مركز أ . د . احمد المنشاوى
للنشر العلمى والتميز البحثى
مجلة كلية التربية
=======

ثبات القياس في نموذج دوافع الإنجاز في اختبار بيزا الدولي ٢٠١٨ (PISA) في عينة الأردن

إعـــــــــــــداد

د/ماجد محمود الجوده
أستاذ مشارك- قسم التربية وعلم النفس
كلية التربية والآداب- جامعة تبوك
تبوك- المملكة العربية السعودية
majed_jodeh@hotmail.com

د/نايل عوده الكعابنه
أستاذ مساعد- قسم التربية وعلم النفس
كلية التربية والآداب- جامعة تبوك
تبوك- المملكة العربية السعودية
nalkaabneh@ut.edu.sa

﴿المجلد الأربعون– العدد الثانى عشر– ديسمبر ٢٠٢٤م﴾

http://www.aun.edu.eg/faculty_education/arabic

الملخص

لا يقل قياس الجوانب الوجدانية والعاطفية لدى الطلبة أهمية عن قياس الجوانب المعرفية، لذا هدفت الدراسة الحالية لتقصي ثبات القياس في مقياس دوافع الإنجاز لدى الطلبة في عينة الأردن تبعاً لمتغير نوع المدرسة في اختبار بيزا (PISA) الدولي 2018 وظهر نموذج القياس لدوافع الإنجاز في ثلاثة عوامل كامنة هي: الاتجاهات نحو المنافسة، الدوافع للإتقان، والخوف من الفشل، توزعت عليها (١٠) فقرات من نوع تدريج ليكرت الرباعي(أوافق بشدة، أوافق، لا أوافق، لا أوافق بشدة) وبلغ حجم عينة الأردن بعد تصفيتها بالمعالجات الإحصائية المناسبة ٨٢٣٧ استجابة توزعت على ثلاثة أنواع من المدارس هي: مدارس الذكور، مدارس الإناث، والمدارس المختلطة. وتوصلت الدراسة إلى وجود ثبات قياس شكلي وعددي أي ملاءمة نموذج قياس دوافع الإنجاز بين أنواع المدارس المختلفة شكلياً، ولم يتحقق ثبات القياس القياسي وثبات البواقي، وهذه النتيجة تشير إلى صعوبة إجراء مقارنات إحصائية تبعاً لمتغير نوع المدرسة على مقياس دوافع الإنجاز في اختبار بيزا (PISA) الدولي، وتم تقديم مجموعة من التوصيات.

الكلمات المفتاحية: دوافع الإنجاز، ثبات القياس، نوع المدرسة، بيزا PISA 2018

Measurement Invariance of Achievement Motives Model in the international PISA test 2018 in the Jordanian sample

Dr.Majed Mahmoud Aljodeh

Associate Professor, Department of Education and Psychology, Faculty of Education and Arts, University of Tabuk, Tabuk, Saudi Arabia

majed_jodeh@hotmail.com

Dr. Nayel Odeh Alkaabnh

Assistant Professor, Department of Education and Psychology, Faculty of Education and Arts, University of Tabuk, Tabuk, Saudi Arabia.

nalkaabneh@ut.edu.sa

Abstract

Measuring the affective aspects of students is no less important than measuring the cognitive aspects, so the current study aimed to investigate the measurement invariance of the achievement motivation among students in the Jordanian sample according to the school type variable in the international PISA test 2018, and the measurement model for achievement motivation questionnaire in the PISA appeared in three latent factors: attitudes toward competition, motivation for mastery, and fear of failure, included (10) items of the four-point Likert scale type (strongly agree, agree, disagree, strongly disagree). The size of the Jordanian sample, after filtering it with appropriate statistical treatments, reached 8237 divided into three types of schools: male schools, female schools, and mixed schools. The study concluded that there was configural and metrical measurement invariance, but the scaler and residuals measurement invariance were not achieved. This result indicates the difficulty of making statistical comparisons according to the school type variable on the achievement motivation scale in the international PISA test, and recommendations were presented in this study.

Key words: Achievement motives, measurement invariance, type of school, PISA2018

المقدمة

أصبح الاهتمام بتقييم مخرجات التعلم ركيزة أساسية من أجل تطور وتقدم التعليم على المستوى المحلي والعالمي، فأصبح قياس مستويات المعرفة والمهارات والقيم التي يمتلكها الطلبة منطلقاً لصناع القرار والمختصين التربويين لدراسة السياسات والبرامج التي تسعى إلى تطوير وتحسين عملية التعليم والتعليم في جميع المراحل الدراسية، وأصبحت الاختبارات التحصيلية بجميع أشكالها والتي تطبق في المؤسسات التعليمية تحظى بالحظ الأوفر من بين أدوات التقييم الأخرى، كونها تمثل مرجعاً مهماً لمعظم القرارات التربوية، وللمقارنات في المؤسسات التربوية على المستوى المحلي والعالمي (مخائيل، ٢٠١٦).

وتعد الاختبارات الدولية إحدى أهم أدوات التقييم التي تتيح للمؤسسات التعليمية تقييم أداء طلبتها ومراقبة نجاحهم وتحسن ادائهم ومقارنته بأداء الطلبة في المؤسسات التعليمية المحلية والدولية. ومن بين تلك الاختبارات اختبارات برنامج التقييم الدولي للطلاب (PISA) Program for International Student Assessment والذي يوفر بيانات مهمة للكشف عن نجاح الطلاب في المدارس، وتحديد العوامل التي تؤثر على أدائهم التحصيلي، كما يوفر بيانات للمقارنات مع الدول الأخرى في مجال التعليم. ويتم تطبيق هذا البرنامج لتقييم الطلبة الذين تبلغ أعمارهم ١٥ سنة لقياس مدى تحصيلهم للمعارف الرئيسة والمهارات الأساسية في ثلاث مواد: القراءة، والرياضيات، والعلوم، وقدرتهم على توظيفها في مجالات الحياة المختلفة.

وبالإضافة إلى تطبيق الاختبارات المعرفية لمعرفة مدى امتلاك الطلاب للمعارف والمهارات اللازمة للمشاركة الفاعلة في المجتمعات التي ينتمون إليها، فإنه يتم تطبيق استبيانات لتقييم العوامل الثقافية لخلفية الطلاب، وعوامل مستوى المدرسة، والعوامل غير المعرفية التي تتعامل مع مواضيع عامة بدلاً من المواضيع الخاصة بالمجال المعرفي والتي تسمى المتغيرات التصريفية والمتغيرات التي تركز على المدرسة مثل معتقدات التعلم، والمواقف تجاه المدرسة ودوافع الإنجاز، وكذلك قياس المتغيرات السلوكية للطلبة من خلال تلك الاستبانات؛ والمتغيرات السلوكية هي السياقات القائمة على الشخصية والتي تتضمن أساليب الطلبة في التعلم أو تجنبهم له، مثل دوافع الإنجاز التي لها دور كبير في إثارة ميل الفرد نحو التعلم والقدرة التنافسية، والخوف من الفشل، وإتقان العمل، والرفاه الذاتي، والمثابرة، وهذه الدوافع ناتجة من التنشئة الاجتماعية من قبل الآباء والمعلمين والمؤثرين على الطلبة، وكذلك البيئة الثقافية التي ينتمي إلها، وتظهر هذه الدوافع بشكل قوي مع مرور الوقت، وهذه المتغيرات مهمة للتنبؤ بدوافع الإنجاز المتمثلة بإتقان العمل والقدرة التنافسية والخوف من الفشل وتجنبه (منظمة التعاون الاقتصادي والتنمية، ٢٠١٩).

ويتم قياس الحاجة إلى الإنجاز والخوف من الفشل من خلال ثلاثة عوامل؛ الإتقان، والعمل، والقدرة التنافسية، ولقد أشار العلماء أن عوامل الإتقان والعمل مترابطة بشكل كبير وتشترك في الوصول إلى اتقان العمل بالشكل التام، بالإضافة إلى القدرة التنافسية بين الأفراد، وهذا يخلق نموذجًا ثنائي الأبعاد، إتقان العمل والقدرة التنافسية، للحاجة إلى الإنجاز، ولقد طور (Franken & Brown, 1995)مقياساً تكون من خمسة عوامل (الرغبة في الفوز، والرضا الناتج عن تحسين الأداء، والدافع لبذل الجهد في المواقف التنافسية، والرضا الناتج عن الأداء الجيد، والتفضيل للمهام الصعبة)، ترتبط العوامل الثلاثة الأولى بالقدرة التنافسية، ويتعلق العاملان الأخيران بدوافع إتقان العمل.

والمتغير الأخير لدوافع الإنجاز هو الخوف من الفشل، قد يكون للخوف من الفشل أهمية كبيرة في مدى تحقيق الطالب لما يصبو إليه ومدي قدرته علي إنجاز الأهداف والطموحات التي وضعها لنفسه وكذلك تجنب ضعف دوافعه وكفاءته، كما أن خبرات الفشل المتكررة تجعل الطلاب يشعرون بالقلق وعدم الاهتمام، وهو أمر مزعج للغاية للأفراد الذين لديهم خوف كبير من الفشل، وقد أثبتت الدراسات أنه مرتبط بميول التجنب والانسحاب، فهو ميل يركز على تجنب عواقب الفشل، على عكس الحاجة إلى الإنجاز، ويُعبر عن الخوف من الفشل على أنه ميل لتجنب الأخطاء المحتملة والإخفاقات من أجل حماية أنفسهم، وشيوع القلق والتوجس والحذر من كل مهام الإنجاز (Hangen & Elliot, 2016).

وخلال الدورات السابقة لاختبارات PISA تم تقييم عوامل مماثلة تؤثر على الدافعية للانجاز، ولكن هذه العوامل تمت مراجعتها وإعادة بنائها في اختبارات PISA لعام ٢٠١٨ باعتبارها عوامل تؤثر على دوافع الإنجاز، ففي اختبارات PISA السابقة تم استخدم متغير قلق الاختبار، ولكن الخوف من الفشل استُخدِم بدلاً منه في اختبارات PISA لعام 2018، ويذكر أن الخوف من الفشل هو ميل أكثر عمومية لتجنب الأخطاء والإخفاقات المحتملة؛ لأنها تُعَد محرجة، ويمكن أن يتنبأ هذا بالإنجاز المعرفي في مواقف الحياة الواقعية أكثر من قلق الاختبار (منظمة التنمية والتعاون الاقتصادي، ٢٠١٩).

وتؤثر دوافع الإنجاز على أداء الطلاب وتحصيلهم بشكل مباشر أو غير مباشر، وعندما تتم مقارنة النتائج بين المجموعات الناتجة من هذه المتغيرات أو المتعلقة بها، فمن الخطأ أن نعزو الاختلافات بين المجموعات فقط إلى خصائصها، لأن هذه الاختلافات بين المجموعات قد تكون بسبب أداة القياس وليس بسبب خصائص المجموعات، وهنا يشير (Cheung & Rensvold, 2002) إلى أنه ليس من المؤكد ما إذا كان أي اختلاف بين المجموعات يرجع إلى

اختلافات حقيقية أو اختلافات سيكومترية، وقد تكون الاختلافات في الدرجات ناتجة عن العديد من المتغيرات المتداخلة، مثل القدرة على الاستجابة لفقرات الاختبار، وتكييف الاختبار، والعديد من العوامل الاجتماعية والثقافية الأخرى، وأن خاصية تكافؤ او ثبات القياس Measurement Invariance خاصية مهمة لأي أداة من أدوات القياس وأن أي خلل في تكافؤ القياس لمحتوى معين يضعف من قدرتها على أن تستخدم في مقارنات بين مجموعات على بنية المقياس موضع الاهتمام، كما أنه لو لم يوجد دليل يشير إلى وجود أو غياب تكافؤ القياس حينئذ لن يكون هناك أساس للتوصل لاستنتاجات علمية دقيقة؛ وبناء على ذلك ستكون النتائج المتعلقة بدراسة الفروق بين الأفراد والمجموعات غامضة التفسير، ويشير (Wu, Li & Zumbo, 2007) أن ثبات القياس يحدث عندما تكون العلاقة بين الدرجات الملاحظة والعوامل الكامنة هي نفسها عبر المجموعات، وعندما يكون لدى الأفراد من مجموعات مختلفة نفس الدرجات على العامل الكامن نفس الدرجات الملاحظة.

ويمكن فحص ثبات القياس بعدة طرق مختلفة، ويشير (Khorramdel, Pokropek & van Rijn ,2020) إلى أن بعض الباحثين المهتمين بالتقاليد والثقافات في المجتمعات المختلفة قد أولوا اهتمامهم لثبات القياس في المقاييس غير المعرفية باستخدام إطار المتغير الكامن والتحليل العاملي التوكيدي متعدد المجموعات multigroup confirmatory factor analysis، والذي يعمل على فحص عدد كبير من القضايا من خلال عملية واحدة بدلاً من العديد من العمليات المنفصلة، ولكن هناك عيوباً لهذا التحليل عندما يكون عدد المجموعات وحجم العينة في البيانات كبيراً (Ding, Yang Hansen & Klapp, 2023)

ويتم فحص ثبات القياس إجرائياً باستخدام التحليل العاملي التوكيدي متعدد المجموعات من خلال اختبار أربعة نماذج أو فرضيات هرمية متداخلة:

أولاً: الثبات الشكلي (التكويني) (configural invariance): هو الشكل الأساس والخطوة الأولى للثبات، بحيث يتم اختبار ما إذا كانت العوامل الثابتة والمتغيرة لها نفس نمط التشبع عبر المجموعات، كما أن الأفراد في المجموعات المختلفة يستخدمون نفس الإطار المفاهيمي عند الإجابة على فقرات المقياس، بلغة أخرى يعتقد بأن هذا النوع من التكافؤ يحدد ما إذا كان البناء العاملي مكونا من نفس العدد من العوامل بين المجموعات موضع المقارنة، ويتم اختبار هذا النموذج من خلال تقييد البنية العاملية لتكون متطابقة عبر المجموعات (Cheung & Rensvold, 2002).

ثانياً: الثبات المتري (metric invariance): هو اختبار المساواة لوحدات القياس عبر المجموعات، فهو يحدد ما إذا كان تشبع العناصر على العوامل هي نفسها عبر المجموعات؛ وتشبع العوامل هي منحنيات الانحدار التي تربط المتغيرات المقاسة بالمتغيرات الكامنة ذات الاهتمام وبالتالي يمثل مقدار التغيير المتوقع في المتغير المقاس لوحدة واحدة من التغيير في المتغير الكامن، حيث يختبر هذا النموذج فيما أذا كانت المجموعات المختلفة من المستجيبين تستجيب على الاختبار بنفس الطريقة أم لا، وعندما يتحقق تكافؤ القياس المتري يمكن المقارنة بين التقديرات عبر المجموعات المختلفة (Wu et al., 2007).

ثالثاً: الثبات القياسي (scalar invariance): هو اختبار المساواة لمقاطع معادلات الانحدار للدرجات الملاحظة على المتغيرات الكامنة عبر المجموعات، أي أن الدرجات الملاحظة مرتبطة بشكل متعادل أو متكافئ مع الدرجات الكامنة بنفس الطريقة في كل مجموعة من المجموعات الفرعية، بحيث يتم اختبار ما إذا كانت الاختلافات المتوسطة في الدرجات الملاحظة تُعزى إلى الاختلافات المتوسطة للمتغيرات الكامنة (Campbell, Barry, Joe) .& Finney, 2008

رابعاً: ثبات البواقي (residual invariance): هو اختبار المساواة للتباينات المفردات غير المشتركة عبر المجموعات، حيث يتم اختبار ما إذا كانت الاختلافات المتوسطة أو التباين في الدرجات المقاسة تعزى إلى الاختلافات أو التباين في المتغيرات الكامنة، بعبارة أخرى؛ ثبات البواقي يعني أن مجموع التباين النوعي (تباين المفردة غير المشترك مع العامل) وتباين الخطأ(خطأ القياس) يكون متماثل عبر المجموعات (Khademi, 2020).

مشكلة الدراسة:

سواء كانت اختبارات PISA أو أي اختبارات دولية أخرى، فإن نتائج هذه الاختبارات تساعد في تحسين وتطوير سياسات التعليم على المستوى المحلي والعالمي، ولكي تكون نتائج الاختبارات ذات مغزى وفائدة فيجب ضمان ثبات القياس لأدوات القياس (الاختبارات والاستبيانات) المستخدمة في البحث بين الفئات الفرعية مثل الجنس والمستوى الاجتماعي والاقتصادي ونوع المدرسة والثقافة؛ وإلا فإن أي مقارنات بين المجموعات لن تكون دقيقة موثوقة، وكذلك لن تكون ذات مغزى (Vandenberg & Lance, 2000). وثبات القياس سيوفر دليلاً على صحة أداة القياس المستخدمة وتحديد ما إذا كانت مقارنات المجموعة ذات مغزى وفقًا للدرجات التي تم الحصول عليها مما يُسهم في تفسير أكثر دقة لنتائج تطبيق هذه الأدوات. وعليه فإن فحص ثبات القياس لنموذج دوافع الإنجاز في مجموعات عمرية مختلفة،

يجب أولاً وقبل كل شيء فحص النماذج المعمول بها والتي تتكون من عدة مقاييس في اختبارات PISA للحصول على دليل ثبات القياس عبر المجموعات قبل استخدامها لغرض المقارنة بين المجموعات بناءً على المتغيرات الديموغرافية مثل الجنس ونوع المدرسة، ويعد فحص ثبات القياس لنموذج دوافع الإنجاز من حيث الجنس ونوع المدرسة أمرًا مهمًا للمساواة في الخدمات التعليمية التي تقدمها الدولة في جميع المناطق سواء كانت في المدن أو الريف أو البادية.

وهدف هذه الدراسة فحص ثبات القياس لنموذج دوافع الإنجاز الذي تم بناؤه من خلال المواقف تجاه المنافسة والدافع لإتقان المهام والخوف الفشل في الاستبيان المقدم إلى الطلبة ضمن اختبارات PISA لعام ٢٠١٨، فيما يتعلق بنوع المدرسة في عينة الأردن.

وبالتحديد ستحاول هذه الدراسة الإجابة عن الأسئلة التالية:

- ما مستويات ملاءمة نموذج دوافع الإنجاز مع البيانات التي تم الحصول عليها من المجموعات الفرعية للمجموعة ككل ونوع المدرسة؟
- هل يتمتع نموذج دوافع الإنجاز بثبات القياس عبر المجموعات الفرعية لنوع المدرسة؟

فرضيات الدراسة:

- لا تختلف مؤشرات جودة الملائمة عن القيم المقبولة لها لبيانات نموذج دوافع الإنجاز في المجموعة الكلية، وفي المجموعات الفرعية حسب نوع المدرسة.
- لا يختلف ثبات القياس لنموذج دوافع الإنجاز باختلاف نوع المدرسة.

هدف الدراسة:

يكمن هدف هذه الدراسة في فحص ثبات القياس لنموذج دوافع الإنجاز الذي تم بناؤه من خلال المواقف تجاه المنافسة والدافع لإتقان المهام والخوف الفشل في الاستبيان المقدم للطلبة ضمن اختبارات PISA لعام ٢٠١٨، ودراسة المتغيرات المتعلقة بنوع المدرسة في عينة طلبة الأردن.

محددات الدراسة:

- اقتصرت هذه الدراسة على الاستجابات المقدمة لنموذج دوافع الإنجاز في استبيان الطلاب في اختبارات PISA لعام ٢٠١٨، فيما يتعلق بالمواقف تجاه المنافسة والدافع لإتقان المهام والخوف من الفشل.
- أقتصرت هذه الدراسة على تقدير ثبات القياس لنموذج دوافع الإنجاز بمجموعة الطلاب في سن ١٥ عامًا في الأردن.

مصطلحات الدراسة:

تكافؤ (ثبات) القياس: يشير مفهوم القياس إلى تعيين الأعداد بشكل منتظم على متغيرات تمثل خصائص الأشخاص أو الموضوعات أو الأحداث، وعندما نُقارن المجموعات ذات العلاقة بهذه الخصائص فإننا بذلك نفحص ما يسمى بتكافؤ القياس equivalent measurements الذي يشير إلى أن العلاقة بين الدرجات الملاحظة والأبنية الكامنة متطابقة عبر المجموعات ذات العلاقة(Vandenberg& Lance, 2000).

الدراسات السابقة

لقد حظيت اختبارات PISA بكثير من الدراسات التي تناولت معظم القضايا التي تتعلق بها، ومن هذه الأمور ثبات القياس من أجل المقارنات بين مجموعات المفحوصين عند تطبيق أدوات القياس سواءً كانت اختبارات لقياس مدى امتلاك الطلاب للمعارف والمهارات اللازمة للمشاركة الفاعلة في المجتمعات التي ينتمون إليها أم استبانات لقياس دافعية الإنجاز لدى الطلبة، وبهذا الصدد أجرى Adsul& Kamble (2008) دراسة هدفت للتحقيق من آثار الجنس والخلفية الاقتصادية والاختلافات الطبقية على دافع الإنجاز الذي يمتلكه طلاب الجامعات على أساس التحول المجتمعي، ولتحقيق أهداف الدراسة قام الباحثان على عينة مكونة من (٢٩٢) طالبا وطالبة من طلبة جامعة سانجلي في الهند، تم اختيارهم عشوائياً وفقًا لخطة البحث، تم اختيار ٤٨ موضوعًا من كل مجموعة طبقية أي الطبقات المتقدمة، والطبقات الأقل تقدماً، وطبقات القبائل البدوية على أساس نسبة الذكور إلى الإناث ١:١، وثلاثة مستويات من الخلفية الاقتصادية للأسرة. تم استخدام اختبار الدافع للإنجاز (ACMT) الذي طوره Bhargave لجمع البيانات من العينة. تم حساب اختبار "t" واختبار Duncan's Multiple Range وANOVA ثلاثي الاتجاهات. وأشارت النتائج أن هناك فرقًا كبيرًا بين طلاب القبائل البدوية وطلاب الطبقات الأخرى وبين الطلاب والطالبات، إذ يتمتع طلاب الطبقات المتقدمة بدافعية عالية للإنجاز بينما يتمتع طلاب مجموعات القبائل الأقل تقدماً والبدوية الأخرى بدافع متوسط للإنجاز، وكذلك يتمتع الطلاب الذكور بدافع مرتفع للإنجاز بينما تتمتع الطالبات بدافع أقل من المتوسط للإنجاز. والنتيجة الأكثر أهمية هي أن نسبة التفاعل المحسوبة لم تكن ذات دلالة إحصائية، مما يُشير إلى أن الطبقة والجنس والخلفية الاقتصادية للأسرة لا تؤثر بشكل مشترك على دافع الإنجاز لدى طلاب الجامعات.

كما أجرى He, Anwyll, Glanville & Opposs (2014) دراسة هدفت للتحقق من مستوى تكافؤ القياس في اختبارات العلوم الوطنية للمرحلة الأساسية (KS2) للعام (٢٠١٠، ٢٠١١)، الذي يرجع إلى متغير اللغة الإنجليزية إن كانت لغة أولى أو ثانية، وللكشف عن مستوى التكافؤ في القياس تم استخدام التحليل العامل الاستكشافي، ونموذج راش لتحليل بيانات الفقرات للكشف عن الأداء التفاضلي للتحقق من طريقة عمل فقرات الاختبارات عبر المجموعات الفرعية. وأشارت نتائج الدراسة أن الطلبة الذين يتحدثون اللغة الإنجليزية كلغة أولى كان أداؤهم أفضل من الذين يتحدثون اللغة الإنجليزية كلغة ثانية، كما أشار التحليل العاملي الاستكشافي نمط مماثل عبر المجموعات الفرعية والعينة الإجمالية في للعامين الدراسيين (٢٠١٠، ٢٠١١)، وكذلك ظهر مستوى عال من اللاتغير في القياس في معالم الفقرات عبر المجموعات.

كما أجرى Uyar & Doğan (2014) دراسة هدفت إلى فحص نموذج لاستراتيجيات التعلم في جزء "التعلم بالاستراتيجيات" من استبيان الطلاب في اختبارPISA لعام ٢٠٠٩ ، ولتحقيق أهداف الدراسة قام الباحثان بتحليل استجابات (٤٣٤٠) طالباً وطالبة ممن تقدموا إلى اختبار وفحص ثبات القياس للنموذج عبر متغيري الجنس ونوع المدرسة. ونظرًا لأن نموذج استراتيجيات التعلم قد يتم تفسيره بشكل مختلف عبر المجموعات الفرعية، فقد تم فحص ثبات النموذج فيما يتعلق بالجنس ونوع المدرسة وتسمية الوحدات الإقليمية للإحصاء (NUTS ١٢) في هذه الدراسة، كما تم فحص المتغيرات باستخدام اختبار الفروق بين شكل الثبات الأكثر تقييدًا والشكل الأساس لتحديد ما إذا كان الثبات للمعلمة عبر المجموعات متشابه. وأشارت نتائج الدراسة إلى أن النموذج قدم فقط شروط الثبات التكويني والمتري في مجموعات الجنس وأنواع المدارس، إلا أنه قدم جميع شروط الثبات بين المناطق.

كما أجرى Tsaousis (2015) دراسة هدفت للتحقق من البناء العاملي للاختبار التحصيلي المعد من قبل المركز الوطني للقياس والتقويم في السعودية لقبول الطلبة في الجامعات الحكومية، وبيان الاختلاف في الأداء الذي يعزى لجنس المستجيب، والذي استجاب له (٦٣٣٨٠) طالباً وطالبة، حيث تم اختبار ثلاثة نماذج لعوامل التأكيد (عامل واحد، وعاملان، وأربعة عوامل). وأشارت النتائج أن نموذج ثنائي العامل كان هو الأفضل ملاءمة مقارنة بالنماذج الأخرى، كذلك كان الأفضل ملاءمة لكلا الجنسين، كما أشارت النتائج إلى أن اختبار متوسطات الفرق الكامنة بين الجنسين كان لصالح الطالبات، وكذلك أن متوسطات أداء الطالبات أعلى مقارنة بالطلاب في موضوعي البيولوجيا والكيمياء، كذلك كان متوسط أداء الطلاب أعلى من متوسط أداء الطالبات فيما يتعلق بالمفاهيم الكامنة في الفيزياء والرياضيات

وكذلك الدراسة التي أجرها السمكي (٢٠١٧) لفحص اللاتغير في القياس عبر متغير موقع المدرسة وجنس الطالب على اختبار (TIMSS) الذي طبق عام ٢٠١١م في مقرري الرياضيات والعلوم للصف الثامن الأساسي في الأردن، باستخدام التحليل العاملي التوكيدي (CFA) والتحليل العاملي التوكيدي متعدد المجموعات (MGCFA)، وأظهرت النتائج ملاءمة نموذج العوامل الأربعة المفترض لاختبار مقرر الرياضيات، ونموذج العوامل الأربعة المفترض لاختبار مقرر العلوم، كما أظهرت النتائج ثبات القياس الشكلي والمتري والقياسي لفقرات اختباري العلوم والرياضيات عبر متغير موقع المدرسة، وكانت الأفضلية لمدارس المدن مقارنةً بمدارس الريف، كما أظهرت النتائج ثبات القياس الشكلي والمتري والقياسي عبر متغير جنس الطالب، وكانت الأفضلية لصالح الطالبات مقارنةً بالذكور.

كما أجرى Alqabbaa (2017) دراسة هدفت إلى بيان البنية العاملية لاختبار (SAAT) باستخدام التحليل العاملي الاستكشافي (EFA) والتحليل العاملي التوكيدي (CFA). ولتحقيق أهداف الدراسة قام الباحث باستخدام بيانات الاختبار المطبق على عينة مكونة من (١٤٠٠٣) طالباً وطالبة بالمدارس الحكومية والخاصة في السعودية ممن تقدموا لاختبار (SAAT) لعام ٢٠١٦؛ منهم (٧٢٤٨) طالباً، و(٦٧٥٥) طالبة. وتم اختبار التحليل العاملي التأكيدي لنموذج مكون من أربعة عوامل (النموذج المفترض) لمعرفة ما إذا كان النموذج مناسبًا أم لا بالإضافة إلى النموذج الناتج عن التحليل العاملي التوكيدي. أظهر التحليل العاملي التوكيدي أن اختبار (SAAT) يقيس عاملين فقط (علم الأحياء، والكيمياء)، وليس العوامل الأربعة التي افترضها مطورو الاختبار، كذلك ظهر أن التحليل العاملي التوكيدي كان ملائم للنموذج المكون من عاملين، كما أشارت النتائج للتحليل العاملي التوكيدي أن نموذج العوامل الأربعة المفترض يناسب البيانات بشكل جيد، وبالتالي تم اختيار النموذج المفترض باعتباره الأكثر ملاءمة. وبناءً على نموذج العوامل الأربعة CFA، تم فحص ثبات القياس (التكويني والمتري والقياسي) عبر نوع المدرسة (حكومية، خاصة) والجنس (ذكر، أنثى) على الاختبار، حيث أشارت النتائج أن نموذج الثبات المتري يناسب البيانات بشكل أفضل مقارنة بالنماذج الأخرى. كما تم اختبار فروق المتوسطات الكامنة باستخدام تحليل التباين ثنائي الاتجاه في جميع المواد ذات العوامل الأربعة (الأحياء والكيمياء والفيزياء والرياضيات)، كما أشارت النتائج أن الإناث في المدارس الثانوية حققن نتائج أفضل من الذكور في جميع الأقسام الأربع لاختبار SAAT، وكذلك لم يحقق الطلاب الذكور في المدارس الحكومية نتائج جيدة في الاختبار مقارنة بالذكور في المدارس الثانوية الخاصة. كما لم يحقق الطلاب الذكور في المدارس الحكومية نتائج جيدة مقارنة بالإناث في كلتا المدرستين.

وأجرى Ardıç& Gelbal (2017) دراسة هدفت إلى فحص ثبات القياس للعناصر المرتبطة بالاهتمام والدافعية المضمنة في الاستبانة المقدمة للطلبة في اختبارات PISA لعام ٢٠١٢ في عينة طلبة تركيا، المرتبط بالجنس ونوع المدرسة والمناطق الإحصائية وتحديد العناصر التي تظهر اختلافًا في الأداء على فقرات الاستبيان عبر المجموعات، ولتحقيق أهداف الدراسة قام الباحثان بإجراء تحليل العوامل التوكيدية متعدد المجموعات لفحص ثبات القياس. وعندما تم التحقيق من الثبات فيما يتعلق بالجنس، تم فحص تحيزات الفقرات المحتملة، حيث لم يتم استيفاء المعايير المستخدمة في تقييم ملاءمة النموذج، وتم استخدام تقنيات Mantel-Haenszel وpoly-SIBTEST ونسبة احتمالية نظرية استجابة الفقرة (IRT-LR) لتحديد الفقرات التي أظهرت اختلافا في الأداء عليها. وأشارت نتائج اختبار الثبات الذي أجري بناءً على نوع المدرسة والمناطق الإحصائية أن النماذج تلبي جميع شروط الثبات، على عكس متغير الجنس، فقد أظهرت النتائج الفشل في تحقيق ثبات القياس وفقًا للجنس، وعند فحص نتائج الاختلاف في الأداء على الفقرات وفقًا للجنس، فقد كان هناك اختلاف في أداء الطلبة على ستة فقرات في المستوى الأول حسب إحصائيات Mantel-Haenszel، وكذلك فقرة واحدة في المستوى الأول وفقرتين في المستوى الثاني، وحسب إحصائيات poly-SIBTEST كان هناك فقرة واحدة على مستوى A، وفقرتين على مستوى B، وثلاث فقرات على مستوى C، وحسب إحصائيات IRT-LR كان هناك فقرتان على مستوى C.

كما أجرى Karaman & Smith (2019) دراسة هدفت إلى تكييف وتقدير الخصائص السيكو مترية للنسخة التركية من مقياس دوافع الإنجاز، ولتحقيق أغراض الدراسة قام الباحثان بتطبيق مقياس دوافع الإنجاز على عينة مكونة من (٣٣٦) طالباً وطالبة بالجامعة. وأشارت النتائج المتعلقة بالمؤشرات الإحصائية للتحليل العاملي التوكيدي والتحليل الارتباطي لبناء العوامل وثبات مقياس دوافع الانجاز أنها جميعها مقبولة، مما يؤكد ملاءمة نموذج ثنائي العامل المكون من ١٣ عنصرًا، وكان لمقياس دافع الإنجاز ثبات جزئي في القياس يعزى الجنس. كما أن نتائج هذا المقياس تتفق مع نتائج المقياس دافع الإنجاز الأصلي.

من خلال ما توصلت إليه الدراسات السابقة التي اهتمت بثبات القياس من أجل المقارنات بين المجموعات المختلفة بناءً على أداء الطلبة على أدوات القياس (اختبارات، استبانات) المقدمة إليهم ضمن قياس المهارات المعرفية أو المهارية وكذلك العناصر المرتبطة بالاهتمام والدافعية المضمنة في الاستبيان المقدم لهم، وكذلك ثبات القياس الشكلي والمتري والقياسي، وكذلك ثبات البواقي، عبر متغير جنس الطالب، أو نوع المدرسة التي ينتمي إليها، ومكان المدرسة التي يدرسون بها (مدنية، ريف، بادية)، وأن هناك تباين في نتائج المقارنات بين مجموعات الطلبة

الذين يخضعون للاختبارات أو للاستبيانات، حيث كانت نتائجهم تختلف من دولة لأخرى، ومن منطقة لأخرى داخل نفس الدولة، وهذا يرجع إلى طبيعة الأنظمة التعليمية المتبعة في كل دولة، ومدى تقدمها وتطورها في المناطق الجغرافية المختلفة داخل الدولة نفسها.

الطريقة والإجراءات:

مجتمع الدراسة وعينتها:

تكونت عينة الدراسة من جميع طلبة الصف العاشر الأساسي بمتوسط أعمار (١٥) عاماً من جميع مدارس الأردن التي شاركت في الاختبار الدولي PISA في عام ٢٠١٨، حيث بلغ عدد المدارس المشاركة (٣١٣) مدرسة، وبلغ عدد الطلبة الذين شاركوا في الاختبار ٨٩٦٣ طالباً وطالبة، علماً أن عدد الدول التي شاركت في هذا الاختبار ٧٩ دولة، وتكونت العينة في صورتها النهائية بعد تصفية البيانات ومعالجتها من ٨٢٣٧ طالباً وطالبة، وتوزعت هذه الاستجابات حسب نوع المدرسة كما هي في الجدول رقم ١

الجدول رقم (١) توزيع عينة الدراسة حسب متغير نوع المدرسة

نوع المدرسة	العدد	النسبة %
ذكور	3347	40.6
إناث	2286	27.8
مختلطة	2604	31.6
المجموع	8237	100.0

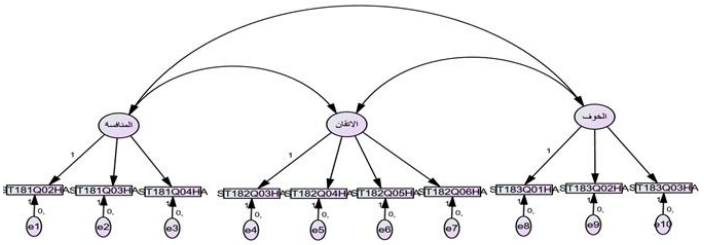
بيانات الدراسة:

تم الحصول على بيانات الدراسة من خلال استجابات الطلبة على مقياس دوافع الإنجاز في بيزا الدولي ٢٠١٨، وهذه البيانات منشورة على الموقع الإلكتروني (OECD PISA) https://www.oecd.org/pisa/data2018database/. ومقياس دوافع الإنجاز يتكون من ١٠ فقرات موزعة على ثلاثة عوامل هي: الاتجاهات نحو المنافسة، الدوافع للإتقان، والخوف من الفشل. وتسير الفقرات وفق تدريج ليكرت الرباعي: موافق بشدة، موافق، غير موافق، غير موافق بشدة. والملحق رقم (١) يوضح فقرات مقياس دوافع الإنجاز وأرقام الفقرات والكود الخاص بكل فقرة، وكل محور.

معالجة البيانات والتحليلات الإحصائية:

تم الحصول على ملف البيانات بصيغة بيانات برنامج SPSS، وفي البداية تم معالجة البيانات المفقودة بالحذف، حيث تبين أن معظم هذه الفقرات مفقودة بشكل كامل على خيارات الاستجابة لبعض الفقرات، حيث تم حذف ٧٢٦ استجابة. ثم تم معالجة القيم المتطرفة باستخدام (Mahala Nobis Distance) باستخدام برنامجٍ SPSS v26، وحسب جميع مسافات Mahala Nobis لجميع الاستجابات وحسبت الدلالة الإحصائية لها باستخدام اختبار كاي تربيع بدرجات حرية ١٠ بعدد فقرات المقياس، وتبين أن قيم مسافات Mahala Nobis لجميع الاستجابات في العينة كاملة بعد حذف القيم المفقودة غير دالة، حيث تراوحت قيم الدلالة الإحصائية لمسافات Mahala Nobis بين ٠.٩ إلى ١، وبالتالي استقرت الاستجابات على ٨٢٣٧ استجابة.

وبعد معالجة البيانات تم بناء نموذج ثبات القياس، واستخدام التحليل العاملي التوكيدي Confirmatory Factor Analysis (CFA) والتحليل العاملي التوكيدي متعدد المجموعات تبعاً لمتغير نوع المدرسة Multi Group Confirmatory Factor Analysis(MGCFA) لفحص جودة مطابقة البيانات للنموذج النظري عبر المجموعات الفرعية لمتغير نوع المدرسة، وللعينة كلها، وللتحقق من ثبات القياس بمستوياته المختلفة. والنموذج النظري لمقياس دوافع الإنجاز الذي تم تحليله في هذه الدراسة تكون من ثلاثة عوامل كما في الشكل رقم ١ الآتي:



شكل (١) النموذج النظري لمقياس دوافع الإنجاز الذي تم تحليله في الدراسة.

وتم استخدام برنامج أموس (AMOS, v22) لتنفيذ المعالجات الإحصائية، وتم فحص ثبات القياس باستخدام MGCFA وتضمنت التحليلات الإحصائية سلسلة من الاختبارات الإحصائية لمستويات ثبات القياس الهرمية، حيث تم فحص جودة مطابقة البيانات للنموذج النظري عبر المجموعات الفرعية لمتغير نوع المدرسة، وللعينة كملة، ثم التحقق من ثبات القياس بمستوياته المختلفة. ولتقييم جودة المطابقة بين النموذج والبيانات تم استخدام المؤشرات: chi-square χ2 ، الجذر التربيعي لمتوسط مربعات أخطاء التقدير Root Mean Squared Error of Approximation (RMSEA)(Steiger, 1989)، الجذر التربيعي لمتوسط مربعات البواقي المعيارية Standardized Root Mean Square Residual (SRMR)(Bentler, 1995)، مؤشر المطابقة المقارن Comparative Fit Index (CFI)(Bentler,1990) ، مؤشر توكر لويس Tucker-Lewis Index(TLI)(Tucker & Lewis, 1973) ، ومؤشر النسبة المركزية Relative Centrality Index (RNI)(McDonald & Marrsh,1990)

إن قيمة χ^2 هي دالة تعتمد بشكل كبير على حجم العينة، فمن الطبيعي استخدام حجم عينة كبير سيؤدي إلى رفض الفرضية الصفرية المتعلقة بجودة المطابقة بين البيانات والنموذج فلذلك يرى الباحثون أنها بحد ذاتها غير كافية للحكم على جودة المطابقة (Wu et al.,2007; Finney& DiStefano, 2006; Kaplan, 2000; Sekercioglu, 2018)، ونظر البعض الآخر إلى اعتماد بعض المؤشرات تكفي للحكم على جودة المطابقة مثل: RMSEA, SRMR, TLI, RNI (Vandenberg & Lance, 2000) ونظراً لأن الدراسة الحالية استخدمت حجوم عينات كبيرة سواء على مستوى حجم العينة كاملة أو على حجم العينات الفرعية، اعتمدت هذه الدراسة بعض المؤشرات التي يوفرها برنامج AMOS وحول الحدود المقبولة لهذه المؤشرات، فالجدول رقم ٣ يظهر هذه الحدود.

الجدول رقم (٣): الحدود المقبولة لمؤشرات جودة المطابقة بين البيانات والنموذج في التحليل العاملي التوكيدي.

المؤشر	المدى المقبول
RMSEA	0.05-0.08
CFI	أعلى أو يساوي 0.8
TLI	أعلى أو يساوي 0.8
SRMR	أقل أو يساوي من 0.09

(Hooper et al., 2008; Hu & Bentler, 1999; Tabchnick & Fidell, 2013; Vandenberg & Lance, 2000; Hair et al., 1995; Brown & Cudeck, 1993)

وحول مؤشرات ثبات القياس في المستوى المتري، والقياسي، وثبات البواقي فإنه يتم اعتماد الفرق بين مؤشرات جودة المطابقة بين مستويات القياس الهرمية، بمعنى ينظر إلى الفرق في قيمة هذه المؤشرات بين مستوى ثبات القياس قيد التحليل والمستوى الذي قبله، ولنفس سبب ارتفاع حجم العينة فإن مؤشر الفرق في χ^2 ($\Delta\chi^2$) يتأثر بحجم العينة واعتماده بحد ذاته مؤشر على ثبات القياس لا يكفي (Yandi et al., 2017) ، بعض الباحثين اقترح بعض المؤشرات للحكم على ثبات القياس مثل الفرق في مؤشر جاما Δ Gamma hat ومؤشر Δ McDonald's (Cheung & Rensvold, 2002) ولكن هاذين المؤشرين لا يوفرهما برنامج التحليل الإحصائي المستخدم في هذه الدراسة (AMOS, v22) ، لذلك تم اعتماد فروق المؤشرات الأخرى، وحول القيم المقبولة لفروق المؤشرات، فإنه يفضل أن يقل ΔCFI عن ٠.٠١ و ΔRMSEA عن ٠.٠١٥ وΔRSMR عن ٠.٠٣ (Chen, 2007)

النتائج:

- للإجابة عن فرضية الدراسة الأولى: لا تختلف مؤشرات جودة الملائمة عن القيم المقبولة لها لبيانات نموذج دوافع الإنجاز في المجموعة الكلية، وفي المجموعات الفرعية حسب نوع المدرسة.

لقد تم تقييد البناء النظري في الشكل رقم ١ في العينة ككل وفي العينات الفرعية تبعاً لمتغير نوع المدرسة أثناء استخدام التحليل العاملي التوكيدي، ويظهر الجدول رقم ٤ مؤشرات جودة المطابقة لمقياس دوافع الإنجاز في العينة كاملة، وفي العينات الفرعية تبعاً لمتغير نوع المدرسة.

الجدول رقم(٤): مؤشرات المطابقة لمقياس دوافع الإنجاز في العينة كاملة وفي المجموعات الفرعية تبعاً لمتغير نوع المدرسة.

المجموعة	χ^2	df	χ^2/df	RMSEA	SRMR	CFI	TLI
كامل العينة	1015.125	32	31.7	.061	0.034	0.970	0.957
مدارس الذكور	918.460	32	28.702	0.079	0.05	0.941	0.916
مدارس الإناث	101.286	32	3.165	.031	0.019	0.991	0.988
المدارس المختلطة	151.925	32	4.748	0.038	0.018	0.986	0.980

يلاحظ من الجدول رقم ٤ عند استثناء مؤشر χ^2 فإن جميع مؤشرات المطابقة وقعت ضمن الحدود المقبولة سواء على مستوى العينة كاملة، أو على مستوى المجموعات الفرعية تبعاً لمتغير نوع المدرسة، وظهر أفضلها في مدارس الإناث إذا أخذنا بعين الاعتبار مؤشر χ^2 ، حيث ظهرت أقل قيمة لها، وعند قسمة المؤشر على درجات الحرية فإن الناتج ٣ تقريباً. ونظراً لارتفاع حجم العينة في هذه الدراسة، فإن اعتماد المؤشرات الواردة في الجدول رقم ٤ يبرر مطابقة البيانات والنموذج النظري لدوافع الإنجاز ويدلل على ثبات القياس الشكلي Configural Invariance لمقياس دوافع الإنجاز عند تقييد البناء النظري له عبر المجموعات المختلفة لمتغير نوع المدرسة.

- وللإجابة عن فرضية الدراسة الثانية: لا يختلف ثبات القياس لنموذج دوافع الإنجاز باختلاف نوع المدرسة.

تم أولاً إجراء تحليل عاملي توكيدي متعدد المجموعات MGCFA ضمن متغير نوع المدرسة دون إجراء أي تقييد على الشكل رقم ١ سواء في أوزان المتغيرات على السمات الكامنة في النموذج أو على ثوابت المعادلات الانحدارية والبواقي، وسجلت النتائج في الجدول رقم ٥

الجدول رقم ٥ : نتائج مطابقة ثبات القياس الشكلي وفق متغير نوع المدرسة

المتغير	χ^2	df	χ^2/df	RMSEA	SRMR	CFI	TLI
نوع المدرسة	1171.622	96	12.204	0.037	0.035	0.966	0.952

بعيداً عن الأخذ بعين الاعتبار لمؤشر χ^2 بسبب ارتفاع حجم العينة، فإن النتائج في الجدول رقم ٥ تشير إلى قيم مقبولة لمؤشرات المطابقة في التحليل العاملي التوكيدي متعدد المجموعات وفقاً لمتغير نوع المدرسة، وهكذا يؤكد النتيجة السابقة عند تقييد النموذج في كل مجموعة على حدى، وبالتالي يمكن القول أن الثبات الشكلي لمقياس دوافع الإنجاز قد تحقق عبر مجموعات نوع المدرسة، بعبارة أخرى، يستخدم الأفراد في مجموعات أنواع المدارس المختلفة نفس الإطار المفاهيمي عند الإجابة على فقرات المقياس.

وللتحقق من أنماط ثبات القياس الهرمية الأخرى كالمتري Metric Invariance وثبات القياس القياسي Scaler Invariance وثبات القياس الصارم أو ثبات قياس البواقي Residual Invariance تم إجراء تحليل متعدد المجموعات من خلال ايقونة التحليل في برنامج AMOS والذي يتم فيه اجراء سلسلة من التحليلات، بحيث يتم اجراء افتراض تقييد شروط مستوى ثبات

القياس المطلوب للمقارنة مع مستوى ثبات القياس الذي يسبقه من خلال مؤشرات المطابقة، فمثلاً إذا أردنا التحقق من ثبات القياس المتري يتم تقييد أوزان المتغيرات على السمات الكامنة في النموذج ويتم تنفيذ التحليل واستخراج مؤشرات المطابقة ومقارنتها مع مؤشرات القياس الشكلي أو التكويني، والجدول رقم ٦ يلخص مؤشرات ثبات القياس بأنماطه وفقاً لمتغير نوع المدرسة.

الجدول رقم ٦: مؤشرات ثبات القياس بأنماطه وفق لمتغير نوع المدرسة.

نمط ثبات القياس	χ^2	df	χ^2/df	RMSEA	SRMR	CFI	TLI
الشكلي	1171.622	96	12.204	0.037	0.035	0.966	0.952
المتري	1314.834	110	11.953	0.036	0.042	0.962	0.953
القياسي	1665.614	122	13.653	0.039	0.087	0.951	0.946
الصارم أو البواقي	2046.158	142	14.410	0.040	0.094	0.939	0.942

وتم حساب الفروق بين المؤشرات التي تم اعتمادها في هذه الدراسة للتحقق من ثبات القياس بأنماطه الهرمية المختلفة، والجدول رقم ٧ يبين ذلك.

الجدول رقم ٧ : فروقات مؤشرات ثبات القياس في مقياس دوافع الإنجاز تبعاً لمتغير نوع المدرسة.

نمط ثبات القياس	$\Delta\chi^2$	ΔRMSEA	ΔRSMR	ΔCFI
الشكلي	-	-	-	-
المتري	143.212	-0.001	0.007	-0.004
القياسي	350.78	0.003	0.045	-0.011
الصارم أو البواقي	380.544	0.001	0.007	-0.012

وإذا تمعنا في قيم المؤشرات في الجدول رقم ٦، ٧ نجد أن ثبات القياس المتري قد تحقق، بعيداً عن الأخذ بعين الاعتبار مؤشر χ^2 ومؤشر $\Delta\chi^2$ نظراً لارتفاع حجم العينات في هذه الدراسة، فقيم مؤشرات المطابقة وفروقاتها مع قيم مؤشرات القياس الشكلي ضمن الحدود المقبولة لها، أي أن أحمال العوامل في النموذج مقبولة لتكون هي نفسها بالنسبة لمجموعات نوع المدرسة كما هو الحال في هياكل العوامل في النموذج. ولم يتحقق الثبات القياسي، وبالتالي عدم تحقق ثبات القياس الصارم أو ثبات قياس البواقي، نظراً لهرمية أنماط ثبات القياس.

مناقشة النتائج والاستنتاجات:

تقييم الجوانب الوجدانية لا يقل أهمية عن تقييم الجوانب المعرفية في الاختبارات الدولية، وخصوصاً عندما نتكلم عن دوافع الإنجاز، فهي الطريق إلى تحقيق الجوانب المعرفية، لذلك تلعب دوراً مهماً في تشكيل سلوك الطلاب في المستقبل، فيجب على المعلمين الانتباه إلى الخصائص الوجدانية للطلبة، وتطوير المناهج بما يتلاءم مع مجالات التعلم الثلاثة المعرفية والنفسية الحركية والوجدانية، ويجب أيضًا قياس مستوى اكتساب هذه المعرفة والمهارات والخصائص الوجدانية والتخطيط لسياسات التعليم وفقًا لذلك. وقبل إجراء القياسات في المجال الوجداني، يجب إثبات ثبات القياس لأدوات القياس التي سيتم استخدامها. وبهذه الطريقة، يمكن تحديد ما إذا كانت النتائج التي تم الحصول عليها ترجع إلى أداة القياس أم لا.

فإلى جانب اهتمام الاختبارات الدولية PISA بقياس وتشخيص الجوانب المعرفية لدى الطلبة إلا أنها لم تغفل عن قياس وتشخيص الجوانب الوجدانية إيماناً بأهمية الدور الذي تلعبه في خصائص سلوك الطلبة مستقبلاً، ومن بين الخصائص العاطفية التي تم قياسها في اختبار PISA لعام ٢٠١٨ مقاييس المواقف تجاه المنافسة، والدافع لإتقان المهام، والخوف من الفشل، والتي تندرج تحت نموذج دوافع الإنجاز، وعندما يكون من المرغوب فيه فحص الخصائص الوجدانية للأفراد أو إجراء دراسات تتعلق بهذه الخصائص، يجب أولاً وقبل كل شيء ضمان ثبات القياس لأدوات القياس التي تقيس هذه الخصائص. يعد ثبات القياس مهمًا لأنه يمكن أن يوفر دليلاً على ما إذا كانت الاختبارات أو الاستبيانات تقيس نفس العامل بنفس الطريقة في مجموعات مختلفة.

ونظرًا إلى الفروق في نسب التحصيل بين أنواع المدارس في الأردن، فإن النتائج المتعلقة بنموذج دوافع الإنجاز التي تم الحصول عليها دون مراعاة ثبات القياس في نوع المدرسة قد لا تكون صالحة وموثوقة للمقارنات حسب نوع المدرسة، ولن تكون ذات مغزى إذا لم يتحقق ثبات القياس بشكل كامل، ففي هذه الدراسة، يُظهر نموذج دوافع الإنجاز أن هناك ثباتاً جزئي (الشكلي، المتري) فقط، وعليه صعوبة المقارنات الإحصائية بين الأنواع المختلفة للمدارس، كمقارنة الأوساط الحسابية، حيث لم يتحقق الثبات القياسي، وبالتالي عدم تحقق ثبات القياس الصارم أو ثبات قياس البواقي، نظراً لهرمية أنماط ثبات القياس، وهذه النتيجة تتفق مع دراسة (Uyar & Doğan, 2014)، وتتفق بشكل جزئي مع دراسة (Ardıç&Gelbal,2017، كما تتفق مع نتائج دراسة (Alqabbaa, 2017) الذي قام بفحص ثبات القياس بين مجموعات المستجيبين على اختبار (SAAT) عند مقارنة مجموعات المستجيبين بالنسبة لمتغير نوع المدرسة (ذكور، إناث، مختلطة).

التوصيات:

من خلال نتائج الدراسة يوصي الباحثان بما يلي:

- التأكد من ثبات القياس للمقاييس أو النماذج التي يتم من خلالها المقارنة بين مجموعات المستجيبين قبل استخدامها وتعميم النتائج، أن كانت تلك المقاييس أو النماذج تستخدم على المستوى المحلي أو الإقليمي او العالمي.
- يمكن إجراء دراسات مماثله باستخدام مجموعات أخرى ونماذج أو مقاييس مختلفة أو في مناطق مختلفة أو حسب متغيرات أخرى.
- تم استخدام طريقة التحليل العاملي التوكيدي متعدد المجموعات MGCFA في هذه الدراسة، ويمكن استخدام طرق أخرى ومقارنتها لفحص ثبات القياس لأن MGCFA لديها قيود في اختبار ثبات القياس عندما يكون عدد المجموعات وحجم العينة كبيرين.

المراجع

السمكي، لطفي .(٢٠١٧). اللاتغير في القياس لاختبار TIMSS لعام ٢٠١١ في الرياضيات والعلوم عبر موقع المدرسة وجنس الطالب .أطروحة دكتوراه غير منشورة، جامعة اليرموك.

مخائيل، أمطانيوس (٢٠١٦). بناء الاختبارات، والمقاييس النفسية والتربوية وتقنينها. عمان، الإعصار للتوزيع والنشر.

Adsul, R. K., Kamble, V., & Sangli, K. W. (2008). Achievement motivation as a function of gender, economic background and caste differences in college students. *Journal of the Indian Academy of Applied Psychology*, 34(2), 323-327.

Alqabbaa, M. (2017). *Factor Invariance of the Standard Achievement Admission Test by Gender and School Types*. Electronic Theses and Dissertations, University of Denver. 1331.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Bentler, P. M. (1995). *EQS structural equations program manual*. Multivariate Software.

andenberg, M. W., & Cudeck, R. (1993). *Alternative ways of assessing model fit*. In K. A. Bollen and J. S. Long (Eds.), Testing structural equation models (pp. 136-162). Newbury Park, CA: Sage.

Campbell, H. L., Barry, C. L., Joe, J. N., & Finney, S. J. (2008). Configural, metric, and scalar invariance of the modified achievement goal questionnaire across African American and white university students. *Educational and Psychological Measurement*, 68(6), 988-1007.

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

504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling: *A Multidisciplinary Journal*, 9(2), 233–255.

Ding, Y., Yang Hansen, K., & Klapp, A. (2023). Testing measurement invariance of mathematics self-concept and self-efficacy in PISA using MGCFA and the alignment method. *European Journal of Psychology of Education*, 38(2), 709-732.

Finney, S. J., & DiStefano, C. (2006). *Nonnormal and Categorical Data in Structural Equation Models*. In G. R. Hancock, & R. O. Mueller (Eds.), A Second Course in Structural Equation Modeling (pp. 269-314). Greenwich, CT: Information Age.

Franken, R. E., & Brown, D. J. (1995). Why do people like competition? The motivation for winning, putting forth effort, improving one's performance, performing well, being instrumental and expressing forceful/aggressive behavior, Personality and Individual Differences, 19(2), 175-184.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Hangen E. J., & Elliott A.J. (2016) *Achievement motives*. In: Zeigler-Hill V., Shackelford T. (eds), Encyclopedia of personality and individual differences. (pp. 1-3). Springer International Publishing.

He, Q., Anwyll, S., Glanville, M., & Opposs, D. (2014). An investigation of measurement invariance of the Key Stage 2 National Curriculum science sampling test in England. *Research Papers in Education*, 29(2), 211-239.

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. *The Electronic Journal of Business Research Methods*, 6, 53-60.

http://www.qiyas.sa/MAndAssesment/Researches/Pages/Researches.aspx

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. https://doi.org/10.1080/10705519909540118

Kaplan, D. (2000). Structural Equation Modeling: Foundations and Extensions. Sage.

https://doi.org/10.4135/9781412984256

Karaman, M. A., & Smith, R. (2019). Turkish adaptation of achievement motivation measure. *International Journal of Progressive Education*, 15(5), 185-197.

Khademi, A. (2020). An investigation of fit criteria within MG-CFA for examining non-negligible measurement invariance [Unpublished doctoral dissertation, University of Massachusetts Amherst]. Doctoral Dissertations.

Khorramdel, L., Pokropek, A., & van Rijn, P. (2020). Special Topic: establishing comparability and measurement invariance in large-scale assessments, part I. *Psychological Test and Assessment Modeling*, 62(1), 3-10.

McDonald, R. P., & Marsh, H. W. (1990). Choosing a multivariate model: Noncentrality and goodness of fit. *Psychological Bulletin*, 107(2), 247–255. https://doi.org/10.1037/0033-2909.107.2.247

OECD. (2019). PISA 2018 Results (Volume I): What Students Know and Can Do, PISA, Paris: OECD Publishing. https://doi.org/10.1787/5f07c754-en.

Şekercioğlu, G. (2018). Measurement invariance: Concept and implementation. *International Online Journal of Education and Teaching* (*IOJET)*, 5(3). 609-634. http://iojet.org/index.php/IOJET/article/view/439/257

Steiger, J. H. (1989). *EzPATH: Causal modeling*. SYSTAT. https://www.statpower.net/Steiger%20Biblio/EzPath%20Manual.pdf

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

Tsaousis, I. (2015). Factorial invariance and latent mean differences of scores on SAAT across gender.

http://www.qiyas.sa/MAndAssesment/Researches/Pages/Researches.aspx

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38(1), 1–10. https://doi.org/10.1007/BF02291170

Uyar, Ş., & Doğan, N. (2014). Examining the measurement invariance of learning strategies model in different groups in PISA 2009 Turkey sample. *International Journal of Turkish Educational Sciences*, 2014(3), 30-43.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational research methods*, 3(1), 4-70.

Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. Practical Assessment, *Research and Evaluation*, 12(3), 1-26.