طرق الكشف عن اللاتغير في القياس	18
اللاتغير في القياس من منظور التحليل العاملي التوكيدي (CFA)	18
اللاتغير في القياس من منظور نظرية استجابة الفقرة (IRT)	32
طريقة المواءمة (Alignment Method)	40
الاختبار التقييمي لمبحث الرياضيات للصف الثالث الأساسي	42
مشكلة الدراسة وأسئلتها	45
أهمية الدراسة	48
التعريفات الإصطلاحية والإجرائية	50
محددات الدراسة	53
الفصل الثاني: الدراسات السابقة	54
الدراسات التي تناولت اللاتغير في القياس (Measurement Invariance)	54
الدراسات التي تناولت طريقة المواءمة (Alignment Method)	71
التعقيب على الدراسات السابقة	79

الموضوع	الصفحة
الفصل الثالث: الطريقة والإجراءات	82
منهج الدراسة	82
مجتمع الدراسة	82
عينة الدراسة	82
أداة الدراسة	84
المعالجات الإحصائية	87
الفصل الرابع: نتائج الدراسة	93
النتائج المتعلقة بالسؤال الأول	93
النتائج المتعلقة بالسؤال الثاني	96
النتائج المتعلقة بالسؤال الثالث	102
النتائج المتعلقة بالسؤال الرابع	108
الفصل الخامس: مناقشة النتائج والتوصيات	121
مناقشة النتائج المتعلقة بالسؤال الأول والسؤال الثاني والسؤال الثالث	121
التوصيات	128
المراجع	129
المراجع العربية	129
المراجع الأجنبية	132
الملاحق	146
الملخص باللغة الإنجليزية	153

قائمة الجداول

الجدول	الصفحة
جدول 1: توزيع عينة الدراسة من طلبة الصف الثالث الأساسي على مديريات التربية والتعليم	83
جدول 2: توزيع فقرات الاختبار على مهارات مبحث الرياضيات السبع التي يقيسها الاختبار التقييمي	86
جدول 3: مؤشرات جودة المطابقة وعلامة القطع لكل منها	88
جدول 4: مؤشرات جودة المطابقة لنموذج مهارات الرياضيات	95
جدول 5: مؤشرات جودة مطابقة النموذج المُقترح للبيانات عبر مديريات التربية والتعليم	98
جدول 6: مؤشرات جودة مطابقة النموذج المُقترح للبيانات عبر مديريات التربية والتعليم مع فرض قيود اللاتغير المتري	99
جدول 7: الفرق بين مؤشرات جودة مطابقة النموذج (التكويني والمتري) المُقترح للبيانات عبر مديريات التربية والتعليم	100
جدول 8: مؤشرات جودة مطابقة النموذج المُقترح للبيانات عبر مديريات التربية والتعليم مع فرض قيود اللاتغير العددي	101
جدول 9: الفرق بين مؤشرات جودة مطابقة النموذج (المتري والعددي) المُقترح للبيانات عبر مديريات التربية والتعليم	102
جدول 10: مؤشرات جودة مطابقة النموذج المُقترح للبيانات عبر جنس الطالب	104
جدول 11: مؤشرات جودة مطابقة النموذج المُقترح للبيانات عبر جنس الطالب مع فرض قيود اللاتغير المتري	105
جدول 12: الفرق بين مؤشرات جودة مطابقة النموذج (التكويني والمتري) المُقترح للبيانات عبر جنس الطالب	106
جدول 13: مؤشرات جودة مطابقة النموذج المُقترح للبيانات عبر جنس الطالب مع فرض قيود اللاتغير العددي	107
جدول 14: الفرق بين مؤشرات جودة مطابقة النموذج (المتري والعددي) المُقترح للبيانات عبر جنس الطالب	107
جدول 15: نتائج اللاتغير (تغير) التقريبي في القياس لتقاطعات العوامل عبر مديريات التربية والتعليم الـ(45) حسب فقرات الاختبار الـ(25)	109

جدول 16: نتائج اللاتغير (تغير) التقريبي في القياس لتشبعات العوامل عبر مديريات التربية والتعليم الـ(45) حسب فقرات الاختبار الـ(25)	111
جدول 17: مقارنة وسط العامل العام (مهارات الرياضيات) عبر المجموعات عند مستوى الدلالة الإحصائية (5%) مرتبة ترتيبًا تنازليًا	113
جدول 18: مقارنة وسط العامل عبر المجموعات عند مستوى الدلالة الإحصائية (5%)	115
جدول 19: تحليل اللاتغير في القياس: تقاطعات الفقرة (1)	118

قائمة الأشكال

الشكل	الصفحة
شكل 1: نموذج قياس يتكون من عاملين كامنين، وست فقرات	19
شكل 2: النموذج العاملي المُقترح للاختبار التقييمي من الدرجة الثانية	94
شكل 3: النموذج العاملي المُقترح للاختبار التقييمي عبر مديريات التربية والتعليم الأردنية الـ(45)	97
شكل 4: النموذج العاملي المُقترح للاختبار التقييمي عبر جنس الطالب (ذكر، أنثى)	103

قائمة الملاحق

الموضوع	الصفحة
خطاب تسهيل مهمة من جامعة اليرموك	150
خطاب تسهيل مهمة لإدارة الامتحانات والاختبارات في وزارة التربية والتعليم	151
نموذج الاختبار التقييمي لمهارات الرياضيات للصف الثالث الأساسي للعام الدراسي 2017/2018	152

الملخص

السعيد، أريج حسن. فاعلية طريقة المواءمة في الكشف عن اللاتغير في أداء طلبة الصف الثالث الأساسي على الاختبار التقييمي لمهارات الرياضيات عبر مديريات التربية والتعليم الأردنية. (2020). (المُشرِف. أ. د. نضال كمال الشريفين).

هدفت هذه الدراسة لدراسة فاعلية طريقة المواءمة في الكشف عن اللاتغير في القياس (ثبات القياس) للاختبار التقييمي للصف الثالث الأساسي لمهارات الرياضيات في ضوء متغير مديريات التربية والتعليم وعددها (45) مديرية، ومتغير الجنس. تكونت عينة الدراسة من (73448) طالباً وطالِبَةً، (37272 طالبًا، و36176 طالِبَةً)، تم اختيارهم بشكل عشوائي من طلبة الصف الثالث الأساسي الملتحقين في المدارس للعام الدراسي (2018/2017)، والذين طُبق عليهم الاختبار التقييمي لمهارات الرياضيات.

ولتحقيق أهداف الدراسة تم استخدام المنهج الوصفي التحليلي في تحليل نتائج الطلبة على الاختبار التقييمي لمهارات الرياضيات لضبط نوعية التعليم المُعَد من قِبَل إدارة الامتحانات والاختبارات في وزارة التربية والتعليم الأردنية للعام الدراسي (2018/2017)، وبعد أن تمّ التحقق من دلالات صدق، وثبات الاختبار، تم تطبيقه على عينة الدراسة.

وبالتحديد تم تحليل استجابات (37272) طالبًا، و(36176) طالبة على (25) فقرة اختبارية من نوع اختيار من متعدد، وتم تحليل بيانات الفقرات باستخدام أسلوبي التحليل العاملي التوكيدي (CFA) من الدرجة الأولى والثانية، والتحليل العاملي متعدد المجموعة (MG−CFA) من خلال برنامجي الإحصاء (Mplus v8.4, Amos v26)، وللحكم على ملاءمة البناء العاملي للاختبار تم استخدام خمس مؤشرات حُسن المُطابقة وهي: (χ^2)، (RMSEA)، (SRMR)، (TLI)، (CFI)، وأظهرت النتائج ملاءمة النموذج المُقترح بمهاراته السبع لمحتوى الاختبار التقييمي

لمهارات الرياضيات للصف الثالث في الأردن، كما وأظهرت قيم مؤشرات الفرق وهي: (ΔRMSEA)، (ΔSRMR)، (ΔTLI)، (ΔCFI) تحقق خاصية اللاتغير في القياس (ثبات القياس) بمستوياته المختلفة (اللاتغير التكويني (ثبات الشكل)، واللاتغير المتري (الثبات المتري)، واللاتغير العددي (ثبات التدريج)) في فقرات الاختبار الـ(25) وفقًا لمتغيري مديريات التربية والتعليم الأردنية الـ(45) مديرية، وجنس الطالب.

كما تم استخدام طريقة المواءمة للكشف عن الفقرات المحققة لخاصية اللاتغير في القياس باستخدام مؤشر (R^2)، وأظهرت النتائج أن كافة الفقرات تحقق خاصية اللاتغير في القياس لمعلمة تقاطعات العامل، ومعظمها تُحقق خاصية اللاتغير في القياس لمعلمة تشبعات العامل ما عدا ست فقرات (5، 6، 13، 14، 17، 22) فهي بحاجة لمراجعة، كما تم مقارنة الأوساط الحسابية للعوامل الكامنة (مهارات الرياضيات السبع) المُقترحة في النموذج البنائي، وتباياناتهم عبر مديريات التربية والتعليم، مما نتج عنه ترتيب للمديريات ترتيبًا تنازليًا؛ حيث حصلت مديرية التعليم الخاص على المرتبة الأولى حسب أداء طلبتها على الاختبار التقييمي لمهارات الرياضيات للصف الثالث الأساسي، وفي المقابل حصلت مديرية الأغوار الجنوبية على المرتبة الأخيرة.

الكلمات المفتاحية: اللاتغير في القياس، ثبات القياس، الاختبار التقييمي لمهارات الرياضيات، مؤشرات حُسن المطابقة، التحليل العاملي التوكيدي، طريقة المواءمة، مديريات التربية والتعليم.

الفصل الأول

خلفية الدراسة وأهميتها

المقدمة

تُنفذ المؤسسات والهيئات التربوية العديد من الأبحاث والدراسات بشكل دوري ومستمر بهدف تحديد قدرات المفحوصين، ومهاراتهم المعرفية، والتعليمية ضمن عينة الدراسة، وكذلك تحديد خصائص الأدوات، والاختبارات التي تُطبق عليهم، ويتم عادةً اعتماد اختبارات مختلفة يتم إعدادها لخدمة أغراض الدراسة، وتُساعد في إتخاذ القرارات، كاختبارات ضبط نوعية التعليم، واختبارات القدرة، والاختبارات التقييمية، وغيرها، وبعض هذه الأبحاث والدراسات يعتمد على النظرية التقليدية في القياس (CTT)، في حين معظمها الآخر يعتمد على النظرية الحديثة في القياس (IRT)، وقد اهتمت الدراسة الحالية بتناول الاختبار التقييمي لمهارات الرياضيات لطلبة الصف الثالث الأساسي للعام الدراسي (2018/2017)، وذلك لعدد من الاعتبارات: أنه يُعد اختبارًا يُقاس من خلاله مهارات التعلم الخاصة بمبحث الرياضيات لدى طلبة الصف الثالث الأساسي الذي يعتبر نهاية المرحلة الأساسية الأولى، ويُساعد المعلمين في تقويم الطلبة، ومتابعتهم، وتحديد مسار تقدمهم، وكذلك يُقدم معلومات عن جودة التعليم، وبشكل سنوي، وتحديد جوانب القوة والضعف في أداء الطلبة لتوجيه المناهج الدراسية على أسس واقعية، ومعلومات صحيحة، ولتحديد التحديات التي تواجه عملية التعلم في هذه المرحلة المهمة لما لها من إسهام بشكل أساسي في تدعيم وتطوير "الحساب بفهم"، كما ويُزود بنتائج للطلبة بحجوم عينات كبيرة، كما أنه يُعد من الاختبارات المتوازنة، والمتضمنة أسس بناء الاختبار الجيد بضمان أن يطبق بمنتهى الجدية على الفئات الطلابية المستهدفة؛ مما يدفعهم لبذل جهدهم وطاقاتهم للإجابة عن الأسئلة بمنتهى الموضوعية، والدقة، والجدية مما يعكس تقدير قدرة حقيقية لهم.

كما تطورت طرق القياس التربوي والنفسي بشكل سريع مع الثورة المعرفية، وطرق تصميم الاختبارات والمقاييس وتحليل فقراتها، فرغم أن النظرية الكلاسيكية في القياس (Classical Test Theory) (CTT) اشتهرت في بناء مختلف أنواع الاختبارات والمقاييس، إلا أنها لم تتغلب على الكثير من المشكلات السيكومترية، وظهرت جوانب قصور لها في الكثير من التطبيقات العملية المستحدثة، مثل: إنشاء بنوك الأسئلة، وتصميم الاختبارات التكيفية، والكشف عن تحيز الفقرات، واختلاف معالم الفقرات باختلاف عينة الأفراد المستخدمة في استخراج هذه المعالم (علام، 2005).

كما أن النظرية الكلاسيكية تفترض تساوي تباين أخطاء القياس لجميع المفحوصين الذين يطبق عليهم الفقرات الاختبارية، وهذا غير مقبول من الناحية العملية، ولا توفر تقديرات دقيقة لقدرات الأفراد ذوي القدرات العليا والدنيا، كما أن الأداء على الفقرة أو مجموعة جزئية من الفقرات في النظرية الكلاسيكية لا يقدم معلومات كافية عن مستوى أداء الفرد، ولا يساعد على التنبؤ بأداء الأفراد على أي فقرة من فقرات الاختبار؛ مما أدى لصعوبة بناء بنوك الأسئلة، بالإضافة إلى أن مفهوم الثبات في النظرية الكلاسيكية يقوم على أساس إمكانية تطوير صور متوازية، يصعب تحقيقها في الواقع العملي، كما أن الخصائص السيكومترية للفقرات في أي اختبار غير متحررة من خصائص الأفراد، والعكس صحيح؛ لذا ظهرت النظرية الحديثة في القياس، أو ما يعرف بنظرية استجابة الفقرة (Item Response Theory) (IRT) أو نظرية السمات الكامنة (Latent Trait Theory) كنظرية بديلة، نظراً للمزايا العديدة التي تتمتع بها إضافة إلى الافتراضات القوية التي اعتمدت عليها، والتي يفترض من خلالها تلافي عيوب النظرية الكلاسيكية ومنها مشكلة العدالة في التوزيع، والتماثل، واللاتغير في القياس (الشريفين، 2019).

وسميت هذه النظرية بنظرية السمات الكامنة (Latent Trait Theory) لأنها تفترض أنه يمكن التنبؤ بأداء المفحوصين، أو تفسير أدائهم في اختبار نفسي أو تربوي معين، في ضوء خاصية أو خصائص مميزة لهذا الأداء تكون على شكل مجموعة من العوامل (Factors) تسمى بالسمات (Traits)، وتحاول تقدير درجات المفحوصين في هذه السمات التي يصعب ملاحظتها ملاحظة مباشرة؛ أي أنها كامنة (Latent)، لذلك يجب تقديرها (أو الاستدلال عليها) من أداء المفحوص الملاحظ على مجموعة من فقرات الاختبار.

وتعد نظرية استجابة الفقرة (Item Response Theory) (IRT) تطورًا حديثًا ومهمًا في القياس النفسي، إذ إن هذه النظرية تمكن المختص من الإجابة عن أي سؤال عن الفقرة أو الاختبار أو المفحوص لأنها تعتمد على أداء المفحوص على كل فقرة في الاختبار لإعطاء تقدير إحصائي لقدرة المفحوص التي يقيسها الاختبار، ومن أكثر التطبيقات شيوعًا لهذه النظرية ما يتعلق بفقرات الاختيار من متعدد في اختبارات القدرة والتحصيل، كما لها تطبيقات في أنواع أخرى من الفقرات لقدرتها على معالجة الكثير من قضايا القياس بشكل أكثر فاعلية من النظرية الكلاسيكية (الشريفين، والكيلاني، 2003).

وكما يرى كل من يانج وكاو (Yang & Kao, 2014) أن نظرية استجابة الفقرة (IRT) تمثل طريقة هامة لتقييم صدق أدوات القياس، وقد أولاها الباحثون الاهتمام لكونها استخدمت لتطوير المقاييس العالمية واسعة النطاق (International Large-Scale Assessments)، والاختبارات الموحدة مثل اختبارات الكفاءة الدراسية، وفي مجال التعليم فقد استخدمت في تطوير اختبارات قياس مستوى تحصيل الطلبة في مختلف المباحث؛ مثل الاختبار التقييمي لمهارات الرياضيات موضوع الدراسة.

ولقد بيّن هامبلتون وسوامينيثان (Hambleton & Swaminathan, 1985) أن مزايا نظرية السمات الكامنة (Latent Trait Theory) تتمثل بقدرتها على توفير تقديرات لقدرات الأفراد مستقلة عن عينة الفقرات التي تطبق عليهم؛ أي أن قدرات الأفراد متحررة من الفقرات (Item–Free) عند افتراض وجود مجموعة كبيرة نسبيًا من فقرات اختبار تقيس نفس السمة، كما أنها توفر تقديرات للخصائص السيكومترية للفقرات (مثل معاملات الصعوبة والتمييز) مستقلة عن عينة الأفراد التي استخدمت في تقدير هذه الخصائص؛ أي لا تتأثر معالم الفقرات بعينة الأفراد المختارة من المجتمع (Person–Free)، كما تقدم هذه النظرية إحصائي (Statistic) لتقدير درجة الدقة في قياس القدرة لكل فرد مثل الخطأ المعياري في التقدير، وقد يختلف هذا الإحصائي من فرد لآخر.

وتمثلت افتراضات نظرية استجابة الفقرة (IRT) كما وضحها هامبلتون وسوامنثان (Hambleton & Swaminathan, 1985) بالآتي:

1- أحادية البُعد (Unidimensionality): يشير مصطلح البُعد أو الأبعاد (Dimensionality) إلى عدد السمات الكامنة (العوامل) (Factors) المسؤولة عن أداء المفحوص في الاختبار، والسمة مفهوم يستخدم لوصف سلوك الفرد؛ وهي عبارة عن تجمع من السلوك المتداخل، والمترابط بصورة متكاملة؛ أي أنها ليست صفة منفردة، وإنما تُعد مفهومًا مجردًا غير ملموس، لذلك تفترض النظرية الحديثة (نظرية السمات الكامنة) وجود سمة (قدرة) واحدة كامنة أو عدة سمات (قدرات) كامنة تُفسر أداء المفحوص على الاختبار وتنبثق عنها نماذج مختلفة؛ ففي حالة افتراض وجود سمة واحدة فإن النماذج تكون نماذج أحادية البعد (Unidimensionality)، أما في حالة افتراض وجود أكثر من سمة واحدة تفسر أداء المفحوص فتنبثق عنها نماذج متعددة الأبعاد (Multidimensional Models)، ومن الجدير

بالذكر أنه من الممكن للاختبار أن يكون أحادي البعد في مجتمع ومتعدد الأبعاد في مجتمع آخر، ومثال على ذلك عندما يكون الاستيعاب القرائي عاملًا مهماً في حل المسائل الرياضية؛ ففي مجتمع يكون جميع أفراده قادرين على فهم المسائل الرياضية تكون السمة الوحيدة التي تؤثر في الأداء على الاختبار هي القدرة الرياضية، أما في مجتمع آخر جميع أفراده غير قادرين على استيعاب المسائل الرياضية في الاختبار، في هذه الحالة فإن مهارة الاستيعاب القرائي ستدخل كسمة ثانية تُؤثر على أداء الفرد على الاختبار، ومثال آخر على ذلك الاختبار المشبع بالعامل الثقافي، سيبدو هذا الاختبار أحادي البعد عندما يطبق على مجموعات ثقافية لديها الخلفيات الثقافية نفسها، ولكن عندما يطبق على مجموعات ثقافية مختلفة فسيظهر بأن هناك أكثر من عامل (سمة) يفسر الأداء على الاختبار، مما سيؤدي لاختلاف التوزيع الاحتمالي في المجتمعين عند مستوى معين من القدرة، ولفحص تحقيق البيانات لهذا الافتراض هناك طرق إحصائية متعددة من أهمها أسلوب التحليل العاملي بأنواعه.

2- **الاستقلال الموضعي للفقرة (Local Item Independence):** هذا الافتراض يعني أن استجابات المفحوصين الذين لديهم نفس القدرة للإجابة عن فقرة ما مستقلة استقلالًا إحصائيًا (Statistically Independent)، أي أن استجابة المفحوص على فقرة ما يجب ألا تؤثر ولا تتأثر باستجابته على أي فقرة أخرى في الاختبار، وعند تحقق هذا الافتراض يكون احتمال الحصول على أي نمط معين لدرجات مفحوص ما مساويًا لحاصل ضرب احتمالات حدوث هذا النمط لكل فقرة من فقرات الاختبار التي أجاب عليها، فعلى سبيل المثال إذا كان نمط إجابة مفحوص على (4) فقرات كالتالي: (u= 1, 0, 1, 1) حيث (1) تعني إجابة صحيحة للفقرة، و(0) تعني إجابة خاطئة لها، فإن احتمال الحصول على هذا النمط من الاستجابة هو:

$$(P_1)(1-P_2)(P_3)(P_4)$$

حيث P_i: احتمال إجابة المفحوص إجابة صحيحة عن الفقرة i.

$(1 - P_i)$: احتمال إجابة المفحوص إجابة خاطئة عن الفقرة i.

كما يرى بيجار (Bejar, 1980) أن هذا الافتراض مُكافئ لافتراض أحادية البعد للسمة الكامنة (الافتراض الأول)، فإذا لم تكن الفقرات مستقلة إحصائيًا عند مستوى قدرة معين؛ فهذا يعني أن العلامات المتوقعة لبعض المفحوصين على الاختبار ستكون أعلى من العلامات المتوقعة لمفحوصين آخرين لهم مستوى القدرة نفسه، مما يدل على وجود أكثر من قدرة لتفسير أداء المفحوص على هذا الاختبار، وهذا انتهاك لافتراض أحاية البعد؛ لذلك من المهم الإشارة إلى أن افتراض الاستقلال الموضعي للفقرة لا يتضمن أن الفقرات مستقلة إحصائيًا عبر المجموعة الكلية للمفحوصين، ولكن عند مستوى قدرة محدد.

3- منحنى خصائص الفقرة (ICC) (Item Characteristic Curve): يمكن تمثيل العلاقة الوتيرية بين تحصيل المفحوص على الفقرة وقدرته المقاسة بالاختبار بيانيًا باقتران تراكمي صاعد (متزايد) يعرف باسم منحنى خصائص الفقرة (ICC) أو منحنى استجابة الفقرة (IRF) (Item Response Function)، وهذا المنحنى يعرض احتمال إجابة المفحوصين إجابة صحيحة على الفقرة على المحور العمودي (المحور الصادي) مقابل مستويات القدرة المختلفة مرتبة تصاعديًا على المحور الأفقي (المحور السيني)؛ لذلك فالمنحنى الناتج سيكون تراكميًا صاعدًا لأن احتمالية إجابة المفحوص إجابة صحيحة على الفقرة من المتوقع أن تزداد مع إزدياد قدرة المفحوص، وفي حالة الاختبارات المصممة لقياس سمة واحدة (أحادية البُعد) تكون منحنياتها بدلالة معلمة واحدة (صعوبة الفقرة)، أو معلمتين (صعوبة الفقرة وتمييزها)، أو ثلاث معالم (الصعوبة والتمييز والتخمين) أو أكثر، لذلك يرجع الاختلاف الرئيسي بين نماذج السمات الكامنة إلى اختلاف منحنى خصائص الفقرة (ICC) فيما بينها.

4- **التحرر من السرعة في الأداء (Speediness):** تفترض النظرية الحديثة أن عامل السرعة لا يؤثر في الإجابة على فقرات الاختبار، أي أن إخفاق المفحوصين في إجاباتهم على فقرات الاختبار يرجع إلى انخفاض قدراتهم وليس لتأثير عامل السرعة على إجاباتهم، إن هذا الافتراض له تأثير غير مباشر على إحتمالية إجابة المفحوصين إجابة صحيحة على فقرات الاختبار؛ أي أنه يُؤثر بشكل غير مباشر على شكل المنحنى المميز للفقرة (ICC)، فهذا العامل قد يجعل المفحوصين يلجأون للتخمين في الإجابة؛ مما يجعل عامل سرعة الأداء إحدى السمات التي يقيسها الاختبار.

كما تعتمد نظرية استجابة الفقرة على نماذج رياضية تهدف لتحديد العلاقة بين احتمالية إجابة المفحوص على فقرة ما إجابة صحيحة وبين السمة الكامنة التي تكمن وراء هذا الأداء، ويمكن تصنيف هذه النماذج إلى ثلاثة أنواع؛ يستخدم النوع الأول منها في حالة الفقرات ثنائية الاستجابة (Dichotomous response) وفيه تكون استجابة المفحوص إما إجابة صحيحة فيأخذ العلامة (1) أو إجابة خاطئة فيأخذ العلامة (0)، أما النوع الثاني فيُستخدم مع الاستجابات المتعددة (Polytomous response)، وفيه تكون استجابة المفحوص مُتدرجة، فقد يحصل المفحوص على علامة جزئية من العلامة الكلية، مثل العلامة (5) من (6)، أو العلامة (3) من (6)، أما النوع الثالث فيكون خليط من النوعين الأول، والثاني.

وهذه النماذج المختلفة التي تم تطويرها في نظرية السمات الكامنة (IRT) عبارة عن دوال رياضية احتمالية تختلف الصيغة الرياضية بينها باختلاف منحنى خصائص الفقرة (ICC)، وعدد معالم الفقرة المأخوذة بعين الاعتبار لتكوين بنائها الرياضي، وتحديد منحنى خصائص الفقرة، حيث أن كل هذه النماذج تهدف لتحديد العلاقة بين أداء المفحوص على الاختبار وبين القدرات (السمات) الكامنة وراء هذا الأداء وفق معادلة رياضية (Hullin, Drasgow & Parsons,

1983)، وأكثر هذه النماذج شيوعًا هي: النماذج اللوجستية (Logistic Models)، ونماذج المنحنى الطبيعي التراكمي (Normal Ogive Models)، وفي كلا النوعين من النماذج يتم استخدام معلمة واحدة تمثل درجة صعوبة الفقرة ويرمز لها بـ(b_i)، أو معلمتين تمثل إحداهما درجة الصعوبة (b_i)، والأخرى درجة التمييز ويرمز لها بـ(a_i)، أو ثلاث معالم وهي: درجة الصعوبة (b_i)، ودرجة التمييز (a_i)، وبالإضافة لدرجة تخمين المفحوص على الفقرة ويرمز لها بالرمز (c_i)؛ والتي تمثل خط التقارب الأدنى (Lower Asymptotic Line) لمنحنى خصائص الفقرة (ICC) ويكون موازيًا لمحور القدرة التي يرمز لها بـ(θ)، أو أربع معالم وذلك بإضافة معلمة الخط التقاربي الأعلى (Upper Asymptote Line) والتي يرمز لها بـ(d_i)، وغالبًا ما تستعمل النظرية الحديثة (IRT) النماذج اللوجستية أكثر من نماذج المنحنى الطبيعي التراكمي وذلك لخصائصها الرياضية، وتقسم النماذج اللوجستية حسب عدد المعالم المستخدمة فيها كما أوردها هامبلتون وسوامينثان وروجرز (Hambelton, Swaminathan and Rogers, 1991)، وأمبريتسون وريز (Embretson and Reise, 2000) إلى أربعة نماذج هي:

1- النموذج اللوجستي أحادي المعلمة (One Parameter Logistic Model-1PLM):

ويُعد هذا النموذج حالة خاصة من النموذج الثنائي والثلاثي والرباعي المعلمة، كما يعتبر من أوسع النماذج المستخدمة في نظرية استجابة الفقرة لسهولة التعامل معه، وقلة المشاكل التي تواجهها عملية تقدير المعالم، كما أنه عندما تتطابق البيانات مع هذا النموذج فإن معالم الفقرات يمكن تقديرها بشكل مستقل عن عينة المفحوصين المستخدمة، كما يمكن تقدير قدرات المفحوصين بشكل مستقل عن درجة صعوبة الفقرات (Masters, 1984)، ويُسمى بنموذج راش (نسبة للعالم Rash)، أو نموذج المعلمة الواحدة، والذي يفترض أن جميع الفقرات لا تختلف عن بعضها إلا بمعلمة صعوبة الفقرة (b_i) التي يتم تقديرها بعدد المفحوصين الذين أجابوا إجابة صحيحة على هذه

الفقرة (i)، كما يتم استخدام مجموع علامات المفحوص على هذه الفقرات كإحصائي كافي لتقدير قدرة المفحوص (θ)، كما يفترض هذا النموذج إنعدام التخمين في استجابة المفحوص عن فقرات الاختبار $(c=0)$، وأن تمييز الفقرات متساوية تقريبًا، ويتخذ النموذج الصيغة الرياضية كما في المعادلة (1) لتعبر عن احتمالية الإجابة الصحيحة (p_i) على الفقرة (i).

$$p_i(\theta) = \frac{1}{1 + e^{-D(\theta - b_i)}} \qquad i = 1, 2, 3, \dots, n \qquad (1)$$

حيث أن:

$p_i(\theta)$: احتمالية الإجابة الصحيحة على الفقرة (i) من قبل مفحوص تم اختياره عشوائيًا من مستوى القدرة (θ)،

b_i: معلمة الصعوبة للفقرة (i)، وتعرف بأنها قدرة المفحوص التي تلزم كي يجيب عن الفقرة (i) إجابة صحيحة باحتمال (0.5) في غياب التخمين، أو $(1+c)/2$ في حالة وجود التخمين،

n: عدد فقرات الاختبار،

e: الأساس اللوغاريتمي الطبيعي ويساوي (2.718)،

θ: قيمة مستوى القدرة،

D: عامل التدريج (Scaling Factor) الذي يجعل القدرة الناتجة من هذا النموذج مساوية للقدرة الناتجة من نموذج المنحنى الطبيعي التراكمي (Normal Ogive Models)، ويساوي (1.7).

2- النموذج اللوجستي ثنائي المعلمة (Two Parameter Logistic Model-2PLM):

وهذا النموذج يفترض أن الفقرات تختلف في معلمتي الصعوبة (b_i)، والتمييز (a_i) في حين تقترب معلمة التخمين (c_i) من الصفر، وتُعطى احتمالية الإجابة الصحيحة على الفقرة بالصيغة الرياضية كما في المعادلة (2):

$$p_i(\theta) = \frac{1}{1 + e^{-Da_i(\theta - b_i)}} \qquad i = 1, 2, 3, \dots, n \qquad (2)$$

حيث أن:

$p_i(\theta)$: احتمالية الإجابة الصحيحة على الفقرة (i) من قبل مفحوص تم اختياره عشوائيًا من مستوى القدرة (θ)،

b_{i}: معلمة الصعوبة للفقرة (i)، وتعرف بأنها قدرة المفحوص التي تلزم كي يجيب عن الفقرة (i) إجابة صحيحة باحتمال (0.5) في غياب التخمين، أو $(1+c)/2$ في حالة وجود التخمين،

n: عدد فقرات الاختبار،

a_{i}: معلمة التمييز للفقرة (i)،

e: الأساس اللوغاريتمي الطبيعي ويساوي (2.718)،

θ: قيمة مستوى القدرة،

D: عامل التدريج (Scaling Factor) الذي يجعل القدرة الناتجة من هذا النموذج مساوية للقدرة الناتجة من نموذج المنحنى الطبيعي التراكمي، ويساوي (1.7).

3- **النموذج اللوجستي ثلاثي المعلمة (Three Parameter Logistic Model-3PLM)**:

ويتعامل النموذج اللوجستي ثلاثي المعلمة مع الخط التقاربي الأدنى (Lower Asymptote Line) الذي يمثل معلمة التخمين (Guessing Parameter) (c_{i})؛ أي أن المفحوصين منخفضي القدرة يجيبوا عن فقرات الاختبار إجابة صحيحة إما بالتخمين، أو بشكل عشوائي، أو على أساس معرفتهم، أما في النموذجين الأحادي والثنائي المعلمة فيفترضان أن منخفضي القدرة لا يجيبوا عن الفقرات الصعبة إجابة صحيحة، وأن احتمالية إجابتهم عن الفقرة تساوي $(p_i=0)$، أما ذوي القدرة المرتفعة يجيبون عن الفقرات السهلة إجابة صحيحة، واحتمالية إجابتهم تساوي $(p_i=1)$، ولكن هذا الافتراض لا يكون مقبولًا دائمًا؛ لأن المفحوص الذي لا يعرف شيئًا لا تزال لديه فرصة في اختيار الإجابة الصحيحة في اختبارات الاختيار من متعدد، وهذا ما يُفسره النموذج ثلاثي

المعلمة ليصبح احتمالية الإجابة الصحيحة $p_i(\theta)$ على الفقرة (i) بالصيغة الرياضية كما هي في المعادلة (3):

$$p_i(\theta) = c_i + (1 - c_i)\frac{1}{1 + e^{-Da_i(\theta - b_i)}} \qquad i = 1, 2, 3, \dots, n \qquad (3)$$

حيث أن:

$p_i(\theta)$: احتمالية الإجابة الصحيحة على الفقرة (i) من قبل مفحوص تم اختياره عشوائيًا من مستوى القدرة (θ)،

b_i: معلمة الصعوبة للفقرة (i)، وتُعرف بأنها قدرة المفحوص التي تلزم كي يجيب عن الفقرة (i) إجابة صحيحة باحتمال (0.5) في غياب التخمين، أو $(1+c)/2$ في حالة وجود التخمين،

n: عدد فقرات الاختبار،

a_i: معلمة التمييز للفقرة (i)،

c_i: معلمة التخمين للفقرة (i)،

e: الأساس اللوغاريتمي الطبيعي ويساوي (2.718)،

θ: قيمة مستوى القدرة،

D: عامل التدريج (Scaling Factor) الذي يجعل القدرة الناتجة من هذا النموذج مساوية للقدرة الناتجة من نموذج المنحنى الطبيعي التراكمي، ويساوي (1.7).

4- النموذج اللوجستي رباعي المعلمة (Four Parameter Logistic Model-4PLM):

وهنا أضاف بارتون ولورد (Barton & Lord, 1981) معلمة الخط التقاربي الأعلى (Upper Asymptote Line) التي تُسمى بمعلمة اللامبالاة، ويُرمز لها بالرمز (d_i) التي تفترض أن ذوي القدرة المرتفعة لا يجيبون إجابة صحيحة على فقرات الاختبار السهلة لاعتمادهم (اللامبالاة) في اختيار الإجابة، أو قد تكون معلوماتهم غير التي يفترضها كاتب فقرة الاختبار، أو قد يخطيء

المفحوص كتابيًا عن الفقرات المقالية، أو لصعوبة هذه الفقرة بالنسبة للمفحوص، فهذا ما يفسره النموذج رباعي المعلمة، لتصبح احتمالية الإجابة الصحيحة $p_i(\theta)$ على الفقرة (i) بالصيغة الرياضية كما هي في المعادلة (4):

$$p_i(\theta) = c_i + (d_{\mathrm{i}} - c_i)\frac{1}{1 + e^{-Da_i(\theta - b_i)}} \qquad i = 1, 2, 3, \dots, n \qquad (4)$$

حيث أن:

$p_i(\theta)$: احتمالية الإجابة الصحيحة على الفقرة (i) من قبل مفحوص تم اختياره عشوائيًا من مستوى القدرة (θ)،

b_{i}: معلمة الصعوبة للفقرة (i)، وتعرف بأنها قدرة المفحوص التي تلزم كي يجيب عن الفقرة (i) إجابة صحيحة باحتمال (0.5) في غياب التخمين، أو $(1+c)/2$ في حالة وجود التخمين،

n: عدد فقرات الاختبار،

a_{i}: معلمة التمييز للفقرة (i)،

c_{i}: معلمة التخمين للفقرة (i)، وتمثل الخط التقاربي الأدنى لمنحنى خصائص الفقرة (i) الذي يمثل احتمالية إجابة المفحوص ذي القدرة الأقل إجابة صحيحة على الفقرة.

d_{i}: معلمة اللامبالاة، وتمثل الخط التقاربي الأعلى لمنحنى خصائص الفقرة (i)،

e: الأساس اللوغاريتمي الطبيعي ويساوي (2.718)،

θ: قيمة مستوى القدرة،

D: عامل التدريج (Scaling Factor) الذي يجعل القدرة الناتجة من هذا النموذج مساوية للقدرة الناتجة من نموذج المنحنى الطبيعي التراكمي، ويساوي (1.7).

وانبثقت هذه النماذج عن نظرية استجابة الفقرة (IRT) التي أُستخدمت في بناء وتطوير المقاييس، والاختبارات التي يُمكن من خلالها الحصول على مؤشرات إحصائية للفقرة لا تعتمد على

خصائص المفحوصين، ولا تعتمد على صعوبة فقرات المقياس، كما تتسم بالدقة، والموضوعية، والكفاءة؛ لذلك تم الاهتمام بالخصائص السيكومترية للمقياس من صدق وثبات عاليين للتقليل من أخطاء القياس، واعتماد النتائج التي يتم الحصول عليها من المقياس من قِبل صُناع القرار، فصدق المقاييس يُعد من المعايير الأساسية التي يتم استخدامها من أجل الحكم على جودة المقاييس النفسية والتربوية، وهي خاصية للأداة نفسها، ويُعرف بأنه "الدقة التي تقوم الأداة من خلالها بقياس ما يجب أن يتم قياسه، أو أن تقيس الأداة لما أُعدت لأجله" (الحجاج، 2013)، وقد ثبت أنه لا يمكن الاستفادة من مزايا نظرية الاستجابة للفقرة (IRT) إذا لم يتم التحقق من افتراضاتها في البيانات المراد تحليلها، كما يجب التأكد من اختيار النموذج المناسب لتحليل هذه البيانات (Zhu, 2009).

وأكثر التعريفات تداولًا في البحوث السيكومترية لمفهوم أحادية البُعد هو: وجود عامل (Factor) واحد يكمن وراء الأداء في المقياس، وهذا العامل هو القدرة أو السمة المقاسة التي تسبب التباين بين المفحوصين في الأداء على فقرات المقياس (Hambleon, Swaminathan & Rogers, 1991).

وقد أكد أبو هاشم (2006)، والخياط (2012) على أن أحادية البُعد تعني تجانس فقرات المقياس فيما بينها وتقيس نفس السمة، وأن أية فقرة من الفقرات تتطلب للإجابة عليها نفس الإجراءات والعمليات السلوكية؛ أي نفس السمة (القدرة)، ولكنها تختلف فيما بينها من حيث تدرج صعوبتها فقط، ويؤخذ على التعريف اقتصاره على اختلاف الفقرات في معلمة الصعوبة فقط؛ إذ قد تتحقق أحادية البُعد أيضًا عند اختلاف الفقرات فيما بينها من حيث معلمة التمييز، أو التخمين، أو اللامبالاة، وهذا ما أدى لظهور نماذج أحادية البُعد، وثنائية، وثلاثية، ورباعية المعلمة، كما أُخذ على كاظم (1988) أيضًا حينما أشارت إلى أن نموذج راش يُعد النموذج الوحيد أحادي البُعد.

ويُميز سيك (Sick, 2010) بين أحادية البُعد السيكولوجية، وأحادية البُعد السيكومترية، فيعتبر أن أحادية البُعد السيكومترية تتطلب أن يكون منحنى خصائص الفقرة لكافة الفقرات متسقة داخل نموذج ضمني؛ بمعنى أن التباين المنتظم بين المفحوصين يمكن عزوه لبعد واحد، والاستجابات الغير متوقعة تحدث بشكل عشوائي وغير متكرر، وذلك لأن الأداء على المقياس يتطلب مهارات، ومعارف، وعمليات، واستراتيجيات متعددة، أما إذا كان من الممكن التنبؤ بالاستجابات غير المتوقعة لمجموعة من الفقرات أو مجموعة من الأفراد فإن ذلك يعني وجود بُعد سيكومتري آخر يؤثر في هذه الاستجابات.

كما أشار هامبلتون وآخرون (Hambleton et al., 1991) إلى أنه يصعب تحقق أحادية البُعد في أغلب الأحوال؛ لوجود عوامل أخرى تؤثر في استجابات المفحوصين على فقرات المقياس، مثل العوامل المعرفية، أو الشخصية، أو المتعلقة بأداء المفحوصين على الاختبار نفسه مثل مستوى الدافعية، وقلق الاختبار، والسرعة في الإجابة، واتباع التعليمات، وحكمة الاختبار، كما أكد علّام (2005) أن افتراض أحادية البُعد في الاختبارات العقلية، وبعض المقاييس الوجدانية لا يُعد افتراضًا صحيحًا كليةً، ويترتب على عدم تحقق أحادية البُعد استخدام نماذج متعددة الأبعاد، كما أوضح لي (Lee, 2012) إمكانية تحقق أحادية البُعد لمقياس ما عند تطبيقه على عينة من المفحوصين، وعدم تحققها لنفس المقياس إذا تم تطبيقه على عينة أخرى، ومن هنا جاءت أهمية خاصية اللاتغير في القياس (ثبات القياس) (Measmurent Invariance) في تقدير معالم الفقرات بين مختلف مجموعات المفحوصين، ووُصفت هذه الخاصية بأنها أهم الخصائص في نظرية استجابة الفقرة (IRT) (Lord, 1980)، حيث أن أهم ما يُميز هذه النظرية الموضوعية في القياس، والتي تَعني اللاتغير في تقدير خصائص الفقرات بتغير المفحوصين، واللاتغير في تقدير قدرات المفحوصين بتغير عينة الفقرات (Hambleton & Swaminathan, 1985)، إلّا أن

علّام (2005) أشار إلى هذه الخاصية بأنه ليس لها ما يبررها في الواقع العملي، وهي ممكنة في إطار نظري إحصائي، كما أكّد عبابنة (2008) في دراسته أن افتراض اللاتغير في القياس هو افتراض نظري ويصعب تحقيقه في الواقع العملي، ومن هنا تأتي مشكلة الدراسة الحالية في محاولة لفحص هذا الافتراض في ضوء بيانات واقعية.

اللاتغير في القياس (ثبات القياس) (Measurement Invariance) (MI)

يُعد جورسكوج (Jöreskog, 1971) أول من كتب عن تكافؤ البناء العاملي الذي يُطلق عليه اللاتغير في القياس بعد تطويره للتحليل العاملي التوكيدي (CFA) (Confirmatory Factor Analysis) للتحقق من صدق البناء (Construct Validity) إعتمادًا على مصفوفة السمات المتعددة–الأساليب المتعددة (Multi Trait–Multi Method Matrix) (MTMM) (Jöreskog, 1969)، كما يُطلق على اللاتغير في القياس باللاتغير العاملي (Windaman, Ferrer, & Conger, 2010)، أو ثبات القياس، أو تكافؤ القياس (Measurement Invariance or Equivalence Measurement)؛ وهو خاصية إحصائية تُشير إلى أن البناء (Structure) أو التركيب الذي يتم قياسه هو نفسه عبر المجموعات المختلفة عند استخدام أداة القياس موضع اهتمام الدراسة؛ لذلك يتم استخدام التحليل العاملي التوكيدي للمجموعات المتعددة (Multiple Groups Confirmatory Factor Analysis) (MG–CFA)؛ لأن خاصية اللاتغير في القياس مرتبطة بالمجموعات المتعددة، وليس المجموعة الواحدة، وقد يتم استخدام نماذج المعادلة البنائية (Structural Equation Modeling) (SEM) لاختبار العلاقات بين المتغيرات من منظور توكيدي، وليس استكشافي؛ أي أن الباحث يستخدم التحليل الإحصائي على البيانات بعد بناء النموذج النظري، وذلك للتأكد من مطابقة البيانات التي يجمعها مع النموذج

المفترض عكس المنظور الاستكشافي الذي يفتقر فيه الباحث لعدد العوامل وطبيعتها (القهوجي، 2018).

ويتمثل هدف النمذجة بالمعادلة البنائية (SEM) بتحديد مدى مطابقة النموذج النظري للبيانات الميدانية؛ أي المدى الذي يتم فيه تأييد النموذج النظري بواسطة بيانات العينة، فإذا دعمت بيانات العينة النموذج النظري، فمن الممكن بعد ذلك افتراض نماذج نظرية أكثر تعقيدًا، أما إذا لم تدعم البيانات النموذج النظري، فإما أن يتم تعديل النموذج الأصلي، أو أن يتم تطوير نماذج نظرية أخرى واختبارها (Byrne, 2011).

ويُعد اللاتغير في القياس أحد الطرق الإحصائية التي تمكن الباحث من التحقق إحصائيًا ما إذا كانت أداة القياس تعكس بصورة دقيقة الفروق بين المفحوصين في بناء فرضي ما (السمة)، حيث أن اللاتغير في القياس يُعد متطلب قبلي لإجراء المقارنات بين المجموعات؛ لأن الاستنتاجات المشتقة من تحليلات المقارنة قد تكون متحيزة، أو غير صادقة إذا كانت القياسات ليست لها نفس المعنى عبر المجموعات؛ بمعنى أن الاستنتاجات المتعلقة بفروق المتوسطات بين المجموعات على مقياس ما تكون غير صادقة إذا كان المفهوم الكامن (السمة) الذي يُفترض أنه يُقاس يكون مختلف عبر المجموعات (Vandenberg & Lance, 2000).

ونظرًا للتقدم الذي شهده علم القياس والتقويم التربوي والنفسي، والتوسع في الاعتماد على نماذج نظرية الاستجابة للفقرة (IRT)، وكذلك الاهتمام الكبير بتطوير المقاييس والاختبارات النفسية والتربوية من قبل علماء القياس والباحثين، للوصول لدقة، ومصداقية، وكفاءة عالية؛ من خلال الجهود المبذولة في تصميمها، وإعدادها، والتحقق من صدقها، وثباتها، ويُعد بناء، أو تطوير أدوات القياس والتقويم من التوجهات والاهتمامات الأساسية لعلماء القياس والتقويم، وعلم النفس التربوي، وأصحاب القرار، ومن المواضيع المتعلقة بها، والتي زاد الاهتمام بها في الآونة الأخيرة؛ موضوع

اللاتغير في القياس للأدوات والمقاييس للتحقق من صدقها، إلى جانب الاهتمام بالطرق الإحصائية المستخدمة للكشف عنه، وخاصةً ضمن إطار نمذجة المعادلة البنائية (SEM) (Little, 2000).

فخاصية اللاتغير في القياس تُعتبر أسلوب إحصائي لإظهار أن أداة القياس تقيس السمة الكامنة (العامل) نفسها عبر مجموعات مختلفة من المفحوصين لهم خلفيات ثقافية متنوعة، أو في أزمان مختلفة (Finch, 2014)، كما يُعرف بأنه مدى إمكانية تعميم السمات الكامنة (العوامل Factors) في السياقات الاجتماعية، والثقافية المختلفة التي تم اختبار اللاتغير في القياس فيها، كما أُشير إليه بأنه الحالة التي تكون فيها مصادر التحيز، والخطأ أقل ما يمكن، كما أنه الحالة التي لا تؤثر فيها الاختلافات الثقافية على قياس السمات الكامنة؛ أي أن هذه السمات الكامنة متشابهة لدى المجموعات المختلفة؛ لأن التباينات الفريدة للفقرات مستقلة عن الثقافات المختلفة (Little, 1997)، كما تُعرفه بايرن و واتكنز (Byrne & Watkins, 2003) بأنه ثبات العمليات، أو الوظائف التي من المُفترض أن تقيسها أداة القياس عبر المجموعات المختلفة؛ بمعنى أن إدراك، وتفسير محتوى فقرات أداة القياس تكون متكافئة إذا ما تم فحصها لدى مجموعات مختلفة من المفحوصين على المقياس.

وبالرغم من أن تاريخ الاهتمام باختبار اللاتغير في القياس يمتد لأكثر من (50) عامًا، لكن الفنيات الإحصائية لاختبار اللاتغير في القياس أصبحت أكثر إتاحة لمجتمع الباحثين حديثًا، فبحلول القرن الحادي والعشرين، وجه المتخصصون اهتمامهم بصورة متزايدة بدلالة، وأهمية اللاتغير في القياس، وخصوصًا داخل إطار نمذجة المعادلة البنائية (SEM)، كما تم مراجعة أدبيات اللاتغير في القياس، ورسم توجيهات للباحثين خطوة بخطوة لتنفيذ اختبارات اللاتغير في القياس مما زاد عدد البحوث التي اهتمت بالتحقق من خاصية اللاتغير في القياس للمقاييس، وهذه الزيادة استمرت بشكل أكبر في العقد التالي، إلا أن تلك الزيادة لم يصاحبها نصائح كافية، أو

ممارسة جيدة، أو فهم كافٍ (Putnick & Bornstein, 2016)، هذا حال البحوث الأجنبية، ففي حدود علم الباحثة أن الاهتمام بالتحقق من اللاتغير في القياس لم ينتقل إلى كافة البحوث العربية، وذلك لعدم معرفة الكثير من الباحثين في الوطن العربي بهذا المفهوم، وعدم معرفة البعض الآخر بالخطوات التي يجب اتباعها لاختبار تحقق اللاتغير في القياس (Vandenberg, 2002).

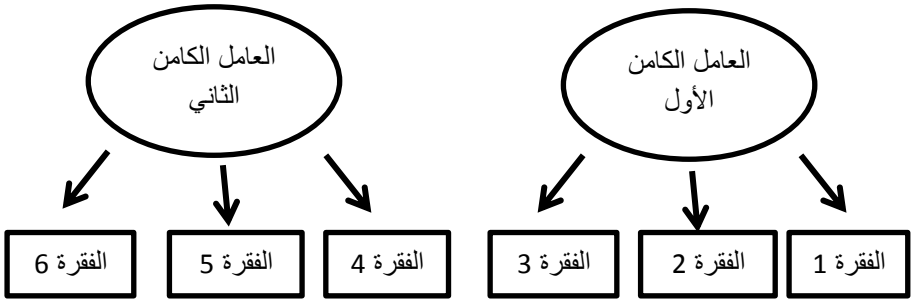
طرق الكشف عن اللاتغير في القياس (ثبات القياس)

ظهر الاهتمام باللاتغير في القياس من الناحية النظرية منذُ أكثر من خمسين عامًا، في حين ازداد الاهتمام بالطرق الإحصائية المتعلقة باختبار اللاتغير في القياس للتحقق من صدق المقاييس منذ فترة قصيرة (Little, 2000)، وقد أوضح فاندينبرغ (Vandenberg, 2002) أن هناك مدخلين يمكن من خلالهما التحقق من اللاتغير في القياس وهما: مدخل التحليل العاملي التوكيدي (CFA)، ومدخل النظرية الحديثة في القياس (IRT).

اللاتغير في القياس من منظور التحليل العاملي التوكيدي (CFA)

يُعد التحليل العاملي التوكيدي (CFA) أسلوبًا إحصائيًا يستخدم للتحقق من البنية العاملية لعدد من المتغيرات الملاحظة (Observed Variables) وبناءاتها الافتراضية الكامنة التي عُرفت إحصائيًا بالعوامل (Factors) ومن منظور النظرية الحديثة بالسمة الكامنة، أو القدرة (Latent trait)؛ أي أن التحليل العاملي التوكيدي يهدف إلى معرفة مدى جودة مطابقة نموذج عاملي اُشتق من نظرية معينة مستندة على أساس نظري لمجموعة من البيانات تُساعد على تحديد الطريقة التي تتشبع بها المتغيرات الملاحظة على تلك العوامل؛ لذلك يُعتبر التحليل العاملي التوكيدي أسلوبًا إحصائيًا مرنًا، وقويًا يُركز على نمذجة العلاقة بين المتغيرات الملاحظة، والسمات الكامنة معتمدًا على نمذجة المعادلة البنائية (Structural Equation Modeling) (SEM) (Stewart, 1981)، حيث يتم فيه التعامل مع نماذج القياس التي تصف العلاقة بين المتغيرات الملاحظة؛

وهي المؤشرات المتمثلة بفقرات الاختبار، والمتغيرات الكامنة؛ وهي العوامل، وعند تجسيد نموذج القياس باستخدام الرسم البياني، فإن المتغيرات الملاحظة تكون داخل المستطيل، والعوامل الكامنة داخل الشكل البيضاوي، والسهم ذا الرأس الواحد يحدد العلاقة بين المتغيرين الملاحظ، والكامن، كما في الشكل (1).



شكل 1: نموذج قياس يتكون من عاملين كامنين، وست فقرات

كما ويوجد في هذا النموذج عدد من المكونات هي (Vandenberg & Lance, 2000):

- **العامل المشترك (العامل الكامن الأول) (Common Factor):** هو المتغير الكامن (السمة الكامنة) (Latent trait) الذي يمثل بناء نظري لا يمكن ملاحظته مباشرة (داخل الشكل البيضاوي).
- **المتغيرات الملاحظة (Observed Variables):** هي فقرات المقياس التي تعمل كمؤشر للمتغير الكامن (داخل المستطيل).
- **تشبع العامل (Factor Loading):** هي العلاقة الارتباطية بين الفقرات والمتغير الكامن، وتمثل معاملات الانحدار، التي تُعبر عن حجم التغير المتوقع في الفقرات لكل تغير في المتغير الكامن (الأسهم أحادية الاتجاه بين المتغير الكامن والفقرات).
- **تقاطع الفقرة (Item Intercept):** وهي نقطة الأصل، أو قيمة البداية للمقياس الذي يستند إليه العامل في البيانات الثنائية، كما يُسمى بالعتبات الفارقة (Thresholds) في البيانات الفئوية، وهو نفس تقاطع الانحدار (لا يتم تمثيله بصريًا في النموذج).

- **الوسط الحسابي للعامل الكامن (Latent Factor Mean)**: الوسط الحسابي للعامل الذي يتم قياسه بواسطة الفقرات (لا يتم تمثيله بصريًا في النموذج).
- **تباين العامل (Factor Variance)**: يُعرف بالخطأ المتبقي، أو خطأ البواقي، وهو الخطأ العام في التنبؤ بالعامل الكامن باستخدام الفقرات الخاصة به (الأسهم الأحادية التي تشير للعامل الكامن).
- **التباينات، والتباينات المشتركة (Variance and Covariance)**: خطأ القياس المرتبط بكل متغير مشاهد، ويُعرف بالخطأ (الأسهم الأحادية التي تشير للفقرات).

ويعرض الشكل (1) نموذج القياس لاختبار مكون من ست فقرات، وعاملين كامنين، فيه الفقرات الثلاث الأولى تتشبع (ترتبط) بالعامل الأول، والفقرات الثلاثة الأخيرة تتشبع على العامل الكامن الثاني، وللحصول على مثل هذا النموذج يتم الرجوع للنظريات، ثم يتم اختبار حُسن مطابقته مع نموذج بيانات العينة، حيث يتم تقدير المعالم المجهولة للنموذج باعتماد مصفوفة التباين، والتغاير لبيانات العينة؛ والتي تكون إما حرة التقدير، أو مثبتة، أو مقيدة، وحتى نستطيع الحكم على حُسن مطابقة النموذج نستخدم مؤشرات حُسن المطابقة المختلفة (Goodness of Fit Indices) كدليل من أدلة الصدق العاملي على أن الفقرات تقيس بناء فرضي مُعين (Harrington, 2009; Raykov and Marcoulides, 2006; Tabachnick and Fidell, 2013; Brown, 2015).

خطوات اختبار النموذج العاملي التوكيدي

يتم اختبار النموذج النظري المُفترض عند توظيف طريقة التحليل العاملي التوكيدي في خمس مراحل أساسية وهي (تيغزة، 2012):

المرحلة الأولى: بناء النموذج العاملي: يتم بتوظيف النظريات، والأطر المرجعية في تطوير نموذج عاملي مُفترض، وغالبًا ما يكون النموذج العاملي عُرضةً لبعض أخطاء التحديد، مثل افتقار النموذج لمتغير أو أكثر، أو أن يُعاني النموذج من عدد زائد في المتغيرات المدرجة؛ مما قد يؤدي إلى عرقلة دور المتغيرات المهمة في النموذج، وأخطاء التحديد تُؤثر على صدق النموذج، وتعرقل قدرته على المطابقة، وتقوي من التحيز، وتضخم من أخطاء القياس، والتباين غير المفسر (Field, 2005).

المرحلة الثانية: تعيين النموذج العاملي: يتم بتقدير معالم النموذج المختلفة باستخدام البيانات المتوفرة من العينة على شكل مصفوفة التباين والتغاير للمجتمع، بحيث يكون حل وحيد ومحدد للمعالم الحرة للنموذج العاملي المُفترض، فإذا افتقر النموذج للتعيين، فإنه يستحيل تقدير قيمة محددة وحيدة لكل معلمة من المعالم الحرة للنموذج المُفترض، فيكون لكل معلمة عدد كبير من القيم التي تمثله، وبالتالي يستحيل انتقاء الحل الأنسب لكل معلمة (Byrne, 2011).

المرحلة الثالثة: تقدير معالم النموذج العاملي المُفترض: تتم هذه المرحلة بإيجاد قيم عددية للمعالم الحرة في النموذج المُفترض بحيث تكون مصفوفة البيانات المشتقة من النموذج (مصفوفة التباين، والتغاير) أقرب ما يمكن من بيانات العينة؛ أي من مصفوفة التباين والتغاير للعينة التي تمثل الإطار المرجعي الذي ينبغي أن يُعيد النموذج المفترض إنتاجها بدقة، لكي يكون نموذجًا نظريًا متطابقًا مع بيانات العينة، ويمكن تقدير المعالم حاسوبيًا بطرق مختلفة منها: طريقة الأرجحية العظمى (Maximum Likelihood) (ML)، أو طريقة المربعات الصغرى غير الموزونة (Unweighted Least Square)، أو طريقة المربعات الصغرى (Two Stage Least Squares).

المرحلة الرابعة: حُسُن مطابقة النموذج العاملي المُفترض: يتم في هذه المرحلة إنتاج مصفوفة الارتباطات، أو التغاير من النموذج المُفترض بحيث تُماثل بدرجة كافية مصفوفة الارتباطات، أو التغاير المُشتقة امبريقيًا من العينة، وتم اقتراح مؤشرات عديدة لدراسة جودة مطابقة النموذج المُفترض لبيانات العينة (Joseph, at al., 2017).

مؤشرات جودة مُطابقة النموذج للبيانات (Model–Data Fit Indicators)

هي مؤشرات إحصائية، أو وصفية تُساعد الباحث على تحديد مدى جودة النموذج المُفترض نظريًا بمقارنته بالنموذج المشتق من البيانات (Gadelrad, 2004)، وهناك العديد منها لاختبار جودة مُطابقة البيانات لنموذج من نماذج المعادلة البنائية (SEM)، وأهمها (Chen, 2007; Thompson & Green, 2013):

أولًا: الجذر التربيعي لوسط الخطأ التقريبي (Root Mean Square Error of Approximation) (RMSEA): يُعد من أفضل مؤشرات جودة المُطابقة، حيث أظهر تفوقه وأداؤه الجيد في دراسات النمذجة البنائية، فإذا كانت قيمته أقل من (0.05) يدل على مطابقة جيدة، والقيم التي تقع بين (0.05) و(0.08) تدل على وجود خطأ تقارب معقول في المجتمع، أما القيم التي تتراوح من (0.08) إلى (0.1) تدل على مطابقة غير كافية، وإذا زادت قيمته عن (0.1) دل على سوء المطابقة؛ أي أن قيمة الصفر تدل على أفضل ملاءمة ممكنة، وكلما زادت قيمته قلت جودة المُطابقة، وزادت سوءًا (Cheung & Rensvold, 2002) كما يتم اعتماد مقدار التغير في قيمته لتحديد ما إذا كانت النماذج المتداخلة متكافئة، ويُفضل أن تقل قيمة مؤشر الفرق (ΔRMSEA) عن (0.015) (Chen, 2007).

ثانيًا: مُؤشر المُطابقة المُقارن (Comparative Fit Index) (CFI): قيم هذا المؤشر تتراوح بين الصفر والواحد الصحيح، فإذا كان قيمته تتعدى (0.9) يدل على مُطابقة معقولة للنموذج

المُقترح، ويُستخدم كمحك أكثر قابلية للتطبيق من اختبار مربع كاي (χ^2) المعروف بتأثره الكبير بحجم العينة، حيث أن مقدار التغير في قيمة مؤشر المُطابقة المقارن (CFI) لتحديد ما إذا كانت النماذج المتداخلة متكافئة من الناحية التطبيقية؛ أي أنه إذا زادت قيمة الفرق في مؤشر جودة المُطابقة المقارن (ΔCFI) عن (0.01) بين النموذجين، فإن النموذج الأكثر تقيدًا يتم رفضه؛ حيث أن إضافة قيود على النموذج أحدثت مطابقة أسوأ من الناحية التطبيقية، لكن عندما لا يتحقق هذا المحك وبعض المعالم (مثل تشبعات العوامل، أو التقاطعات) لم يتم تحديدها على أنها متساوية بين المجموعتين، يمكن في هذه الحالة النظر بعين الاعتبار لنموذج يحقق خاصية اللاتغير الجزئي في القياس (Cheung and Rensvold, 2002).

ثالثًا: مؤشر تاكر لويس (Tuker–Lewis Index): ينطبق عليه ما ينطبق على (CFI)، ويُطلق عليه أحيانًا مؤشر المطابقة غير المعياري (Non–Normal Fit Index) (NNFI).

رابعًا: مربع كاي (Chi Square) (χ^2): يُعد مؤشرًا لسوء المطابقة؛ لأنه كلما زادت قيمته كلما تدهورت مطابقة النموذج المُقترح للبيانات، وهو أعرق مؤشر لمدى جودة مُطابقة النموذج المُقترح للبيانات، ولكن بسبب عيوبه الكثيرة، وتأثره بحجم العينة يُنصح باستعماله مع مؤشرات أخرى، وعند اختبار اللاتغير في القياس يتم استخدام فرق مربع كاي ($\Delta\chi^2$)، فإذا كانت قيمته دالة إحصائيًا فإنه يدل على وجود معلمة مقيدة واحدة على الأقل في النموذج الأكثر تقييدًا تختلف بصورة دالة إحصائيًا عبر المجموعات، في المقابل عندما يكون ($\Delta\chi^2$) غير دال إحصائيًا تكون مطابقة البيانات للنموذج الأكثر تقييدًا لا تختلف بصورة دالة إحصائيًا عن مطابقة البيانات للنموذج الأقل تقيدًا؛ أي أن المعالم المُقيدة تكون مُستقلة عبر المجموعات، ويفتح المجال للبحث عن المزيد من خصائص اللاتغير في القياس (ثبات القياس) (Millsap & Yun–Tein, 2004).

وبسبب أهم عيوب اختبار مربع كاي (χ^2) وهو تأثره بحجم العينة؛ حيث أنه في حالة العينات كبيرة الحجم يُعطي نتائج دالة إحصائيًا، لذلك ينحصر استخدامه كمؤشر على التحسينات الدالة فقط (Milfont & Fischer, 2010)، وفي المقابل يكون مناسب مع الحجوم المتوسطة التي تكون ما بين (200–100) وقد تصل لـ(400)، وإن زادت حجوم العينات عن ذلك فيتم رفض كل فرضية صفرية (Shook, Hult & Kancmar, 2004; Karkay–Ozyer & Aksu–Dunya, 2018).

خامسًا: جذر وسط البواقي المعياري (Standardized Root Mean Residual) (SRMR): ينطبق عليه ما ينطبق على مربع كاي؛ أي أنه مؤشر لسوء المُطابقة، وهو من المؤشرات المهمة، فإذا قلت قيمته عن (0.1) يدل على مطابقة جيدة، وإذا ساوى الصفر تكون المُطابقة تامة، ويتم اعتماد مؤشر الفرق منه (ΔSRMR) لتكافؤ النماذج المتداخلة بحيث تقل قيمته عن (0.03) (Chen, 2007).

سادسًا: مؤشرات التعديل (Modification Indices): يتم وضع نموذج مُقترح مُشتق من الأدبيات السابقة أو نظرية البحث، ثم يتم محاولة اختبار مُطابقة النموذج للبيانات التجريبية سابقة الذكر، فإذا كانت المطابقة غير كافية، فإن الإجراء الشائع هو تعديل النموذج من خلال حذف المعالم، أو المسارات غير الدالة إحصائيًا، أو إضافة معالم تُحسن مُطابقة النموذج للبيانات، لذلك تُقدم معظم برامج النمذجة مؤشرات التعديل لكل معلمة لتساعد الباحث على تحديد خياراته في تحسين مُطابقة النموذج المُقترح بتعديله؛ أي جعل المعلمة المرتبطة بالمؤشر حرة، لينخفض مؤشر حُسن المطابقة (مربع كاي)، ثم يتم إعادة تقدير النموذج المُقترح بعد تعديله، وهكذا دواليك، مع مراعاة الباحث أن يكون لديه أساس منطقي لهذا الإجراء التجريبي، ويستمر بهذه العملية للوصول لمطابقة نموذج مقبولة تتفق مع المنطق النظري للبحث (McArdle, 2005).

ويُعد التحليل العاملي التوكيدي متعدد المجموعة (Multiple Group Confimatory Factor Analysis) (MG-CFA) الأكثر استخدامًا وتطورًا من أساليب التحليل العاملي التوكيدي لتقييم اللاتغير في القياس ضمن إطار نمذجة المعادلة البنائية (SEM)، وقد تم تطوير مدخل التحليل العاملي التوكيدي متعدد المجموعة ليناسب البيانات المتصلة، كما تم توسيعه ليشمل البيانات الفئوية أيضًا (Millsap & Yun-Tein, 2004)، ويتم فيه تحليل بيانات مجموعتين، أو أكثر في نفس الوقت، ومقارنة النماذج المتداخلة بالنسبة لمطابقة بيانات النموذج العام، وفحص ما إذا كانت المعالم المقيدة المضافة للنموذج الأقل تعقيدًا تكون متكافئة عبر المجموعات، وأن نموذج القياس في كل مجموعة له نفس التنظيم من العناصر المساوية للصفر، وغير المساوية للصفر في مصفوفة تشبعات الفقرات، والعوامل الكامنة، وبذلك يكون اللاتغير التكويني (Configural Invariance) متحقق، وأثناء تلك الخطوة يتم اختبار مطابقة بيانات النموذج، ودقة بناء العوامل، وإذا تكافأت مصفوفة تشبعات الفقرات، والعوامل عبر المجموعات، فإن اللاتغير المتري (Metric Invariance) يكون قد تحقق؛ أي أن في هذه الخطوة يتم التحقق من عدد العوامل والفقرات التي تشبعت على كل عامل أنها كانت نفسها عبر المجموعات، وفي الخطوة النهائية يتم التحقق من اللاتغير العددي (Scalar Invariance) عن طريق مطابقة تقاطعات الفقرات (Items Intercepts) عبر المجموعات، الذي يُعد متطلبًا سابقًا لمقارنة أوساط العوامل الكامنة أو المعالم البنائية عبر المجموعات، وهو هدف المقاييس التربوية والنفسية.

وقد قام فاندينبرغ و لايس (Vandenberg and Lance, 2000) بمراجعة حثيثة لأدبيات اللاتغير في القياس، ولخصا عددًا من اختبارات النماذج المتداخلة الموصى بها لاكتشاف عدم تحقق اللاتغير في القياس، حيث يُجري الباحثون عادةً اختبارًا عامًا لتساوي مصفوفة التباين المشترك (Covariance Matrix)؛ فإذا أظهر الاختبار العام عدم وجود فروق في مصفوفات

التباين المشترك بين فئات البيانات، فإن شروط اللاتغير في القياس تكون قد تحققت لتلك البيانات، لكن شكك بعض الباحثين في فائدة هذا الاختبار النوعي على أساس أن هذا الاختبار يمكن أن يشير إلى أن اللاتغير في القياس متحقق بصورة منطقية عندما تجد الاختبارات الأكثر نوعية، أو الأكثر تحديدًا خلاف ذلك، بصرف النظر عما إذا كان الاختبار العام لتساوي مصفوفة التباين المشترك يشير إلى عدم تحقق اللاتغير في القياس، فإن سلاسل من اختبارات فرق مربع كاي ($\Delta\chi^2$) للنماذج المتداخلة يمكن إجراؤها لتحديد المصادر المحتملة للفروق.

كما يحدد اللاتغير في القياس من منظور التحليل العاملي التوكيدي متعدد المجموعة (MG−CFA) من خلال مستويات اللاتغير المتمايزة، مع قيود مساواة إضافية على معالم النموذج عبر المجموعات تضاف للمستوى الأعلى، على سبيل المثال: يَفترض اللاتغير التكويني (Configural Invariance) أن بنية العوامل فقط تكون نفسها عبر المجموعات، أما اللاتغير المتري (Metric Invariance) فيتطلب تكافؤ تشعبات العوامل (Factor Loading) عبر المجموعات، أما اللاتغير العددي (Scalar Invariance) فيشترط تكافؤ أو ثبات كل من تشبعات العوامل، والعتبات الفارقة (Thresholds) عبر المجموعات، أما اللاتغير الصارم أو التام في القياس (Strict/Full Invariance) فيتطلب تكافؤ، أو ثبات تباينات البواقي (Items Variances) للفقرات عبر المجموعات بالإضافة لتشبعات العوامل، والعتبات الفارقة.

وقد قام مارش وآخرون (March et al., 2009) بتوصيف تصنيف مكون من ثلاثة عشر نموذجًا متداخلًا بصورة جزئية تتمايز فيه نماذج اللاتغير في القياس من التكويني إلى التام؛ حيث فُرضت قيود المساواة على تشبعات العوامل، وتقاطعات الفقرات، وتباين بواقي الفقرات، بالإضافة إلى المتوسطات، والتباينات، والتباينات المشتركة للعوامل الكامنة، ولكن أكّد ميلفونت

وفيشر (Milfont and Fischer, 2010) إلى أنه يوجد أربع نماذج شائعة في حالة نماذج القياس وهي:

1- **اللاتغير التكويني (ثبات الشكل) (Configural Invariance)**: ويُسمى باللاتغير الشكلي، أو البنائي، أو النمطي، وهو الخطوة الأولى، والأقل تقيدًا في سلم اختبارات اللاتغير في القياس، وفيها يجب أن تكون العوامل الكامنة تحمل نفس النسق من التشبعات الحرة، والمُقيدة (التي تم تقديرها، أو تثبيتها لتساوي صفر) عبر المجموعات؛ أي أن التنظيم القاعدي (البنية العاملية) للعوامل الكامنة تم تثبيتها في المجموعات محل المقارنة؛ بمعنى أن عدد العوامل الكامنة في المجموعات المختلفة يرتبط به نفس عدد الفقرات في كافة المجموعات، وفي حالة عدم تحقق اللاتغير التكويني يكون تنظيم تشبعات المفردات على العوامل الكامنة يختلف في المجموعات محل المقارنة؛ أي أنه في أحد المجموعات يوجد على الأقل فقرة واحدة تشبعت على عامل مختلف، أو تشبعت بصورة مركبة على عاملين أو أكثر؛ أي أن السمة الكامنة المراد قياسها بأداة القياس مختلفة عبر المجموعات، والفقرات لا تعبر عن نفس مستويات السمة الكامنة التي يهدف المقياس لقياسها.

وعدم تحقق اللاتغير التكويني يتضمن بديلين: إما إعادة تعريف البناء العاملي (مثل حذف بعض الفقرات، وإعادة اختبار النموذج)، أو افتراض عدم تكافؤ المفهوم المراد قياسه عبر المجموعات، وعدم الاستمرار باختبارات اللاتغير في القياس، وعدم إمكانية اختبار الفروق بين المجموعات، وفي حالة إعادة تعريف البناء العاملي الفرضي (من حذف فقرات، أو إجراء أي تعديلات بعدية للنموذج Post-Hoc Alternatives) في أي خطوة من خطوات اختبار اللاتغير في القياس تكون استراتيجية استكشافية لأنها محكومة بالبيانات (Data-Driven) أكثر من النظرية (Theory-Driven)، لذلك يجب التأكد من مصداقية تعديلات النموذج بأخذ

عينات مماثلة، والعمل على تكرار الدراسة الميدانية، وهذه العملية تُسمى بعبر المصداقية (Cross Validation) (Putnick & Bornstein, 2016).

2- **اللاتغير المتري (الثبات المتري) (Metric Invariance):** بعد تحقق اللاتغير التكويني تكون الخطوة التالية اختبار اللاتغير المتري؛ أي ثبات تشبعات الفقرات على العوامل (Factor Loading) بمعنى أن كل فقرة تسهم بالعوامل الكامنة بدرجة مماثلة عبر المجموعات، وهي تُمثل قوة العلاقة بين الفقرات، والعوامل، ويتم اختبار اللاتغير المتري بجعل تشبعات العوامل متكافئة عبر المجموعات في نموذج قياس مُقيد، ثم يتم مقارنة النموذج المُقيد العوامل بالنموذج المُتكافئ من ناحية التكوين لتحديد المطابقة، وإذا كانت مطابقة النموذج أسوأ بصورة دالة في نموذج اللاتغير المتري مقارنةً بنموذج اللاتغير التكويني، فإن هذا يشير إلى أن تشبع واحد على الأقل غير متكافئ عبر المجموعات، وأن اللاتغير المتري غير مُتحقق، أما إذا كانت مُطابقة النموذج العام غير سيئة بصورة دالة في نموذج اللاتغير المتري مُقارنة بنموذج اللاتغير التكويني، فإن هذا دليل على أن افتراض تساوي التشبعات على العوامل عبر المجموعات لم يؤثر بصورة دالة على جودة النموذج، ومن ثم فإن هذا يُثبت أن النموذج تحقق فيه اللاتغير المتري؛ مما يعني أن استجابات المفحوصين على الفقرات هي بنفس الطريقة عبر المجموعات؛ لأن قوة العلاقة بين الفقرات، والعوامل الكامنة هي نفسها.

وفي حالة عدم تحقق اللاتغير المتري يكون هناك ثلاثة اختيارات: أولها: فحص مصدر عدم تحقق اللاتغير في القياس عن طريق تحرير (باستخدام نهج الرجوع للخلف Backward approach) أو إضافة (باستخدام نهج التقدم للأمام Forward approach) قيود على تشبعات العوامل، وإعادة اختبار النموذج إلى أن يتحقق اللاتغير الجزئي في القياس، ثانيها: حذف الفقرات ذات التشبعات غير المتكافئة عبر المجموعات، وإعادة اختبار النموذجين

التكويني، والمتري؛ وثالثها: افتراض أن البناء العاملي لم يحقق خاصية اللاتغير في القياس، وعدم مواصلة اختبار اللاتغير في القياس، واختبار الفروق بين المجموعات (Jung & Yoon, 2016).

3- **اللاتغير العددي (ثبات التدريج) (Scalar Invariance)**: بعد التأكد من تحقق اللاتغير المتري (أي تقييد تشبعات العوامل) نحتاج لتقييد تقاطعات الفقرات (Items Intercepts)، وهي نقطة بداية الفقرة؛ أي درجة الفقرة عندما تكون درجة العامل صفر؛ بمعنى أنه لمقارنة أوساط السمات الكامنة المراد قياسها من تطبيق أداة القياس يجب أن يتحقق اللاتغير العددي لترتبط الدرجات الملاحظة بالدرجات الكامنة، وبذلك يكون لمن لهم نفس الدرجة على السمة الكامنة يحصلون على نفس الدرجة الملاحظة بصرف النظر عن المجموعة التي ينتمون إليها، ويتم التحقق من ذلك بتثبيت (تقييد) تقاطعات الفقرات عبر المجموعات؛ أي تصبح واحدة عبر المجموعات (Milfont and Fischer, 2010).

وإذا أظهرت النتائج أن مطابقة النموذج العددي أسوأ بصورة دالة إحصائيًا مقارنةً مع نموذج اللاتغير المتري، فإن هذا يدل على أنه يوجد تقاطع لفقرة واحدة على الأقل تختلف عبر المجموعات، وأن البيانات لا تحقق خاصية اللاتغير العددي؛ أي أن اختلاف الدرجات الملاحظة للمفحوصين ليست بسبب اختلاف مستويات السمة الكامنة لديهم بل بسبب إنتماؤهم لمجموعة معينة، وتأثير عوامل خارجية ليس لها علاقة بالسمة الكامنة لديهم، وفي هذه الحالة تكون درجة الفقرة (4) مثلًا ليس لها نفس المعنى عبر المجموعات؛ أي أن الفقرة (4) تحتوي على الأداء التفاضلي (Differential item functioning) (DIF)، في المقابل عند تحقق خاصية اللاتغير العددي، فإن هذا مؤشر على أن تقييد تقاطعات الفقرات عبر المجموعات لا تؤثر بشكل دال إحصائيًا على مطابقة النموذج، وأن البيانات تحقق الخاصية.

وإذا لم يتحقق اللاتغير العددي فإن الباحث يجد نفسه أمام ثلاث اختيارات: الأول: فحص مصدر عدم اللاتغير في القياس من خلال التخلص المتتابع، أو الإضافة من تقاطعات الفقرات، وإعادة اختبار النموذج حتى نحصل على نموذج يحقق اللاتغير الجزئي في القياس، الثاني: حذف الفقرات ذات التقاطعات غير المتكافئة، وإعادة اختبار نماذج اللاتغير التكويني، والمتري، والعددي، أو الثالث: افتراض أن البناء العاملي (السمة الكامنة) لا تحقق خاصية اللاتغير في القياس، وعدم مواصلة اختبار اللاتغير في القياس، ومقارنة الفروق بين المجموعات (Putnick and Bornstein, 2016)، هناك من يعتبر هذا المستوى من اللاتغير في القياس كافٍ لعمل المقارنات بين المجموعات المختلفة، وبالتالي المقياس يقيس نفس السمة الكامنة وبنفس الطريقة، لكن باحثين آخرين (Milfont and fischer, 2010; Wu et al, 2007) يذكرون بأنه غير كافٍ، ويجب توفر المستوى الأعلى.

4- **اللاتغير الصارم في القياس (Strict Invariance)**: ويُطلق عليه لاتغير البواقي (Residual Invariance)، وهو الخطوة الأخيرة في اختبار اللاتغير في القياس بعد تحقق اللاتغير العددي، ويتمثل في اختبار بواقي الفقرات المتكافئة في المستوى المتري، والعددي، ويُقصد بلاتغير البواقي أن يكون مجموع التباين النوعي (تباين الفقرة غير المشترك مع العامل)، وتباين الخطأ (خطأ القياس) متماثل عبر المجموعات، وهو مُتطلب للوصول للاتغير التام (Full Invariance)، ولكنه ليس مُتطلب قَبلي لاختبار فروق الأوساط بين المجموعات، لأن البواقي ليست جزءًا من العامل الكامن، وليس له صلة بتفسير فروق أوساط السمات الكامنة (Vandenberg & Lance, 2000)؛ لذلك يتم حذف هذه الخطوة من قبل الكثير من الباحثين، إلا أنها تذكر في تقرير بحوث اللاتغير في القياس.

ويتم اختبار لاتغير البواقي من خلال تقييد بواقي الفقرات لتصبح متماثلة عبر المجموعات، وهذا القيد يضاف للفقرات المحققة للاتغير المتري، والعددي، والفقرات التي لم تحقق فيها خاصية اللاتغير المتري، أو العددي يُسمح لها بأن لا تتقيد في نموذج لاتغير البواقي، وكما هو الحال في المستويات السابقة يُقارن النموذج المُقيد بنموذج اللاتغير العددي لتحديد المطابقة، فإذا كانت مُطابقة النموذج أسوأ بصورة دالة إحصائيًا، فإن هذا مؤشر على أنه يوجد على الأقل بواقي فقرة واحدة مختلف عبر المجموعات، وأن اللاتغير الصارم لم يتحقق، والعكس صحيح. وإذا لم يتحقق اللاتغير الصارم يُمكن للباحث أن يفحص مصدر عدم تحقق لاتغير البواقي من خلال التخلص التدريجي (باستخدام نهج الرجوع للخلف)، أو إضافة (باستخدام نهج التقدم للأمام) قيود بواقي الفقرات، وإعادة اختبار النموذج للوصول للاتغير الجزئي في القياس، أو قبول عدم تحقق اللاتغير الصارم، واستكمال اختبارات فروق أوساط العوامل الكامنة عبر المجموعات (Putnick and Bornstein, 2016).

5- **اللاتغير الجزئي في القياس (Partial Measurement Invariance)**: يتم عندما لا يمكن تحقق اللاتغير العددي للمقياس بدءًا من المستوى الثاني (اللاتغير المتري)، فإنه يمكن للباحث اختبار النموذج بتحرير بعض المعالم المقيدة، ويرى مدى تحسن النموذج، فتحسن النموذج يدل على خلل في هذه المعالم، وأنها أدت لعدم تحقق اللاتغير التام في القياس، وتُسمى هذه الإجراءات باللاتغير الجزئي في القياس، والتي طورها بايرن وآخرون (Byrne et al., 1989)؛ وهي تقوم على إفتراض وجود بعض الفقرات التي لا تحقق خاصية اللاتغير في القياس، ولكن لا تُؤثر على إمكانية المقارنة بين المجموعات، وبالتالي لا تؤدي لاختلاف السمة الكامنة بين المجموعات، كما أن القيام بهذا الإجراء يسمح بمعرفة الفقرات التي فيها خلل،

وتستدعي عرضها على الخبراء، والمحكمين لدراسة محتواها لمعرفة سبب تباينها عبر المجموعات (Brown, 2015; Byrne, 2012; Wu et al., 2007).

6- **اللاتغير التقريبي في القياس (Approximate Measurement Invariance)**: يتم عندما يكون من غير الممكن تحقق اللاتغير التام في القياس، وإجراءات اللاتغير الجزئي في القياس تكون مرهقة، ولا يمكن تطبيقها، فيسمح اللاتغير التقريبي في القياس ببعض التباين بين معالم النموذج المُفترض عبر المجموعات بحيث تكون هذه التباينات صغيرة وغير دالة إحصائيًا (Zercher at al., 2015).

اللاتغير في القياس من منظور نظرية الاستجابة للفقرة (IRT)

على الرغم من أن التحليل العاملي متعدد المجموعة هو الطريقة الأكثر استخدامًا للتحقق من اللاتغير في القياس في البحوث النفسية والتربوية، إلا أن مناهج نظرية استجابة الفقرة استخدمت أيضًا لنفس الغرض، وفي بعض الحالات قدمت معلومات أكثر فائدة في التحقق من اللاتغير في القياس (Raju, Laffitte & Byrne, 2002)، وتُشتق من النظرية الحديثة نماذج رياضية تُفسر العلاقة بين السمات الكامنة (الخصائص أو الصفات غير المشاهدة) ومظاهرها (أي الأداء أو الاستجابات أو النواتج المشاهدة)، وتُنشيء رابطة بين خصائص الفقرات التي يُجيب عليها المفحوصين، والسمة الكامنة موضوع القياس، كما تُقدم هذه النظرية أساس للطرق الإحصائية التي تستخدم في سياقات مثل: إعداد الاختبارات، وتحليل الفقرات، وبنوك الأسئلة، والاختبارات المتوافقة مع الكمبيوتر، كما تُستخدم النظرية الحديثة لتحقيق واحد أو أكثر من الأهداف التالية: تقويم الخصائص السيكومترية للمقاييس، واختبار اللاتغير في القياس في استجابات المفحوصين على المقاييس، ثم تطبيقها عبر المجتمعات المختلفة، وربط أداتين أو أكثر يقيسان مجالات متشابهة على تدريج مشترك، وإعداد اختبارات متوافقة تُقدر مكانة المفحوص على تكوين فرضي ما

بأقل عدد ممكن من الفقرات، كما تفترض نظرية السمات الكامنة نموذج علاقة منحنية وليست خطية لوصف العلاقة بين استجابات المفحوص المشاهدة، ومستوى السمة الكامنة التي يُرمز لها بالرمز (θ)، وتتحدد الطبيعة الدقيقة لهذا النموذج بواسطة فئة من معالم الفقرة التي تكون فريدة لكل فقرة، ويعتمد تقدير السمة الكامنة لمفحوص معين على استجابات الفقرات المشاهدة بمعلومية معالم الفقرات، ويُوجد معلمتان عادًة ما يتم تقديرهما لكل فقرة، التمييز الذي يمثل ميل خط مسار الفقرة (ICC) الذي يحدد العلاقة بين السمة الكامنة والدرجة المشاهدة، والنوع الثاني من المعالم هو موقع الفقرة أو مستوى الصعوبة؛ أي أن العلاقة بين احتمال أن يختار شخص ما ذو قدرة ما بديل استجابة لفقرة معينة يمكن التعبير عنه بصورة بيانية باستخدام منحنى خصائص الفقرة (علام، 2005).

يرتبط معامل التمييز في النظرية الحديثة بصورة رياضية بتشبعات العوامل (Foctor Loading) في منهج التحليل العاملي التوكيدي (McDonald, 1989)، في المقابل لا يوجد معلمة تُماثل بشكل واضح معلمة معامل الصعوبة في التحليل العاملي التوكيدي، على الرغم أن معلمة تقاطعات الفقرات (Intercepts) في التحليل العاملي التي تمثل القيمة الملاحظة للفقرة عندما تكون قيمة السمة الكامنة (Factor) مساوية للصفر هي الأكثر مشابهة لمعلمة معامل الصعوبة من الناحية المفاهيمية، كما أن هناك ثابت واحد فقط يمكن تقديره لكل فقرة في التحليل العاملي التوكيدي، إلا أن هناك عدة معالم صعوبة يمكن تقديرها لكل فقرة ضمن نظرية استجابة الفقرة، وكذلك اختبارات النظرية الحديثة للاتغير في القياس تختلف اعتمادًا على النمط النوعي للمنهجية المستخدمة، وأكثر الطرق استخدامًا هي اختبار نسبة الأرجحية (Likelihood Ratio Test)، كما يطبق اختبار الأرجحية العظمى (Maximum Likelihood Test) على مستوى الفقرة مثل منهج التحليل العاملي التوكيدي، ويستخدم تقدير الأرجحية العظمى (Maximum

Likelihood Estimation) لتقدير معالم الفقرة التي تنتج قيمة لمطابقة النموذج، تُعرف بدالة المطابقة (Fit Function) وهذه الدالة تُعد مؤشر على جودة مطابقة النموذج المُقترح للبيانات كنتيجة لإجراءات تقدير الأرجحية العظمى المستخدمة في تقدير معالم الفقرة (Camilli & Shepard,1994).

يتضمن اختبار الأرجحية العظمى (Maximum Likelihood Test) مقارنة مطابقة نموذجين؛ النموذج المدمج (Compact Model)، والنموذج الزائد المقارن، أولًا، يُقيم النموذج القاعدي (Baseline Model) حيث تُقدر معالم جميع الفقرات تحت قيد أن معالم الفقرة للفقرات المتشابهة (مثل الفقرة الأولى في المجموعة الأولى، والفقرة الأولى في المجموعة الثانية) تكون متساوية عبر المجموعات، أما النموذج المدمج (Compact Model) فيُقدم قيمة أرجحية قاعدية لمطابقة معالم الفقرة للنموذج، ثم يلي ذلك، تختبر كل فقرة واحدة في كل مرة للأداء التفاضلي للفقرة (Differential Item Functioning) (DIF) بصورة منفصلة (Thissen, Steinberg & Wainer, 1988).

في نظرية استجابة الفقرة (IRT) يُعرف تحليل اللاتغير في القياس عادةً على أنه اكتشاف الأداء التفاضلي للفقرة (DIF)، ويُوجد أداء تفاضلي للفقرة عندما تكون العلاقة بين درجة المفحوص الملاحظة على الفقرة، والسمة الكامنة ليست متطابقة عبر المجموعات الفرعية (Drasgow, 1984)، ويُمثل الأداء التفاضلي للاختبار (Differential Test Functioning) بالأداء التفاضلي التراكمي لكافة الفقرات على مقياس ما، حيث العلاقة بين الدرجة على الاختبار والسمة الكامنة غير متطابقة عبر المجتمعات الفرعية (Penfield & Camilli, 2007)

ولاختبار الأداء التفاضلي للفقرة، يتم إجراء تحليل منفصل لكل فقرة حيث تُقيد جميع تقديرات الفقرات المتماثلة لتصبح متساوية عبر المجموعات، باستثناء معالم الفقرة موضع الاختبار

للأداء التفاضلي، يُقدم النموذج الزائد قيمة أرجحية مصاحبة لتقدير معالم الفقرة بصورة منفصلة لكل مجموعة، هذه القيمة للأرجحية يمكن مقارنتها مع قيمة الأرجحية للنموذج المدمج حيث المعالم لجميع الفقرات المتماثلة مُقيدة لتكون متساوية عبر المجموعات، وتعطي المعادلة التالية صيغة اختبار نسبة الأرجحية:

$$LR_i = \frac{L_C}{L_{Ai}}$$

حيث L_C: تمثل دالة الأرجحية للنموذج المدمج (الذي يحتوي على عدد أقل من المعالم)، أما L_{Ai} فيرمز لدالة الأرجحية للنموذج الزائد (الذي يسمح للفقرة محل الاهتمام أن تختلف عبر المجموعات)، وعند استخدام اختبار نسبة الأرجحية يتم حساب قيمة مربع كاي لكل فقرة في الاختبار، كما في المعادلة التالية:

$$\chi^2(M) = -2ln(LR) = -2ln(L_C) + 2ln(L_{Ai})$$

حيث M: تساوي الفرق في عدد معالم الفقرة المُقدرة في النموذج المُدمج مقابل النموذج الزائد (أي الفرق في درجة الحرية بين النموذجين)، وهذا يُعد مؤشر على سوء المطابقة، لأنه إذا كان اختبار مربع كاي دال إحصائيًا، فإن النموذج المدمج يتطابق بصورة أكثر سوءً مع النموذج الزائد؛ أي أن الفقرات ذو قيم مربع كاي الدالة فإنها تُظهر أداءً تفاضليًا (أي استخدام تقديرات معالم فقرة مختلفة يحسن المطابقة العامة للنموذج)، واختبار نسبة الأرجحية يتم استخدامه عن طريق عدة خطوات متسلسلة ولكن بطريقة آلية وهذا ما جعلها أسهل من الطرق التقليدية بالنسبة للباحثين، كما أنها قادرة على اكتشاف الأداء التفاضلي للفقرة عبر المجموعات أيضًا، وأثناء ذلك وُجد أن معدلات الخطأ من النوع الأول لمعظم النماذج التي تم فحصها تقع ضمن المستويات المقبولة (Cohen, Kim, and Wollack, 1996; Thissen, 2001).

أوجه التقارب والاختلاف بين المنظورين

اللاتغير في القياس في التحليل العاملي التوكيدي، والأداء التفاضلي للفقرة في نظرية استجابة الفقرة يحملان تشابهات، واختلافات، حيث يُعبر الأداء التفاضلي للفقرة عن عدم تحقق اللاتغير في القياس، الذي يختبر ما إذا كان المفحوصين ذوي القدرة المتساوية لهم احتمالات مختلفة للإجابة بشكل صحيح على الفقرة، فإذا كانت الفقرة تعمل بشكل مختلف عبر المجموعات، فيُوضع عليها علامة على أنها فقرة متحيزة، فطرق الكشف عن الأداء التفاضلي للفقرة حسب النظرية الحديثة تبحث ما إذا كانت الفقرة تُفسر بطريقة متماثلة عبر المجموعات، وهذا يشبه اللاتغير المتري والعددي في التحليل العاملي التوكيدي متعدد المجموعة؛ أي أنه يمكن استخدام الأداء التفاضلي في النظرية الحديثة كمصطلح مُرادف للاتغير في القياس في التحليل العاملي التوكيدي (Suh & Cho, 2014; Tay, Meade & Cao, 2015) إلا أنه يوجد فروق في الخطوات التطبيقية بينهما، حيث أن التحليل العاملي التوكيدي ونظرية استجابة الفقرة لاختبار اللاتغير في القياس مختلفان من الناحية التاريخية.

فالهدف غالبًا من الكشف عن اللاتغير في القياس باستخدام التحليل العاملي التوكيدي يكون بمحاولة إثبات أن الدرجة الكلية الملاحظة عبر المجموعات (الدرجات الخام للمقياس) قابلة للمقارنة، ويمكن تفسير الفروق بينها تفسيرات جوهرية عبر المجموعات، كما أنه يستخدم كمتطلب قَبْلي لمقارنة أوساط المجموعات على المتغير موضع الاهتمام، لذلك يكون التركيز على مستوى بنية البناء الإفتراضي أكثر من أداء المفحوصين على فقرات الاختبار (Mencl at al., 2012).

في المقابل الهدف من اختبار اللاتغير في القياس باستخدام النظرية الحديثة لتقييم اللاتغير في القياس للفقرات، أو بشكل أدق عدم اللاتغير في القياس للفقرات، وذلك لأن النظرية الحديثة تم تطويرها لتحديد فقرات الاختبار النوعية التي قد تُظهر تحيز ضد مجموعة معينة (خاصية هامة

لبنوك الأسئلة)، بنفس الوقت يستطيع الباحث دراسة اللاتغير في القياس على مستوى المقياس ككل وليس فقط على مستوى الفقرة، ويتم ذلك بتحديد التأثير المتراكم للأداء التفاضلي للفقرات على الفروق في أوساط الدرجات الملاحظة عبر المجموعات (أي الأداء التفاضلي للاختبار)، وذلك عندما يكون منحنى خصائص الاختبار مختلف عبر المجموعات، لأن النظرية الحديثة تهتم بالبناء التطبيقي للاختبار؛ أي أن الاهتمام بالفروق في البناء العاملي الافتراضي بين المجموعات سيكون أقل عند عدم تحقق اللاتغير في القياس، أي أن الأداء التفاضلي للفقرة يستخدم في حالات تحديد القضايا المرتبطة بمحتوى وصياغة الفقرة (Hambleton, Swaminathan & Rogers, 1991).

كما أن التأكيدات المختلفة بين نظرية استجابة الفقرة، والتحليل العاملي التوكيدي قادت لاختلافات في المفاهيم والإجراءات رغم وجود التشابهات الضمنية بينهما، ومن أبرز الفروق يكون في خطوات الكشف عن اللاتغير في القياس، فيتم الكشف عن اللاتغير في القياس في التحليل العاملي التوكيدي بشكل تدريجي، بحيث يتم دراسة أشكال اللاتغير من التكويني إلى المتري ثم العددي (Vandenberg & Lance, 2000)، وفي كل خطوة توضع قيود أكثر على نموذج ما، وإذا وُجد أن المطابقة أصبحت أسوأ بشكل دال إحصائيًا، فهذا يعني أن شكل اللاتغير موضع الاهتمام لم يتحقق، ويتم ربط أشكال اللاتغير بتفسيرات جوهرية، على سبيل المثال، عند عدم تحقق اللاتغير التكويني لمقياس ما فهذا يشير إلى أن المفهوم محل الاهتمام يحمل تعريفات مختلفة بصورة جوهرية عبر المجموعات، وهذا ما يقابل افتراض أحادية البُعد في نظرية استجابة الفقرة، وهذا ما يجعل إحدى طرق التحقق من أحادية البُعد تطبيق بعض خطوات طريقة التحليل العاملي التوكيدي (Cheung & Rensvold, 2002).

أما للكشف عن اللاتغير في القياس باستخدام نظرية استجابة الفقرة فيتم إجراء تحليل على مستوى الفقرة، ويتم ذلك بمقارنة نماذج النظرية الحديثة باستخدام أحد المدخلين: الخط الأساسي الحر (Free-baseline) أو الخط الأساسي المُقيد تمامًا (Fully constrained baseline approach)، وكلا المدخلين يمكن تطبيقهما عند استخدام التحليل العاملي التوكيدي (Stark, Chernyshenko & Drasgow, 2006).

الفرق الجوهري الثاني، هو أنه عند البحث عن الأداء التفاضلي للفقرة يكون للكشف عن وجوده وتأثيره على الأداء التفاضلي للاختبار الذي له تأثير جوهري على مقارنة درجات الاختبار الملاحظة، وليس لدراسة طبيعة الأداء التفاضلي للفقرة أو نوعيته (Stark et al., 2004)، وفي العديد من إجراءات نظرية استجابة الفقرة للكشف عن الأداء التفاضلي للفقرة ضعف في الكشف عن الأداء التفاضلي غير المنتظم (Nonuniform DIF) بسبب عدم تأثيره على الدرجات الملاحظة للاختبار؛ لأنه قد يتم تعويض الدرجات المنخفضة على سمة كامنة ما بالدرجات المرتفعة على مستوى سمة كامنة أخرى، أما عند استخدام التحليل العاملي التوكيدي فإن الاهتمام الجوهري بمقارنة السمات الكامنة وتأثرها بالعلامات الملاحظة يتم الاهتمام به جنبًا إلى جنب عند دراسة اللاتغير في القياس (Finch & French, 2007; Tay, Vermunt & Wang, 2013).

الفرق الأخير، لاكتشاف اللاتغير في القياس باستخدام نماذج نظرية استجابة الفقرة يتم استخدام مؤشرات مستوى الفقرة، وهي أكثر اتساعًا من اختبار مطابقة معالم الفقرة (Lord, 1980)، وتشمل استراتيجيات مختلفة منها: المساحة بين منحنيات الاستجابة للفقرة، والانحدار اللوجستي، وفروق العلامة الملاحظة (Holland & Thayer, 1988; Narayanan & Swaminathan, 1994)، في المقابل عند استخدام التحليل العاملي التوكيدي للكشف عن

اللاتغير في القياس يتم استخدام مؤشرات مطابقة النموذج للبيانات المستخدمة في نمذجة المعادلة البنائية (SEM)، وهي: (CFI, TLI, RMSEA) (Cheung & Rensvold, 2002).

ولضمان قابلية مقارنة أوساط العامل (Factor) بين المجموعات فإن اللاتغير العددي يجب أن يتحقق (Millsap, 2011)، ولكن غالبًا ما يتم رفض النموذج العاملي المقترح لعدم تحقيقه لشروط اللاتغير العددي، لذلك يتم اللجوء لدراسة اللاتغير المتري وتخفيف القيود المفروضة على معالم النموذج العددي؛ أي أنه يتم اللجوء لدراسة اللاتغير الجزئي (Sörbom, 1989).

ويجب التحقق من خاصية اللاتغير في القياس قبل استخدام مقاييس واسعة النطاق مع مجموعة متنوعة من الأهداف مثل: مقاييس مقارنات الإنجاز للطلبة الدوليين (Programme for International Student Assessment) (PISA)، ومقاييس اتجاهات الطلبة نحو دراسة العلوم والرياضيات (Trends in International Mathematics and Science Study) (TIMSS)، ومقاييس دراسة التقدم في القراءة والكتابة الدولية (Progress in International Reading Literacy Study) (PIRLS)، أو قبل استخدام أي من المقاييس المختلفة المستخدمة في الدراسات الاجتماعية التي تهتم بقياس السمات الكامنة عبر الثقافات أو المجموعات مثل: برنامج المسح الاجتماعي الدولي (International Social Survey Programme) (ISSP)، والمسح الاجتماعي الأوروبي (European Social Survey) (ESS)، لأن من أهداف هذه الدراسات مقارنة أوساط العامل (Factor) الذي يمثل السمات الكامنة خلف أداء المفحوصين، لذلك يجب التأكد من قابلية مثل هذه المقارنات، وذلك بدراسة اللاتغير في القياس لمثل هذه الأدوات المستخدمة، ولكن يجب العلم أنه قد تكون طريقة التحليل العاملي التوكيدي متعدد المجموعة (MG-CFA) مرهقة للغاية مع كثرة المجموعات، وكثرة الانتهاكات المحتملة للاتغير في القياس، وتطبيق مجموعة متتالية من مؤشرات التعديل (McArdle, 1996).

ويعتبر أسلوب التحليل العاملي متعدد المجموعة إحصائي لتحليل البيانات، والكشف عن اللاتغير في القياس، لكنه غير قادر على الكشف عن مكامن الخلل فيها، فعلى الباحث أخذ بعض المحاذير بعين الاعتبار للحصول على أدق النتائج أهمها: يتطلب التحليل العاملي التوكيدي معاينة جيدة للمفحوصين، وأن تكون العينة كبيرة وممثلة، كما أنه يجب أن تكون المتغيرات الملاحظة من مستوى القياس الفئوي، بالإضافة لأن يجب أن تكون نمذجة البُنى الافتراضية للنماذج ممثلة للنظريات المستقاة منها، أما الإحصائيات المستخدمة في التحليل العاملي التوكيدي على فحص مطابقة النموذج المُقترح للبيانات فقط ولا تساعد في اختيار النموذج الأنسب (McArdle, 1996).

طريقة المواءمة (Alignment Method)

هي طريقة إحصائية جديدة بديلة لطرق التحليل العاملي التوكيدي متعدد المجموعة (MG-CFA) متخصصة في الكشف عن اللاتغير في القياس، ومقارنة أوساط وتباينات العوامل عبر المجموعات، وذلك بتقدير تباينات وأوساط عامل مجموعة محددة، دون اللجوء لاختبار اللاتغير في القياس الصارم (Strict)؛ أي أنها تُحسن الفروق الدالة إحصائيًا في أقل عدد ممكن من الفقرات لتحافظ غالبية الفقرات الأخرى باختلافات عادية مسموح بها (غير دالة إحصائيًا) عبر المجموعات (Asparouhov & Muthén, 2014, p. 497).

وهي تشبه في طريقة عملها طريقة عمل التدوير في التحليل العاملي الاستكشافي (Exploratory Factor Analysis) (EFA) لتحسن المواءمة بين معالم الفقرات عبر المجموعات، بنفس الوقت تأخذ درجة جودة مطابقة النموذج التكويني للبيانات، ولا تفرض قيودًا على تساوي تقديرات المعلمة بل تسمح بوجود اختلافات مسموحة قدر الإمكان بين معالم النموذج المُقترح عبر كافة المجموعات حتى يتم الوصول لأفضل نموذج تكويني ممكن بتقدير أوساط وتباينات العامل (Factor) لكل مجموعة، وأثناء عملية التعديل لإيجاد النموذج التكويني الأمثل يتم

تقليل الفروق بين معالم النموذج المُقترح (تقاطعات، وتشبعات العامل) لكل زوج من المجموعات الممكنة ما أمكن.

وتتم هذه العملية بخطوتين: الأولى: تسمى بتحسينات المواءمة الثابتة (Fixed Alignment Optimization)، وفيها يتم تقدير معالم النموذج التكويني، بحيث تكون معاملات التشبع، والتقاطع للعامل (Factor) حرة (Free) عبر المجموعات، أما أوساط العامل في كافة المجموعات مثبتة عند قيمة الـ(0)، وتباينات العامل في كافة المجموعات مثبتة عند قيمة الـ(1)، والنموذج الناتج هو نموذج الأساس (Baseline Model)، ويكون أفضل نموذج ملائم للبيانات من بين جميع النماذج الممكن الحصول عليها باستخدام طريقة التحليل العاملي التوكيدي متعدد المجموعة (MG-CFA)، أما الخطوة الثانية: فتُعرف بتحسينات المواءمة الحرة (Free Alignment Optimization)، ويتم فيها تحرير أوساط العامل، وتبايناته، واختيار قيمهم التي تُقلل من الفروق الكلية بين معالم نموذج الأساس (Baseline Model) عبر كافة المجموعات، وذلك بتطبيق مقارنات زوجية من المجموعات المختلفة الممكنة، وبذلك نحصل على نموذج المواءمة النهائي (Align Model) (Jennrich, 2006).

بهذه الطريقة نحصل على نموذج نهائي (Align Model) مطابق للبيانات محقق لخاصية اللاتغير في القياس دون الحاجة لإجراء كافة خطوات التحليل العاملي متعدد المجموعة، كما أن هذه الطريقة تدعم أعدادًا كبيرة من المجموعات قد تصل لـ(100) مجموعة، والنموذج الذي يتم الحصول عليه يتم بطريقة آلية لا تحتاج لباحث متخصص لتطبيقها، وتفسير نتائجها، كما تُحدد المعالم التي لا يوجد بينها فروق دالة إحصائيًا؛ أي تحقق خاصية اللاتغير في القياس عبر المجموعات، وتُرتب هذه المجموعات حسب أوساط العامل؛ أي أنها تقارن المجموعات مباشرةً، وتميز المجموعات التي تختلف معالمها عن وسط مجموعة المقارنة بشكل دال إحصائيًا بأنها

تظهر بين أقواس، ومن هنا تظهر أهمية طريقة المواءمة (Alignment Method) في تحديد الفقرات المحققة لخاصية اللاتغير في القياس عبر المجموعة باستخدام اختبار الملاءمة (R^2) حيث يتم تفسير قيمته كلما اقترب من (1) تكون معالم الفقرة أكثر تحقيقًا لخاصية اللاتغير في القياس عبر المجموعات (Asparouhov & Muthén, 2014).

وقد تم اختيار طريقة المواءمة (Alignment Method) للكشف عن اللاتغير في القياس في هذه الدراسة بسبب إنطباق معايير ميوثن واسباروف (Muthén & Asparouhov, 2017) في الكشف عن اللاتغير في القياس على الاختبار التقييمي لمهارات الرياضيات لطلبة الصف الثالث وهي: أن عدد المؤشرات (الفقرات) الدالة على العامل (Factor) ليست كبيرة جدًا، أما عدد المجموعات فهي تناسب حتى (100) مجموعة بأي عدد مفحوصين، كما أن طريقة المواءمة تقدم معلومات حول مجموعات المقارنة، ولا تتأثر باختلاف توزيعات المعالم، ومن أهم المعايير التي تحققها طريقة المواءمة هي: التعامل مع البيانات الناتجة من التصاميم المعقدة للمسوح، وكذلك تمتاز بالسرعة الحسابية لأنها تعتمد على أتمتة كافة العمليات الحسابية، ولا تحتاج لباحثين مختصين لاستخدامها.

الاختبار التقييمي لمبحث الرياضيات للصف الثالث الأساسي للعام الدراسي (2017-2018)

نظرًا لأهمية الاختبارات وبنوك الأسئلة في مجال قياس السلوك الإنساني والنفسي وما ينطوي عليها من قرارات ونتائج وتعديل وتطوير، تعددت مجالات البحث المتعلقة بها وبكيفية بنائها وكيفية تطويرها، كما تعددت استخداماتها في المجالات النفسية والتربوية، وأصبح يعتمد على العديد منها في اتخاذ القرارات الضرورية في كافة القطاعات العامة والخاصة، وأهمها قطاع وزارة التربية والتعليم، إذ يُعتمد عليها في العديد من التطبيقات التربوية كاختبارات القدرة المتنوعة، واختبارات الكفاءة وغيرها من الاختبارات المختلفة.

فتسعى وزارة التربية والتعليم الأردنية لإكساب الطلبة أعلى مستويات المهارات التعليمية، وخصوصًا طلبة المراحل الأولى المتمثلة بالصفوف الثلاثة الأولى الأساسية؛ بوصفها المرحلة الأساس التي تحدد مستوى الطالب الدراسي في المراحل التعليمية اللاحقة التي تهدف إلى رفع مستوى الطالب في جميع المهارات الواجب على الطالب اكتسابها.

وإيمانًا من مديرية الاختبارات بأهمية التقويم ودوره في الكشف عن جوانب الضعف لدى المتعلمين أعدت مديرية الاختبارات في العام الدراسي (2018/2017) اختبارًا تقييميًا في مبحث الرياضيات يهدف للكشف عن مدى اكتساب طلبة الصف الثالث الأساسي المهارات الرئيسة في مبحث الرياضيات، وطُبق الاختبار على عينة من طلبة الصف الثالث الأساسي، وحُللت النتائج، وأُصدر تقرير بالنتائج على مستوى كل من المديرية، والمدرسة، وعُمم على المشرفين التربويين والمعلمين، وطُلب منهم دراسة التقرير والوقوف على جوانب الضعف التي أظهرتها النتائج، وإعداد الخطط العلاجية على مستوى كل من المديرية والمدرسة لمعالجة جوانب الضعف لدى الطلبة (وزارة التربية والتعليم الأردنية، 2018، ص6).

كما يُعد التقويم أحد المداخل الأساسية لتطوير عملية التعليم، فهو يُساعد على الكشف عن سير العملية أيضًا، وعلى تحديد مواطن القوة والضعف في العملية التعليمية، ومن ثم التوصل إلى طرق العلاج المناسبة لدى المتعلمين، ويهدف الاختبار إلى: تعّرف المستوى التعليمي لطلبة الصف الثالث الأساسي في المهارات، والقدرات الأساسية في مبحث الرياضيات، ولتزويد المعلمين بمعلومات عن جوانب القوة والضعف في أداء طلبتهم تُعينهم على متابعة الطلبة، ووضع الخطط العلاجية المناسبة وفقًا لنتائج الاختبارات، كما يُساعد في اتخاذ القرارات المتعلقة بالمناهج الدراسية على أسس واقعية، ومعلومات صحيحة بحيث يتم التركيز في المناهج الدراسية على تلك الجوانب،

والمهارات التي أظهرت نتائج الاختبار ضعفًا في أداء الطلبة عليها (وزارة التربية والتعليم الأردنية، 2018، ص9).

واختبار الرياضيات هو الاختبار الذي سيتم دراسة مدى تحقيقه لخاصية اللاتغير في القياس عبر المجموعات في هذه الدراسة، وهذا الاختبار يتكون من (25) فقرة من نوع الاختيار من متعدد موزعة على محاور التعلم في مبحث الرياضيات وهي: الأعداد والعمليات، والجبر (الأنماط)، والقياس، والهندسة، والإحصاء والاحتمالات، أما المهارات التي شملها فهي: تمثيل الأعداد وقراءتها وكتابتها، مقارنة الأعداد وتقريبها، وإجراء العمليات الحسابية على الأعداد، وإكمال النمط، وتحديد وحدات القياس المختلفة، وتمييز الأشكال الهندسية والمجسمات والعلاقات بينها، وجمع البيانات وتسجيلها وتنظيمها بجدول الإشارات والصور، وتطبيق مفهومي التغير والنسبة (وزارة التربية والتعليم الأردنية، 2018، ص12).

ولإعداد الاختبار شُكلت لجنة لكتابة فقرات اختبار الصف الثالث الأساسي في مبحث الرياضيات، من مشرفي إدارة الامتحانات والاختبارات، وعدد من مشرفي الميدان اتبعت فيها الخطوات الآتية:

١- **بناء الفقرات:** تحليل محتوى كتاب الصف الثالث الأساسي في مبحث الرياضيات، ودراسة نتاجات التعلم الخاصة بالصف الثالث الأساسي في المبحث المستهدف في الاختبار، وتحديد محاور ومهارات التعلم التي سيقيسها الاختبار، وإعداد جدول مواصفات للاختبار، وبناء تجمع من الفقرات الاختبارية وطباعتها.

٢- **تجريب الفقرات:** جُربت الفقرات على عينة تجريبية من الطلبة، وأجري التحليل الإحصائي لاستجابات الطلبة على الفقرات، للتعّرف على مقروئيتها، وخصائصها السيكومترية

مثل: فاعلية البدائل، ومستوى صعوبة الفقرات، وتمييزها، ثم أُختيرت الفقرات الملائمة، وعُدلت صياغة بعض الفقرات، وحُذفت الفقرات التي كان تمييزها سالبًا، أو قريبًا من الصفر.

٣- **إعداد الاختبار في صورته النهائية:** أُنتقيت الفقرات النهائية من تجمع الفقرات التي بُنيت، بحيث كان الانتقاء ممثلاً للأوزان والنسب في جدول المواصفات الذي طور ليشمل مؤشرات الأداء (وزارة التربية والتعليم الأردنية، 2018، ص13).

مشكلة الدراسة وأسئلتها

إن التغيرات المستمرة في مجال القياس والتقويم أدت إلى ظهور منهجيات إحصائية حديثة للمقارنة بين المدارس، أو المناطق المختلفة باستخدام أدوات قياس على شكل اختبارات معرفية، أو مقاييس شخصية، وكان يجب التأكد من معقولية افتراض عدم تحيز هذه الأدوات عند تطبيقها في ظروف، وبيئات مختلفة، وخصوصًا عند استخدام الاختبارات الوطنية عبر السياقات التعليمية واسعة الاختلاف، هذه القضية بارزة بشكل خاص عند العديد من الدول مثل الأردن التي تطبق الاختبارات على مستوى واسع النطاق، مثل اختبار التميس (TIMSS) والبيزا (PISA)، والاختبار التقييمي لطلبة الصف الثالث لمهارات الرياضيات، حيث يتم إدارة هذه التقييمات على المستوى الوطني، ومختلف المدارس الابتدائية، والثانوية، ومن المؤكد أن ترغب المدارس، والمديريات في معرفة ترتيبها في تعزيز هذه المهارات، لذلك تم اختيار متغير مديريات التربية والتعليم الأردنية وعددها (45) مديرية محددة بتوزيع المناطق الجغرافية في الأردن كمتغير رئيس لهذه الدراسة، ولصحة هذه المقارنات يجب تحقق خاصية اللاتغير في القياس (ثبات القياس) للاختبار، حيث أن هذه الخاصية تعني أن خصائص الاختبار لا تتغير عبر السياقات، أو المجموعات المختلفة، وهي تفترض ضمنًا أنه يتم استخدام نماذج قياس للإجابة عن أسئلة الاختلاف في السمة (المهارات

الحسابية) عبر الزمن، أو المجموعات (المديريات)، فمن الضروري أن يكون الاختبار المُستخدم يقيس نفس السمة (المهارات) عبر المديريات المختلفة؛ أي أن اختلاف المجموعات (المديريات) لا تؤثر في طريقة عمل الاختبار؛ لأنه إذا كانت خصائص الاختبار تتغير عبر المجموعات، فلا يمكننا معرفة مصدر الاختلاف إن كان اختلاف في السمة المقيسة لدى المفحوصين، أم الاختبار نفسه؛ لذلك يتم فحص تحقق خاصية اللاتغير في القياس (ثبات القياس)، والكشف عنها قبل إجراء عملية المقارنات المختلفة بين المفحوصين باستخدام البيانات الناتجة من الاختبار.

وكثيرةٌ هي الدراسات التي تناولت التحقق من اللاتغير في القياس لمقاييس مختلفة في ضوء متغيرات متنوعة؛ كالجنس، والعمر، والمرحلة الدراسية، والعرق، والسكن، وغيرها من المتغيرات المحدودة المجموعات، إلا أن الباحثة لم تجد أي دراسة عربية، أو أجنبية – في حدود علم الباحثة– تناولت موضوع اللاتغير في القياس (ثبات القياس) للاختبار التقييمي لمهارات الرياضيات، حيث يُعد اختبارًا له أثره في توجيه أصحاب القرار في وزارة التربية والتعليم الأردنية، التي تبحث عن البرامج التعليمية الفاعلة عالية التأثير على ممارسات الطلبة ومهاراتهم، والتدخلات العملية لإنجاحهم في المراحل التأسيسية الأولى من أعمارهم، كما أن الباحثة لاحظت أن الدراسات العربية المختلفة تطرقت بشكل محدود لاستخدام طرق كشف مختلفة عن طريق التحليل العاملي متعدد المجموعة (MG–CFA) عبر مجموعات محدودة لا تتجاوز (3) مجموعات، وبما أن الاختبار التقييمي موضع اهتمام الدراسة يعتبر من الاختبارات واسعة النطاق؛ الذي يحتاج للكشف عن تحقيقه لخاصية اللاتغير في القياس إجراء فحص لأجزاء عديدة من نموذج القياس المُقترح؛ وهذه العملية متعددة الخطوات، ولا تخلو من الصعوبات بسبب العدد الكبير للمجموعات؛ الذي سيؤدي لصعوبة الوصول للنموذج الملائم لمثل هذه البيانات (MacCallum, Roznowski, & Necowitz, 1992).

لذلك ستستخدم الباحثة طريقة جديدة على الدراسات العربية لفحص اللاتغير في القياس للاختبار التقييمي لمهارات الرياضيات، وهي طريقة المواءمة (Alignment Method) للحصول على نتائج أكثر تفاؤلًا، والوصول بشكل أسرع، وأبسط للنموذج المُلائم للبيانات الناتجة من الاختبار التقييمي، كما يمكن استخدامها على عدد كبير من المجموعات قد يصل لمئة مجموعة، كما أنها يمكن تطبيقها على بنوك الأسئلة؛ لأنها تجمع بين النظرية الكلاسيكية والحديثة من ناحية المنشأ والإجراءات.

كما سيتم بيان إجراءات طريقة المواءمة كطريقة جديدة عندما تكون البيانات ثنائية، وهو أمر شائع في اختبارات التحصيل، والاختبارات المهارية، والتقييمات النفسية، وبيان طريقة تفسير نتائجها؛ لأنها تُزود ببيانات تفصيلية عن فقرات الاختبار، ومجموعات المقارنة، ثم سيتم تطبيق طريقة المواءمة للكشف عن اللاتغير في أداء طلبة الصف الثالث الأساسي على الاختبار التقييمي لمهارات مبحث الرياضيات عبر مديريات التربية والتعليم الأردنية، حيث أظهرت دراسات المحاكاة الأجنبية التي درست طريقة المواءمة أنها تعمل بشكل جيد مع المقاييس ذات الفقرات المتصلة (Continuous Items)، وكذلك المقاييس واسعة النطاق، حيث تقوم طريقة المواءمة بأتمتة تقدير نماذج القياس، ومن ثم تحدد المجموعات والفقرات غير الملائمة للنموذج، والتي من شأنها أن تجعل العديد من المقارنات الجماعية أكثر بساطة وسهولة (Asparouhov & Muthén, 2014).

لذلك فإن جزءً من أهداف هذه الدراسة تعليمي في المقام الأول، وذلك بتوجيه أنظار الباحثين لأهمية هذا المفهوم، وكيف يمكن اختباره باستخدام طريقة المواءمة على مقاييس واختبارات تتكون من فقرات ثنائية الاستجابة (Dichotomus) ومطبقة على عينة واسعة النطاق من المفحوصين، ولتحقيق ذلك سيتم تطبيق طريقة المواءمة للكشف عن اللاتغير في أداء الطلبة

على الاختبار التقييمي لمهارات الرياضيات لطلبة الصف الثالث الأساسي عبر مديريات التربية والتعليم الأردنية، وتوضيح خطواتها خطوة خطوة.

وبالتحديد فإن هذه الدراسة سعت للإجابة عن الأسئلة التالية:

السؤال الأول: هل البناء العاملي المُقترح لمحتوى اختبار مهارات الرياضيات للصف الثالث الأساسي يُطابق البيانات بشكل كافٍ؟

السؤال الثاني: هل يختلف البناء العاملي للاختبار التقييمي لمهارات الرياضيات للصف الثالث الأساسي عبر مديريات التربية والتعليم الأردنية؟

السؤال الثالث: هل يختلف البناء العاملي للاختبار التقييمي لمهارات الرياضيات للصف الثالث الأساسي عبر جنس الطالب؟

السؤال الرابع: ما مدى فاعلية طريقة المواءمة في الكشف عن اللاتغير في أداء طلبة الصف الثالث الأساسي على الاختبار التقييمي لمهارات الرياضيات عبر مديريات التربية والتعليم الأردنية؟

أهمية الدراسة

تتمثل أهمية الدراسة في عاملين: الأهمية النظرية التي من أهمية خاصية اللاتغير في القياس، وحداثتها، كما ترجع لندرة البحوث العربية التي اهتمت بهذا المفهوم، وإنعدام البحوث العربية – حسب علم الباحثة – التي تطرقت لطريقة المواءمة (Alignment Method)، وكذلك عدد وكم البحوث التي تهتم بعملية المقارنة، حيث أن هذا النوع من البحوث هو الأكثر انتشارًا بين أنواع البحوث الأخرى، كما أن التحقق من اللاتغير في القياس لا يقتصر على البحوث الوصفية، ولكن يمتد للدراسات التجريبية، سواء تلك التي تستخدم منهج المجموعة الواحدة، أو المجموعتين، كما أن اللاتغير في القياس يكون له دور هام في الدراسات عبر الثقافات، وواسعة النطاق، لأن التحقق من اللاتغير في القياس للأدوات المستخدمة في عملية المقارنة قبل استخدام الدرجات الخام

عملية سوف تسهم في تحسين جودة، ونتائج البحوث، وجعلها أكثر دقة، كما وتُقدم الدراسة اختبارًا تقييميًا لمهارات الرياضيات مطبق على نطاق واسع، وعلى عينات كبيرة من طلبة الصف الثالث الأساسي لفحص تحقيقه لخاصية اللاتغير في القياس، وإلقاء الضوء على طبيعة فقراته، واكتشاف نقاط القوة والضعف فيها، وذلك للحكم عليها من حيث مدى ملاءمتها لتحقيق الأهداف التي وضعت من أجلها، كما تُقدم الدراسة الحالية إطارًا نظريًا يتضمن مجموعة من الحقائق النظرية، والإحصائية المتعلقة بخاصية اللاتغير في القياس لمقاييس واختبارات أخرى، وشرح مُبسط لخطوات طريقة المواءمة، وتفسير نتائجها، وهذا كله يعكس الأهمية النظرية للدراسة.

أما الأهمية العملية فتتمثل فيما يلي: 1- توفير المزيد من مؤشرات الصدق للاختبار التقييمي لمهارات الرياضيات لدى مجتمع طلبة الصف الثالث الأساسي في الأردن.

2- تشجيع المسؤولين، ومصممي، وواضعي فقرات الاختبار التقييمي لمهارات الرياضيات للصف الثالث الأساسي للإرتقاء بها إلى الأفضل من خلال الاستفادة من هذه الدراسة ونتائجها عند تطوير الاختبار.

3- توجيه بناة الاختبارات الوطنية لدراسة تحقق خاصية اللاتغير في القياس كمتطلب قبلي لدراسات المقارنة عبر المجموعات أو الأزمان.

4- توفير مصدر معلوماتي يستفيد منه الباحثون في مجال تطوير، وتقييم المقاييس المختلفة من خلال توفير الأدب التربوي المناسب.

5- عرض طريقة حديثة في الكشف عن خاصية اللاتغير في القياس عبر المجموعات، وإجراء المقارنات بشكل سريع ومبسط.

التعريفات الإصطلاحية والإجرائية

اللاتغير في القياس (ثبات القياس) (Measurement Invariance) (MI)

يُقصد باللاتغير في القياس أن نفس التكوين الفرضي (السمة) يُقاس بطريقة مماثلة عبر المجموعات، والسؤال الرئيسي خلف اختبارات اللاتغير في القياس عبر المجموعات هو ما إذا كانت المعالم التحليلية العاملية مثل: التشبعات، والتقاطعات، وتباينات الخطأ، وتباينات العوامل، والتباين المشترك للعوامل، وأوساط العوامل يُمكن افتراض أنها متكافئة عبر المجموعات، ويتحقق اللاتغير في القياس عندما تقيس أداة ما نفس التكوين الفرضي (السمة) بنفس الطريقة عبر المجموعات المختلفة من المفحوصين؛ وفي هذه الحالة يُفترض أن علامات المقياس تحمل نفس المعنى عبر المجموعات (Terluin et al., 2016)، وتُعرفه الباحثة إجرائيًا بأنه درجة مُطابقة النموذج المُقترح الذي يُعبر عن العلاقة بين العلامات المشتقة من أدوات القياس (الاختبار)، والتكوينات الافتراضية (المهارات) التي يتم قياسها عبر المجموعات (المديريات)؛ أي أنه تم إدراك وتفسير محتوى فقرات الاختبار بشكل متماثل، أو متكافيء من قبل مختلف المفحوصين في مختلف المجموعات.

التحليل العاملي الاستكشافي (Exploratory Factor Analysis) (EFA)

هو أسلوب من أساليب التحليل العاملي يهدف لاستكشاف المجموعة المثلى التي يمكن أن تتضمن المتغيرات الكامنة ودون اعتبار مسبق لصياغة الفروض النظرية.

التحليل العاملي التوكيدي (Confirmatory Factor Analysis) (CFA)

هو إجراء إحصائي متعدد المتغيرات يستخدم للتحقق من جودة تمثيل عدد من المتغيرات الملاحظة (الفقرات) لعدد من التكوينات الفرضية (المهارات)، ويُعد أسلوب إحصائي مرن وقوي، يركز على نمذجة العلاقة بين المتغيرات الملاحظة والتكوينات الافتراضية الكامنة، من أشهر أنواعه:

التحليل العاملي التوكيدي من الدرجة الأولى (First Order Factor)

سُمي بذلك لوجود عامل واحد كامن، أو عدة عوامل كامنة ترتبط بمؤشراتها الخاصة بها، وعدم احتوائها على عامل عام تندرج العوامل الأخرى تحته (القهوجي، 2018)، وفي هذه الحالة لا يفترض الباحث وُجود عامل عام تنطوي العوامل السابقة تحته؛ أي أن عوامل الدرجة الأولى تصبح غير كافية بذاتها، بل تشتق قدرًا كبيرًا من دلالتها ومغزاها من هذا العامل العام، والنموذج العاملي التوكيدي من الدرجة الأولى قد يحتوي على عامل واحد ترتبط فيه جميع المؤشرات المُقاسة، وهذا النموذج يُسمى بالنموذج العاملي وحيد البُعد، وقد يحتوي على بنية عاملية تتكون من عاملين، أو عدة عوامل، وحينئذ يُسمى بالنموذج العاملي متعدد الأبعاد (تغيزة، 2012).

التحليل العاملي التوكيدي من الدرجة الثانية (Second Order Factor Analysis)

أهم ما يُميز التحليل العاملي التوكيدي من الدرجة الثانية هو وجود عامل هرمي عام، ويتم فيه استبدال العلاقات الارتباطية بين العوامل بمسارات تدل على تأثير العامل العام، ويُعد افتراض النموذج العاملي من الدرجة الثانية مرحلة متقدمة في التنظير، لكون الباحث قد اهتدى إلى تصور وُجود بنية هرمية بين العوامل الكامنة، فافتراض وُجود عامل كامن جامع (عامل من الدرجة الثانية)، وعوامل كامنة فرعية له (عوامل كامنة من الدرجة الأولى)، وبالتالي فلا يتم افتراض وُجود ارتباطات، أو علاقات تغاير بين عوامل الدرجة الأولى؛ وذلك لأنه تم تفسير مصدر العلاقات بين العوامل، وعزوها إلى تأثير متغير كامن عام (تيغزة، 2011).

التحليل العاملي متعدد المجموعات (Multiple Group Confirmatory Factor Analysis) (MG-CFA)

هو أسلوب إحصائي متعدد المتغيرات يستخدم لمقارنة البناء العاملي الذي يمثل العلاقة بين عدد من المتغيرات الملاحظة والتكوينات الافتراضية الكامنة التي يتم قياسها عبر المجموعات المختلفة.

نظرية استجابة الفقرة (Item Response Theory) (IRT)

يُطلق عليها النظرية الحديثة، ونظرية السمات الكامنة (Latent Trait Theory) وهي عبارة عن مجموعة من النماذج الرياضية التي تحاول أن تُفسر العلاقة بين السمات الكامنة (الصفات أو الخصائص غير الملاحظة) ومظاهرها (الأداء أو الاستجابات الملاحظة)، وهي تُمثل مدخل يُستخدم لبناء وتقويم وتطوير أدوات القياس باستخدام نماذجها التي تصف العلاقة بين استجابات المفحوص لفقرات أداة القياس ومكانه أو موقعه على متصل السمة الكامنة.

طريقة المواءمة (Alignment Method)

هي طريقة جديدة يتم استخدامها كوسيلة بديلة لتقدير معالم نماذج العامل التوكيدي عبر العديد من المجموعات لقياس اللاتغير في القياس عبر المجموعات باستخدام نماذج النظرية الحديثة في القياس (IRT)، وتستخدم لتقدير الأوساط والتباينات لعامل المجموعة المحددة دون الحاجة لقياس دقيق للاتغير في القياس (Asparouhov & Muthén, 2014).

الاختبار التقييمي للصف الثالث الأساسي لمبحث الرياضيات

اختبار تُعده وزارة التربية والتعليم ممثلة بمديرية الاختبارات سنويًا لقياس مدى امتلاك الطلبة في الصف الثالث الأساسي لمهارات العلم الخاصة بمبحث الرياضيات ومستويات أداء الطلبة على المهارات السبع: (تمثيل الأعداد وقراءتها وكتابتها، مقارنة الأعداد وتقريبها، وإجراء

العمليات الحسابية على الأعداد، وإكمال النمط، وتحديد وحدات القياس المختلفة، وتمييز الأشكال الهندسية والمجسمات والعلاقات بينها، وجمع البيانات وتسجيلها وتنظيمها بجدول الإشارات والصور، وتطبيق مفهومي التغير والنسبة)، وفي هذه الدراسة تم استخدام نتائج الطلبة على الاختبار المبني للعام الدراسي (2018/2017).

محددات الدراسة

تحدّد تعميم نتائج هذه الدراسة في ضوء المحددات الآتية:

- اقتصار عينة الدراسة على طلبة الصف الثالث الأساسي الملتحقين في المدارس الأردنية للعام الدراسي (2018/2017) والمُطبق عليهم الاختبار التقييمي لمهارات الرياضيات.
- اقتصار الاختبار على مبحث الرياضيات للصف الثالث الأساسي للعام الدراسي (2018/2017).
- اقتصار المتغيرات الكامنة في مرحلة النمذجة البنائية على مهارات مبحث الرياضيات للصف الثالث الأساسي السبع.
- اقتصار متغيرات الدراسة على متغيري مديريات التربية والتعليم الأردنية، وجنس الطالب في مرحلة الكشف عن خاصية اللاتغير في القياس.
- استخدام طريقة المواءمة في الكشف عن اللاتغير في القياس للاختبار التقييمي.
- اقتصرت معايير جودة مطابقة النموذج للبيانات على مؤشرات جودة مطابقة وهي: (CFI، RMSEA، RMSR، χ^2).
- اقتصرت طريقة المواءمة على استخدام مؤشر اللاتغير في القياس (R^2) لتحديد الفقرات المحققة لخاصية اللاتغير في القياس عبر المجموعات للاحتفاظ بها، أو التي لم تحقق لمراجعتها .

الفصل الثاني

الدراسات السابقة

بعد الإطلاع على الدراسات السابقة المتعلقة بموضوع فاعلية طريقة المواءمة في الكشف عن اللاتغير في أداء طلبة الصف الثالث الأساسي على الاختبار التقييمي لمهارات الرياضيات عبر مديريات التربية والتعليم الأردنية، تبين للباحثة وجود عدد قليل من الدراسات ذات العلاقة، وذلك نظرًا لحداثة هذا الموضوع من جهة، وحداثة وجود برمجيات حاسوبية تتعامل مع بنوك الأسئلة والاختبارات واسعة النطاق من جهة أخرى. وقد صُنفت الدراسات المتعلقة بالموضوع التي توافرت للباحثة في مجموعتين:

1. الدراسات التي تناولت اللاتغير في القياس.
2. الدراسات التي تناولت طريقة المواءمة في الكشف عن اللاتغير في القياس.

وتم عرض الدراسات منطقيًا حسب أهميتها بالنسبة لموضوع الدراسة كما يلي:

أولًا: الدراسات التي تناولت اللاتغير في القياس (Measurement Invariance)

أجرى التقي (1992) دراسة بهدف فحص مدى تحقق افتراضين أساسيين من افتراضات نظرية السمات الكامنة في الاختبارات متعددة التدريج (الفقرات المقالية) وهما: افتراض اللاتغير في درجات صعوبة الأسئلة من خلال استخدام عينات مختلفة من الأفراد، وافتراض اللاتغير في قدرات الأفراد من خلال التقدم لاختبارات تتضمن أسئلة متفاوتة في درجات الصعوبة باستخدام نموذجي التقدير الجزئي، وسلم التقدير، وتكونت أداة الدراسة من اختبار مقالي في الرياضيات صُمم بهدف قياس قدرة (458) طالبًا وطالبة من طلبة الصف الأول الثانوي العلمي من مدارس مديرية التربية والتعليم قصبة عمان للعام الدراسي (1992/1991) على حل المعادلات الجبرية والمثلثية في مجهول واحد وفي مجهولين. وأشارت النتائج إلى تحقق الافتراضين عند استخدام نموذج التقدير

الجزئي، بينما لم يتحقق هذان الافتراضان عند استعمال نموذج سلم التقدير؛ لذلك أوصت الدراسة بالاعتماد على نموذج التقدير الجزئي عند فحص اللاتغير في القياس للاختبارات التحصيلية ذات الفقرات المقالية.

وأوضح جلوكنر- ريست وهوجتنكر (Glöckner-Rist & Hoijtink, 2003) في دراستهما أنه يمكن استخدام كل من نمذجة المعادلة البنائية (SEM) ونظرية استجابة الفقرة (IRT) لتحليل عوامل استجابات المفحوصين على الفقرات ثنائية الاستجابة (Dichotomous Item)، وفي هذه الحالة، تكون نماذج القياس لكلا النهجين متكافئة، وفي هذه الدراسة تم التركيز على أمرين هامين أولهما: عرض أوجه التشابه بين نماذج (IRT) و(SEM)، وثانيهما: توضيح كيفية استخدام كلا المنهجين، ودمجهما للتغلب على أغلب التحديات في تحليل البناء العاملي، والكشف عن اللاتغير في القياس للبيانات الناتجة عن استجابات المفحوصين على فقرات مقياس مسحي للهوية الوطنية الألمانية ثنائية الاستجابة، وتم إجراء جميع التحليلات الإحصائية باستخدام برمجية (Mplus) على عينة من (933) مفحوصًا من الجنسين المولودين قبل الحرب العالمية الثانية وبعدها، والحاصلين على التعليم الثانوي أو الأساسي، التي سمحت بتطبيق متكامل لكلا النهجين في إطار نمذجة متغير كامن (Factor) واحد، وهدفت هذه الدراسة لتعزيز نشر أساليب وتقنيات ومهارات القياس المفيدة للبحوث الاجتماعية التجريبية، وأن الدمج بين المنهجين يخدم عمليات البحث بشكل أكبر؛ حيث تتلافى نقاط ضعفهما، وتضاعف من أثر نقاط قوتهما.

وهدفت دراسة آدم وجاري (Adam & Gary, 2004) لتوضيح الظروف التي يحبذ فيها استخدام أسلوب التحليل العاملي التوكيدي (CFA) للكشف عن اللاتغير في القياس/تكافؤ القياس (Measurement Equivalence/ Measurement Invariance)، ولتحقيق ذلك أُستخدم في الدراسة مجموعة من البيانات المولدة والمختلفة فيما بينها بتشبعات العوامل (تحميلات العامل)

(Factor Loadings) على شكل ثلاث مجموعات بأحجام (150، 500، 1000)، وأداة القياس كانت عبارة عن مقياس ليكرت مكون من (6–12) فقرة، وأشارت النتائج إلى أن استخدام التحليل العاملي التوكيدي (CFA) جيدة لاختبار اللاتغير في القياس في ظل ظروف مثالية، وتوفر أحجام عينات كبيرة، وعدد كافٍ من المؤشرات (الفقرات) الواضحة، وقلة من الاختلافات بين معالم الفقرات.

كما قدم شرايبر، ونورا، وستيج ، وبارلو، وكينج (Schreiber, Nora, Stage, Barlow & King, 2006) مراجعة لنتائج التحليل العاملي التوكيدي (CFA)، ونمذجة المعادلات البنائية (SEM) في الدراسات التي استخدمت إحدى الطريقتين للكشف عن اللاتغير في القياس للمقاييس المختلفة، وتضمنت الدراسة مجموعة أساسية من المبادئ التوجيهية، وتوصيات للمعلومات التي يجب تضمينها في أي دراسة تستخدم إحدى الطريقتين باعتبارها التحليل الإحصائي الأساسي، وكما شملت المراجعة مقدمة لكلتا الطريقتين، وعينات التحليل، وتوصيات لإعداد التقارير وتقييم المقالات في المجلة التربوية المستخدمة لهذه التقنيات، والملاحظات الختامية، ولتحقيق ذلك أُستخدم مقياس المعتقدات المعرفية، وعمليات التعلم المكون من (6) فقرات على تدريج ليكرت، طُبق على عينة من (206) طالبًا وطالبة من الطلبة المسجلين لمساق التربية التعليمي (Teacher Education) في الجامعة المتوسطة الحكومية، وأسفرت النتائج على أنه عند دراسة اللاتغير في القياس يجب الاهتمام بعدة جوانب نظرية، وأخرى عملية؛ ولُخصت الجوانب النظرية بالتالي: الاهتمام بالأسئلة البحثية، ثم توضيح الخطوات المتبعة في جزء الإجراءات، كما يجب الاهتمام بمبررات استخدام هذه الطريقة، وكذلك الإطار النظري للطريقة المستخدمة، وتنظيم إدراج الصور، والأشكال، والجداول الواضحة، والمناسبة للإحصاءات الوصفية، كما يجب إدراج عرض بياني للنموذج المقترح، والنهائي أو كليهما، والآثار المترتبة على النتائج، أما الجوانب التحليلية، أو التقنية فتم تلخيصها

بالآتي: يجب تضمين كافة خطوات التحليل، إما بجداول، أو فقرات نصية تحوي على العناصر الأساسية التالية: حجم العينة، ويجب الاهتمام بها؛ بسبب أثرها على استقرار تقدير المعالم، وإظهار كفاية البيانات للتحليل، وطريقة التعامل مع البيانات المفقودة، وعرض مؤشرات جودة المطابقة التي تم الاعتماد عليها ودرجات القطع الخاصة بكل منها، ومؤشرات التعديل المستخدمة إن وُجدت.

وأجرى محمد (2011) دراسة هدفت إلى قياس اللاتغير العاملي لمقياس فورست للكمالية متعدد الأبعاد المكون من (35) فقرة باستخدام التحليل العاملي التوكيدي للمجموعات المتعددة، ولتحقيق ذلك الهدف تم تطبيق المقياس على عينة مكونة من (149) طالبًا وطالبة بكلية التربية – جامعة المنيا، وإجراء المعالجة الإحصائية المناسبة. وقد أظهرت النتائج أن مقياس الكمالية متعدد الأبعاد يتكون من أربعة أبعاد بدلًا من ستة أبعاد، وهي: (التنظيم، والتوقعات الوالديه، والمعايير الشخصية، والحساسية تجاه الأخطاء والشكوك تجاه الأفعال)، وعدد فقراته (22) فقرة بعد حذف (13) فقرة، وأن معاملات الثبات تراوحت بين (0.62–0.82)، كما أكدت النتائج حُسن مطابقة العوامل الأربعة، وتوافر اللاتغير في القياس عبر النوع مما يدل على إمكانية تعميم بنية الكمالية على الذكور والإناث.

وبين كل من هو وفان دي فيجفر (He & Van de Vijver, 2012) أن مصطلحي التحيز (Bias)، واللاتغير في القياس (Invariance Measurment) مصطلحات أساسية في الدراسات عبر الثقافات (Cross–Cultural)، حيث أوضحا أن التحيز (Bias) عبارة عن مصطلح عام لأي تحد يواجه إمكانية مقارنة البيانات بين المجتمعات مختلفة الثقافات، والتحيز موضع اهتمام هذه الدراسة هو التحيز المنتظم المتوقع تكراره عند إعادة الدراسة وليس التحيز العشوائي، فالتحيز (Bias) قد يؤدي لاستنتاجات غير صحيحة، وفي المقابل تحقق خاصية اللاتغير في القياس (عدم الانحياز) (Lack of Bias) هو شرط قَبْلي أساسي لأي مقارنة عبر

الثقافات، وليست خاصية لأداة القياس فقط، بل هي خاصية للمقارنات بين الثقافات؛ لذلك بدايةً: يجب وصف الغرض المراد تحقيقه عند اختيار أدوات القياس في أي دراسة تُطبق على المفحوصين عبر الثقافات، وخصوصًا إذا كانت هناك مفاضلة بين أداة قياس حالية وأخرى جديدة، ثم تحديد المفاهيم ومصادر التحيز الممكنة (بناء الأداة، وطريقة التطبيق، وتحيز الفقرة)، وكذلك مستويات اللاتغير في القياس التي يجب تحقيقها (التكويني، المتري، العددي)، كما عرضت الدراسة استراتيجيات تعمل على تقليل مصادر التحيز؛ أي أنها تدعم تحقيق اللاتغير في القياس، وتستخدم هذه الاستراتيجيات إما في مرحلة التصميم، أو التطبيق، أو التحليل الإحصائي الذي هو موضع اهتمام الدراسة، وأكدا على أهمية دمج كافة الاستراتيجيات في كافة المراحل في الدراسات بين الثقافات لزيادة صدق الاستنتاجات المتعلقة بأوجه الشبه والاختلاف بين الثقافات المختلفة، وأوصت الدراسة بالامتناع عن المقارنات عبر الثقافات إذا لم يتم التحقق من خاصية اللاتغير في القياس بدايةً.

كما أجرى ميوثن واسباروف (Muthen & Asparouhov, 2013) دراسة للكشف عن اللاتغير في القياس عبر مجموعات متعددة، أو فترات زمنية مختلفة باستخدام طريقة جديدة مقترحة من قبلهم، وهي نوع خاص من نمذجة المعادلات البنائية عُرفت بنمذجة المعادلة البنائية لبايزون (Bayesian Structural Equation Modeling) (BSEM) لتحديد مصادر التباين عبر المجموعات، وهي تشبه عمل مؤشرات التعديل مع تقدير الأرجحية العظمى (Maximum Likelihood Estimation) (ML)، ولكن امتازت على الأرجحية العظمى بأنها تُطبق على نماذج المتغير الكامن متعدد الأبعاد (high-dimensional latent variable models) تحت ظروف محددة، كما تعمل طريقة (BSEM) مقارنات مناسبة لأوساط، وتباينات العوامل الكامنة عبر المجموعات باستخدام اللاتغير في القياس التقريبي (Approximate Measurement

Invariance)، ودون التقليل من خصوصية اللاتغير في القياس، أو حذف الفقرات المتباينة، وذلك باستبدال قيد تثبيت الفروق بين معالم النموذج عبر المجموعات عند الصفر ليتم استخدام الصفر التقريبي للمعالم (zero-mean) وتباين قليل (small-variance) باستخدام تحليل بايزين (Bayesian analysis)، وهذا النهج (BSEM) مهتم بالتحميلات المتقاطعة (cross-loading) وارتباطات البواقي في التحليل العاملي التوكيدي (CFA) مع المتغيرات المشتركة، والتي يُشار إليها باسم نمذجة المؤشرات المتعددة والأسباب المتعددة (Multiple Indicators and Multiple Causes) (MIMIC)، ولتحقيق ذلك تم استخدام بيانات مولدة من أحجام مختلفة (1000، 100)، وباستخدام (10) فقرات على مقياس ليكرت، وأسفرت النتائج على أن طريقة (BSEM) يمكن دمجها بالطرق المختلفة للكشف عن اللاتغير في القياس لتجاوز كافة التحديات الإحصائية التي تواجهها طريقة التحليل العاملي التوكيدي في الدراسات التجريبية عبر الثقافات، وبوجود عدد كبير من مؤشرات التعديل، ولضمان الحصول على تقديرات صحيحة تم إقتراح إجراء تحليل (Bayesian) من خطوتين، الأولى: تحديد المعلمات المتباينة (non-invariance parameters)، أما الثانية: فتحرير تلك المعلمات.

وقام سكوت، كلويتمان، توميرز، لاغتيغ، هوكس، وميوثن (Schoot, Kluytmans, Tummers, Lugtig and Muthén, 2013) بالتأكيد على أن الكشف عن اللاتغير في القياس شرط مسبق لمقارنة درجات المتغير الكامن (Latent variable) عبر المجموعات، كما قاموا بتقديم مفهوم اللاتغير التقريبي في القياس (Approximate Measurement Invariance)، واستخدام طريقة (BSEM) لاستبدال قيد تثبيت الفروق بين أوساط العوامل عبر المجموعات عند القيمة صفر في اللاتغير التام في القياس بأصفار تقريبية مع الإبقاء على جودة مُطابقة النموذج، كما تم مناقشة اللاتغير التقريبي في القياس، وفاعليته في التوصل لنموذج مطابق للبيانات حتى في

الحالات التي فشل فيها تحقق اللاتغير التام في القياس، كما أنهم أكدوا على الباحث المهتم بمقارنة أوساط السمة الكامنة عبر المجموعات، أو عبر الأزمان بأن أمامه أربع خيارات: الخيار الأول: فرض اللاتغير العددي في القياس (Scalar MI)، وذلك عند الحصول على نموذج مطابق للبيانات قابل للتفسير متوافق مع التفسيرات المنطقية والأدب النظري، الخيار الثاني: فرض اللاتغير الجزئي في القياس (Partial MI) وذلك عند الحاجة لتحرير بعض تشبعات العوامل، أو تقاطعاته ذات الفروق الأعلى باستخدام مؤشرات التعديل مع الإبقاء على مطابقة جيدة للنموذج مع البيانات، أما الخيار الثالث: هو فرض عدم تحقق اللاتغير في القياس، وفي هذه الحالة يُستبعد قابلية مقارنة أوساط العامل عبر المجموعات؛ لأن التباينات بين السمات الكامنة عبر المجموعات تكون ناتجة عن عدة أسباب لا يمكن تفسيرها أو حصرها، ولكن مع طريقة بايزين للاتغير التقريبي في القياس أصبح هناك خيار رابع هو: فرض اللاتغير التقريبي في القياس، أو دمج اللاتغير التقريبي، واللاتغير الجزئي في القياس لتقدير معالم النموذج لتكون أقرب ما يمكن لبعضها البعض عبر المجموعات، وخصوصًا عندما يؤدي استخدام مؤشرات التعديل لإفشال مؤشرات أخرى، ولتحقيق ذلك تم تطبيق الطريقة على بيانات مولدة مكونة من سبع مجموعات، كل مجموعة تحوي على (1000) مفحوص، تختلف فيما بينها بأوساط المتغير الكامن، باستخدام أربع فقرات من تدريج ليكرت، وأوصت الدراسة باستخدام اللاتغير التقريبي في القياس في الدراسات واسعة النطاق.

وقام الزبون (2013) بدراسة هدفت إلى الكشف عن اللاتغير في القياس في تقدير معالم فقرات الاختيار من متعدد عندما تطبق في سياقات مختلفة، ولتحقيق أهداف الدراسة تم الاعتماد على اختبار مكون من (113) فقرة من نوع الاختيار من متعدد في مادة الرياضيات، بالإضافة إلى (15) فقرة مشتركة طُبقت على عينة تكونت من (2511) طالبًا وطالبة من طلبة الصف الثامن الأساسي المنتظمين في المدارس الحكومية الأردنية، وباستخدام البرامج المناسبة في تحليل البيانات

تبين وجود فروق ذات دلالة إحصائية لصالح النموذج الثلاثي مقارنة مع النموذج الثنائي، كما أظهرت قيم معامل الارتباط بين تقديرات معالم الفقرة: التمييز، والصعوبة، والتخمين التي تمت معايرتها بصورة مستقلة وضمن الإطار الكلي للاختبار أنها كانت عالية، ودالة إحصائيًا، كما تبين أن قيم معاملات الارتباط كانت هي الأعلى عندما تمت معايرة الفقرات كعينة مستقلة (لوحدها)، ويمكن التوصل من ذلك إلى أنه يمكن الحصول على دقة أفضل عند معايرة عينة من الفقرات بشكل مستقل مقارنة بمعايرتها ضمن الإطار الكلي للاختبار، وبناءً على ذلك تُوصي الدراسة باستخدام عينة ممثلة بعدد قليل من الفقرات في الاختبارات التكيفية، وإجراء المزيد من الدراسات باستخدام دالة المعلومات، كما تُوصي باستخدام النموذج الثلاثي عند معايرة الفقرات.

كما قامت راتكوسكي وسفيتينا (Rutkowski and Svetina, 2014) بدراسة هدفت لتقييم طريقة التحليل العاملي التوكيدي متعدد المجموعة (MG−CFA) في التحقق من فرضية اللاتغير في القياس في المسوحات الدولية واسعة النطاق التي تحوي عددًا كبيرًا من المجموعات، وحجوم عينات كبيرة لهذه المجموعات، ولتحقيق ذلك استخدمت بيانات مولدة مكونة من (10، 20) مجموعة بأحجام عينات (600، 6000) مفحوص لكل مجموعة، وبتطبيق عينة (5، 6) فقرات من مسح التعليم والتعلم الدولي (Teaching and Learning International Survey)، وأسفرت النتائج على أنه لتقييم طريقة التحليل العاملي التوكيدي متعدد المجموعة (MG−CFA) في جودتها بإيجاد نموذج قياس مناسب للبيانات المستخدمة عن طريق استخدام اختبارات جودة المطابقة وهي: (χ^2، CFI، TLI، SRMR، RMSEA)؛ وبسبب كِبَر حجوم العينات المستخدمة، وأعداد المجموعات، فتُوصي الدراسة بالاعتماد على اختبارات جودة المطابقة جميعها باستثناء اختبار مربع كاي لحساسيته لكِبر حجوم العينات.

وكذلك أوضحت دراسة سيوتشي، ودافيدوف، وشميدت، والجاشايمر، وشوارتز (Cieciuch, Davidov, Schmidt, Algesheimer and Schwartz, 2014) أن التحليل العاملي التوكيدي متعدد المجموعة (MG-CFA) يُعد الطريقة الأكثر استخدامًا في الكشف عن اللاتغير في القياس، وقد تم اقتراح طريقة جديدة للكشف عن اللاتغير التقريبي في القياس وهي: (Bayesian MG-CFA) التي تسمح لوجود تباينات صغيرة غير دالة إحصائيًا بين معلمات النموذج عبر المجموعات بدل أن تكون متساوية تمامًا في طريقة (MG-CFA)، وجاءت هذه الدراسة لمقارنة نتائج الطريقتين باستخدام بيانات ناتجة من تطبيق استبيان القيم الشخصية (Portrait Values Questionnaire) (PVQ) المكون من (19) قيمة، على عينات مختلفة ذات أحجام متفاوتة بين (201-527) عبر ثمان دول، وأسفرت النتائج على أنه عند استخدام طريقة التحليل العاملي متعدد المجموعة تحقق مستوى اللاتغير العددي في القياس لِـ(10) قيم من (19) قيمة، بينما عند استخدام طريقة (Bayesian MG-CFA) فقد تحقق اللاتغير التقريبي في القياس على كافة الـ(19) قيمة، وعلى ذلك أوصت الدراسة بما يلي: أولًا: النهج التقريبي يُوفر نتائج أكثر تشجيعًا للمهتمين بالمقارنات عبر الثقافات، مع حاجة هذا الاستنتاج لتعميمه، والتحقق من صحته في البحوث المستقبلية باستخدام البيانات الواقعية، ثانيًا: استخدام الطرق المختلفة للكشف عن اللاتغير التقريبي في القياس لأن تحققه أكثر احتمالية من اللاتغير العددي في القياس، على الرغم من أن هناك حاجة للمزيد من دراسات المحاكاة لتحديد التباين المسموح به بين معلمات النموذج المقترح تطابقه مع البيانات عبر المجموعات.

كما أجرى ايونيس (Ioannis, 2014a) دراسة في البناء العاملي للبيانات حسب جنس الطالب (ذكر، أنثى) لاختبار القدرات العامة (General Aptitude Test) (GAT) المُعد من قبل المركز الوطني للقياس والتقويم (National Center for Assessment) (NCA) في

المملكة العربية السعودية، طُبقت هذه الدراسة على (8272) طالبًا وطالبة من المرحلة الثانوية؛ منهم (7409) إناث، و(863) ذكور، تم دراسة نموذجين، النموذج الأول مكون من سبع مجالات والثاني من مجالين، لكل منهما بعدين (اللفظي، والعددي)، وأسفرت النتائج على أن كلا النموذجين يناسب البيانات، ويُقدم دليلًا على صحة البناء العاملي لاختبار القدرات العامة (GAT)، وتحقق فيهما وجود اللاتغير في القياس من المستويات الثلاث (التكويني، والمتري، والعددي)، أما بالنسبة للفروق في الأوساط الكامنة بين الجنسين، فقد تبين أن الأوساط للذكور أعلى من الإناث في البُعد الرقمي، وأن الأوساط للإناث أعلى من الذكور في البُعد اللفظي، بينما في المجالات الفرعية كانت أوساط الإناث أعلى من الذكور في مجالات معنى الكلمة، وإتمام الجملة، بينما أوساط الذكور أعلى من الإناث في مجالات الحساب، والهندسة، ولم يوجد أي فروق بين الجنسين في مجالات القراءة، والفهم، والتحليل الكامن.

كما أجرى ايونيس (Ioannis, 2014b) دراسة أخرى في البناء العاملي على نفس البيانات، ولكن تم اختيار متغير نوع المدرسة (خاصة، حكومية) في هذه الدراسة لاختبار القدرات العامة (General Aptitude Test) (GAT) المُعد من قبل المركز الوطني للقياس والتقويم في المملكة العربية السعودية، شارك في هذه الدراسة (8272) طالبًا وطالبة من المرحلة الثانوية، تم دراسة نموذجين، النموذج الأول مكون من سبع مجالات والثاني من مجالين، لكل منهما بعدين (اللفظي، والعددي)، وأظهرت النتائج أن كلا النموذجين يناسب البيانات، ويقدم دليلًا على صحة البناء العاملي لاختبار القدرات العامة (GAT)، كما تحققت خاصية اللاتغير في القياس بمختلف مستوياتها (التكويني، والمتري، والعددي) للنموذجين، كما وُجد أن أوساط الطلاب في المدارس الخاصة أعلى منها في المدارس العامة في المجالات الكامنة (اللفظية، والعددية) على حد سواء في اختبارات (GAT)، كما أن أوساط الطلاب في المدارس الخاصة أعلى من الطلبة في المدارس

الحكومية في كافة المجالات الكامنة لاختبارات (GAT)، باستثناء مجالات الحساب، والهندسة الكامنة، فلم يكن هناك فروق ذات دلالة إحصائية بين هذين المجالين عبر النوعين من المدارس.

واستخدم بوجاكش، وفيتيرسو، هوتا، وكازمارك (Bujacz, Vittersø, Huta & Kaczmarek, 2014) تقدير بايزون (Bayesian estimation) لفحص اتساق بنية العوامل الخاصة بمقياس الدوافع (Hedonic and Eudamonic Motives for Activities) (HEMA) عبر الثقافات، وبينوا أن طريقة بايزون تمتاز بالمرونة لسماحها باستبدال الفروق بين معلمات النموذج الصفرية (exact zero costraints) لتصبح تقريبية (approximate zero constraints)، كما تم اختبار اللاتغير في القياس لمقياس الدوافع عبر عينتين وطنيتين (البولندية، وأمريكا الشمالية) باستخدام كل من التقدير التقليدي (CFA) والبايزون (BCFA)، وأسفرت النتائج عما يلي: أولاً، تم تأكيد إنطباق النموذج ثلاثي العوامل في كلتا العيّنتين، ثانيًا، تم اختبار تحقق اللاتغير المتري في القياس، ورفضه باستخدام النهج التقليدي (CFA) مع تقدير الأرجحية العظمى (Maximum Likelihood estimation)، كما تم قبول اللاتغير المتري الجزئي في القياس (Partial metric invariance) باستخدام تقدير بايزون (BCFA)، وبما أنه لم يتحقق اللاتغير العددي في القياس، فإن مقارنة العوامل الكامنة غير ممكنة عبر المجموعات، وأوصت هذه الدراسة باستخدام تقدير بايزون للكشف عن اللاتغير في القياس؛ لأنه أكثر إفادة من النهج التقليدي، ولأنه يسمح بمزيد من المرونة في خصائص النموذج المقترح.

ولاحظ لومين، وشوت، وانجلهارد (Lommen, Schoot and Engelhard, 2014) أن الدراسات التي تتضمن تقييمات متعددة لمقياس معين مطبق على نفس العينة من المجتمع تفترض أن هذا المقياس يقيس نفس السمة الكامنة عبر الزمن؛ لذلك قاموا بدراسة تحقق خاصية اللاتغير في القياس لاستبيان يقيس أعراض الإجهاد اللاحق للصدمة (Post traumatic Stress

Symptoms) (PSS) قبل وبعد تعرض المفحوصين لحدث مؤلم لعينة من الجنود الهولنديين قبل وبعد أن انتقلوا إلى أفغانستان (n=249)، وأظهرت النتائج أن نموذج القياس المُقترح ثلاثي الأبعاد قبل نشر الجند مُختلف عن نموذج القياس المُقترح ثلاثي الأبعاد بعد النشر؛ بسبب عدم تحقق خاصية اللاتغير في القياس على عتبات الفقرات (item thresholds)، ثم تم تطبيق المقياس مرة أخرى على عينة من الجنود الذين تم نشرهم في العراق (n=305)، ونظرًا لافتقار المقياس لخاصية اللاتغير في القياس؛ بسبب عدم استقرار غالبية الفقرات، فمن المتوقع استنتاج أن البنية العاملية الأساسية لنظام (PSS) غير مستقرة مع مرور الوقت في حالة وقوع أحداث مؤلمة متعلقة بمنطقة الحرب؛ أي أنه من وجهة نظر إحصائية، لا يمكن مقارنة الدرجات بمرور الوقت عندما يكون هناك انتهاك لافتراض اللاتغير في القياس، وأوصت الدراسة الباحثين الذين يستخدمون مقاييس الإجهاد بعد الصدمة (PSS) في الدراسات الطولية بأنه يجب ألا يأخذوا تحقق خاصية اللاتغير في القياس كأمر مُسَلّم به، بل يجب عليهم استخدام البيانات الناتجة من تطبيق المقياس ما قبل، وما بعد الأعراض كعوامل (سمات) مختلفة لكل نقطة زمنية في مرحلة التحليل الإحصائي.

كما قارن النعيمي (2015) معالم الفقرات والأفراد وخاصية اللاتغير في القياس في الاختبارات الوطنية لضبط جودة التعليم في الأردن بين النظرية الكلاسيكية، والنظرية الحديثة في القياس، ولتحقيق ذلك استخدمت بيانات الاختبار الوطني لضبط نوعية التعليم في الأردن للعام الدراسي (2008/2007)، وخصوصًا بيانات طلبة الصف الثامن الأساسي لمبحثي اللغة العربية واللغة الإنجليزية، وأُخذت منها عينات فرعية بحسب الجنس، وعينات عشوائية مكونة من (1000) طالبٍ وطالبة لكل منها، وإحصائيًا تم مقارنة المعالم باستخدام معامل الارتباط بيرسون، وأظهرت نتائج الدراسة المتعلقة بقدرات الأفراد أنه توجد علاقة قوية، وطردية بين توزيع العلامات الخام في النظرية الكلاسيكية، وتوزيع القدرة في نماذج نظرية استجابة الفقرة (الأحادية، والثنائية، والثلاثية)،

أما النتائج المتعلقة بمعاملات التمييز فأظهرت أن العلاقة بين معاملات التمييز المقدرة باستخدام نظرية القياس التقليدية (الكلاسيكية)، ونظيراتها المقدرة باستخدام النموذج الثنائي قوية نوعًا ما، وضعيفة في النموذج ثلاثي المعالم، كما تمتعت معاملات الصعوبة والتمييز بخاصية اللاتغير في القياس في كلٍ من النظرية الكلاسيكية، ونماذج نظرية استجابة الفقرة (IRT).

وأوضح كريش (2015) أهمية التحليل العاملي التوكيدي متعدد المجموعات (MG–CFA) كأسلوب للقياس في ميدان علم النفس الجنائي، حيث تعتبر الاختبارات النفسية من أدوات القياس الموضوعية للحصول على معلومات مهمة حول شخصية المتقاضين في المحاكم، والتي يتم إصدار أحكام وقرارات مصيرية في حقهم على أساسها؛ لذلك على المُمارس أن يضمن قدر الإمكان دقة هذه المعلومات التي يستمدها من درجات الاختبارات المطبقة، ومن أشهر الاختبارات النفسية اختبار مينيسوتا متعدد الأوجه للشخصية (Minnesota Multiphasic Personality) (MMPI) التي يطبقها الممارسون في ميدان علم النفس الجنائي في الجزائر، فقد تم بناؤها في الدول الغربية مما يستدعي تكييفها عند نقلها إلى بيئة ثقافية جديدة لجعلها تتلاءم معها، ويعتبر التحليل العاملي التوكيدي متعدد المجموعات (MG–CFA) من التقنيات الإحصائية المهمة لاختبار اللاتغير في القياس للاختبارات النفسية في البيئة الأصلية، والجديدة عند نقلها، وذلك لإعطاء دليل كمي على أن الفقرات لها نفس المعنى في البيئتين الثقافتين، وأنها تقيس نفس التكوين الفرضي (البناء العاملي)، وبنفس الطريقة، وبهذا نضمن أن درجات الاختبار تُعبر فعلًا عن مستويات التكوين الفرضي (السمات الكامنة)، وهذا دليل على صدق الاستدلال بها، مما يعني دقة المعلومات التي نحصل عليها بواسطتها، وصلاحية استخدام هذا الاختبار في البيئة الثقافية الجديدة؛ لذلك هدفت هذه الدراسة لاستعراض أهمية هذه التقنية الإحصائية وخطواتها للتأكد من اللاتغير في القياس للاختبارات النفسية، والتحقق من مدى صلاحية استخدامها في غير بيئتها

الأصلية بعد تكييفها، وأوصت بالكشف عن اللاتغير في القياس لأدوات القياس المختلفة قبل استخدامها وخصوصًا المترجمة، أو المطبقة في غير بيئتها الأصلية.

وأجرى ايونيس (Ioannis, 2015) دراسة في البناء العاملي لبيانات الاختبار (Standard Achievement Admission Test) (SAAT) التحصيلي المعياري للقبول، الذي يُدار من قبل المركز الوطني للقياس والتقويم في التعليم العالي في الرياض، المملكة العربية السعودية، حيث تقدم للاختبار (63380) طالبًا وطالبة (36277 ذكور، 27041 إناث، 62 طالبًا، وطالبة لم يُحددوا جنسهم)، تم دراسة ثلاثة نماذج مقترحة للبناء العاملي لاختبار (SAAT) كان النموذج الثاني أفضل من النموذجين الآخرين، وقد أسفرت النتائج عن أن جميع النماذج تُناسب البيانات، وتُقدم دليلًا على صحة البناء العاملي لاختبار (SAAT) التحصيلي للقبول، كما ثبت وجود اللاتغير في القياس عبر الجنس، كما وأظهرت النتائج أن الأوساط للإناث أعلى منها للذكور في علم الأحياء، والكيمياء، ومجالات تحقيق العالمية، بينما كانت أوساط الذكور أعلى منها للإناث في مجالات الفيزياء، والرياضيات.

كما درس زيرشر، وشميت، وسيسيوش، ودافيدوف (Zercher, Schmidt, Cieciuch and Davidov, 2015) قابلية مقارنة القيمة العالمية (universalism value) بمرور الزمن، وعبر الدول في "المسح الاجتماعي الأوروبي" (European Social Survey) (ESS): باستخدام اللاتغير التام الكلاسيكي في القياس مقابل اللاتغير التقريبي في القياس، فمنذ القدم، تم جمع البيانات الدولية الكبيرة مثل: المسح الاجتماعي الأوروبي (ESS)، ودراسة القيم الأوروبية (European Social Survey) (EVS)، ومسح القيم العالمية (World Value Srvey) (WVS)، لمقارنة أوساط القيم عبر نقاط زمنية متعددة، أو عبر دول متعددة، ومع ذلك، فإن تحليل بيانات المسح المُقارن يتطلب تحقيق افتراضات محددة، وهي: مدى قابلية مقارنة هذه القيم عبر

الزمن، أو عبر الدول، ونظرًا للعدد الكبير من المجموعات التي يمكن مقارنتها عند تطبيق المسوحات الوطنية، فإن تحقيق خاصية اللاتغير في القياس يعتبر صعبًا، حيث أن الدراسات التي قيّمت اللاتغير في القياس في أغلب الأحيان فشلت في تحقيق مستويات عالية منه وهو: اللاتغير العددي في القياس؛ لتصبح المقارنات مسموحة؛ لذلك قدمت هذه الدراسة، أولاً: النهج التقريبي المطوَّر حديثًا استنادًا إلى نمذجة المعادلات البنائية البايزونية (Bayesian Structural Equation Modeling) (BSEM) لتقييم اللاتغير في القياس عبر مجموعات الدول، أو الزمن، ومقارنة النتائج مع نتائج اختبار اللاتغير التام الكلاسيكي في القياس، حيث يفحص (BSEM) ما إذا كانت معلمات القياس متكافئة تقريبًا (وليست متساوية)، لذلك تم تطبيق (BSEM) على عينة من الفقرات التي تقيس قيمة الشمولية من استبيان قيم الصور الشخصية (PVQ) في المسح الاجتماعي الأوروبي (ESS)، حيث تم اختبار اللاتغير في القياس لهذه القيمة في (وقت واحد) عبر (15) دولة من دول (ESS) عبر ست جولات من (ESS) مع (173071) مستجيبًا، و(90) مجموعة بالمجمل، وأسفرت النتائج عن أن النهج التقليدي (CFA) أجاز مقارنة الأوساط الكامنة لـ(37) مجموعة، بينما النهج التقريبي (BSEM) أجاز مقارنة الأوساط الكامنة لـ(73) مجموعة، وكانت هذه أول دراسة يتم فيها تطبيق نهج (BSEM) على عينة من مجموعات كبيرة.

كما حقق دافيدوف وآخرون (Davidov et al., 2015) في قابلية مقارنة الاتجاهات نحو الهجرة (Attitudes Toward Immigration) في المسح الاجتماعي الأوروبي (ESS): باستخدام اللاتغير التام الكلاسيكي في القياس مقابل اللاتغير التقريبي في القياس، حيث تم تحليل بيانات المسح الدولي بتواتر متزايد لاستقصاء، ومقارنة الاتجاهات نحو الهجرة، وفحص العوامل الكامنة التي تُشكل هذه الاتجاهات، فالمقارنات الدولية للبنى التجريدية، والنفسية تتطلب قياسات مكافئة؛ بمعنى أنها يجب أن تقيس نفس المفهوم على نفس أداة القياس، وتؤدي أغلب الطرق

التقليدية لتقييم اللاتغير في القياس لاستنتاج أن القياسات لا يمكن مقارنتها عبر الدول، وقد تم انتقادها لكونها صارمة للغاية؛ لذلك قدمت هذه الدراسة طريقة بايزون كبديل تقريبي عن الطريقة الكلاسيكية لتقييم اللاتغير في القياس، وهذه الطريقة تسمح بتباينات صغيرة في معلمات القياس عبر المجموعات، وتم ذكر التحليل العاملي التوكيدي متعدد المجموعة كنقطة انطلاق، حيث تم اختبار اللاتغير التقريبي في القياس، والكلاسيكي على مقياس الاتجاهات نحو الهجرة الذي نُفذ في المسح الاجتماعي الأوروبي (ESS)، وتم اختبار اللاتغير في القياس عبر المجموعة الكاملة من (27120) مفحوصًا في (35) دولة على ست جولات، وكانت نتائج كل من الطريقة الكلاسيكية والتقريبية مختلفة تمامًا، فالطريقة التقريبية أثبتت تحقق خاصية اللاتغير في القياس في كل جولات المسح الاجتماعي الأوروبي (ESS)، وبالتالي السماح للباحثين بمقارنة أوساط الدرجات وعلاقاتها بالبناءات النظرية الأخرى، أما الطريقة الكلاسيكية فنتج عنها عدم قابلية البيانات للمقارنة عبر الدول، أو الفترات الزمنية المتعددة؛ بسبب عدد المجموعات الكبير.

وأجرى السمكي (2017) دراسة لفحص اللاتغير في القياس لاختبار تمس (2011) في الرياضيات والعلوم للصف الثامن الأساسي في الأردن عبر كل من متغير موقع المدرسة (مدينة، ريف) وجنس الطالب (ذكر، أنثى)، ولتحقيق ذلك تم استخدام المنهج التحليلي الوصفي في تحليل نتائج اختبار تمس في الرياضيات والعلوم لطلبة الصف الثامن الذين شاركوا في دراسة تمس في عام (2011)، وبالتحديد فقد تم تحليل استجابات (547) طالبًا وطالبةً في كراستي الاختبار (Two Booklet) الثامنة والتاسعة للرياضيات والعلوم، ولتحليل فقرات الاختبارين تم استخدام أسلوبي التحليل العاملي التوكيدي (CFA) والتحليل العاملي التوكيدي متعدد المجموعة (MG–CFA) من خلال البرنامج الإحصائي (AMOS 22)، وقد أظهرت النتائج ملاءمة نموذج العوامل الأربعة المُقترح لمحتوى اختبار الرياضيات (الأعداد، الجبر، الهندسة، الإحصاءات والاحتمالات)،

وملاءمة نموذج العوامل الأربعة المقترح لمحتوى اختبار العلوم (البيولوجيا، الفيزياء، الكيمياء، وعلوم الأرض)، كما وأظهرت النتائج تحقق اللاتغير في القياس بمستوياته الثلاث: (التكويني، والمتري، والعددي) في فقرات كل من اختبار الرياضيات والعلوم وفقًا لمتغيري موقع المدرسة، وجنس الطالب، وأظهرت النتائج أفضلية للإناث على الذكور، وأفضلية لمدارس المدن على مدارس الريف.

كما قام الزعبي وسوالمه (2017) ببناء اختبار لقياس مهارات التفكير الناقد باستخدام نظرية استجابة الفقرة؛ وذلك بكتابة (228) فقرة من نوع الاختيار من متعدد، تغطي مهارات التفكير الناقد، ومن خلال التحكيم والتجريب تم الاحتفاظ بـ(89) فقرة، وطُبق الاختبار على (444) طالبًا وطالبة في المرحلة الأساسية العليا، و(610) طالبًا وطالبة في المرحلة الثانوية، و(120) طالبًا وطالبة في المرحلة الجامعية، وتم التحقق من افتراض أحادية البعد، والصدق، والثبات للاختبار، كما تم التحقق من مطابقة الاستجابات عن فقرات الاختبار لنماذج نظرية استجابة الفقرة، وقد اعتمدت الدراسة النموذج ثلاثي المعلمة، لكون تقديرات القدرة وفقًا له أكثر دقة منها للنموذجين أُحادي، وثنائي المعلمة خاصة عند مستويات القدرة العليا، واتصفت معالم الفقرات، ومعلمة القدرة بتحقيقها لخاصية اللاتغير في القياس، وقد قُدرت معالم الفقرات، ودالة المعلومات لها، كما حُسبت بعض إحصائيات الاختبار.

وقارن صديق ولايتوب (Seddig & Leitgöb, 2018) في دراستهم بين التحليل العاملي التوكيدي التقليدي (الصارم)، والنهج التقريبي (Approximate) لتقييم اللاتغير في القياس باستخدام (وسط حسابي صفر، وتباينات صغيرة أساسية) لنمذجة بايزين (Bayesian) التي تسمح باختلافات طفيفة في المعلمات المُقدرة عبر الزمن، وذلك مع الحفاظ على قابلية المقارنة بين البنى العاملية الأساسية، ولتحقيق ذلك تم استخدام البيانات المستقاه من تطبيق استبيان علم الاجتماع

والجريمة الألمانية (German sociological and criminological) على أربع جولات بعينة عمرية (14، 16، 18، 20)، وبحجم (1002) مفحوص، وتم توضيح الإجراءات المُتبعة لأول مرة مع بيانات طولية (متغيرة عبر الزمن) عن تفضيلات الشباب لزيادة المتعة، والاستمتاع بالحياة، وأشارت النتائج إلى أن النهج التقليدي لقياس اللاتغير التام في القياس فشل في تحقيق اللاتغير العددي في القياس عبر الزمن، وأن المقارنات بين الأوساط الكامنة غير ممكنة؛ لذلك تم استخدام اللاتغير العددي التقريبي في القياس، وأسفرت النتائج عن بيان إجراءات اللاتغير التقريبي في القياس الذي رصد قيم مقبولة لمطابقة النموذج المُقترح للبيانات، واُعتبرت التباينات صغيرة بشكل غير دال إحصائيًا عبر الزمن، وأصبحت المقارنات ممكنة.

المحور الثاني: الدراسات التي بحثت في طريقة المواءمة (Alignment Method)

تناول ميوثن واسباروف (Muthen & Asparouhov, 2013b) في دراستهم طريقتين جديدتين للكشف عن اللاتغير في القياس عبر المجموعات وهما: طريقة المواءمة (Alignment Method) التي تتعامل مع تباينات معالم النموذج عبر المجموعات كنمط تباين ثابت، وطريقة التقاطع والتحميل العشوائي ذي المستويين (Random-intercept, Random-loading two-level Method) التي تعتبر تباينات معالم النموذج عبر المجموعات كنمط تباين متغير، وتم تطبيق الأرجحية العظمى (Maximum Likelihood) وتحليل (Bayesian)، على مسح مكون من (50000) مفحوص في (26) دولة؛ لدراسة افتراضات الطريقتين التي يجب مراعاتها، كما تم استخدام دراسات المحاكاة (Monte Carlo)، وأسفرت النتائج عن تقديم قائمة معايير للاختيار بين الطريقتين للكشف عن اللاتغير في القياس.

وفي دراسة أخرى لميوثن واسباروف (Muthen & Asparouhov, 2014) قدما فيها طريقة جديدة من طرق التحليل العاملي التوكيدي متعدد المجموعات (MG-CFA)، سُميت بطريقة

المواءمة (Alignment Method)، أُستخدمت لتقدير أوساط وتباينات العوامل الكامنة للمجموعة دون الحاجة لقياس دقيق للاتغير في القياس، وكَمُنَت قوة هذه الطريقة في تقدير النماذج المناسبة لعدة مجموعات كما هو الحال عند مقارنة عدة دول، كما تُعد هذه الطريقة بديلًا رائعًا لأساليب التحليل العاملي التوكيدي متعدد المجموعات (MG-CFA) المستخدمة حاليًا لدراسة اللاتغير في القياس، والتي تحتاج عادةً للكثير من التعديلات اليدوية باستخدام مؤشرات التعديل (Modification Indeces) مما يجعل منها طريقة غير عملية مع المجموعات المتعددة، خصوصًا عند سوء مطابقة النموذج المُقترح للبيانات، ومحاولة تعديله باستخدام عدد كبير من مؤشرات التعديل، في المقابل طريقة المواءمة تعتمد على النموذج التكويني، وتقوم بالحسابات آليًا، كما تقوم بتبسيط تحليل اللاتغير في القياس، وتوفر هذه الطريقة حساب تفصيلي لمعالم اللاتغير في القياس لكل معلمة في النموذج المقترح، ولكل مجموعة، وفي هذه الدراسة تم التركيز على تطبيقات طريقة المواءمة ضمن نظرية (IRT)، وذلك بتطبيقها على مسحين منفصلين ذات فقرات ثنائية (binary items) طُبق في عدة دول عددها (28) مجموعة بحجم عينة (90000) مفحوص من الطلبة ذوي أعمار (14) سنة، ولتحقيق ذلك تم استخدام دراسات محاكاة (مونت كارلو)، وأسفرت النتائج عن رفض النموذج المقترح للاتغير المتري، والعددي في القياس عند استخدام طريقة التحليل العاملي متعدد المجموعة (MG-CFA)، في المقابل عند استخدام طريقة المواءمة توصلت لنموذج مُطابق للبيانات بدرجة مقبولة جعلت المقارنة بين أوساط العوامل الكامنة ممكنة عبر المجموعات (الدول)، وعبر المسحين أيضًا.

كما قاس كل من دي بوندت وفان بيتيجم (De Bondt and Van Petegem, 2015) اللاتغير التقريبي لاستبان الإفراط في استطلاعه الثاني (Over Excitability Questionnaire) (OEQ-II) وتأثير درجة، وطبيعة الإفراط في الإثارة، مما ساعد في تحديد الإمكانات التنموية

للفرد وفقًا لنظرية دابروفسكي في التفكك الإيجابي، وقد تم دراسة صدق المقياس سابقًا باستخدام التحليل العاملي التوكيدي (CFA)، ورُفض النموذج المُقترح حتى مع السلسلة الطويلة من مؤشرات التعديل التي تم استخدامها، إلا أن نمذجة المعادلة البنائية لبايزون (Baysien Structural Equation Modiling) (BSEM) التي استخدمت في هذه الدراسة سمحت بالتباينات الصغيرة، وتغايرات البواقي (residual covariances) بين معلمات النموذج المُقترح عبر المجموعات التي يمكن تفسيرها نظريًا، وأدت لملاءمة أفضل للنموذج، وتقدير أنسب لارتباطات العامل الكامن عبر المجموعات؛ لذلك ففي هذه الدراسة تم تطبيق تحليل (BSEM) على عينة من (516) طالبًا من التعليم العالي، وأسفرت النتائج عن تحقق صدق المقياس (OEQ-II)، وتحقق اللاتغير التقريبي في القياس مما سمح بإمكانية مقارنة أوساط السمات الكامنة الذي نتج عنها تسجيل الإناث لدرجات أعلى بشكل ملحوظ في فرط الإثارة العاطفية والحسية، وأقل بشكل ملحوظ في فرط الإثارة الحركية من الذكور.

وعرض بايرون وفان دي فيجفير (Byrne and Vijver, 2017) في دراستهم تطبيق نموذجي لطريقة المواءمة باستخدام تقدير الأرجحية العظمى (Maximum Likelihood Estimation) (ML) لاختبار اللاتغير التقريبي في القياس عبر الثقافات، والتي هدفت إلى: وصف وتوضيح الصعوبات الشائعة التي واجهت الباحثون عند استخدامهم التحليل العاملي متعدد المجموعة التقليدي (MG-CFA) لاختبار اللاتغير في القياس مع عدد كبير من المجموعات، كما أوضحت طريقة المواءمة (Alignmant Method) باستخدام تقدير الأرجحية العظمى (ML) في اختبار اللاتغير التقريبي في القياس، ولتحقيق ذلك تم استخدام بيانات مستمدة من دراسة سابقة لأداء الأُسرة عبر (30) ثقافة، تتضمن استجابات (5842) طالبًا جامعيًا من (27) دولة من الثقافات الـ(30) على مقياس القيم العائلية، استندت التحليلات على برنامج (Mplus v7.4)،

وكان من نتائج الدراسة أن طريقة (MG-CFA) أظهرت (108) معلمة متباينة تحتاج لتعديل، بينما طريقة المواءمة أثبتت الموثوقية الكامنة للتقديرات، وإمكانية مقارنتها بين المجموعات، ولقد خلُصت الدراسة إلى أن طريقة المواءمة لاختبار اللاتغير التقريبي في القياس تُوفر إجراءً آلياً لتحرير المعلمات وتحسين مطابقة النموذج المُقترح للبيانات يمكنه التغلب على القيود المفروضة من تحليل (CFA) التقليدي في المقارنات واسعة النطاق.

وقارن كيم، وتساو، ووانغ، ونجوين (Kim, Cao, Wang & Nguyen, 2017) خمس طرق لاختبار قياس اللاتغير في القياس عبر المجموعات المتعددة؛ بسبب الاستخدام المتزايد لبيانات المسوح الدولية، وخاصة في الدراسات متعددة الثقافات، والمتعددة الجنسيات، مما زاد من الاهتمام بفحص تحقق خاصية اللاتغير في القياس (MI) عبر عدد كبير من المجموعات، وفي هذه الدراسة تم المقارنة بين خمس طرق إحصائية للكشف عن اللاتغير في القياس وهي: (التحليل العاملي التوكيدي متعدد المجموعة (MG-CFA)، والتحليل العاملي التوكيدي متعدد المستويات (Multilevel-CFA) (ML-CFA)، ونمذجة العامل الخليط متعدد المستويات (Multilevel Factor Mixture Modeling) (ML-FMM)، واختبار (MI) التقريبي لبايزون (Bayesian approximate MI testing)، وطريقة المواءمة (Alignment Method))، وتم عرض أوجه التشابه والاختلاف بين هذه الطرق من حيث النماذج العاملية، والإجراءات الإحصائية، كما تم إجراء دراسة محاكاة (مونت كارلو) مكونة من أعداد مجموعات (25، 50)، لها حجوم مختلفة (50، 100، 1000) لكل مجموعة؛ للتأكد من فاعلية الطرق الخمسة في الكشف عن اللاتغير في القياس عبر العديد من المجموعات باستخدام معايير مطابقة مختلفة، وأظهرت الطرق الخمسة أداءً معقولاً في تحديد مستوى اللاتغير في القياس، كما وفرت الدراسة إرشادات عامة في اختيار الطريقة المناسبة.

كما استخدم إيفرز وآخرون (Evers et al., 2017) طريقة المواءمة في دراستهم لاختبار ممارسات وسلوكات علماء النفس نحو الاختبارات وفحوصاتها؛ هدفت الدراسة لجمع البيانات من أجل فهم أفضل لحالة الاختبار النفسي في جميع أنحاء العالم من خلال تطبيق استبيان على (20476) عالم نفس محترف من (29) دولة، ولتحقيق ذلك بُنيت خمسة مقاييس تهتم بـ: الاستخدام الخاطئ للاختبار، وتعليمات تطبيق الاختبار، ومدى معرفة وتدريب الفاحص قبل استخدام الاختبار، والإنترنت وإمكانية حوسبة الاختبار، ورفع أهمية الاختبارات، حيث تم تقييم اللاتغير في القياس عبر الدول باستخدام طريقة المواءمة (Alignment Method)، وأثبتت أربعة مقاييس تحقيقها لمستويات مقبولة من اللاتغير في القياس، كما تم استخدام تحليل متعدد المستويات لتحديد كيفية ارتباط النتائج بالعمر، والجنس، والتخصص، وكذلك كيفية تنوع الدرجات بين الدول، كما أظهرت النتائج ارتفاع أهمية الاختبارات بشكل عام، وانخفاض أهمية الإنترنت والاختبار المحوسب، كما أظهرت تباينًا منخفضًا بين الدول، رغم أن التباينات بين الدول في تقارير المقاييس الأخرى أكبر بكثير مما توصلت إليه هذه الدراسة، وهذا يعني الحاجة لبعض التحسينات الشاملة عند تطبيق الاختبار على المستوى الوطني، وكذلك الالتزام بتعليمات تطبيق الاختبار الخاصة بكل دولة على حدا.

كما استخدمت فليك وماك كاوش (Flake & McCoach, 2018) طريقة المواءمة (Alignment Method) كبديل للتحليل العاملي متعدد المجموعة (MG-CFA) لتقدير نماذج القياس، واختبار اللاتغير في القياس عبر المجموعات، فأظهرت دراسات المحاكاة التي تقيم أداء طريقة المواءمة في تقدير نماذج القياس عبر المجموعات نتائج واعدة للمؤشرات (الفقرات) المستمرة (Continuous items)، وقد استخدمت هذه الدراسة بيانات مُستقاة من أبحاث سابقة للتحقق من مُطابقة النماذج المُقترحة للبيانات التي تم تقديرها باستخدام طريقة المواءمة مع مؤشرات (فقرات)

متعددة التدريج (Polytomous Items)، كما تم التحقق من خاصية اللاتغير الجزئي في القياس، وقدمت تقييمًا لإجراءات الاختبار، والذي لم يتم مناقشتها سابقًا في دراسات المحاكاة السابقة، وأسفرت النتائج على أن طريقة المواءمة تُقدر معلمات النموذج المُقترح في حالات وُجود تباينات قليلة، ومتوسطة غير دالة إحصائيًا بين تلك المعلمات عبر المجموعات، وتكون مطابقة النموذج المقدر للبيانات المقبولة؛ لذلك أوصت الدراسة باستخدام طريقة المواءمة للحصول على نموذج مطابق للبيانات مهما كان نوع الفقرات المستخدمة في المقاييس.

وكذلك قارن ميوثن واسباروف (Muthen & Asparouhov, 2018) في دراستهم الأساليب الحديثة في تحري اللاتغير في القياس عبر عدة مجموعات وهي: (طريقة المواءمة (Alignment Method)، والتأثيرات العشوائية (Random Effects))، تم في هذه الدراسة وصف الطريقتين ومقارنتهما، فطريقة المواءمة تعتبر المجموعات ذات تباين ثابت النمط، في حين أن طريقة التقاطع/التحميل العشوائية من المستوى الثاني تعتبر المجموعات ذات تباينات عشوائية النمط، وللمقارنة بينهما تم تطبيق تقدير بايزون (Bayesian Estimation)، والأرجحية العظمى (Maximum Likelihood Estimation) على بيانات مُستقاة من مسح عينته (50000) مفحوص في (26) دولة، كما تم دراسة الطريقتين باستخدام محاكاة (مونت كارلو)، واقتراح قائمة من المعايير للاختيار بين الطريقتين عند دراسة اللاتغير في القياس.

كما طبقت لومازي (Lomazzi, 2018) طريقة المواءمة لاختبار اللاتغير في القياس للمسح الدولي للقيم (World Value Suvey) (WVS) لدراسة اتجاهات الأفراد عبر (59) دولة بحجم عينة (89320) مفحوص، حيث شملت الدراسة مقارنة الطريقتين الأكثر شيوعًا: (MG-CFA، وطريقة المواءمة)، ولوحظ أن الطريقة الجديدة واعدة للغاية رغم ندرة استخدامها، وأسفرت النتائج عن أنه عند استخدام الطريقة الأولى (MG-CFA): تحقق اللاتغير العددي الجزئي في

القياس لـ(27) دولة، أما عند استخدام طريقة المواءمة: فتحقق درجة مقبولة من اللاتغير في القياس لـ(35) دولة؛ لذلك أوصت الدراسة باستخدام طريقة المواءمة كطريقة بديلة لـ (MG-CFA) عند دراسة اللاتغير في القياس عبر عدد كبير من المجموعات.

وكذلك أجاب مارش آخرون (March et al., 2018) على سؤال ما يجب فعله عند فشل اللاتغير العددي في القياس؟ بأن طريقة المواءمة (Alignment Method) هي الحل لتحليل العامل متعدد المجموعة لمقارنة أوساط العوامل الكامنة عبر عدة مجموعات؛ لأن اللاتغير العددي في القياس لا يمكن تحقيقه؛ لذلك عمليًا على الأغلب يتم تقريبه فقط باستخدام أساليب مختلفة مثل اللاتغير الجزئي في القياس (Partial Measurement Invariance) استنادًا للاختيار التدريجي لتقدير المعلمات مع الكثير من مؤشرات التعديل.

تُوضح الدراسة (1): قوة ومرونة طريقة المواءمة في مقارنة أوساط العامل الكامن في الدراسات واسعة النطاق (Large-Scale Studies) لـ(30) دولة في منظمة التعاون والتنمية في الميدان الاقتصادي، واحتوت أداة القياس على (8) عوامل، (44) فقرة، و(N=249840) مفحوص، والتي عادةً لا يتم فيها تحقق اللاتغير العددي في القياس عند استخدام النهج التقليدي لتحليل العامل التوكيدي للاتغير في القياس (CFA)؛ لذلك تم استخدام طريقة المواءمة (Alignment Method) على البيانات نفسها للقيام بالتالي: اختبار اللاتغير في القياس، تحديد تباينات العامل (Factor Variance)، وتغايره (Factor Covariance)، كما تم استخدامها مع نماذج (MIMIC) (Multiple Indicators and Multiple Causes Models) متعددة المجموعات، واكتشاف التباينات في أوساط العوامل الكامنة عبر المجموعات، كما سمحت أيضًا بإجراء مقارنة أوساط العامل عبر الجنس، وعبر (30) دولة؛ أي تحليل (60) مجموعة ناتجة عن ضرب (30 دولة × 2 جنس).

أما الدراسة (2): كانت دراسة محاكاة للتأكيد على القضايا التي أُثيرت في الدراسة (1)، حيث أوضحت أن تقدير أوساط العامل الكامن بطريقة المواءمة أعلى دقة مقارنة مع اللاتغير العددي في القياس، وخصوصًا مع نماذج اللاتغير العددي الجزئي في القياس المستندة لاستراتيجية الانتقاء التدريجي لمؤشرات التعديل المُنتقدة بشدة.

وخلصت الدراسة إلى التأكيد على أن طريقة المواءمة تُوفر للباحثين التطبيقيين من مختلف التخصصات مرونة كبيرة في التعامل مع القضايا المهمة عندما لا يتناسب النموذج العاملي المقترح مع البيانات باستخدام (CFA).

كما أكد بوكروبك وديفادوف وسشميدت (Pokropek, Davidov & Schmidt, 2019) في دراستهم على أنه تم تطوير العديد من استراتيجيات نمذجة المعادلات البنائية (SEM) لتقييم اللاتغير في القياس (MI) عبر المجموعات، ومحاولة التخفيف من افتراضات (MI) التام لافتراضيات اللاتغير الجزئي أو التقريبي في القياس، إلا أنه لا يزال الباحثون التطبيقيون لا يعرفون تحت أي ظروف يمكنهم استخدام هذه الاستراتيجيات، وأي البيانات تكون صالحة للمقارنات بين المجموعات في مسوحات المقارنة واسعة النطاق؛ لذلك في هذه الدراسة تم تطبيق دراسة محاكاة (مونت كارلو) لتقييم الظروف التي تُمكّن الباحثون من الاستخدام الأمثل لمختلف أساليب (SEM) لتقدير أوساط المتغيرات الكامنة، ومعاملات المسار، وتبايناتها عبر المجموعات على الرغم من أن معاملات المسار لنماذج (SEM) قوية نسبيًا لانتهاكات افتراضات اللاتغير في القياس، ويُمكن تقديرها بشكل فعّال، إلا أن تقدير أوساط المتغير الكامن، وتصنيفات مجموعاتهم قد تكون صعبة، وأوصت الدراسة بأن اللاتغير الجزئي في القياس قد يُقدر معاملات المسار، وأوساط المتغير الكامن بشكل فعّال عندما تكون غالبية الفقرات متكافئة عبر المجموعات، وأكدت على أنه رغم صعوبة تقدير أوساط المتغير الكامن باستخدام طرق اللاتغير التقريبية، والجزئية في القياس، ولكن من

الممكن تحت ظروف محددة، وباستخدام النماذج المناسبة تقديرهم بشكل جيد؛ لذلك ينصح باستخدام طريقة المواءمة لتقدير أوساط المتغير الكامن في الحالات التي تكون فيها بعض معلمات النموذج المُقترح متباينة عبر المجموعات.

التعقيب على الدراسات السابقة

ومن خلال استعراض الباحثة للدراسات السابقة في محوريها والتي ورد ذكرها وجدت أن البعض منها:

- هدف للكشف عن اللاتغير في القياس بالطريقة الكلاسيكية التقليدية باستخدام التحليل العاملي التوكيدي (CFA)، مثل دراسة (Adam & Gary 2004، محمد 2011، الزبون 2013، Rutkowski and Svetina 2014، Ioannis 2014a، Ioannis 2014b، Ioannis 2015، السمكي 2017).
- البعض الآخر هدف لدراسة اللاتغير الجزئي أو التقريبي في القياس باستخدام طرق التقدير المختلفة، مثل دراسة (التقي 1992، Muthen & Asparouhov 2013، Schoot et al. 2013، Cieciuch et al. 2014، Zercher et al. 2015، Davidov et al., 2015).
- معظم الدراسات قارنت بين طريقتين أو أكثر من طرق الكشف عن اللاتغير في القياس الكلاسيكي (التام) مع الجزئي أو التقريبي، وأوصت باستخدام الجزئي أو التقريبي في حالات فشل تحقق اللاتغير التام في القياس، مثل دراسة (Schoot et al. 2013، Cieciuch et al. 2014، Zercher et al. 2015، Davidov et al., 2015).
- بعض الدراسات بحثت أهمية اختبار اللاتغير في القياس لمقارنة البيانات عبر الثقافات أو السياقات أو الزمن والتأكد من تحققه وأنه شرط قبلي وأساسي لإجراء هذه المقارنات ومعقولية تفسيرها، والآثار الناتجة عن انتهاك هذا الشرط، مثل دراسة (Lommen et al. 2014).

- أظهرت الدراسات السابقة التي درست طريقة المواءمة ومقارنتها مع الطرق الأخرى أن طريقة المواءمة هي الطريقة الأفضل في التعامل مع القضايا المهمة عندما لا يتناسب النموذج المقترح مع البيانات وهذا متوقع في المسوحات واسعة النطاق.
- انعدام الدراسات العربية التي تبحث في طريقة المواءمة كطريقة لقياس اللاتغير في القياس رغم أنها طريقة عملية بديلة عن طرق التحليل العاملي التوكيدي متعدد المجموعة، أما في الدراسات الأجنبية، فتم فيها توضيح لطريقة المواءمة بتطبيقها على بيانات محاكاة وفقرات ثنائية، وفقرات متعددة، وأثبتت جدارتها وفعاليتها في الكشف عن اللاتغير في القياس.
- كما لم يتم دراسة الاختبار التقييمي لمهارات الرياضيات للصف الثالث الأساسي نهائيًا.

لذلك اختلفت الدراسة الحالية عن الدراسات السابقة في كونها استخدمت طريقة المواءمة في الكشف عن اللاتغير في القياس للاختبار التقييمي لمهارات الرياضيات للصف الثالث الأساسي عبر مديريات التربية والتعليم الأردنية، حيث أن استخدام هذه الطريقة لها المزايا الآتية:

- تقدير النماذج الملائمة للعديد من المجموعات.
- القدرة على تقدير أوساط المتغير الكامن (Factor means).
- مقارنة المجموعات عند عدم إمكانية ذلك باستخدام التحليل العاملي.
- استخدامها في حالات المجموعات الكبيرة والعينات المتنوعة.
- تقدم مجموعة من التفاصيل عن المعلمات المتكافئة لكل معلمة في النموذج، ولكل مجموعة، وهذا يُساعد في تطوير الاختبارات وتكييفها.
- تبسيط العمليات الحسابية وأتمتتها.

وقد جاءت هذه الدراسة لتأخذ نوعًا من الخصوصية في كونها تناولت الكشف عن اللاتغير في القياس في أداء طلبة الصف الثالث الأساسي عبر مديريات التربية والتعليم الأردنية، حيث

استخدمت الباحثة اختبارًا تم بناؤه لضبط نوعية التعليم في وزارة التربية والتعليم الأردنية من قبل إدارة الامتحانات والاختبارات في الوزارة، ولم يتم العثور على دراسة تناولت هذا الاختبار، وذلك حسب علم الباحثة من خلال استعراضها للدراسات السابقة.

ولقد أثرت الباحثة الأدب النظري لدراستها الحالية من خلال الرجوع للدراسات السابقة، كما استفادت منها في إجراء المقارنات بين النتائج التي توصلت إليها.

الفصل الثالث

الطريقة والإجراءات

يتناول هذا الفصل وصفًا لمنهج الدراسة، ومجتمع الدراسة، وطريقة اختيار عينة الدراسة، وبناء أداة الدراسة، والإجراءات التي اتبعت للتأكد من صدق أداة الدراسة، وثباتها، كما يتضمن وصفًا للمعالجات الإحصائية التي أُجريت.

منهج الدراسة

المنهج الملائم للكشف عن اللاتغير في القياس للاختبار التقييمي لمهارات مبحث الرياضيات لطلبة الصف الثالث الأساسي عبر مديريات التربية والتعليم لعام (2018/2017م) هو المنهج الوصفي التحليلي الذي عرفه أبو النصر (2004) بأنه: "المنهج الذي يهتم بتحديد الواقع وجمع الحقائق عنه، وتحليل بعض جوانبه، بما يساهم في العمل على تطويره" (ص. 131)، حيث قامت الباحثة باستخدام طريقة المواءمة، ثم اجراء الوصف للتحليلات الاحصائية في الكشف عن اللاتغير في القياس للاختبار التقييمي.

مجتمع الدراسة

تألف مجتمع الدراسة من جميع طلبة الصف الثالث الأساسي في مدارس المملكة الأردنية الهاشمية الحكومية، والخاصة، والمدارس التابعة لوكالة الغوث، والمدارس التابعة لمديرية التعليم والثقافة العسكرية للعام الدراسي (2018/2017م) منهم (72630) طالبًا، و(74500) طالبةً.

عينة الدراسة

قُسّم مجتمع الدراسة إلى عينتين، وبشكل عشوائي، حيث تقدمت العينة الأولى إلى اختبار اللغة العربية فيما تقدمت العينة الثانية (عينة الدراسة) إلى اختبار الرياضيات (وزارة التربية والتعليم الأردنية، 2018، ص10-12).

ويبين الجدول الآتي توزيع عينة الدراسة من طلبة الصف الثالث الأساسي على مديريات التربية والتعليم حسب الجنس.

جدول 1: توزيع أفراد مجتمع الدراسة من طلبة الصف الثالث الأساسي على مديريات التربية والتعليم

اسم المديرية	عدد الطلبة الذكور	نسبة الذكور للمديرية	عدد الطلبة الإناث	نسبة الإناث للمديرية	المجموع	نسبة المديرية إلى مجتمع الدراسة	رمز المديرية
التعليم الخاص	10304	56%	8200	44%	18504	12.58%	10
قصبة عمان	2733	46%	3165	54%	5898	4.01%	11
الجامعة	1177	41%	1661	59%	2838	1.93%	12
القويسمة	2553	43%	3341	57%	5894	4.01%	13
ماركا	3537	44%	4537	56%	8074	5.49%	14
وادي السير	1039	46%	1200	54%	2239	1.52%	15
ناعور	639	44%	798	56%	1437	0.98%	16
سحاب	529	41%	768	59%	1297	0.88%	17
الجيزة	575	49%	587	51%	1162	0.79%	18
الموقر	532	46%	635	54%	1167	0.79%	19
مادبا	1578	50%	1550	50%	3128	2.13%	20
ذيبان	327	47%	366	53%	693	0.47%	21
قصبة اربد	3400	46%	3932	54%	7332	4.98%	22
بني عبيد	1347	51%	1317	49%	2664	1.81%	23
المزار الشمالي	663	52%	624	48%	1287	0.87%	24
الطيبة والوسطية	511	46%	604	54%	1115	0.76%	25
الكورة	992	46%	1185	54%	2177	1.48%	26
بني كنانة	1014	50%	1018	50%	2032	1.38%	27
الرمثا	1026	44%	1303	56%	2329	1.58%	28
الأغوار الشمالية	669	47%	769	53%	1438	0.98%	29
جرش	2099	50%	2078	50%	4177	2.84%	30
عجلون	1816	52%	1677	48%	3493	2.37%	31
قصبة المفرق	2071	50%	2065	50%	4136	2.81%	32
البادية الشمالية الشرقية	943	47%	1052	53%	1995	1.36%	33
البادية الشمالية الغربية	965	48%	1027	52%	1992	1.35%	34
الزرقاء الأولى	5340	52%	5023	48%	10363	7.04%	35
الزرقاء الثانية	2336	50%	2309	50%	4645	3.16%	36
الرصيفة	3435	50%	3475	50%	6910	4.70%	37
قصبة السلط	1983	50%	1965	50%	3948	2.68%	38

اسم المديرية	عدد الطلبة الذكور	نسبة الذكور للمديرية	عدد الطلبة الإناث	نسبة الإناث للمديرية	المجموع	نسبة المديرية إلى مجتمع الدراسة	رمز المديرية
دير علا	722	51%	703	49%	1425	0.97%	39
الشونة الجنوبية	588	53%	523	47%	1111	0.76%	40
عين الباشا	1800	56%	1438	44%	3238	2.20%	41
قصبة الكرك	1177	50%	1161	50%	2338	1.59%	42
المزار الجنوبي	926	50%	939	50%	1865	1.27%	43
القصر	507	50%	508	50%	1015	0.69%	44
الأغوار الجنوبية	667	49%	685	51%	1352	0.92%	45
الطفيلة	751	51%	728	49%	1479	1.01%	46
بصيرا	321	52%	300	48%	621	0.42%	47
معان	346	49%	355	51%	701	0.48%	48
البتراء	307	47%	346	53%	653	0.44%	49
الشوبك	144	49%	149	51%	293	0.20%	50
البادية الجنوبية	554	44%	704	56%	1258	0.86%	51
العقبة	1618	49%	1684	51%	3302	2.24%	52
وكالة الغوث	5393	48%	5832	52%	11225	7.63%	53
الثقافة العسكرية	676	76%	214	24%	890	0.60%	54
المجموع	72630	49.4%	74500	50.6%	147131	100%	

أداة الدراسة: تكونت أداة الدراسة من:

- تكوّن الاختبار من (25) فقرة من نوع الاختيار من متعدد موزّعة على محاور التعلم في مبحث الرياضيات لطلبة الصف الثالث الأساسي للعام الدراسي (2018/2017) وهي: الأعداد والعمليات، والجبر (الأنماط)، والقياس، والهندسة، والإحصاء والاحتمالات.
- أما المهارات التي شملها اختبار الرياضيات فهي: تمثيل الأعداد وقراءتها وكتابتها، مقارنة الأعداد وتقريبها، وإجراء العمليات الحسابية على الأعداد، وإكمال النمط، وتحديد وحدات القياس المختلفة، وتمييز الأشكال الهندسية والمجسمات والعلاقات بينها، وجمع البيانات وتسجيلها وتنظيمها بجدول الإشارات والصور، وتطبيق مفهومي التغير والنسبة (وزارة التربية والتعليم الأردنية، 2018، ص12).

خطوات إعداد الاختبار

كما وضحتها وزارة التربية والتعليم الأردنية في تقريرها كانت كالتالي: شُكلت لجنة لكتابة فقرات اختبار الصف الثالث الأساسي في مبحث الرياضيات من مشرفي إدارة الامتحانات والاختبارات، وعدد من مشرفي الميدان أُتبعت فيها الخطوات الآتية:

1- **بناء الفقرات:** تحليل محتوى كتاب الصف الثالث الأساسي في مبحث الرياضيات، ودراسة نتاجات التعلّم الخاصة بالصف الثالث الأساسي في مبحث الرياضيات المُستهدف في الاختبار، وتحديد محاور ومهارات التعلّم التي سيقيسها الاختبار، وإعداد جدول مواصفات للاختبار، وبناء تجمع الفقرات الاختبارية وطباعتها.

2- **تجريب الفقرات:** جُربت على عينة تجريبية من الطلبة، وأُجري التحليل الإحصائي لاستجابات الطلبة على الفقرات، للتعرّف على مدى مقروئيتها، وخصائصها السيكومترية مثل: فاعلية البدائل، ومستوى صعوبة الفقرات وتمييزها، ثم اُختبرت الفقرات الملائمة، وعُدلت صياغة بعض الفقرات، وحُذفت الفقرات التي كان تمييزها سالبًا، أو قريبًا من الصفر.

3- **إعداد الاختبار في صورته النهائية:** انتُقيت الفقرات النهائية من تجمع الفقرات التي بُنِيت، بحيث كان الانتقاء مُمثِلًا للأوزان والنسب في جدول المواصفات الذي طُوِّر ليشمل مؤشرات الأداء (وزارة التربية والتعليم الأردنية، 2018، ص13).

والجدول (2) التالي يبين توزيع فقرات الاختبار على مهارات مبحث الرياضيات السبع التي يقيسها الاختبار التقييمي في صورته النهائية، وحسب جدول المواصفات الذي تم بناؤه:

جدول 2: توزيع فقرات الاختبار على مهارات مبحث الرياضيات السبع التي يقيسها الاختبار التقييمي

رقم المهارة	اسم المهارة	أرقام الفقرات
1	تمثيل الأعداد وقراءتها وكتابتها، وتمثيل الكسور العادية والأعداد الكسرية باستخدام المحسوسات والأشكال	1، 2، 3، 13، 14
2	مقارنة الأعداد وتقريبها	4، 5
3	إجراء العمليات الحسابية على الأعداد، وعمليتي الجمع والطرح على الكسور (خطوتان على الأكثر)	6، 7، 8، 9، 10، 11، 12
4	إكمال النمط وتوضيح قاعدة النمط	15، 16، 17، 18
5	تحديد وحدات القياس المختلفة	19، 20، 21
6	تمييز الأشكال الهندسية والمجسمات والعلاقات بينها	22 ، 23
7	جمع البيانات وتسجيلها وتنظيمها بجدول الإشارات والصور	24، 25

متغيرات الدراسة: اشتملت الدراسة على المتغيرات التالية:

1- المتغيرات المستقلة (Independent Variables): مديريات التربية والتعليم الأردنية (45 مديرية)، وجنس الطالب (ذكر، أنثى).

2- المتغيرات التابعة (Dependent Variables): أداء طلبة الصف الثالث الأساسي على الاختبار التقييمي لمهارات الرياضيات للعام الدراسي (2018/2017).

ثبات الاختبار

تم تقدير ثبات اختبار الرياضيات للصف الثالث الأساسي باستخدام معامل الثبات "كرونباخ ألفا" وبلغت قيمته (0.88)، وتُعد هذه القيمة مناسبة لإجراء هذه الدراسة.

إجراءات جمع البيانات

1- عُقد اجتماع مع رؤساء أقسام الامتحانات والاختبارات في مديريات التربية والتعليم جميعها لتوضيح أهداف الاختبار التقييمي في الرياضيات للصف الثالث الأساسي، وإجراءات تطبيقه، وموعد تسليم أوراق الإجابة.

2- تسليم مغلفات الأسئلة، وأوراق الإجابة (القارئ الضوئي) للمديريات.

3- تسليم المديريات تعليمات، وإرشادات التطبيق للاختبار التقييمي.

4- تطبيق الاختبار وفق الموعد المحدد.

5- استلام أوراق الإجابة (القارئ الضوئي) من المديريات.

6- إدخال إجابات الطلبة للحاسوب باستخدام القارئ الضوئي، وتصحيح الاختبار آليًا.

7- التحليل الإحصائي لنتائج الاختبار، وإيجاد كل مما يأتي:

- نِسَبْ الطلبة على مستوى المملكة حسب مستويات الأداء.

- المتوسط الحسابي لأداء الطلبة على مستوى المملكة، والمديرية، والمدرسة.

- المتوسط الحسابي لأداء الطلبة على محاور التعلم، ومهاراته على مستوى المملكة، والمديرية، والمدرسة.

- المتوسط الحسابي لأداء الطلبة في كل مستوى وفقًا لمستويات الأداء (وزارة التربية والتعليم الأردنية، 2018، ص14).

المعالجات الإحصائية

لمطابقة استجابات طلبة الصف الثالث الأساسي على الاختبار التقييمي لمهارات الرياضيات ثنائي التدريج (Dichotomos) مع النموذج السباعي المقترح من قبل واضعي الاختبار، اعتمدت المعالجة الإحصائية على النتائج التي تم الحصول عليها من برمجية (Mplus v8.4)، وبرمجية (Amos v26)، إذ تم استخدام بعض المحكات الإحصائية للحكم على مدى جودة مطابقة النموذج للبيانات كما يلي:

للإجابة على السؤال الأول؛ تم استخدام أسلوب التحليل العاملي التوكيدي (CFA) من الدرجة الثانية في برمجية (Mplus v8.4) للتأكد من مُطابقة النموذج المُقترح المكون من سبع عوامل، وعامل عام للبيانات بشكل كافٍ، اعتمادًا على مؤشرات مُطابقة النموذج المُقترح للبيانات

(Model-Data Fit Indicators)، وهي مؤشرات إحصائية، أو وصفية تُساعد الباحث على تحديد مدى جودة مطابقة النموذج المُقترح للبيانات المستخدمة عن طريق مقارنته بنموذج آخر (Gadelrad, 2004)، ومن أهم هذه المؤشرات كما وضحها جورج واريني (George, & Irini. 2002) ولاتن وكارول وجرين (Lattin, Carrol & Green. 2003):

1- مؤشر مربع كاي (χ^2) (Chi-Square).

2- مؤشر المطابقة المقارن (CFI) (Comparative Fit Index).

3- مؤشر جذر وسط مربع الخطأ التقريبي (RMSEA) (Root Meam Square Error of Approximatin).

4- مؤشر الجذر المعياري لوسط البواقي (SRMR) (Standardized Root Mean Residual).

5- مؤشر تاكر لويس (TLI) (Tucker-Lewis Index).

وتم اعتماد هذه المؤشرات الإحصائية لحسن المطابقة، والحكم عليها وفق علامة قطع متفق عليها في الأدب النظري، ويبين الجدول (3) المؤشرات، وعلامة القطع لكل منها.

جدول 3: مؤشرات جودة المطابقة، وعلامة القطع لكل منها

اسم المؤشر	الرمز	علامة القطع	المصدر
مربع كاي	χ^2	غير دال	Cangur & Ercan, 2015
مربع كاي/درجات الحرية	χ^2/df	أقل من 5	Cangur & Ercan, 2015
مؤشر المطابقة المقارن	CFI	الحد الأدنى 0.8	Hair, Anderson, Tatham, & Bentler, 1995
مؤشر جذر وسط مربع الخطأ التقريبي	RMSEA	الحد الأعلى 0.08	Hu & Bentler, 1999
مؤشر الجذر المعياري لوسط البواقي	SRMR	الحد الأعلى 0.09	Hu & Bentler, 1999
مؤشر تاكر لويس	TLI	**1-0**	Bentler, 1990

للإجابة على السؤال الثاني؛ تم استخدام أسلوب التحليل العاملي التوكيدي متعدد المجموعة (MG-CFA) في برمجية (Amos v26) لفحص اللاتغير في القياس للبناء العاملي للاختبار التقييمي عبر مديريات التربية والتعليم الأردنية.

للإجابة على السؤال الثالث؛ تم استخدام أسلوب التحليل العاملي التوكيدي متعدد المجموعة (MG-CFA) في برمجية (Amos v26) لفحص اللاتغير في القياس للبناء العاملي للاختبار التقييمي عبر جنس الطالب.

للإجابة على السؤال الرابع؛ تم استخدام طريقة المواءمة (Alignment Method) في برمجية (Mplus v8.4) بخطوتيها: الحرة (FREE)، ثم الثابتة (FIXED)؛ للكشف عن اللاتغير في القياس في أداء طلبة الصف الثالث على الاختبار التقييمي لمهارات الرياضيات عبر مديريات التربية والتعليم الأردنية، بسبب إنطباق المعايير التي وضحها ميوثن واسباروف (Muthén & Asparouhov, 2017) للمفاضلة بين الطرق المختلفة في الكشف عن اللاتغير في القياس عبر مجموعات متعددة، وهي:

1- عدد المؤشرات الدالة على العامل: (25) فقرة وهذا يناسب استخدام طريقة المواءمة (Alignment Method).

2- عدد المجموعات: بين (30-100) مجموعة، والمجموعات موضع الاهتمام (45) مديرية تتناسب مع المعايير المتبعة.

3- حجم المجموعة: متفاوتة من (141-9127).

4- معلومات حول المجموعات والفقرات التي قد تسبب عدم تحقق خاصية اللاتغير في القياس: تُوفر طريقة المواءمة تفصيلات حول المجموعات، والفقرات التي تُساعد في تطوير، وتحسين فقرات الاختبار، وإجراءات تطبيق الاختبار على مجتمع الدراسة.

5- طبيعة توزيعات المعالم: لا يحتاج تطبيق طريقة المواءمة لأن يكون توزيع الفقرات طبيعيًا، فهي تتعامل مع أي توزيع كان للبيانات.

6- تعقيدات البيانات المسحية: في الأدوات المسحية يتم اختيار العينات باستخدام الاحتمال النسبي بالنسبة للحجم (Probabilities Proportional to Size) (PPS)، وأوزان الحجوم، والتقسيم الطبقي، وجميعها يمكن لطريقة المواءمة التعامل معها.

7- السرعة الحسابية: سرعة تطبيق خطوات طريقة المواءمة، وبساطتها.

وإجراءاتها كالتالي: 1- اختيار النموذج التكويني الملائم للاختبار التقييمي لمهارات الرياضيات للصف الثالث الأساسي، وهذا ما سيتم في إجابة السؤال الثاني، واعتماده حسب المؤشرات الإحصائية لحُسن مطابقة النموذج للبيانات، كخطوة أولى في تطبيق طريقة المواءمة، وهنا تم استخدام النموذج السباعي المُقترح.

2- تطبيق طريقة المواءمة مع خيار تقدير بايزون (Bayesian) للجمع بين اللاتغير التقريبي في القياس، وطريقة المواءمة.

3- إعداد نموذج المواءمة الحر (FREE).

4- إعداد نموذج المواءمة الثابت (FIXED)، وذلك بناءً على تنويه من البرمجية في الخطوة السابقة (3).

5- مقارنة أوساط العامل فيما بينها.

6- تحديد المجموعات والفقرات التي تحقق عبرها خاصية اللاتغير في القياس.

الهدف الأساسي للمواءمة هو توفير مقارنة بين أوساط العوامل، وتبايناتهم عبر المجموعات مع السماح بقياس اللاتغير التقريبي في القياس، وكما يمكن توفير معلومات حول درجة اللاتغير

في القياس، وبعد الانتهاء من تقدير المواءمة يمكن إجراء تحليل مفصل لتحديد معلمات القياس المتكافئة تقريبًا عبر المجموعات، لن نقدم مبررًا نظريًا، بل نقدم تفاصيل حول خوارزمية ما بعد التقدير التي تستخدم لتحديد اللاتغير في القياس، وهذا الإجراء يعمل بشكل جيد للغاية مع البيانات المولدة، حيث تكون المعلمات المتكافئة وغير المتكافئة عبر المجموعات معروفة بالتصميم، وبالتالي نتوقع أن يعمل الإجراء بشكل جيد في التطبيقات العملية كذلك، كما تم ذكر تفاصيل الخوارزمية المنفذة في (Mplus v8.4).

الفكرة وراء الخوارزمية هي كما يلي: 1- لكل معلمة قياس تم العثور على أكبر تجمع ثابت من المجموعات، بحيث تكون معلمة القياس لكل مجموعة في هذا التجمع الثابت لا تختلف بشكل دال إحصائيًا عن وسط قيم تلك المعلمة عبر كل المجموعات في التجمع الثابت، وكل مجموعة ليست في التجمع الثابت فإن معلمة القياس فيها تختلف بشكل دال إحصائيًا عن الوسط.

2- الخوارزمية تعتمد على مقارنات زوجية متعددة للمعلمات عبر المجموعات، مثل: اختبارات متعددة يتم تطبيقها لرفض المجموعات، أو الفقرات التي تختلف بشكل دال إحصائيًا عن وسط التجمع المذكور في الخطوة (1)؛ لذلك تستخدم قيم (p) أصغر من (0.05) للفقرات الثنائية.

خطوات الخوارزمية: 1- الخطوة الأولى في الخوارزمية: هي تحديد تجمع انطلاق من المجموعات المتكافئة.

2- يتم عمل اختبار زوجي لكل زوج من المجموعات، ويتم "بربط" المجموعتين إذا كانت قيمة (p) التي تم الحصول عليها بواسطة اختبار المقارنة الزوجية أكبر من (0.01).

3- ثم يتم تحديد أكبر تجمع مرتبط لتلك المعلمة، وهذه تكون هي بداية تجمع المجموعات، وسيتم تعديل تجمع البداية باستخدام الإجراء التالي:

أولاً، يتم حساب وسط المعلمة باستخدام التجمع الثبات الحالي، ثم بالنسبة لكل مجموعة يتم إجراء اختبار الدلالة الإحصائية لمقارنة قيمة المعلمة لكل مجموعة بوسط المجموعة الحالي، وإذا كانت قيمة (p) أعلى من (0.001)، يتم إضافة هذه المجموعة للتجمع الثابت، وإذا كانت أقل من (0.001)، تتم إزالة تلك المجموعة منها، ثم نكرر هذه العملية حتى تستقر المجموعة الثابتة. ولا تتم إضافة أو إزالة مجموعات من مجموعة الثبات، كما يتم إضافة قواعد أخرى لضمان استقرار العملية بالفعل، وفي حالة الإجراء السابق المستخدم مع طريقة دلتا وتقدير بايزون يتم استخدام التوزيع الخلفي (posterior distribution) للاختبار الإحصائي، ومع تقدير بايزون، يوصى بأن تكون تسلسل (Markov chain Monte Carlo algorithms) (MCMC) أطول للحصول على تقدير أدق لقيم (p) الصغيرة، وقد يكون (1000) تكرار (iteration) لـ(MCMC) كافٍ، كما يستخدم الإجراء أعلاه قيم (p) صغيرة كدرجة قطع، كما أنه من المهم تقدير قيم (p) الصغيرة بدقة، ويمكن إجراء تحليل اللاتغير في القياس لمعلمات الأفراد، ومعلمات فقرات الاختبار، ويتم استخدامها كفقرة ربط (anchor item) في (MG-CFA)؛ لتحديد الفقرات غير المتكافئة التي يمكن بعد ذلك مراجعتها أو إزالتها من أداة القياس.

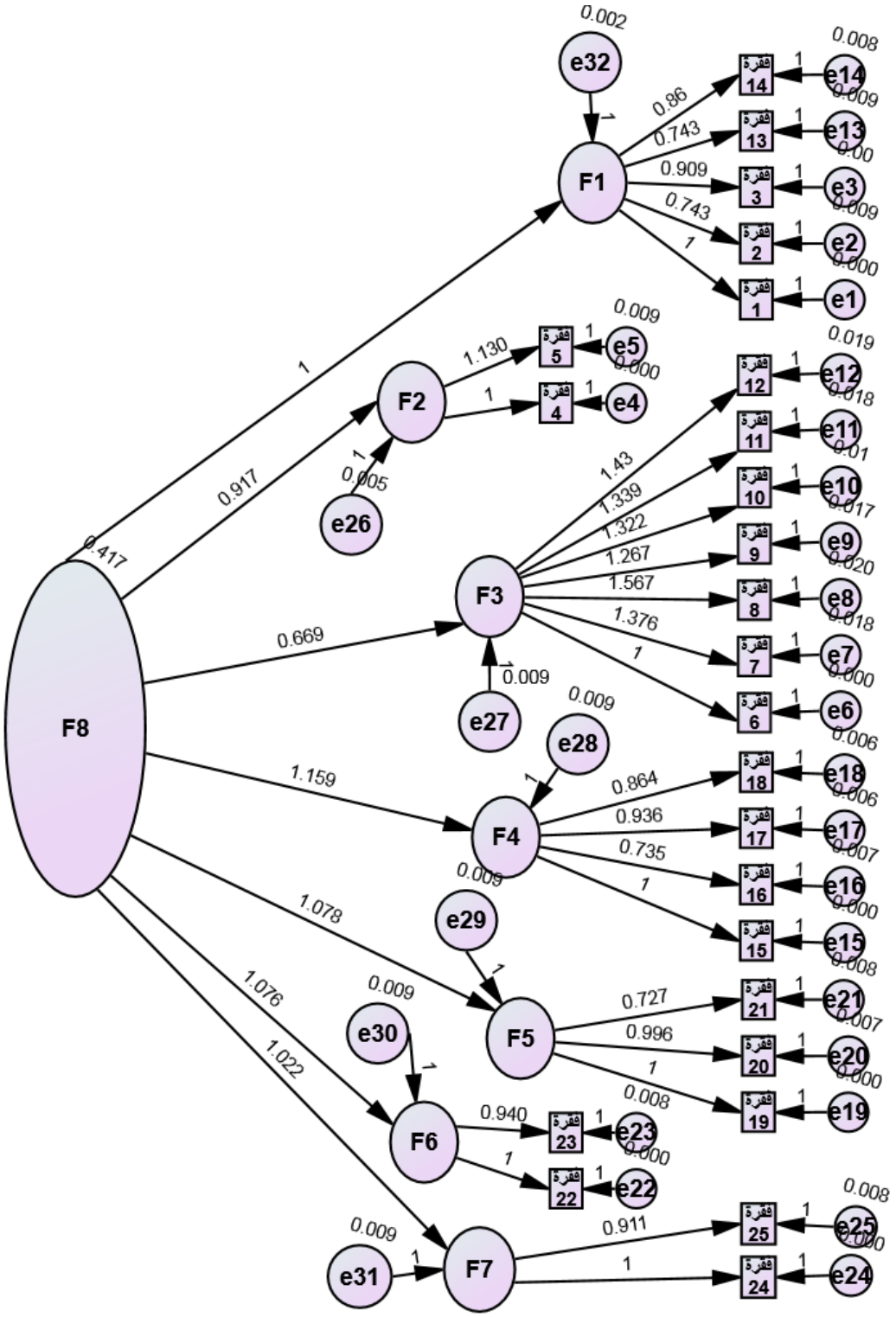
الفصل الرابع
نتائج الدراسة

يتناول هذا الفصل عرضًا للنتائج التي توصلت إليها الدراسة الحالية، والتي هدفت إلى الكشف عن "فاعلية طريقة المواءمة في الكشف عن اللاتغير في أداء طلبة الصف الثالث الأساسي على الاختبار التقييمي لمهارات الرياضيات عبر مديريات التربية والتعليم الأردنية"، وقد عُرضت في ضوء أسئلة الدراسة، وأُتبع في العرض المنهجية العلمية في الإجابة عن الأسئلة حيث عُرض نص السؤال، يليه بيان نوع التحليل الإحصائي المناسب للإجابة عنه، ثم جدولة البيانات وفق هذا التحليل وصولًا إلى التعليق على النتائج والبيانات المتضمنة في الجدول، وفيما يلي تقديم لتلك النتائج التي تم التوصل إليها وفق ترتيب أسئلة الدراسة.

النتائج المتعلقة بالسؤال الأول

نص السؤال الأول على: "هل البناء العاملي المُقترح لمحتوى اختبار مهارات الرياضيات للصف الثالث الأساسي يُطابق البيانات بشكل كاف؟"

للإجابة على السؤال الأول؛ تم استخدام أسلوب التحليل العاملي التوكيدي (CFA) من الدرجة الثانية في برمجية (Mplus v8.4) للتأكد من مطابقة النموذج المُقترح المُكون من سبع عوامل للبيانات بشكل كافٍ، اعتمادًا على مؤشرات مطابقة النموذج للبيانات (Model–Data fit indicators)، ويبين شكل (2) النموذج العاملي السباعي مع عامل عام واحد المُقترح لمحتوى اختبار الرياضيات (25 فقرة).



f1-f7: العوامل السبعة الممثلة لمهارات الرياضيات للاختبار التقييمي الموضحة في الجدول (2).

f8: العامل العام؛ وهو مهارات الرياضيات (اسم الاختبار).

u1-u25: فقرات الاختبار من نوع الاختيار من متعدد.

شكل 2: النموذج العاملي المقترح للاختبار التقييمي من الدرجة الثانية

وللتحقق من مطابقة النموذج العاملي المُقترح للبيانات الناتجة من استجابات طلبة العينة (73448) على الاختبار، تم إجراء التحليل العاملي التوكيدي (Cofirmatory Factor Analysis) (CFA) من الدرجة الثانية على الفقرات الـ(25) التي أُستخدمت في تقييم قُدرة نموذج العوامل على التعبير عن مجموعة البيانات الفعلية في المقارنة بين نماذج العوامل في هذا المجال، وتم استخدام برمجية (Mplus V8.4)، وذلك لاختبار مدى جودة مطابقة النموذج المُقترح للبيانات، والتأكّد من صدق الاختبار، ويبين الجدول (4) نتائج مؤشرات جودة مُطابقة النموذج للبيانات المُستقاة من البرمجية المستخدمة.

جدول 4: مؤشرات جودة المطابقة لنموذج مهارات الرياضيات

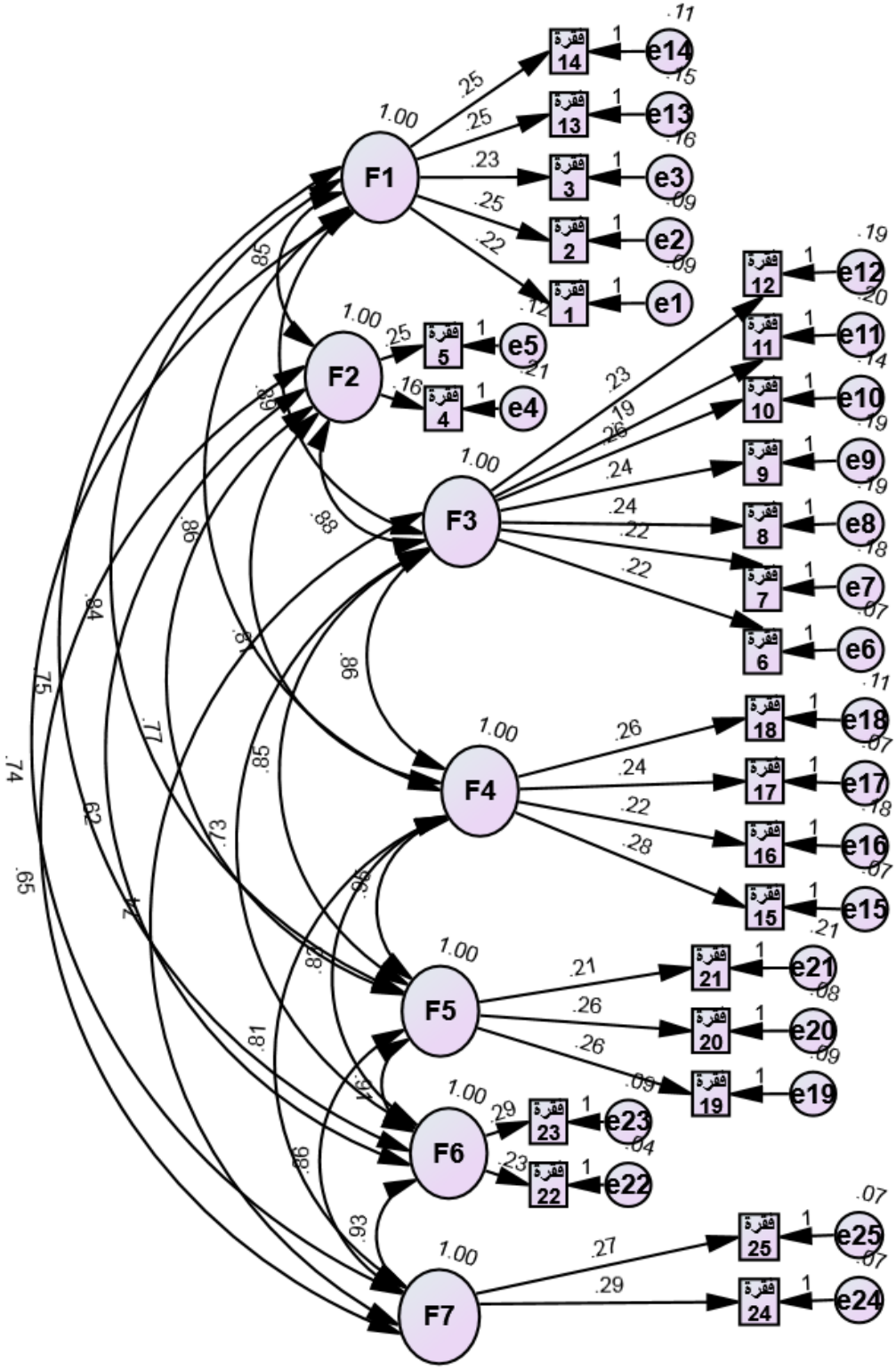
المؤشر	حدود الثقة	القيمة
χ^2	دالة	19573.376
P	$0.05 \geq$	0.00
RMSEA	$0.08 \geq$	0.034
CFI	$0.8 \leq$	0.975
TLI	0−1	0.972
SRMR	$0.09 \geq$	0.037

يُلاحظ من الجدول (4) وجود دلالة إحصائية لقيمة (χ^2) وهذا متوقع، ويمكن تجاهله بسبب حجم عينة الدراسة الكبير، وارتفاع قيم معاملات الارتباط بين العوامل (Rutkowski and Svetina, 2014; Karakaya-Ozyer & Aksu-Dunya, 2018)، أما كافة قيم المؤشرات الأُخرى فتقع ضمن المدى المقبول لمُطابقة النموذج المستخدم للبيانات، والمُوضح في الشكل (3)، وهذا يدل على أن البناء العاملي ذي السبع عوامل المنطوية تحت العامل العام للاختبار مُطابق للبيانات بشكل كافٍ.

النتائج المتعلقة بالسؤال الثاني

نص السؤال الثاني على: "هل يختلف البناء العاملي للاختبار التقييمي لمهارات الرياضيات للصف الثالث الأساسي عبر مديريات التربية والتعليم الأردنية؟"

للإجابة عن السؤال الثاني؛ تم استخدام التحليل العاملي التوكيدي متعدد المجموعة (MG–CFA) من الدرجة الأولى في برمجية (Amos v26) على البيانات كاملة لاختبار مدى تحقق خاصية اللاتغير في القياس (ثبات القياس) للبناء العاملي السباعي المُقترح للاختبار التقييمي لمهارات الرياضيات للصف الثالث الأساسي عبر مديريات التربية والتعليم الأردنية الـ(45) المُوضح في الشكل (3)، وقد تم اختبار اللاتغير التكويني (ثبات الشكل)، واللاتغير المتري (الثبات المتري)، واللاتغير العددي (ثبات التدريج).



F1-F7: العوامل السبعة الممثلة لمهارات الرياضيات للاختبار التقييمي الموضحة في الجدول (2).

شكل 3: النموذج العاملي المُقترح للاختبار التقييمي عبر مديريات التربية والتعليم الأردنية الـ (45)

أولًا: اللاتغير التكويني

تم فحص ملاءمة النموذج المُقترح والموضح في شكل (3)، وتم إجراؤه دون أية قيود (No constraints) على المعالم عبر المديريات، ويبين الجدول (5) قيم مؤشرات جودة مُطابقة النموذج المُقترح للبيانات عبر المديريات تحت شروط اللاتغير التكويني وهي عدم فرض أي قيود على المعالم عبر المديريات، مع حدود الثقة حسب الأدب النظري.

جدول 5: مؤشرات جودة مطابقة النموذج المُقترح للبيانات عبر مديريات التربية والتعليم

المؤشر	حدود الثقة	القيم تحت شروط اللاتغير التكويني (ثبات الشكل)
χ^2	دالة	41483.760
df	–	11430
χ^2/df	$5 \geq$	3.629
P	$0.05 \geq$	0.00
RMSEA	$0.08 \geq$	0.006
CFI	$0.8 \leq$	0.924
TLI	0–1	0.91
SRMR	$0.09 \geq$	0.01

يُلاحظ من الجدول (5) وُجود دلالة إحصائية لقيمة مؤشر (χ^2) وهذا متوقع، ويمكن تجاهله بسبب حجم عينة الدراسة الكبير، وارتفاع قيم معاملات الارتباط بين العوامل (Rutkowski and Svetina, 2014; Karakaya-Ozyer & Aksu-Dunya, 2018)، أما كافة قيم المؤشرات الأخرى فتقع ضمن حدود الثقة لجودة مُطابقة النموذج المُستخدم للبيانات الموضح في الشكل (3)، وهذا يدل على أن البناء العاملي سباعي العوامل المُقترح للاختبار حقق مستوى اللاتغير التكويني وفقًا لمتغير مديريات التربية والتعليم الأردنية الـ(45).

ثانيًا: اللاتغير المتري (الثبات المتري)

تم فحص وجود اللاتغير المتري في فقرات الاختبار وفقًا لمتغير مديريات التربية والتعليم بوضع قيود على النموذج المراد فحصه، حيث تم وضع قيود على جميع الفقرات؛ أي فحص مدى ثبات تشبعات الفقرات على العوامل الكامنة بدرجة متماثلة عبر المديريات، والجدول (6) يبين قيم مؤشرات جودة مطابقة النموذج للبيانات مع فرض قيود اللاتغير المتري.

جدول 6: مؤشرات جودة مطابقة النموذج المُقترح للبيانات عبر مديريات التربية والتعليم مع فرض قيود اللاتغير المتري

المؤشر	حدود الثقة	القيم تحت شروط اللاتغير المتري (الثبات المتري)
χ^2	دالة	47531.310
df	–	12530
χ^2/df	$5 \geq$	3.793
P	$0.05 \geq$	0.00
RMSEA	$0.08 \geq$	0.006
CFI	$0.8 \leq$	0.912
TLI	0–1	0.905
SRMR	$0.09 \geq$	0.013

يُلاحظ من الجدول (6) وُجود دلالة إحصائية لقيمة مؤشر (χ^2) وهذا متوقع، ويمكن تجاهله بسبب حجم عينة الدراسة الكبير، وارتفاع قيم معاملات الارتباط بين العوامل (Rutkowski and Svetina, 2014; Karakaya-Ozyer & Aksu-Dunya, 2018)، أما كافة قيم المؤشرات الأخرى فتقع ضمن حدود الثقة لجودة مُطابقة النموذج المُستخدم للبيانات الموضح في الشكل (3)، وللتحقق من وجود اللاتغير المتري حُسبت المقادير الآتية: (ΔCFI, ΔSRMR, ΔTLI, ΔRMSEA) التي تبين الفرق في قيم (CFI, SRMR, TLI, RMSEA) بين النموذج التكويني والنموذج المتري، ويبين جدول (7) نتائج هذه القيم.

جدول 7: الفرق بين مؤشرات جودة مطابقة النموذج (التكويني والمتري) المُقترح للبيانات عبر مديريات التربية والتعليم

المؤشر	حدود الثقة	الفرق بين المؤشرات
χ^2	دالة	6047.550
df	–	1100
P	≥ 0.05	0.00
Δ RMSEA	≥ 0.015	0.000
Δ CFI	≥ 0.01	0.012
Δ TLI	≥ 0.01	0.006
Δ SRMR	≥ 0.03	0.003

يُلاحظ من الجدول (7) أن قيم الفرق في مؤشرات (ΔCFI, ΔSRMR, ΔTLI, ΔRMSEA) تقع ضمن حدود الثقة بين النموذج التكويني والنموذج المتري، وهذا يدل على تحقق اللاتغير المتري.

ثالثًا: اللاتغير العددي (ثبات التدريج)

تم فحص وجود اللاتغير العددي في فقرات الاختبار وفقًا لمتغير مديريات التربية والتعليم بوضع قيود على النموذج المراد فحصه، حيث تم فحص مدى ثبات تقاطعات الفقرات على العوامل الكامنة بدرجة متماثلة عبر المديريات؛ أي نقطة بداية الفقرة عندما تكون درجة العامل الكامن (0)، والجدول (8) يبين قيم مؤشرات جودة مطابقة النموذج للبيانات مع فرض قيود اللاتغير العددي.

جدول 8: مؤشرات جودة مطابقة النموذج المُقترح للبيانات عبر مديريات التربية والتعليم مع فرض قيود اللاتغير العددي

المؤشر	حدود الثقة	القيم تحت شروط اللاتغير العددي (ثبات التدريج)
χ^2	دالة	49299.865
df	–	13454
χ^2/df	$5 \geq$	3.664
P	$0.05 \geq$	0.00
RMSEA	$0.08 \geq$	0.006
CFI	$0.8 \leq$	0.909
TLI	0–1	0.909
SRMR	$0.09 \geq$	0.013

يُلاحظ من الجدول (8) وُجود دلالة إحصائية لقيمة مؤشر (χ^2) وهذا متوقع، ويمكن تجاهله بسبب حجم عينة الدراسة الكبير، وارتفاع قيم معاملات الارتباط بين العوامل (Rutkowski and Svetina, 2014; Karakaya-Ozyer & Aksu-Dunya, 2018)، أما كافة قيم المؤشرات الأخرى فتقع ضمن حدود الثقة لجودة مُطابقة النموذج المُستخدم للبيانات الموضح في الشكل (3)، وللتحقق من وجود اللاتغير العددي حُسبت المقادير الآتية: (ΔCFI, ΔSRMR, ΔTLI, ΔRMSEA) التي تبين الفرق في قيم (CFI, SRMR, TLI, RMSEA) بين النموذج المتري والنموذج العددي، ويبين جدول (9) نتائج هذه القيم.

جدول 9: الفرق بين مؤشرات جودة مطابقة النموذج (المتري والعددي) المُقترح للبيانات عبر مديريات التربية والتعليم

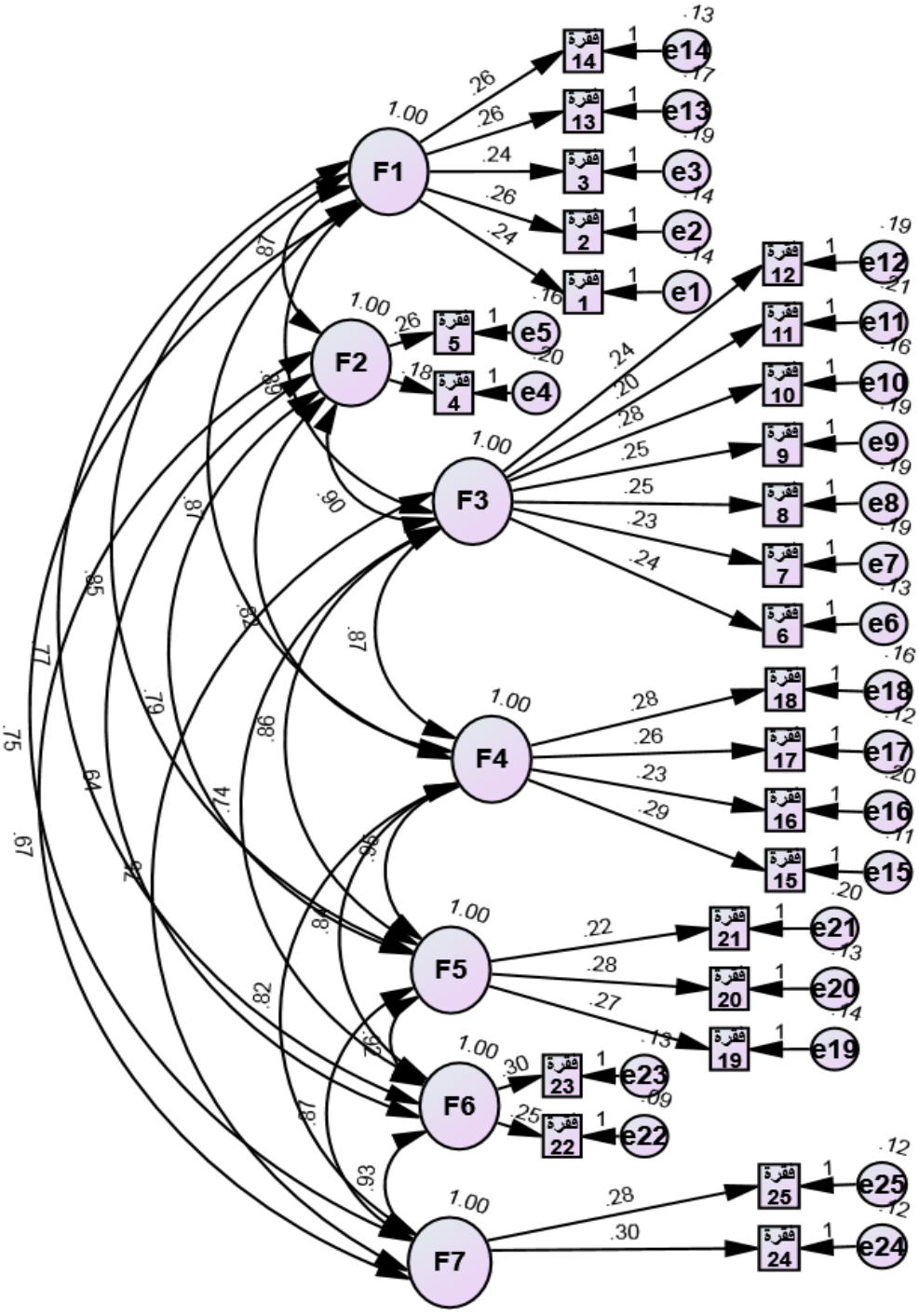
المؤشر	حدود الثقة	الفرق بين المؤشرات
χ^2	دالة	1768.555
df	–	924
P	≥ 0.05	0.00
Δ RMSEA	≥ 0.015	0.000
Δ CFI	≥ 0.01	0.003
Δ TLI	≥ 0.01	−0.004
Δ SRMR	≥ 0.03	0.000

يُلاحظ من الجدول (9) أن قيم الفرق في مؤشرات (ΔCFI, ΔSRMR, ΔTLI, ΔRMSEA) تقع ضمن حدود الثقة بين النموذج المتري والنموذج العددي، وهذا يدل على تحقق اللاتغير العددي.

النتائج المتعلقة بالسؤال الثالث

نص السؤال الثالث على: "هل يختلف البناء العاملي للاختبار التقييمي لمهارات الرياضيات للصف الثالث الأساسي عبر جنس الطالب؟"

للإجابة عن السؤال الثالث؛ تم استخدام التحليل العاملي التوكيدي متعدد المجموعة (MG–CFA) من الدرجة الأولى في برمجية (Amos v26) على البيانات كاملة لاختبار مدى تحقق خاصية اللاتغير في القياس (ثبات القياس) للبناء العاملي السباعي المُقترح للاختبار التقييمي لمهارات الرياضيات للصف الثالث الأساسي عبر جنس الطالب بفئتيه (ذكر، أنثى) المُوضح في الشكل (4)، وقد تم اختبار اللاتغير التكويني (ثبات الشكل)، واللاتغير المتري (الثبات المتري)، واللاتغير العددي (ثبات التدريج).



F1-F7: العوامل السبعة الممثلة لمهارات الرياضيات للاختبار التقييمي الموضحة في الجدول (2).

شكل 4: النموذج العاملي المُقترح للاختبار التقييمي عبر جنس الطالب (ذكر، أنثى)

أولًا: اللاتغير التكويني

تم فحص ملاءمة النموذج المُقترح والموضح في شكل (4)، وتم إجراؤه دون أية قيود (No constraints) على المعالم عبر جنس الطالب، ويبين الجدول (10) قيم مؤشرات جودة مُطابقة النموذج المُقترح للبيانات عبر جنس الطالب تحت شروط اللاتغير التكويني وهي عدم فرض أي قيود على المعالم عبر جنس الطالب ، مع حدود الثقة حسب الأدب النظري.

جدول 10: مؤشرات جودة مطابقة النموذج المُقترح للبيانات عبر جنس الطالب

المؤشر	حدود الثقة	القيم تحت شروط اللاتغير التكويني (ثبات الشكل)
χ^2	دالة	39802.417
df	–	508
P	≥ 0.05	0.00
RMSEA	≥ 0.08	0.006
CFI	≤ 0.8	0.924
TLI	0–1	0.927
SRMR	≥ 0.09	0.027

يُلاحظ من الجدول (10) وُجود دلالة إحصائية لقيمة مؤشر (χ^2) وهذا متوقع، ويمكن تجاهله بسبب حجم عينة الدراسة الكبير، وارتفاع قيم معاملات الارتباط بين العوامل (Rutkowski and Svetina, 2014; Karakaya-Ozyer & Aksu-Dunya, 2018)، أما كافة قيم المؤشرات الأخرى فتقع ضمن حدود الثقة لجودة مُطابقة النموذج المُستخدم للبيانات الموضح في الشكل (4)، وهذا يدل على أن البناء العاملي سباعي العوامل المُقترح للاختبار حقق مستوى اللاتغير التكويني وفقًا لمتغير جنس الطالب.

ثانيًا: اللاتغير المتري (الثبات المتري)

تم فحص وجود اللاتغير المتري في فقرات الاختبار وفقًا لمتغير جنس الطالب بوضع قيود على النموذج المراد فحصه، حيث تم وضع قيود على جميع الفقرات؛ أي فحص مدى ثبات تشبعات الفقرات على العوامل الكامنة بدرجة متماثلة عبر جنس الطالب، والجدول (11) يبين قيم مؤشرات جودة مطابقة النموذج للبيانات مع فرض قيود اللاتغير المتري.

جدول 11: مؤشرات جودة مطابقة النموذج المُقترح للبيانات عبر جنس الطالب مع فرض قيود اللاتغير المتري

المؤشر	حدود الثقة	القيم تحت شروط اللاتغير المتري (الثبات المتري)
χ^2	دالة	39968.847
df	–	533
P	$0.05 \geq$	0.00
RMSEA	$0.08 \geq$	0.026
CFI	$0.8 \leq$	0.938
TLI	0−1	0.930
SRMR	$0.09 \geq$	0.006

يُلاحظ من الجدول (11) وُجود دلالة إحصائية لقيمة مؤشر (χ^2) وهذا متوقع، ويمكن تجاهله بسبب حجم عينة الدراسة الكبير، وارتفاع قيم معاملات الارتباط بين العوامل (Rutkowski and Svetina, 2014; Karakaya-Ozyer & Aksu-Dunya, 2018)، أما كافة قيم المؤشرات الأخرى فتقع ضمن حدود الثقة لجودة مُطابقة النموذج المُستخدم للبيانات الموضح في الشكل (4)، وللتحقق من وجود اللاتغير المتري حُسبت المقادير الآتية: (ΔCFI, ΔSRMR, ΔTLI, ΔRMSEA) التي تبين الفرق في قيم (CFI, SRMR, TLI, RMSEA) بين النموذج التكويني والنموذج المتري، ويبين جدول (12) نتائج هذه القيم.

جدول 12: الفرق بين مؤشرات جودة مطابقة النموذج (التكويني والمتري) المُقترح للبيانات عبر جنس الطالب

المؤشر	حدود الثقة	الفرق بين المؤشرات
χ^2	دالة	166.429
df	–	25
P	≥ 0.05	0.00
Δ RMSEA	≥ 0.015	0.001
Δ CFI	≥ 0.01	0.000
Δ TLI	≥ 0.01	−0.003
Δ SRMR	≥ 0.03	0.000

يُلاحظ من الجدول (12) أن قيم الفرق في مؤشرات (ΔCFI, ΔSRMR, ΔTLI, ΔRMSEA) تقع ضمن حدود الثقة بين النموذج التكويني والنموذج المتري، وهذا يدل على تحقق اللاتغير المتري.

ثالثًا: اللاتغير العددي (ثبات التدريج)

تم فحص وجود اللاتغير العددي في فقرات الاختبار وفقًا لمتغير جنس الطالب بوضع قيود على النموذج المراد فحصه، حيث تم فحص مدى ثبات تقاطعات الفقرات على العوامل الكامنة بدرجة متماثلة عبر جنس الطالب؛ أي نقطة بداية الفقرة عندما تكون درجة العامل الكامن (0)، والجدول (13) يبين قيم مؤشرات جودة مطابقة النموذج للبيانات مع فرض قيود اللاتغير العددي.

جدول 13: مؤشرات جودة مطابقة النموذج المُقترح للبيانات عبر جنس الطالب مع فرض قيود اللاتغير العددي

المؤشر	حدود الثقة	القيم تحت شروط اللاتغير العددي (ثبات التدريج)
χ^2	دالة	40022.884
df	–	554
P	$0.05 \geq$	0.00
RMSEA	$0.08 \geq$	0.025
CFI	$0.8 \leq$	0.938
TLI	0–1	0.933
SRMR	$0.09 \geq$	0.006

يُلاحظ من الجدول (13) وُجود دلالة إحصائية لقيمة مؤشر (χ^2) وهذا متوقع، ويمكن تجاهله بسبب حجم عينة الدراسة الكبير، وارتفاع قيم معاملات الارتباط بين العوامل (Rutkowski and Svetina, 2014; Karakaya-Ozyer & Aksu-Dunya, 2018)، أما كافة قيم المؤشرات الأخرى فتقع ضمن حدود الثقة لجودة مُطابقة النموذج المُستخدم للبيانات الموضح في الشكل (4)، وللتحقق من وجود اللاتغير العددي حُسبت المقادير الآتية: (ΔCFI, ΔSRMR, ΔTLI, ΔRMSEA) التي تبين الفرق في قيم (CFI, SRMR, TLI, RMSEA) بين النموذج المتري والنموذج العددي، ويبين جدول (15) نتائج هذه القيم.

جدول 14: الفرق بين مؤشرات جودة مطابقة النموذج (المتري والعددي) المُقترح للبيانات عبر جنس الطالب

المؤشر	حدود الثقة	الفرق بين المؤشرات
χ^2	دالة	54.037
df	–	21
P	$0.05 \geq$	0.00
Δ RMSEA	$0.015 \geq$	0.001
Δ CFI	$0.01 \geq$	0.000
Δ TLI	$0.01 \geq$	−0.003
Δ SRMR	$0.03 \geq$	0.000

يُلاحظ من الجدول (14) أن قيم الفرق في مؤشرات (ΔCFI, ΔSRMR, ΔTLI, ΔRMSEA) تقع ضمن حدود الثقة بين النموذج المتري والنموذج العددي، وهذا يدل على تحقق اللاتغير العددي.

النتائج المتعلقة بالسؤال الرابع

نص السؤال الرابع على: "ما مدى فعالية طريقة المواءمة في الكشف عن اللاتغير في أداء طلبة الصف الثالث الأساسي على الاختبار التقييمي لمهارات الرياضيات عبر مديريات التربية والتعليم الأردنية؟"

تم استخدام طريقة المواءمة (Alignment Method) في برمجية (Mplus v8.4) بخطوتيها الحرة (FREE)، ثم الثابتة (FIXED) للكشف عن اللاتغير في أداء طلبة الصف الثالث على الاختبار التقييمي لمهارات الرياضيات عبر مديريات التربية والتعليم الأردنية، وإجراءاتها كالتالي:

1- استخدام النموذج المُطابق للبيانات الناتج من التحليل العاملي متعدد المجموعة (MG-CFA) المُستخدم في إجابة السؤال الثاني، والمُحقق لخاصية اللاتغير التكويني في القياس، الذي تم عرضه أثناء الإجابة عن السؤال الثاني في الشكل (2)، واعتماد النموذج سباعي العوامل المُقترح الذي تم إثبات مطابقته للبيانات حسب المؤشرات الإحصائية لجودة مُطابقة النموذج، كخطوة أولى في تطبيق طريقة المواءمة.

2- تطبيق الخطوة الأولى لطريقة المواءمة الحرة (Free Alignment Optimization)، وكان من ضمن توصيات البرمجية أن يتم استخدام الخطوة الثانية (Fixed Alignment Optimization) وهذا ما تم.

3- تطبيق الخطوة الثانية لطريقة المواءمة الثابتة (Fixed Alignment Optimization).

4- تنظيم نتائج طريقة المواءمة؛ لأن مخرجات البرمجية عبارة عن ملف من مئات الصفحات؛ لذلك تم تبسيط النتائج على شكل جداول تتضمن معلومات محددة، فيُبين الجدول (15) نتائج اللاتغير (التغير) التقريبي في القياس لتقاطعات العوامل، ويدل رقم المديرية بين الأقواس على أن معلمة الفقرة لم تحقق خاصية اللاتغير في القياس في هذه المديرية؛ أي أن هناك تباين بشكل كبير في معلمة القياس (التقاطعات/التشبعات) الخاصة بهذه المديرية.

جدول 15: نتائج اللاتغير (تغير) التقريبي في القياس لتقاطعات العوامل عبر مديريات التربية والتعليم الـ(45) حسب فقرات الاختبار الـ(25)

رقم الفقرة	أرقام المديريات كما هي موضحة في الجدول 1	المعدل الموزون	R^2
التقاطعات			
1	(10) 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 (53) 54	0.830-	0.754
2	(10) 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 (33) (34) 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54	0.795-	0.785
3	(10) (11) 12 13 14 15 16 17 18 (19) 20 21 22 23 (24) 25 (26) 27 28 29 30 31 32 33 34 35 (37) 36 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54	0.007	0.700
4	(10) 11 12 13 (14) 15 16 17 18 19 20 (21) 22 23 24 25 26 27 28 (29) 30 31 (32) 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 (53) 54	0.199	0.865
5	(10) 11 12 13 (14) 15 16 17 18 (19) 20 21 22 23 24 25 26 27 28 29 30 31 (32) 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 (53) 54	0.913-	0.831
6	(10) 11 12 13 (14) 15 16 17 18 19 20 21 (22) 23 24 25 26 27 28 (29) 30 (31) 32 33 34 35 (36) 37 38 39 (40) 41 42 43 44 45 46 47 48 49 50 51 (52) 53 54	1.298-	0.759
7	(10) 11 (12) (13) 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 (29) 30 31 (32) (33) (34) 35 36 37 38 39 40 41 42 43 44 (45) 46 (47) 48 49 50 51 52 (53) (54)	0.304-	0.749
8	(10) 11 12 13 14 15 16 17 18 19 20 21 (22) 23 24 25 26 27 28 (29) 30 31 32 33 34 35 (36) 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54	0.191-	0.923
9	(10) 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 (35) 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 (53) (54)	0.282-	0.919
10	(10) 11 (12) 13 14 15 16 17 18 19 20 21 22 23 24 (25) 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 (41) 42 43 44 45 46 47 48 49 50 51 52 53 (54)	0.599-	0.894
11	(10) 11 12 13 (14) 15 16 17 18 19 20 21 (22) 23 24 25 26 27 28 (29) 30 (31) 32 33 34 35 36 37 (38) 39 40 41 (42) 43 44 45 46 47 48 49 50 51 52 (53) 54	0.130-	0.868

رقم الفقرة	أرقام المديريات كما هي موضحة في الجدول 1	المعدل الموزون	R^2
12	36 35 34 (33) 32 31 30 29 28 27 (26) 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 37 38 39 40 (41) 42 43 44 45 46 47 (48) 49 50 51 52 53 (54)	0.493-	0.900
13	36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 (11) (10) 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54	0.451-	0.887
14	36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 (16) 15 14 13 12 11 (10) 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 (53) 54	0.894-	0.893
15	36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 (10) 37 38 (39) 40 41 42 43 44 45 46 47 48 49 50 51 52 53 (54)	1.697-	0.886
16	35 34 33 32 31 30 (29) 28 27 26 25 24 23 22 (21) 20 19 18 17 16 15 14 (13) 12 11 (10) 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54	0.405-	0.803
17	35 34 33 32 31 30 (29) 28 27 26 25 24 23 22 (21) 20 19 18 17 16 15 14 (13) 12 11 (10) 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54	1.545-	0.820
18	34 33 32 (31) (30) 29 28 27 26 25 24 23 22 21 20 (19) 18 17 16 15 14 13 12 11 (10) 37 (35) 36 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54	0.881-	0.874
19	36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 (10) 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 (53) 54	1.241-	0.922
20	34 33 32 (31) 30 29 (28) 27 26 25 24 23 22 21 20 19 18 17 16 15 (14) (13) 12 11 (10) 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 (51) 52 53 54	1.330-	0.901
21	35 34 33 32 31 30 (29) (28) 27 26 25 24 23 (22) 21 20 19 18 17 16 15 14 (13) 12 11 10 36 37 38 39 40 41 42 43 44 45 (46) (47) 48 49 50 51 52 53 54	0.282-	0.836
22	36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 (10) 37 (38) 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54	2.158-	0.792
23	36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 (10) 37 (38) 39 (40) 41 42 43 44 45 46 47 48 49 50 51 52 53 54	1.061-	0.898
24	35 34 33 (32) 31 30 29 28 27 (26) 25 24 23 22 21 20 19 18 17 16 15 (14) 13 12 11 (10) 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 (52) 53 54	1.570-	0.878
25	35 34 33 32 31 30 29 28 (27) 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 (10) 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 (52) (53) 54	1.634-	0.969

يُلاحظ من الجدول (15)، أن معلمة تقاطع العوامل متكافئة عبر معظم المجموعات (المديريات)؛ أي تُحقق خاصية اللاتغير في القياس بعد تطبيق طريقة المواءمة عليها؛ وهي كافة المديريات ما عدا المديريات ذات الأرقام المحاطة بأقواس، فمثلًا معلمة تقاطع العامل للفقرة رقم (1) تتباين بشكل دال إحصائيًا عبر المجموعات (10)، و(53)، بينما في المجموعات الأخرى تكون مُتكافئة تقريبًا، وتُظهر قيم (R^2) مؤشر التباين المُفسر (مؤشر اللاتغير في القياس) لكافة

فقرات الاختبار وهي ضمن حدود الثقة (0.2–0.7)؛ أي أن الفقرات جميعها فسرت أداء الطلبة على الاختبار بشكل كافٍ، بل بعضها كانت ممتازة مثل: الفقرة (25) حيث أن قيمة (R^2=0.969) قريبة جدًا من القيمة العليا (1).

أما الجدول (16) فيبين نتائج اللاتغير (التغير) التقريبي في القياس لتشبعات العوامل،

جدول 16: نتائج اللاتغير (تغير) التقريبي في القياس لتشبعات العوامل عبر مديريات التربية والتعليم الـ(45) حسب فقرات الاختبار الـ(25)

رقم الفقرة	أرقام المديريات كما هي موضحة في الجدول 1	المعدل الموزون	R^2
تشبعات العامل (1)			
1	10 11 12 13 14 15 16 17 18 19 20 21 (22) 23 24 25 26 27 (28) (29) 30 31 32 33 34 35 36 37 38 39 (40) 41 42 43 44 45 46 47 48 49 50 51 52 53 54	0.968	0.373
2	10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 (33) (34) 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54	1.079	0.533
3	(10) 11 12 (13) 14 15 16 17 18 (19) 20 21 22 23 24 25 (26) 27 (28) (29) 30 31 32 (33) (34) 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54	0.783	0.391
13	(10) 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 (29) 30 31 32 33 (34) (35) (36) (37) 38 39 40 41 42 43 44 45 46 47 48 49 50 51 (52) (53) 54	0.997	0.000
14	(10) 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 (29) 30 31 32 33 (34) 35 36 (37) 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54	1.204	0.130
تشبعات العامل (2)			
4	(10) 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 (30) 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 (53) 54	0.610	0.209
5	(10) 11 12 13 (14) 15 16 17 (18) (19) 20 21 22 23 24 25 26 27 28 29 30 31 (32) 33 34 35 (36) 37 38 39 40 41 42 43 44 (45) 46 47 48 49 50 51 52 (53) 54	1.133	0.190
تشبعات العامل (3)			
6	(10) 11 12 13 14 15 16 17 18 19 20 21 (22) 23 24 25 26 27 28 29 30 31 32 33 34 35 (36) 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 (52) 53 54	1.073	0.000
7	(10) 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 (53) 54	0.748	0.224
8	10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 (36) 37 38 39 40 41 (42) 43 44 45 46 47 48 49 50 51 52 53 54	0.801	0.537
9	(10) 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 (35) 36 37 38 39 40 (41) 42 43 44 45 46 47 48 49 50 51 52 53 54	0.823	0.432
10	10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 (54)	0.975	0.589

رقم الفقرة	أرقام المديريات كما هي موضحة في الجدول 1	المعدل الموزون	R^2
11	(10) 11 12 13 14 15 16 17 18 19 20 21 (22) 23 24 25 26 27 28 29 (30) 31 32 33 34 35 36 (37) 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 (53) 54	0.599	0.433
12	(10) 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 (29) 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 (48) 49 50 51 52 53 54	0.764	0.217
تشبعات العامل (4)			
15	10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 (53) 54	1.315	0.462
16	(10) 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 (29) 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 (53) 54	0.684	0.494
17	(10) 11 12 13 14 15 (16) 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 (33) 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 (53) (54)	1.109	0.000
18	10 11 12 13 14 15 16 17 18 19 20 21 22 (23) 24 25 26 27 28 29 30 31 32 33 34 (35) 36 37 38 (39) 40 41 42 43 44 45 46 47 48 49 50 51 52 (53) 54	0.897	0.377
تشبعات العامل (5)			
19	10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 (54)	0.971	0.490
20	10 11 12 (13) 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 (37) 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54	1.010	0.566
21	10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 (40) 41 42 43 44 45 46 47 48 49 50 51 52 (53) 54	0.596	0.630
تشبعات العامل (6)			
22	10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 (46) 47 48 49 50 51 52 53 54	1.361	0.112
23	(10) 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 (31) 32 33 34 35 36 37 (38) 39 (40) 41 42 43 44 45 46 47 48 49 50 51 52 53 54	1.077	0.777
تشبعات العامل (7)			
24	10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 (26) 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 (52) 53 54	1.383	0.272
25	10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 (45) 46 47 48 49 50 51 (52) (53) (54)	1.234	0.821

يُلاحظ من الجدول (16)، أن معلمة تشبع العوامل متكافئة عبر معظم المجموعات (المديريات)؛ أي تحقق خاصية اللاتغير في القياس بعد تطبيق طريقة المواءمة عليها عبر كافة المديريات ما عدا المديريات ذات الأرقام المحاطة بأقواس، فمثلًا معلمة تشبع العامل (1) للفقرة رقم (1) تتباين بشكل دال إحصائيًا عبر المجموعات (22)، (28)، (29)، و(40)، بينما في

المجموعات الأخرى تكون متكافئة تقريبًا، وتُظهر قيم (R^2) مؤشر التباين المُفسر لكافة فقرات الاختبار ضمن حدود الثقة (0.2–0.7)؛ أي أن الفقرات جميعها فسرت أداء الطلبة على الاختبار بشكل كافٍ، ما عدا الفقرات (5، 6، 13، 14، 17، 22) فهي بحاجة لمراجعة.

5- كما تنتج طريقة المواءمة مقارنة بين أوساط العامل المُقدرة في النموذج المُقترح مرتبة بشكل تنازلي، ففي الجدول (17) تم ترتيب المديريات تنازليًا حسب الأوساط الحسابية لعامل مهارات الرياضيات المكتسبة لدى طلبة الصف الثالث الأساسي المُقاسة بفقرات الاختبار التقييمي، وتم تقديرها بتجاهل معلمات القياس الغير متكافئة عبر المجموعات.

جدول 17: مقارنة وسط العامل العام (مهارات الرياضيات) عبر المجموعات عند مستوى الدلالة الإحصائية (5%) مرتبة ترتيبًا تنازليًا

الترتيب	رقم المديرية	وسط العامل	أرقام المديريات ذات الأوساط الأصغر بشكل دال إحصائيًا
1	10	1.178	25 23 38 31 30 26 53 21 41 20 42 22 48 12 29 27 35 24 11 44 49 46 37 43 52 50 17 15 54 47 34 33 18 14 32 16 40 28 51 19 13 39 36 45
2	25	0.958	38 31 30 26 53 21 41 20 42 22 48 12 29 27 35 24 11 44 49 46 37 43 52 50 17 15 54 47 34 33 18 14 32 16 40 28 51 19 13 39 36 45
3	23	0.914	30 26 53 21 41 20 42 22 48 12 29 27 35 24 11 44 49 46 37 43 52 50 17 15 54 47 34 33 18 14 32 16 40 28 51 19 13 39 36 45
4	38	0.832	53 21 41 20 42 22 48 12 29 27 35 24 11 44 49 46 37 43 52 50 17 15 54 47 34 33 18 14 32 16 40 28 51 19 13 39 36 45
5	31	0.832	53 21 41 20 42 22 48 12 29 27 35 24 11 44 49 46 37 43 52 50 17 15 54 47 34 33 18 14 32 16 40 28 51 19 13 39 36 45
6	30	0.779	53 20 41 22 42 48 12 29 27 35 24 11 44 49 46 37 43 52 50 17 15 54 47 34 33 18 14 32 16 40 28 51 19 13 39 36 45
7	26	0.778	53 20 41 22 42 48 12 29 27 35 24 11 44 49 46 37 43 52 50 17 15 54 47 34 33 18 14 32 16 40 28 51 19 13 39 36 45
8	53	0.705	20 22 48 12 29 27 35 24 11 44 49 46 37 43 52 50 17 15 54 47 34 33 18 14 32 16 40 28 51 19 13 39 36 45
9	21	0.668	27 35 24 11 44 49 46 37 43 52 50 17 15 54 47 34 33 18 14 32 16 40 28 51 19 13 39 36 45
10	41	0.655	22 12 29 27 35 24 11 44 49 46 37 43 52 50 17 15 54 47 34 33 18 14 32 16 40 28 51 19 13 39 36 45
11	20	0.628	12 29 27 35 24 11 44 49 46 37 43 52 50 17 15 54 47 34 33 18 14 32 16 40 28 51 19 13 39 36 45

الترتيب	رقم المديرية	وسط العامل	أرقام المديريات ذات الأوساط الأصغر بشكل دال إحصائيًا
12	42	0.620	40 16 32 14 18 33 34 47 54 15 17 50 52 43 37 46 49 44 11 24 35 27 29 12 45 36 39 13 19 51 28
13	22	0.573	13 19 51 28 40 16 32 14 18 33 34 47 54 15 17 50 52 43 37 46 44 11 24 35 45 36 39
14	48	0.545	45 36 39 13 19 51 28 40 16 32 14 18 33 34 47 54 15 17 52 43 37
15	12	0.542	45 36 39 13 19 51 28 40 16 32 14 18 33 34 47 54 15 17 50 52 43 37 46 11
16	29	0.510	45 36 39 13 19 51 28 40 16 32 14 18 33 34 47 54 15 17 52 43 37 46
17	27	0.508	45 36 39 13 19 51 28 40 16 32 14 18 33 34 47 54 15 17 52 43 37 46
18	35	0.503	45 39 36 19 13 51 50 28 32 40 33 18 16 34 54 14 15 52 43 37
19	24	0.473	36 39 13 19 51 28 40 16 32 14 18 33 34 47 54 15 17 52 43
20	11	0.452	45 36 39 13 19 51 28 40 16 32 14 18 33 34 47 54 15 17 52 43 37
21	44	0.447	45 36 39 13 19 51 28 40 16 32 14 18 33 34 54 15 17 52 43
22	49	0.431	45 36 39 13 19 51 28 40 16 32 14 18 33 34
23	46	0.400	45 36 39 13 19 51 28 40 16 32 14 18 33 34 54 15
24	37	0.399	45 36 39 13 19 51 28 40 16 32 14 18 33 34 54 15 17 52
25	43	0.331	45 36 39 13 19 51 28 40 16 32 14
26	52	0.312	45 36 39 13 19 51 28 40 16 32 14
27	50	0.300	45 36 39 13
28	17	0.296	45 36 39 13 19 51 28 40
29	15	0.289	45 36 39 13 19 51 28 40 16 32
30	54	0.277	45 36 39 13 19 51 28
31	47	0.272	45 39 36 19 13
32	34	0.270	45 36 39 13 19 51 28 40
33	33	0.248	45 36 39 13 19 51 28
34	18	0.243	45 36 39 13 19 51 28
35	14	0.238	45 36 39 13 19 51 28
36	32	0.203	45 36 39 13 19 51 28
37	16	0.201	45 36 39 13 19 51 28
38	40	0.140	45 36 39 13
39	28	0.089	45 36
40	51	0.070	45 36
41	19	0.066	45 36
42	13	0.036	45 36
43	39	0.023-	45
44	36	0.026-	45
45	45	0.508-	

ويُوضح الجدول (17) قيمة الوسط الحسابي للعامل العام (مهارات الرياضيات) عبر المجموعات، كما تم تقديره عند استخدام طريقة المواءمة، وترتيب هذه الأوساط الحسابية تنازليًا من

الأكبر إلى الأقل، ومقارنته مع المجموعات التي لها أوساط حسابية مختلفة عن الوسط المُقدر اختلافًا دالًا إحصائيًا عند مستوى الدلالة (5%) لسهولة المقارنة والعرض، وعند مقارنة الترتيب الناتج من طريقة المواءمة، والترتيب الوارد في تقرير وزارة التربية والتعليم الأردنية (وزارة التربية والتعليم الأردنية، 2018، ص29–30) وُجد أن هناك تطابقًا بنسبة (80%)، وهذا تطابق جيد بين الطريقتين رغم وجود بيانات مفقودة.

وكذلك الجدول (18) يُوضح ترتيب المجموعات تنازليًا حسب الأوساط الحسابية للعامل (1) كمثال على السبعة عوامل المُقترحة في النموذج؛ والتي تُمثل المهارات السبع المُوضحة في الجدول (2)، وكما ذُكر سابقًا أنه تم تقديرها بتجاهل معلمات القياس غير المتكافئة عبر المجموعات.

الجدول 18: مقارنة وسط العامل عبر المجموعات عند مستوى الدلالة الإحصائية (5%) مرتبة ترتيبًا تنازليًا

الترتيب	رقم المديرية	وسط العامل	أرقام المديريات ذات الأوساط الحسابية الأصغر بشكل دال إحصائيًا
نتائج العامل (1)			
1	10	1.095	43 37 46 17 44 29 24 27 35 22 11 21 12 41 48 20 42 26 53 30 38 31 23 25 45 39 36 19 13 51 50 28 32 40 33 18 16 34 54 14 15 52 47 49
2	25	0.837	52 47 49 43 37 46 17 44 29 24 27 35 22 11 21 12 41 48 20 42 26 53 30 38 45 39 36 19 13 51 50 28 32 40 33 18 16 34 54 14 15
3	23	0.792	54 14 15 52 47 49 43 37 46 17 44 29 24 27 35 22 11 21 12 41 48 20 42 53 45 39 36 19 13 51 50 28 32 40 33 18 16 34
4	31	0.746	16 34 54 14 15 52 47 49 43 37 46 17 44 29 24 27 35 22 11 21 12 41 20 42 45 39 36 19 13 51 50 28 32 40 33 18
5	38	0.694	16 34 54 14 15 52 47 49 43 37 46 17 44 29 24 27 35 22 11 21 12 41 20 42 45 39 36 19 13 51 50 28 32 40 33 18
6	30	0.692	16 34 54 14 15 52 47 49 43 37 46 17 44 29 24 27 35 22 11 21 12 41 20 42 45 39 36 19 13 51 50 28 32 40 33 18
7	53	0.680	33 18 16 34 54 14 15 52 47 49 43 37 46 17 44 29 24 27 35 22 11 12 41 20 45 39 36 19 13 51 50 28 32 40
8	26	0.678	33 18 16 34 54 14 15 52 47 49 43 37 46 17 44 29 24 27 35 22 11 12 41 20 45 39 36 19 13 51 50 28 32 40
9	42	0.567	51 50 28 32 40 33 18 16 34 54 14 15 52 47 49 43 37 46 17 44 29 24 27 35 45 39 36 19 13

الترتيب	رقم المديرية	وسط العامل	أرقام المديريات ذات الأوساط الحسابية الأصغر بشكل دال إحصائيًا
10	20	0.561	51 50 28 32 40 33 18 16 34 54 14 15 52 47 49 43 37 46 17 44 29 24 27 35 45 39 36 19 13
11	48	0.557	45 39 36 19 13 51 50 28 32 40 33 18 16 34 54 14 15 52 47 49 43 37
12	41	0.543	36 19 13 51 50 28 32 40 33 18 16 34 54 14 15 52 47 49 43 37 46 17 24 27 45 39
13	12	0.540	45 39 36 19 13 51 50 28 32 40 33 18 16 34 54 14 15 52 47 49 43 37 46 17
14	21	0.520	45 39 36 19 13 51 50 28 32 40 33 18 16 34 54 14 15 52 47 49 43 37
15	11	0.508	45 39 36 19 13 51 50 28 32 40 33 18 16 34 54 14 15 52 47 49 43 37 46
16	22	0.486	45 39 36 19 13 51 50 28 32 40 33 18 16 34 54 14 15 52 47 49 43 37 46
17	35	0.480	45 39 36 19 13 51 50 28 32 40 33 18 16 34 54 14 15 52 47 49 43 37 46
18	27	0.441	45 39 36 19 13 51 50 28 32 40 33 18 16 34 54 14 15 52 43 37
19	24	0.423	45 39 36 19 13 51 50 28 32 40 33 18 16 34 54 14 15 52
20	29	0.419	45 39 36 19 13 51 50 28 32 40 33 18 16 34 54 14 15 52
21	44	0.409	45 39 36 19 13 51 50 28 32 40 33 18 16 34 54 14 15 52
22	17	0.393	45 39 36 19 13 51 50 28 32 40 33 18 16 34 54 14 15
23	46	0.379	45 39 36 19 13 51 50 28 32 40 33 18 16 34 14 15
24	37	0.348	45 39 36 19 13 51 50 28 32 40 33 18 16 34 14 15
25	43	0.306	45 39 36 19 13 51 50 28 32 40 33 18
26	49	0.305	45 39 36 19 13 51 28 32
27	47	0.283	45 39 36 19 13 51 28 32
28	52	0.278	45 39 36 19 13 51 28 32 40 33
29	15	0.243	45 39 36 19 13 51 28 32
30	14	0.233	45 39 36 19 13 51 28 32
31	54	0.222	45 39 36 19 13
32	34	0.205	45 39 36 19 13 51 32
33	16	0.195	45 39 36 19 13
34	18	0.182	45 39 36
35	33	0.159	45 39 36
36	40	0.150	45 39 36
37	32	0.118	45 39 36
38	28	0.112	45 39 36
39	50	0.095	45
40	51	0.092	45 39 36
41	13	0.089	45 39 36
42	19	0.051	45
43	36	0.030-	45
44	39	0.051-	45
45	45	0.400-	

يُلاحظ من الجدول (18) أن المديرية التي حصلت على الترتيب الأول في المهارة رقم (1): تمثيل الأعداد وقراءتها وكتابتها، وتمثيل الكسور العادية والأعداد الكسرية باستخدام المحسوسات والأشكال هي مديرية التعليم الخاص، وفي المقابل المديرية التي حصلت على الترتيب الأخير هي الأغوار الجنوبية، وهذا متوافق مع تقرير وزارة التربية والتعليم (وزارة التربية والتعليم، 2018، ص31–32)، ووُجد أن هناك توافقًا عامًا بين نتائج طريقة المواءمة، وتقرير وزارة التربية والتعليم لكافة ترتيب المديريات بنسبة (76%).

6- كما توفر طريقة المواءمة معلومات مُفصلة لكل معلمة؛ حيث تعمل مقارنات زوجية للأوساط الحسابية لكل المجموعات المتاحة، ولكل معلمة، ولكل فقرة من فقرات المقياس، ثم تُلخص معلومات اللاتغير في القياس الخاصة بكل معلمة في نهاية المقارنات لتقدم دليلًا إحصائيًا على صدق أداة القياس، وفي ختام المقارنات، تظهر قيم المعلمات التي تم مواءمتها عبر المجموعات.

ويُوفر الجدول (19) مثالًا مُبسّطًا للمقارنات الزوجية الممكنة لوسط المجموعة (45) (الثقافة العسكرية) الذي يُرمز لها بـ(54) مع جميع أوساط المجموعات الأخرى الممكنة لمعلمة تقاطعات العامل (1) للفقرة رقم (1) فقط كمثال على المقارنات الزوجية؛ وتم أخذ هذه العينة من الجدول لتوضيح طريقة التعامل مع البيانات التي يحتويها؛ فهو عبارة عن جدول كبير جدًا مكون من (990) سطرًا مع (45) مجموعة في دراستنا.

جدول 19: تحليل اللاتغير في القياس: تقاطعات الفقرة (1)

نتائج العامل (1)								
مجموعة 1	مجموعة 2	وسط عامل مجموعة 1	وسط عامل مجموعة 2	الفرق بين الوسطين	تباين الخطأ للفرق	قيمة P	الحد الأنى (%2.5)	الحد الأنى (%2.5)
54	10	1.003-	1.549-	0.546	0.158	0.001	0.805-	0.187-
54	11	1.003-	0.880-	0.123-	0.121	0.118	0.061-	0.406
54	12	1.003-	0.916-	0.087-	0.125	0.225	0.115-	0.376
54	13	1.003-	0.864-	0.139-	0.121	0.092	0.055-	0.421
54	14	1.003-	0.889-	0.114-	0.120	0.141	0.070-	0.393
54	15	1.003-	0.753-	0.249-	0.128	0.007	0.047	0.547
54	16	1.003-	0.925-	0.077-	0.138	0.278	0.169-	0.375
54	17	1.003-	0.901-	0.102-	0.144	0.210	0.155-	0.415
54	18	1.003-	1.060-	0.058	0.174	0.360	0.433-	0.270
54	19	1.003-	0.921-	0.082-	0.156	0.291	0.218-	0.416
54	20	1.003-	0.760-	0.243-	0.126	0.005	0.047	0.533
54	21	1.003-	0.731-	0.272-	0.150	0.019	0.016	0.602
54	22	1.003-	0.822-	0.181-	0.121	0.030	0.005-	0.462
54	23	1.003-	0.847-	0.155-	0.125	0.066	0.041-	0.445
54	24	1.003-	0.697-	0.305-	0.135	0.001	0.084	0.609
54	25	1.003-	0.923-	0.080-	0.140	0.265	0.153-	0.399
54	26	1.003-	0.757-	0.246-	0.127	0.007	0.041	0.540
54	27	1.003-	0.712-	0.291-	0.130	0.002	0.087	0.593
54	28	1.003-	0.750-	0.253-	0.131	0.006	0.045	0.558
54	29	1.003-	0.649-	0.354-	0.137	0.000	0.135	0.668
54	30	1.003-	0.810-	0.193-	0.123	0.023	0.002	0.481
54	31	1.003-	0.963-	0.040-	0.125	0.374	0.163-	0.324
54	32	1.003-	0.841-	0.162-	0.124	0.060	0.036-	0443
54	33	1.003-	0.931-	0.072-	0.130	0.274	0.144-	0.372
54	34	1.003-	0.767-	0.236-	0.131	0.013	0.022	0.539
54	35	1.003-	0.841-	0.161-	0.120	0.050	0.024-	0.438
54	36	1.003-	0.786-	0.217-	0.124	0.013	0.022	0.505
54	37	1.003-	0.749-	0.254-	0.121	0.003	0.063	0.530
54	38	1.003-	0.821-	0.182-	0.122	0.032	0.010-	0.464
54	39	1.003-	0.715-	0.288-	0.141	0.006	0.057	0.612
54	40	1.003-	1.043-	0.040	0.166	0.402	0.380-	0.289
54	41	1.003-	0.823-	0.180-	0.123	0.036	0.014-	0.471
54	42	1.003-	0.761-	0.241-	0.126	0.008	0.039	0.528
54	43	1.003-	0.820-	0.183-	0.129	0.045	0.022-	0.481
54	44	1.003-	0.821-	0.182-	0.133	0.052	0.035-	0.483

مجموعة 1	مجموعة 2	وسط عامل مجموعة 1	وسط عامل مجموعة 2	الفرق بين الوسطين	تباين الخطأ للفرق	قيمة P	الحد الأنى (2.5%)	الحد الأنى (2.5%)
54	45	1.003-	0.663-	0.340-	0.156	0.005	0.078	0.683
54	46	1.003-	0.869-	0.134-	0.131	0.126	0.078-	0.430
54	47	1.003-	0.930-	0.072-	0.161	0.307	0.243-	0.404
54	48	1.003-	1.008-	0.005	0.170	0.489	0.339-	0.336
54	49	1.003-	0.749-	0.253-	0.148	0.017	0.013	0.595
54	50	1.003-	0.583-	0.420-	0.197	0.010	0.066	0.833
54	51	1.003-	0.781-	0.221-	0.136	0.029	0.007-	0.524
54	52	1.003-	0.804-	0.198-	0.123	0.022	0.005	0.484
54	53	1.003-	0.706-	0.297-	0.119	0.000	0.110	0.572

ويُلاحظ من الجدول (19) أن المقارنات طويلة جدًا؛ لذلك تم عرض هذا الجزء فقط كمثال لبقية جداول المقارنة، فنُلاحظ في هذا الجدول مقارنة أوساط العامل (1) لمعلمة تقاطع العامل للفقرة (1) لاختبار تكافؤها عبر المجموعات (المديريات)، ويُختبر إحصائيًا تكافؤهم عبر كل الأزواج الممكنة من المجموعات، فمثلًا: يُقارن السطر الأخير في هذا الجدول معلمة تقاطع للفقرة (1) في المجموعة (54) وفي المجموعة (53)، ويمكننا أن نرى أن الفرق هو (−0.297)، وهو فرق كبير بين الوسطين الحسابيين، ودال إحصائيًا، وعند مراجعة المقارنات الزوجية الأخرى للمجموعة (53) نلاحظ أنها بالأغلب دالة إحصائيًا؛ أي أنها غير متكافئة مع المجموعات الأخرى وهذا يُفسر ظهور رقم (53) بين قوسين في الجدول (7)، أما الفرق بين تقاطعات المجموعتين (54) و(52) فنجد قيمته (−0.198)، وقيمة (p) أكبر من (0.001)؛ أي أنها غير دالة إحصائيًا؛ أي أنها معلمة متكافئة عبر المجموعتين، وهكذا يتم تحديد كافة المجموعات المتكافئة لمختلف المعلمات.

وكما هو مُلاحظ أن الجدول كبيرًا حقًا، نظرًا لوجود كل زوج ممكن من المجموعات، لذلك بالنسبة لـ(45) مجموعة، فيكون هناك (990) سطرًا لكل معلمة، فما يُوفره هذا الجدول معلومات تفصيلية للغاية، وهذه المعلومات تُساعد في تطوير وتحسين فقرات الاختبار ومعرفة أسباب الخلل.

وبسبب حجم الجدول الكبير تقوم البرمجية بتلخيص هذه البيانات بعد كل جدول مقارنات زوجية، حيث يظهر أسفل كل جدول قائمة بالمجموعات التي وُجدت فيها هذه المعلمة ثابتة بعد المواءمة، وهذا إجراء ضروري، ومفيد للغاية، وخصوصًا في الحالات التي يكون فيها العديد من المجموعات؛ لأن المقارنات تكون متعددة كما هو الحال في هذه الدراسة، فتظهر المجموعات التي تحقق فيها اللاتغير التقريبي في القياس كما في الجدول (6) أعلاه؛ وجدول (10) يُفسر النتائج التي تم عرضها في جدول (6)، والكيفية التي تم عن طريقها تحديد المجموعات غير المتكافئة لكل معلمة لكل فقرة من فقرات الاختبار.

وبعد أن تُظهر أسفل كل جدول ملخص للمجموعات الثابتة للمعلمة، يتم حساب المعدل الموزون عبر تلك المجموعات الثابتة (Weighted Average Value Across Invariant Groups)، بالإضافة لذلك، يتم حساب قيمة (R^2) كمؤشر للاتغير في القياس (R-square/Explained variance/Invariance index) لكل فقرة لكل معلمة.

الفصل الخامس
مناقشة النتائج والتوصيات

يتناول هذا الفصل مناقشة لنتائج الدراسة التي هدفت إلى الكشف عن "فاعلية طريقة المواءمة في الكشف عن اللاتغير في أداء طلبة الصف الثالث الأساسي على الاختبار التقييمي لمهارات الرياضيات عبر مديريات التربية والتعليم الأردنية"، وفيما يلي عرض للنتائج التي تناقشها الباحثة وفق ترتيب أسئلة الدراسة، ويتبعها مجموعة من التوصيات:

أولًا: مناقشة النتائج المتعلقة بالسؤال الأول والذي نص على: "هل البناء العاملي المُقترح لمحتوى اختبار مهارات الرياضيات للصف الثالث الأساسي يُطابق البيانات بشكل كاف؟"

أظهرت نتائج السؤال مُطابقة جيدة للنموذج العاملي السباعي (تمثيل الأعداد، مقارنة الأعداد، العمليات الحسابية على الأعداد، الجبر (الأنماط)، القياس، الهندسة، الإحصاء والاحتمالات) المُقترح لمحتوى اختبار الرياضيات والتي تنطوي تحت العامل العام لمهارات الرياضيات، وتدل هذه النتيجة على أن العوامل، والأبعاد السبعة تُمثل وبشكل واضح محتوى اختبار الرياضيات، وهذه نتيجة متوقعة لاختبار وطني على مستوى الأردن، الذي يُشرف عليه إدارة الامتحانات والاختبارات في وزارة التربية والتعليم الأردنية التي تضم عددًا كبيرًا من الخبراء، والمختصين من مشرفي الإدارة والميدان التربوي من ذوي الخبرة في مجال الرياضيات، الذين عملوا على بنائه وفق أسس علمية جيدة، ولا شك أن هذا الاختبار مرّ بمراحل عديدة وطويلة في إعداده وبنائه، وتطويره، وتطبيقه لمرات متعددة جعلت منه اختبارًا وطنيًا موثوقًا به من حيث الإعداد والبناء.

ورغم أنه لم يتم دراسة خصائص الاختبار التقييمي لمهارات الرياضيات مسبقًا كاختبار وطني حسب قراءات الباحثة، إلا أن بعض نتائج الدراسة تتفق مع نتائج الدراسات البحثية السابقة

المتبعة للمعالجات الاحصائية في دراسة خاصية اللاتغير في القياس لاختبارات، أو مسوحات دولية كما سيتم عرضها تاليًا:

تتفق هذه النتيجة مع نتيجة دراسة ايونيس (Ioannis,2014a)، وايونيس (Ioannis, 2014b) لفحص أحادية البُعد، والبناء العاملي لبيانات اختبار القدرات اللفظية العامة (GAT) باستخدام التحليل العاملي التوكيدي (CFA)، وأظهرت النتائج حسن المُطابقة للبناء العاملي لبيانات نموذج (GAT) عند استخدام النموذج الأحادي العام للكفاءة اللفظية والنموذج الثلاثي مُحدد المحتوى للجوانب الكامنة للكفاءة اللفظية، كما اتفقت هذه النتيجة مع دراسة ايونيس (Ioannis, 2015) للبناء العاملي لبيانات اختبار (SAAT) التحصيلي للقبول، الذي يُدار من قبل المركز الوطني للقياس والتقويم في التعليم العالي في الرياض، المملكة العربية السعودية، وبينت النتائج أن جميع النماذج تُناسب البيانات، وتُقدم دليلًا على صحة البناء العاملي لمعيار اختبار (SAAT) التحصيلي للقبول، كما اتفقت مع دراسة السمكي (2017) لفحص اللاتغير في القياس لاختبار تمس (2011) في الرياضيات والعلوم للصف الثامن الأساسي في الأردن، وأظهرت النتائج ملاءمة نموذج العوامل الأربعة المقترح لمحتوى اختبار الرياضيات.

ثانيًا: مناقشة النتائج المتعلقة بالسؤالين الثاني والثالث واللذان ينصان على: "هل يختلف البناء العاملي للاختبار التقييمي لمهارات الرياضيات للصف الثالث الأساسي عبر مديريات التربية والتعليم الأردنية وجنس الطالب؟"

أظهرت نتائج السؤال الثاني والثالث تحقق خاصية اللاتغير في القياس للبناء العاملي السباعي المُقترح للاختبار التقييمي لمهارات الرياضيات باختلاف مديرية التربية والتعليم الـ(45)، وجنس الطالب (ذكر، أنثى) بحيث أظهرت النتائج وجود اللاتغير في القياس بمستوياته الثلاث وفقًا لمتغير مديرية التربية والتعليم، وجنس الطالب؛ مما يدل على عدم اختلاف البناء العاملي السباعي

المُقترح لاختبار الرياضيات، ويمكن عزو ذلك إلى صدق الاختبار، والتزام الهيئات المُشرفة على إعداده بالمعايير المُتفق عليها في المحتوى والعمليات، والتزام الجهات المُشرفة على تطبيقه حسب تعليمات، وإجراءات الاختبار.

وهذا يؤكد أن للاختبار نفس التركيب العاملي لدى فئات الطلبة المختلفة من حيث مديريات التربية والتعليم الـ(45) وجنس الطالب، وهذه النتيجة تعني أن الفروق بين الأوساط الحسابية للأداء على الاختبار من قبل الطلبة في هذه المديريات تعزى لاختلاف خصائص تلك الفئات وقدراتها، وليس للاختلاف في خصائص الاختبار عبر الفئات المختلفة.

وكما ذكرنا سابقًا لم يتم دراسة خصائص الاختبار التقييمي لمهارات الصف الثالث الأساسي المُطبق بالأردن حسب علم الباحثة، إلا أن بعض نتائج الدراسة تتفق مع نتائج الدراسات البحثية السابقة المتبعة للمعالجات الاحصائية في دراسة خاصية اللاتغير في القياس لاختبارات، أو مسوحات دولية كما سيتم عرضها تاليًا:

اتفقت نتائج الدراسة مع نتائج دراسة راتكوسكي وسفيتينا (Rutkowski and Svetina, 2014) التي هدفت لتقييم فرضية اللاتغير في القياس في سياق المسوحات الدولية واسعة النطاق، وتحقيقًا لغاية هذه الدراسة تم استخدام بيانات من دراسة واسعة النطاق لدراسة التدابير النموذجية المستخدمة في التحليل العاملي التوكيدي متعدد المجموعة (MG–CFA) للكشف عن اللاتغير في القياس في مسح نطاق واسع، فأشارت النتائج إلى أن العديد من التدابير النموذجية، والمعايير المرتبطة بها غير مناسبة في المجموعات الكبيرة، وفي حالة تنوع سياق العينات ينبغي ضبطها، كما درس زيرشر، وشميت، وسيسيوش، ودافيدوف (Zercher, Schmidt, Cieciuch, and Davidov, 2015) قابلية مقارنة القيمة العالمية بمرور الوقت، وعبر الدول في "المسح الاجتماعي الأوروبي": باستخدام اللاتغير الدقيق الكلاسيكي مُقابل اللاتغير التقريبي في القياس،

كما تم اختبار تكافؤ هذه القيمة في وقت واحد عبر (15) دولة من دول (ESS) عبر ست جولات من (ESS) مع (173071) مستجيبًا، و(90) مجموعة بالمجمل، وعند استخدام النهج التقليدي أجاز مقارنة الأوساط الكامنة لـ(37) مجموعة من (90) مجموعة، كما اتفقت مع نتائج السمكي (2017) في دراسته لفحص اللاتغير في القياس لاختبار تمس (2011) في الرياضيات والعلوم للصف الثامن الأساسي في الأردن عبر كل من متغير موقع المدرسة (المدينة، والريف)، وجنس الطالب (ذكر، وأنثى)، وتم استخدام التحليل العاملي التوكيدي متعدد المجموعة (MG−CFA) من خلال البرنامج الإحصائي (AMOS 22)، وأظهرت النتائج تحقق خاصية اللاتغير في القياس بمستوياتها الثلاث (التكويني، والمتري، والعددي) في فقرات كل من اختبار الرياضيات والعلوم وفقًا لمتغيري موقع المدرسة، وجنس الطالب.

واختلفت نتائج الدراسة الحالية عن دراسة دافيدوف وآخرون (Davidov et al., 2015) لقياسات الاتجاهات نحو الهجرة في المسح الاجتماعي الأوروبي: باستخدام اللاتغير الدقيق الكلاسيكي بمقابل اللاتغير التقريبي في القياس، حيث تم اختبار اللاتغير في القياس عبر المجموعة الكاملة من (27120) فردًا في (35) دولة على ست جولات، ونتج عنها عدم قابلية البيانات للمقارنة عبر الدول، والفترات الزمنية المتعددة بسبب عددها الكبير.

ثالثًا: مناقشة النتائج المتعلقة بالسؤال الرابع والذي نص على: "ما مدى فعالية طريقة المواءمة في الكشف عن اللاتغير في أداء طلبة الصف الثالث الأساسي على الاختبار التقييمي لمهارات الرياضيات عبر مديريات التربية والتعليم الأردنية؟"

أظهرت نتائج السؤال الرابع أن معظم الفقرات حققت درجة كبيرة من اللاتغير في القياس لمعلمة تقاطعات العامل، وتشبعات العامل، كما يُوضح الجدول (15) أن الفقرات (2)، (8) ليس

لهما قياس غير متكافيء حسب تقاطعات العامل، مما يجعل هاتين الفقرتين مفيدتين بشكل خاص لمقارنة أداء الطلبة على الاختبار التقييمي عبر المديريات.

بالإضافة إلى نموذج المواءمة المُقدر، فإن إجراءات المواءمة تعطي اختبار للاتغير في القياس، والنتائج التي تُنتجها خوارزمية المواءمة تُحدد أكبر مجموعة من المعلمات التي لا يوجد فيها فروق كبيرة ذات دلالة إحصائية بين المعلمات، كما يتم ترتيب المجموعات حسب الاختلافات الكبيرة الناتجة عن المقارنات الزوجية الممكنة التي ظهرت في الجدول (19)، كما ويتم توفير معلومات إضافية حول تقاطع وتشبع العامل لكل فقرة، وقيمة (R^2) لكل معلمة، ولكل فقرة تم تنظيمها في جدول (15)، وهي: عبارة عن إحصائي وصفي مفيد لتحديد درجة اللاتغير في القياس لكل معلمة، حيث يُوضح مقدار اختلاف المعلمات التكوينية عبر المجموعات من خلال أوساط العوامل وتبايناتهم، تشير قيمة (R^2) العالية إلى درجة عالية من اللاتغير في القياس، ومن قيم (R^2) الموضحة في الجدول (15) نلاحظ أن أغلب فقرات الاختبار حصلت على قيم عالية على معلمة تقاطعات العامل للفقرات، وقيم مقبولة على معلمة تحميلات العوامل على الفقرات، وهذا مؤشر إضافي على صدق الاختبار التقييمي لمهارات الرياضيات للصف الثالث الأساسي للعام الدراسي (2018/2017)، ويُؤمل أن يتم مراجعة الفقرات من قبل إدارة الامتحانات والاختبارات الأردنية، وخصوصًا الفقرات التي حصلت على قيم متدنية لمؤشر اللاتغير في القياس (R^2).

وكما ذكرنا سابقًا لم يتم دراسة خصائص الاختبار التقييمي لمهارات الصف الثالث الأساسي المطبق بالأردن حسب علم الباحثة، إلا أن بعض نتائج الدراسة تتفق مع نتائج الدراسات البحثية السابقة المُتبعة للمعالجات الاحصائية في دراسة خاصية اللاتغير في القياس لاختبارات، أو مسوحات دولية استخدمت طريقة المواءمة، ونتج عنها نتائج مُبشرة بانطباق خاصية اللاتغير في القياس، والتخلص من معيقات كِبَر حجم العينة، وتعدد المجموعات، كما سيتم عرضها تاليًا:

اتفقت نتائج الدراسة مع نتائج دراسات ميوثن واسباروف (Muthen & Asparouhov, 2013)، و(Muthen & Asparouhov, 2014)، و(Muthen & Asparouhov, 2018) التي تناولت طرق جديدة للتحليل العاملي ونظرية استجابة الفقرة (IRT) لدراسة اللاتغير في القياس عبر المجموعات، منها طريقة المواءمة كنمط تباين ثابت، وتم تطبيق الأرجحية العظمى وتحليل (Bayesian)، على مسح مكون من (50000) عنصر في (26) دولة، كما تمت دراسة الطريقتين باستخدام محاكاة (Monte Carlo)، ووفرت طريقة المواءمة حساب تفصيلي لمعالم اللاتغير في القياس لمعلمة كل نموذج لكل مجموعة.

كما اتفقت أيضًا مع نتائج دي بوندت وفان بيتيجم (De Bondt and Van Petegem, 2015) في دراستهم للاتغير التقريبي في القياس لاستبان الإفراط في استطلاعه الثاني (OEQ-II) وتأثير درجة وطبيعة الإفراط في الإثارة، ونتج عن تطبيق تحليل (BSEM) على عينة من (516) طالبًا من التعليم العالي نتائج إيجابية فيما يتعلق بصدق العوامل لـ(OEQ-II)، كما اتفقت مع بايرون وفان دي فيجفير (Byrne and Vijver, 2017) في دراستهم لتطبيق نموذجي لطريقة المواءمة، وخلصت الدراسة إلى أن طريقة المواءمة لاختبار اللاتغير التقريبي في القياس توفر إجراءً آلياً يمكنه التغلب على القيود المفروضة من تحليل (CFA) التقليدي في المقارنات واسعة النطاق، كما نتج عن دراسة إيفرز وآخرون (Evers et al., 2017) أن طريقة المواءمة لاختبار ممارسات، وسلوكات علماء النفس نحو الاختبارات، وفحوصاتها؛ حيث تم تقييم اللاتغير في القياس بين الدول باستخدام طريقة المواءمة، وأثبتت أربعة مقاييس تحقيقها لمستويات مقبولة من اللاتغير في القياس.

كما أوصت دراسة فليك وماك كاوش (Flake & McCoach, 2018) باستخدام طريقة المواءمة في اختبار اللاتغير في القياس بناءً على نتائجها، كما اتفقت مع نتائج لومازي

(Lomazzi, 2018) عن تطبيقها لطريقة المواءمة عند اختبار اللاتغير قي القياس في سلوكات الأفراد حسب النوع الاجتماعي عبر (59) دولة، حيث شملت الدراسة مقارنة الطريقتين الأكثر شيوعًا: (MG–CFA)، وطريقة المواءمة، ولوحظ أن الطريقة الجديدة، واعدة للغاية رغم ندرة استخدامها.

وكذلك عندما أجاب مارش وآخرون (March et al., 2018) على سؤال ما يجب فعله عند فشل اللاتغير العددي؟ بأن طريقة المواءمة الموسعة لتحليل العامل متعدد المجموعة لمقارنة أوساط الكامن عبر عدة مجموعات هي الحل؛ لأن اللاتغير العددي مثالي لا يمكن تحقيقه؛ وخلُصت الدراسة إلى التأكيد على أن طريقة المواءمة تُوفر للباحثين التطبيقيين من مختلف التخصصات مرونة كبيرة في التعامل مع القضايا المهمة جوهريًا عندما لا يتناسب النموذج القياسي (CFA–MI) التقليدي مع البيانات، كما أوصى بوكروبك وديفادوف وسشميدت (Pokropek, Davidov & Schmidt, 2019) في دراستهم باستخدام طريقة المواءمة لاستعادة أوساط المتغير الكامن في الحالات التي يوجد فيها بضعة معلمات غير متكافئة عبر المجموعات.

وخُلاصة القول، أن استخدام طريقة المواءمة في الكشف عن اللاتغير في القياس في المسوحات، والاختبارات الوطنية، والدولية أحرز تغيرًا جذريًا في نمط القياس التحليلي، وتطوير أدوات القياس النفسية المختلفة، لأنها زادت من قدرة الباحثين على التحقق من النموذج العاملي لأداة القياس بما يتوافق مع الجانب النظري لها.

التوصيات

في ضوء النتائج التي توصلت لها الدراسة، ومناقشة تلك النتائج، فإن الباحثة توصي بالآتي:

- إجراء دراسات مشابهة أخرى، وعلى موضوعات علمية أخرى، أو مراحل دراسية مختلفة وأدوات قياس متنوعة.
- توظيف طريقة المواءمة في الكشف عن اللاتغير في القياس لسهولة تطبيقها، وفهم نتائجها.
- مراجعة إجراءات أصحاب القرار المسؤولين عن الاختبار التقييمي لمهارات الرياضيات للصف الثالث الأساسي لرفع مصداقية كافة فقرات الاختبار، والحرص على عدم وجود بيانات مفقودة.
- العمل على تقنين كافة الاختبارات الوطنية للوصول بها للمستوى العالمي.
- العمل على دراسة طرق إحصائية جديدة تسهل على الباحثين من دراسة الخصائص السيكومترية للاختبارات وتطويرها.
- توظيف طريقة المواءمة في دراسات المقارنة الطولية والمستعرضة بين المجموعات المختلفة.
- إجراء دراسات مشابهة لتسليط الضوء على طرق إحصائية جديدة في الكشف عن صدق المقاييس.
- التحقق من الخصائص السيكومترية للاختبارات الوطنية قبل اعتمادها من قبل أصحاب القرار لزيادة الثقة بنتائجها.

المراجع

المراجع العربية:

أبو النصر، مدحت. (2004). *قواعد ومراحل البحث العلمي*. القاهرة: مجموعة النيل العربية، 131-132.

أبو هاشم، السيد محمد. (2006). دراسة مقارنة بين النظرية التقلبدية ونموذج راش في اختبار فقرات قائمة مداخل الدراسة لدى طلاب الجامعة. *مجلة كلية التربية بالزقازي*، 52، 1-52.

التقي، احمد محمد عيسى. (1992). *اللاتغير في تقدير معالم الأفراد ودرجات صعوبة أسئلة المقال من خلال نموذجي التقدير الجزئي وسلم التقدير كحالتين خاصتين من نماذج راش*. رسالة دكتوراه، الأردن، الجامعة الأردنية، كلية الدراسات العليا.

تيغزة، محمد بوزيان. (2011). اختبار صحة البنية العاملية للمتغيرات الكامنة في البحوث منحى التحليل والتحقق. بحث علمي محكم، مركز بحوث *كلية التربية، جامعة الملك سعود، السعودية*.

تيغزة، محمد بوزيان. (2012). *التحليل العاملي الاستكشافي والتوكيدي*، دار المسيرة للنشر والتوزيع، عمان-الأردن.

الحجاج، أيمن. (2013). *أثر ترتيب الفقرات حسب خصائصها السيكومترية على أداء الأفراد والخصائص السيكومترية للاختبار*. رسالة ماجستير غير منشورة، جامعة مؤتة، الأردن.

الخياط، ماجد محمد. (2012). درجة مطابقة اختبار تحصيلي وفق نموذج راش أحادي المعلمة في الكشف عن مستوى المعرفة العلمية في المهارات الرياضية لدى طلبة الصف الثامن الأساسي. *مجلة جامعة الأقصى (سلسلة العلوم الإنسانية)*، 16(1)، 87-111.

الزبون، حابس. (2013). اللاتغير في تقديرات معالم فقرات الاختيار من متعدد عندما تطبق في سياقات مختلفة باستخدام نماذج النظرية الحديثة في القياس. *المجلة التربوية*، جامعة الكويت، مج 27 (107) 309-338.

الزعبي، آمال أحمد، سوالمه، يوسف محمد. (2017). بناء اختبار لقياس مهارات التفكير الناقدر باستخدام نظرية استجابة الفقرة. *مجلة اتحاد الجامعات العربية للتربية وعلم النفس*، المجلد الخامس عشر- العدد الثالث، 55-72.

السمكي، لطفي فوزي حمد. (2017). *اللاتغير في القياس لاختبار TIMSS 2011 في الرياضيات والعلوم عبر موقع المدرسة وجنس الطالب*. رسالة دكتوراه، جامعة اليرموك، الأردن.

الشريفين، نضال كمال. (2019). أثر نموذج نظرية الاستجابة للفقرة ذات الاستجابة المتعددة التدريج على دقة تقدير القدرات للأفراد والخصائص السيكومترية للفقرات والاختبار. *المجلة التربوية*. مج. 33، ع. 130، ج. 1، مارس 2019. ص ص. 241-295.

الشريفين، نضال كمال محمد، والكيلاني، عبدالله زيد. (2003). *مدى تحقق معايير الفاعلية في معادلة اختبارين أحدهما ثنائي التدريج والآخر متعدد التدريج وفق نماذج النظرية الكلاسيكية والنظرية الحديثة في القياس*. (رسالة دكتوراه غير منشورة). جامعة عمان العربية، عمان. مسترجع من:

http://search.mandumah.com.ezproxy.yu.edu.jo/Record/572765.

عبابنة، عماد غصاب. (2008). استقصاء اللاتغير في تقدير احصائيات الفقرة المعايرة باستخدام نظرية الاستجابة للفقرة والنظرية الكلاسيكية في القياس، *مجلة علوم انسانية:* السنة السادسة: العدد 39: ص 98.

علام، صلاح الدين. (2005). *نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي*. الطبعة الأولى، القاهرة، دار الفكر العربي.

القهوجي، أيمن وأبو عواد، فريال (2018). *النمذجة بالمعادلة البنائية*. دار وائل للنشر والتوزيع. عمان: الأردن.

كاظم، أمينة محمد. (1988) *دراسة نظرية نقدية حول القياس الموضوعي للسلوك. نموذج راش*. الكويت، مؤسسة الكويت للتقدم العلمي (إدارة التأليف والترجمة والنشر).

كريش، أ. (2015). أهمية التحليل العاملي التوكيدي المتعدد المجموعات كأسلوب للقياس في ميدان علم النفس الجنائي. *المؤتمر العلمي الدولي الأول: علم النفس وقضايا المجتمع الراهنة*. جامعة الجزائر 2، ص: 276-287.

محمد، محمد إبراهيم. (2011). قياس اللاتغير العاملي للكمالية بالتحليل العاملي التوكيدي للمجموعات المتعددة. *مجلة التربية وعلم النفس*، كلية التربية، جامعة المنيا، 114-137.

النعيمي، عز الدين عبدالله. (2015). معالم الفقرات والأفراد وخاصية اللاتغير في الاختبارات الوطنية لضبط جودة التعليم في الأردن-مقارنة بين النظرية الكلاسيكية والنظرية الحديثة في القياس. *مجلة اتحاد الجامعات العربية للتربية وعلم النفس*، المجلد الثالث عشر- العدد الأول، 138-155.

وزارة التربية والتعليم الأردنية. (2018). *نتائج الاختبار التقييمي في مبحثي اللغة العربية والرياضيات للصف الثالث الأساسي للعام الدراسي 2017-2018*. عمان.

المراجع الأجنبية:

Adam W. and Gary J. (2004). A Monte-Carlo Study of Confirmatory Factor Analytic Tests of Measurement Equivalence/ Invariance, *Structural Equation Modeling. 1(1). 60-72.*

Asparouhov, T., & Muthén, B. O. (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling: A Multidisciplinary Journal*, *21, 1–14.*doi:10.1080/10705511.2014.919210

Barton, M. A., Lord, F. M. (1981). *An upper asymptote for the three-parameter logistic item-response model.* Princeton, NJ: Educational Testing Services.

Bejar, I. I. (1980). A procedure for investigating the unidimensionality of achievement tests based on item parameter estimates. *Journal of Educational Measurement.* 17 (4), Winter, P. 283.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York, NY: Guilford.

Bujacz, A., Vittersø, J., Huta, V., & Kaczmarek, L. D. (2014). Measuring hedonia and eudaimonia as motives for activities: cross-national investigation through traditional and Bayesian structural equation modeling. *Frontiers in psychology*, 5, 984. https://doi.org/10.3389/fpsyg.2014.00984.

Byrne, B. M., Shavelson, R. J., & Muthén, B. O. (1989). Testing for the equivalence of factorcovariance and mean structures : The issue of partial measurement invariance.*Psychological Bulletin*, 105(3), 456–466.

Byrne, B. and Watkins, D. (2003). The issue of measurement invariance revisited. *Journal of Cross-Cultural Psychology,* 34, 155-175.

Byrne, Barbara M. (2011). *Structural Equation Modeling with AMOS*. Palo Alto,CA: Consulting Psychologists Press.

Byrne, B. (2012). *Multivariate applications series.Structural equation modeling with Mplus: Basic concepts, applications, and programming*. Routledge/Taylor & Francis Group.

Byrne, B and Vijver, F. (2017). The maximum likelihood alignment approach to testing for approximate measurement invariance: A paradigmatic cross-cultural application. *Psicothema 2017,* 29(4), 539-551. doi: 10.7334/psicothema2017.178.

Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items.* Thousand Oaks, CA: Sage.

Cangur, S. and Ercan, L. (2015). Comparison of Model Fit Indices Used in Structural Equation Modeling Under Multivariate Normality. *Journal of Modern Applied Statistical Methods,* 14 (1), 152-167. DOI: 10.22237/jmasm/1430453580.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3), 464-504.

Cheung, G. W., & Rensvold, R. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equation Modeling*, 9: 233-255.

Cieciuch, J., Davidov, E., Schmidt, P., Algesheimer, R., and Schwartz, S. H., (2014). Comparing results of an exact vs. an approximate (Bayesian) measurement invariance test: across-country illustration with a scale to measure 19 human values. *Frontiers in Psychology| Quantitative Psychology and Measurement 982(5).* doi: 10.3389/fpsyg.2014.00982.

Cohen, A. S., Kim, S. H., & Wollack, J. A. (1996). An investigation of the likelihood ratio test for detection of differential item functioning. *Applied Psychological Measurement,* 20, 15-26.

Davidov, E., Cieciuch, J., Meuleman, B., Schmidt, P., Algesheimer, R., Hausherr, M. (2015). The comparability of measurements of attitudes toward immigration in the European Social Survey: exact versus approximate measurement equivalence. Public Opinion Quarterly, 79(S1), 244-266. DOI: https://doi.org/10.1093/poq/nfv008.

De Bondt, N., and Van Petegem, P., (2015). Psychometric Evaluation of the Overexcitability Questionnaire-Two Applying Bayesian Structural Equation Modeling (BSEM) and Multiple-Group BSEM-Based Alignment with Approximate Measurement Invariance. *Front. Psychol.* 6:1963. doi: 10.3389/fpsyg.2015.01963.

Drasgow, F. (1984). Scrutinizing psychological test: Measurement equivalence and equivalent relations with external variables are the central issues. *Psychological Bulletin*, 95, 134-135.

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

Evers, A., McCormick, C. M., Hawley, L. R., Muñiz, J., Balboni, G., Bartram, D., Dusica Boben, D., Jens Egeland, J., Karma El-Hassan, K., José R. Fernández-Hermida, J. R., Fine, S., Frans, Ö., Gintiliené, G., Hagemeister, C., Halama, P., Iliescu, D., Jaworowska, A., Jiménez, P., Manthouli, M. at al., (2017). Testing Practices and Attitudes Toward Tests and Testing: An International Survey. *International Journal of Testing, 17(2), 158-190.* https://doi.org/10.1080/15305058.2016.1216434.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). Sage Publications, Inc.

Finch, H. (2014). Measurement Invariance. In Michalos, A.C, (ed.), *Encyclopedia of Quality of Life and Well-Being Research,* (pp 3909-3912). Springer Science+Business Media.

Finch, H. & French, B. F. (2007). Detection of Crossing Differential Item Functioning A Comparison of Four Methods. *Educational and Psychological Measurement* 67(4): 565-582.

Flake, J. K., & McCoach, D. B. (2018). An investigation of the alignment method with polytomous indicators under conditions of partial measurement invariance. *Structural Equation Modeling, 25(1), 56-70.* https://doi.org/10.1080/10705511.2017.1374187.

Gadelrad, H. (2004). *The effect of model misspecification on goodness-of-fit indices for Structural equation modeling.* Unpublished PhD, Wayne State University, Detroit, MI.

George, A. & Irini, M. (2002). *Latent Variable & Latent Structure Models,* Lawrence Erlbaum Association, Inc, USA.

Glöckner-Rist, A., & Hoijtink, H. (2003). The best of both worlds: Factor analysis of dichotomous data using item response theory and structural equation modeling. *Structural Equation Modeling, 10*(4), 544–565. https://doi.org/10.1207/S15328007SEM1004_4.

Hair, J., Anderson, R., Tatham, R. and Black, W. (1995). *Multivariate Data Analysis with Reading, Fourth Edition,* New Jersey. Prentice-Hall, Inc.

Hambleton, R.K. and Swaminathan, H. (1985). *Item Response Theory: Principles and applications,* Boston MA: KluwerNyjhoff.

Hambleton, R.K. and Swaminathan, H. and Rogers, H.J. (1991). *Fundamentals of Item Response Theory: International Educational and Professional*, Publisher Newbury park.

Harrington, D. (2009). *Confirmatory factor analysis.* New York: Oxford University Press.

He, J., & van de Vijver, F. (2012). Bias and Equivalence in Cross-Cultural Research. *Online Readings in Psychology and Culture*, 2(2). https://doi.org/10.9707/2307-0919.1111.

Holland, P. W., & Thayer, D. T. (1988). *Differential item performance and the Mantel-Haenszel procedure.* In H. Wainer & H. I. Braun (Eds.), *Test validity* (p. 129-145). Lawrence Erlbaum Associates, Inc.

Hu, L. and Bentler, P. (1999). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications (pp. 76-99).* Thous and Oaks, CA: Sage.

Hulin, C. L., Drasgow, F. & Parsons, C. K. (1983). *Item Response Thoery Application to Psychological Measurement.* DOW JONES-IRWIN: Homewod, Illinois.

Ioannis, T. (2014a*). Factorial Invariance and laten mean differences of scores on GAT across Gender*, Technical Report TR063-2014. National Center for Assessment in Higher Education.

Ioannis, T. (2014b). *Factorial Structure of GAT and its Measurement Invariance across School Types*, Technical Report TR062-2014. National Center for Assessment in Higher Education.

Ioannis, T. (2015). *Factorial Invariance and latenMean Differences of Scores on SAAT across Gender*, Technical Report TR092-2015. National Center for Assessment in Higher Education.

Jennrich, R. (2006). Rotation to simple loadings using component loss functions: The oblique case. *Psychometrika, 71, 173-191.* doi: 10.1007/s11336-003-1136-B.

Jöreskog, K. (1969). A general approach to confirmatory maximum likelihood factor Analysis. *Psychometrical, 34(2), 183-202.*

Jöreskog, K. (1971). Simulaneous factor analysis in several populations. *Psychometrical,* 36, 409-426.

Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika,* 36, 109-133.

Joseph, F., Hair, J., & Marco, J. (2017). *Advanced Issues in Partial Structural Equation Modling*: Conceptual and methodological Norwell, MA: Kluwer Academic.

Jung, E. & Yoon, M. (2016). Comparisons of Three Empirical Methods for Partial Factorial Invariance: Forward, Backward, and Factor-ratio Tests. *Structural Equation Modeling: A Multidisciplinary Journal*, 23:4, 567-584.

Karakaya-Ozyer, K. & Aksu-Dunya, B. (2018). A review of structural equation modeling applications in Turkish educational science literature, 2010-2015. *International Journal of Research in Education and Science (IJRES)*, 4(1), 279-291. DOI:10.21890/ijres.383177

Kim, E., Cao, C., Wang, Y. & Nguyen, D. T. (2017). Measurement Invariance Testing with Many Groups: A Comparison of Five Approaches. *Structural Equation Modeling: A Multidisciplinary Journal, 24(4), 524-544.* DOI: 10.1080/10705511.2017.1304822.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). NewYork, NY: Guilford.

Lattin, J., Carrol, J. and Green, P. (2003). *Analyzing multivariate data.* Brooks/Cole, Pacific Grove, CA. Thompson Learning.

Lee, S. (2012). *The impact of missing data on the dichotomous mixture IRT models*. Unpublished PhD's thesis – The Graduate Faculty of The University of Georgia.

Little, D.T, (1997). Mean and Covariance Structures (MACS) Analyses of Cross-Cultural Data: Practical and Theoretical Issues. *Multivariate Behavioral research,* 32(1), 53-76.

Little, T. (2000). On the comparability of constructs in cross-cultural research: A critique of Cheung and Rensvold. *Journal of Cross-Cultural Psychology.* 31, 213-219.

Lord, F. (1980). *Application of Item Resonse to practical testing problems*. Hillsdale, NJ, Evbaum.

Lomazzi, Vera (2018). Using Alignment Optimization to Test the Measurement Invariance of Gender Role Attitudes in 59 Countries. *Methods, data, analyses 12(1), 2018, pp. 77-104.* doi: 10.12758/mda.2017.09.

Lommen, M. J. J., Schoot, R., and Engelhard, I.M. (2014). The experience of traumatic events disrupts the measurement invariance of aposttraumatic stress scale. *Frontiers in Psychology| Quantitative Psychology and Measurement, 1304(5), 1-7.* doi: 10.3389/fpsyg.2014.01304.

MacCallum RC, Roznowski M, Necowitz LB. (1992). Model modifications in covariance structure analysis: the problem of capitalization on chance. *Psychol. Bull.* 111: 490– 504.

Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J. S., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling*, 16, 439-476.

Marsh, H. W., Guo, J., Nagengast, B., Parker, P. D., Asparouhov, T., Muthén, B., & Dicke, T. (2018). What to do when scalar invariance fails: The extended alignment method for multi- group factor analysis comparison of latent means across many groups. *Psychological Methods, 23* (3), 524–545. https://doi.org/10.1037/met0000113.

Masters, G. N. (1984). Constructing an items bank using partial credit scoring. *Journal of Educational Measurement. 12 (1),* pp 19-32.

McArdle, J.J., (1996). Current Directions in Structural Factor Analysis. *Current Directions in Psychological Science,* 5(1), 11-18.

McArdle, J. J. (2005). *The Development of the RAM Rules for Latent Variable Structural Equation Modeling.* In A. Maydeu-Olivares & J.

J. McArdle (Eds.), *Multivariate applications book series. Contemporary psychometrics: A festschrift for Roderick P. McDonald* (p. 225–273). Lawrence Erlbaum Associates Publishers.

McDonald, R. P. (1989). An index of goodness-of-fit based on noncentrality. Journal of Classification, 6, 97-103.

Menel, J., Tay, L., Schwoerer, C. E., & Drasgow, F. (2012). Evaluting quantitative and qualitative type of change: An analysis of the malleability of general and specific self-efficacy constructs and measures. *Journal of Leadership & Organizational Studies*, 19(3), 378-391.

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research International. *Journal of Psychological Research*, 3(1), 111-130.

Millsap, R. E. (2011). *Statistical approaches to measurement invariance.* Florence, KY: Routledge.

Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. Multivariate Behavioral Research, 39 (3), 479-515.

Muthén, B. O., & Asparouhov, T. (2013). BSEM Measurement Invariance Analysis. *Mplus web notes 1(1).*

Muthén, B. O., & Asparouhov, T. (2013b). *New methods for the study of measurement invariance with many groups.* Retrieved from statmodel.com

Muthén, B. O., & Asparouhov, T. (2014). IRT studies of many groups: The alignment method. *Frontiers in Psychology, 5(September), 1-7.* doi: 10.3389/fpsyg.2014.00978

Muthén, B. O., & Asparouhov, T. (2017). Recent Methods for the Study of Measurement Invariance With Many Groups: Alignment and Random Effects. *Sociological Methods & Research,* 47(4): pp. 637-664. DOI: 10.1177/0049124117701488.

Muthen, B., & Asparouhov, T. (2018). Recent methods for the study of measurement invariance with many groups: Alignment and random effects. *Sociological Methods & Research, 47*(4), 637-664. doi: 10.1177/0049124117701488.

Narayanan, P., & Swaminathan, H. (1994). Performance of the Mantel-Haenszel and simultaneous item bias procedures for detecting differential item functioning. *Applied Psychological Measurement, 18*(4), 315-328. https://doi.org/10.1177/014662169401800403.

Penfield, R. D., & Camilli, G., (2007). Differential item functioning and item bias. In S. Sinharay & C. R. Rao (Eds.), Handbook of Ststistics, 26: Psychometrics (pp. 125-167). New York: Elsevier, 2007.

Pokropek, A., Davidov, E., & Schmidt, P. (2019). A Monte Carlo Simulation Study to Assess The Appropriateness of Traditional and Newer Approaches to Test for Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 26:5, 724-744.* DOI: 10.1080/10705511.2018.1561293.

Putnick, D. L., and Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41: 71-90.

Raju, N. S., Laffitte, L. J., & Byrne, B. M. (2002). Measurement equivalence: A comparison of methods based on confirmatory factor analysis and item response theory. *Journal of Applied Psychology, 87*(3), 517-529. https://doi.org/10.1037/0021-9010.87.3.517

Raykov, T., & Marcoulides, G. A. (2006). *A first course in structural equation modeling(Second.).* Mahwah, NJ: Lawrence Erlbaum Associates.

Rutkowski, L., and Svetina, D. (2014). Assessing the Hypothesis of Measurement Invariance in the Context of Large-Scale International

Surveys. *Educational and Psychological Measurement 74(1) 31–57.* DOI: 10.1177/0013164413498257

Schoot, R., Kluytmans, A., Tummers, L., Lugtig, P., and Muthén, B. (2013). Facing off with Scylla and Charybdis: acomparison of scalar, partial, and the novel possibility of approximate measurement invariance. *Frontiers in Psychology| Quantitative Psychology and Measurement 770(4) 1-15.* doi: 10.3389/fpsyg.2013.00770.

Schreiber , J. B., Nora, A., Stage, F. K., Barlow, E. A. & King, J., (2006). Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review, *The Journal of Educational Research*, 99:6, 323-338, DOI: 10.3200/JOER.99.6.323-338.

Seddig, D., & Leitgöb, H. (2018). Approximate measurement invariance and longitudinal confirmatory factor analysis: concept and application with panel data. *Survey Research Methods*, 12(1), 29-41. https://doi.org/10.18148/srm/2018.v12i1.7210.

Shook, C.L., Ketchen, D.J., Jr., Hult, G.T.M. and Kacmar, K.M. (2004), An assessment of the use of structural equation modeling in strategic management research. *Strat. Mgmt. J.,* 25: 397-404. doi:10.1002/smj.385

Sick, J. (2010). Rasch measurement in language education part 5: assumptions and requirements equirements of Rasch measurement. *SHIKEN: JALT Testing & Evaluation SIG Newsletter*, 14(2), 23 – 29.

Sörbom, D. (1989). Model modification. *Psychometrika*, *54*, 371–384.

Stark, S., Chernyshenko, O. S., & Dragow, F. (2004). Examining the effects of differential item/test functioning (DIF/DTF) on selection decision: When are statistically significant effects practically important? *Journal of Applied Psychology*, 89, 497-508.

Stark, S., Chernyshenko, O. S., & Dragow, F. (2006). Detecting differential item functioning with CFA and IRT: Toward a unified strategy. *Journal of Applied Psychology*, 91, 1292-1306.

Stewart, D. (1981). The Application and Misapplications of Factor Analysis in Marking Research, *Journal of Marking Research,* 18, pp. 51-56.

Suh, Y., & Cho, S. J. (2014). Chi-square difference tests for detecting functioning in a multidimensional IRT model: A Monte Carl study. *Applied Psychological Measurement*, 38, 359-375.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.). Boston, MA: Pearson.

Tay, L., Vermunt, J.K., and Wang, C. (2013). Assessing the item response theory with covariate (IRT-C) procedure for ascertaining differential item functioning. *International Journal of Testing*, 13, 201-222.

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

Terluin, B., Smits, N., Brouwers, M., & Vet, H. (2016). The Four-Dimensional Symptom Questionnaire (4DSQ) in the general population: scale structure, reliability, measurement invariance and normative data: a cross-sectional survey. *Health and Quality of Life Outcomes*, 114-130.

Thissen, D., Steinberg, L., & Wainer, H. (1988). Use of item response theory in the study of group differences in trace lines. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 147-169). Hillsdale, NJ: Erlbaum.

Thissen, D. (2001). IRTLRDIF v.2.0b: Software for the computation of the statistics involved in item response theory likelihood-ratio tests for differential item functioning. Retrieved February 10, 2018, from http://www.unc.edu/~dthissen/dl.html.

Thompson, M. S. and Green, S. B. (2013). Evaluating Between-Group Differences in Latent Variable Means. Pp. 163-218 in

., edited by Gregory R. Hancock and Ralph O. Mueller. Charlotte, NC: Information Age Publishing.

Vendenberg, 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. doi:10.1177/109442810031002.

Vandenberg, R. J. (2002). Toward a further understanding of and improvement in measurement invariance methods and procedures. *Organizational Research Methods*, 5, 139-158.

Widaman, K., Ferrer, E. and Conger, R. (2010). Measurement invariance within Longitudinal structural equation models: Evaluating change in the same construct across time. *Child Development Perspectives,* 4, 10-18.

Wu, Amery D., Li, Zhen and Zumbo, Bruno 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(3). Available online: http://pareonline.net/getvn.asp?v=12&n=3.

Yang, F. M., & Kao, S. T. (2014). Item response theory for measurement validity. *Shanghai Archives of Psychiatry*, *26* (3), 171-177. https://doi.org/10.3969/j.issn.1002-0829.2014.03.

Zercher, F., Schmidt, P., Cieciuch, J. and Davidov, E. (2015). The comparability of the universalism value over time and across countries in the European Social Survey: exact vs. approximate measurement invariance. *Frontiers in Psychology| www.frontiersin.org 733(6), 1-11.* doi: 10.3389/fpsyg.2015.00733.

Zhu, X. (2009). *Assessing fit of item response models for performance assessments using bayesian analysis.* Unpublished PhD's thesis – The Graduate Faculty of School of Education, University of Pittsburgh.

الملاحق

وزارة التربية والتعليم

الرقم ٣/١١٠/٢٢٩٥٧
التاريخ ١٩ شعبان ١٤٤٠
الموافق ٢٠١٩/٠٤/٢٥

السيـــــــد مديـــــــر إدارة الامتحانـــــــات والاختبـــــــارات

الموضوع: البحث التربوي

السلام عليكم ورحمة الله وبركاته، وبعد؛

فأرجو العلم بأن الطالبة أريج حسن السعيد تقوم بإجراء دراسة عنوانها "فاعلية طريقة المواءمة في الكشف عن اللا تغير في أداء طلبة الصف الثالث الأساسي على الاختبار التقييمي لمهارات الرياضيات عبر مديريات التربية والتعليم الأردنية"، استكمالاً لمتطلبات الحصول على درجة الدكتوراه في تخصص القياس والتقويم من جامعة اليرموك، ويحتاج ذلك إلى الحصول على نتائج الاختبار التقييمي لمهارات الرياضيات لطلبة الصف الثالث الأساسي لعـــــــام ٢٠١٧-٢٠١٨.

راجياً تسهيل مهمة الطالبة المذكورة وتقديم المساعدة الممكنة لها، شريطة ألا تستخدم البيانات والمعلومات المتحصلة إلا لأغراض البحث العلمي.

واقبلوا الاحترام

وزيـــــر التربيـــــة والتعليـــــم

عبدالله علي حسونه
مدير السياسات والتخطيط الاستراتيجي

نسخة/ لمدير إدارة التخطيط والبحث التربوي
نسخة/ لمدير البحث والتطوير التربـــــوي
نسخة/ لرئيس قسم البحث التربوي
نسخة/ الملف ١٠/٣

المملكة الأردنية الهاشمية
هاتف: ٥٦٠٧١٨١ ٦ ٩٦٢+ فاكس: ٥٦٦٦٠١٩ ٦ ٩٦٢+ص.ب ١٦٤٦ عمان ١١١١٨ الأردن. الموقع الإلكتروني: www.moe.gov.jo

جامعة اليرموك دائرة رئاسة الجامعة
YARMOUK UNIVERSITY Presidency Dept.

نهضة

Reference: ..... الرقم: ر/ ١٢٢٢ /١٤/١٢٥

Date: ..... التاريخ: ٢٥ / رجب / ١٤٤٠ هـ

الموافق: ١ / ٤ / ٢٠١٩ م

معالي وزير التربية والتعليم الأكرم

الموضوع: تسهيل مهمة الطالبة أريج حسن السعيد

تحية طيبة وبعد،،،

تقوم **الطالبة أريج حسن السعيد**، ورقمها الجامعي (٢٠١٦٢٢٠٠٤٩) بدراسة بعنوان **"فاعلية طريقة المواءمة في الكشف عن اللا تغير في أداء طلبة الصف الثالث الأساسي على الاختبار التقييمي لمهارات الرياضيات عبر مديريات التربية والتعليم الاردنية "**؛ وذلك استكمالاً لمتطلبات الحصول على درجة الدكتوراه في كلية التربية، تخصص القياس والتقويم، ويستدعي معرفة نتائج الاختبار التقييمي لمهارات الرياضيات لطلبة الصف الثالث الأساسي لعام ٢٠١٨/٢٠١٧ الذي عقدته الوزارة.

أرجو التكرم بالاطلاع والموافقة على تسهيل مهمة الطالبة المذكورة أعلاه .

وتفضلوا بقبول فائق الاحترام، ، ،

رئيس الجامعة

أ.د. زيدان عبد الكافي الكفافي

أربد - الأردن هاتف: ٧٢١١١١١ - ٢ - ٩٦٢+ فاكس: ٧٢٧٤٧٢٥ - ٢ - ٩٦٢+
P.O. Box 566, Irbid, Jordan P: +962 - 2 - 7211111 F: +962 - 2 - 7274725 E-mail: yarmouk@yu.edu.jo, president@yu.edu.jo

إدارة الامتحانات والاختبارات

مديرية الاختبارات

اختبار الصف الثالث الأساسي للعام ٢٠١٨/٢٠١٧

(وثيقة محمية / محدودة)

المبحث: الرياضيات رمز المبحث: (٣٢). زمن الاختبار:ساعة واحدة

١- يُقْرَأُ الْعَدَدُ(٩٠٠٤) كَما يَأتي:

أ) أربَعَةٌ وتِسعُون ب) تِسعُ مائَةٍ وأربَعَة

ج) تِسعُ مائَةٍ وأربَعُون د) تِسعَةُ آلافٍ وأربَعَة

٢- الْقِيمَةُ الْمَنْزِلِيَّةُ لِلرَّقْمِ (٣) في الْعَدَدِ (٨٣٠٥) هِيَ :

أ) ٣ ب) ٣٠ ج) ٣٠٠ د) ٣٠٠٠

٣- يُكْتَبُ الْعَدَدُ الذي آحادُهُ (٤) وعَشَراتُهُ (٥) وآلافُهُ (٨) كَما يَأتِي :

أ) ٤٥٨ ب) ٤٥٠٨ ج) ٨٥٤ د) ٨٠٥٤

٤- العَدَدُ الَّذِي تَقْرِيبُهُ (٥٤٠٠) لأَقْرَبِ مِئَةٍ من بَيْنِ الأَعْدَادِ الآتِيَةِ هُوَ :

أ) ٥٤٩٤ ب) ٥٣٧١ ج) ٥٣١٩ د) ٥٤٦٠

٥- العَدَدُ الأَكْبَرُ مِنَ الْعَدَد (٥٥٧٦) مِنْ بَيْنِ الأَعْدَادِ الآتِيَةِ هُوَ :

أ) ٤٥٥٥ ب) ٥٤٩٨ ج) ٥٣٩٩ د) ٥٦٢٢

٦- ناتِجُ (١٢ ÷ ٣) يُسَاوِي :

أ) ٤ ب) ٩ ج) ١٥ د) ٣٦

٧- حافِلَةٌ تَحْتَوِي عَلَى (١٢) صَفًّا مِنَ الْمَقَاعِدِ، إذا عَلِمْتَ أَنَّ فِيْ كُلِّ صَفٍّ يُوْجَدُ (٤) مَقَاعِد.
فَمَا الْعَدَدُ الكُلِّيُ لِرُكَّابِ الْحَافِلَةِ؟

أ) ١٦ ب) ٨ ج) ٤٨ د) ٣٦

٨- أَنْتَجَ مَصْنَعٌ لِلْسَيَارَاتِ فِيْ إِحْدَى السَنَوَاتِ (٢٦٥٧) سَيَّارَةً، وَفِيْ السَنَةِ الَّتِي تَلِيْهَا أَنْتَجَ (٣٦٤٣) سَيَّارَةً. كَمْ سَيَّارَةً أَنْتَجَ المَصْنَعُ فِيْ السَنَتَيْنِ مَعاً ؟

أ) ٥٢٩٠ ب) ٦٢٠٠ ج) ٦٢٩٠ د) ٦٣٠٠

٩- نَاتِجُ ($\frac{٤}{٦} + \frac{٧}{٦}$) يُساوِي :

أ) $\frac{١١}{٦}$ ب) $\frac{١١}{١٢}$ ج) $\frac{١٠}{٦}$ د) $\frac{١٣}{٦}$

١٠- العَدَدُ المُنَاسِب وَضْعُهُ فِيْ ☐ لِتُصْبِحَ العَمَلِيَّةُ (☐ × ٧ = ٢٨٠) صَحِيْحَةٌ هُوَ:

أ) ٣٠ ب) ٤٠ ج) ٥٠ د) ٦٠

١١- نَاتِجُ (٥١٩٥ – ٢٤٦٤) يُساوِي :

أ) ٢٣٣١ ب) ٢٧٣١ ج) ٣٣٣١ د) ٣٥٣١

١٢- اشْتَرَى مَحْمُودُ سَبْعَ قِطَعٍ مِنَ الْحَلْوَى، فَإِذَا أَكَلَ مَحْمُودُ $\frac{١}{٧}$ قِطَعِ الحَلْوَى وَأَكَلَتْ والِدَتُهُ $\frac{٢}{٧}$ مِنْ قِطَعِ الْحَلْوَى، مَا الكَسْرُ الدَّالُ عَلَى قِطَعِ الْحَلْوَى المَأْكُوْلَةِ؟

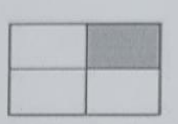
أ) $\frac{٣}{٧}$ ب) $\frac{٤}{٧}$ ج) $\frac{٣}{١٤}$ د) $\frac{٤}{١٤}$

١٣- لَدَى طَالِب مَجْمُوعَةٌ مِنَ الصَّنَادِيقِ كَمَا فِي الشَّكْلِ المُجَاوِرِ.

الكَسْرُ الَّذِي يُمَثِّلُ الصُّنْدُوقَ الأَسْوَدَ هُوَ:

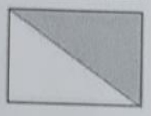
أ) $\frac{١}{٢}$ ب) $\frac{١}{٣}$ ج) $\frac{١}{٤}$ د) $\frac{١}{٥}$

١٤- الجُزْءُ المُظَلَّلُ الَّذِي يُمَثِّلُ الكَسْرَ $\frac{١}{٤}$ فِي الأَشْكَالِ الآتِيَةِ هُوَ:

أ) 

ب)

ج)

د) 

١٥- العَدَدُ الَّذِي يُكَمِّلُ النَّمَط (٦٣١٠ ، ٦٣١٥ ، ٦٣٢٠ ، ... ، ٦٣٣٠) هُوَ :

أ) ٦٣٢١ ب) ٦٣٢٤ ج) ٦٣٢٥ د) ٦٣٢٨

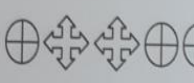
١٦- العَدَدُ الَّذِي يُكَمِّلُ النَّمَط (١٧٧، ١٧٤، ١٧١، ... ، ١٦٥) هُوَ :

أ) ١٧٠ ب) ١٦٨ ج) ١٦٧ د) ١٦٤

١٧- النَّمَط الَّذِي لَهُ نَفْسُ قَاعِدَةِ النَّمَطِ( ) هُوَ :

أ)

ب)

ج) 

د)

١٨- الشَّكْلُ الَّذِي يُكَمِّلُ النَّمَط (...) هُوَ :

أ)

ب)

ج)

د)

١٩- الوِحْدَةُ المُنَاسِبَةُ لِقِيَاسِ المَسَافَةِ بَيْنَ مَدِينَةِ عَمَّانَ و مَدِينَةِ إِربِدَ هِيَ :

أ) م ب) كم ج) سم د) مم

٢٠- يُشِيْرُ الْوَقْتُ في السَّاعَةِ المُجاوِرَةِ إلى :

أ) الثَّالِثَةِ ب) السَّادِسَةِ

ج) السادِسَةِ والنِّصْفِ د) الثَّالِثَةِ والنِّصْفِ

٢١- الوِحْدَةُ المُنَاسِبَةُ لِقِياسِ كُتْلَةِ المِفَكِّ فِيْ الشَّكْلِ المُجَاوِرِ هِيَ :

أ) مل ب) لتر ج) غ د) كغ

٢٢- يُسَمَّى الْمُجَسَّمُ الْمَرْسُوْمُ جانِباً :

أ) مُتَوازِي أَضْلاعٍ ب) اسطُوانَةٌ

ج) مُكعَّبٌ د) هَرَمٌ

٢٣- شَكْلُ الوَجْهِ السُفْلِيِّ لِلْمُجَسَّمِ المُوضَّحِ جَانِبًا هُوَ :

أ) مُسْتَطِيلٌ ب) شِبْهُ مُنْحَرِفٍ ج) مُرَبَّعٌ د) دائِرَةٌ

٢٤- يُمَثِّلُ الجَدْوَل المُجاوِرُ عَدَدَ الخِرَافِ في أَرْبَعِ مَزَارِعَ، ما عَدَدُ الخِرَافِ في المَزْرَعَةِ الثَّانِيَةِ؟

المزارع	عدد الخراف
المزرعة الأولى	٤٠
المزرعة الثانية	٧٠
المزرعة الثالثة	٣٥
المزرعة الرابعة	٧٤

أ) ٣٥ ب) ٤٠ ج) ٧٠ د) ٧٤

٢٥- الْجَدْوَلُ الْمُقابِل يُمَثِّلُ عَدَدَ الألْعابِ الَّتِي يُمَارِسُهَا طَلَبَةُ الصَّفِ الثَّالِثِ خِلالَ الأُسْبُوْعِ، كَمْ عَدَدُ الألْعابِ الَّتِي لَعِبَهَا الطَّلَبَةُ في يَوْمِ الأَحَدِ؟

اليوم	الأحد	الاثنين	الثلاثاء	الأربعاء
الاشارات	𝍸 𝍸 /// 𝍸	𝍸 𝍸 𝍸	𝍸 𝍸 //	/ 𝍸 𝍸

أ) ١٥ ب) ١٨ ج) ١١ د) ١٢

Abstract

Alsaeed, Areej Hasan. 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. (2020).

(Superviser: Professor. Nedal Kamal Alshreefen)

This study aimed to study the effectiveness of the Alignment method in detecting the invariance measurement (measurement consistency) for the third-grade evaluation test of basic mathematics skills in light of the variable of education directorates, which are (45) directorates, and the gender variable. The sample of the study consisted of (73448) male and female students, (37272 male and 36176 female students), who were randomly selected from the third-grade students enrolled in schools for the academic year (2017-2018), and who were subjected to the mathematics skills assessment test.

In order to achieve the objectives of the study, the descriptive and analytical approach was used in analyzing the results of students on the assessment test of mathematics skills to control the quality of education prepared by the Examinations and Examinations Department in the Jordanian Ministry of Education for the academic year (2017-2018), and after verifying the indications of validity and reliability of the test, It was applied to the study sample.

Specifically, the responses of (37272) students and (36176) students were analyzed to (25) test items of a multiple-choice type, and the data of the items were analyzed using the two methods of confirming factor analysis (CFA) and multi-group factor analysis (MG-CFA) through two statistical programs (Mplus v8.4, Amos v26), and to judge the suitability of the global construction of the test, five indicators of goodness of conformity were used: (χ^2), (RMSEA), (SRMR), (TLI), (CFI), and the

results were shown. The suitability of the proposed model with its seven skills to the content of the assessment test of mathematics skills for the third grade in Jordan, as well as the values of the difference indicators, which are: (RMSEA), (SRMR), (TLI), (CFI) achieving the invariance measurement property and the results showed that the invariance characteristic (measurement consistency) was achieved at its various levels (formative invariance (form stability), metric invariance (metric stability), and scalar inveriance (stability of the scale)) In the (25) test items according to the variables of the (45) Jordanian Education Directorates, and the sex of the student.

The alignment method was also used to detect the items that achieved the inveriance measurement using the index (R^2), and the results showed that all the items achieve the inveriance measurement property of the factor intersections parameter, and most of them achieve the inveriance measurement property of the factor saturations parameter except for items (5, 6, 13, 14, 17, 22) they need to be reviewed, and the arithmetic means of the underlying factors (seven mathematics skills) proposed in the structural model, and their variances across education directorates, were compared, resulting in a descending order of the districts. The Directorate of Special Education ranked first according to the performance of its students on the assessment test for mathematics skills for the third grade, and in contrast, the Southern Jordan Valley got the last place.

Key words: invariance measurement, assessment test of mathematics skills, indicators of good conformity, confirmatory factor analysis, Alignment method, education directorates.