Journal of Educational and Psychological Sciences
Volume (4), Issue (21) : 30 يونيو 2020
P: 45 - 67

ISSN: 2522-3399

مجلة العلوم التربوية والنفسية
المجلد (4)، العدد (21) : 30 يونيو 2020 م
ص: 45 - 67

Validity and reliability Measurement of TIMSS 2011 test in mathematics across school location and supervising authority in Jordan

Lutfi Fawzi Al-Samaky
Ministry of Education || Jordan

Abstract: The purpose of this study is to examine the measurement invariance of TIMSS mathematics test for eighth grade across school location and the supervising authority (Ministry of Education, UNRWA) in Jordan. To achieve this, the researcher used the analytical descriptive approach to analyze the test results for Jordanian 8th graders who participated in the TIMSS study in 2011. Specifically, the responses of 547 students to 34 items in booklet eight and nine were analyzed. Seven items were excluded because the skewness index is greater than 2 and the kurtosis index is greater than 7. The data of the remaining 27 items were analyzed by using Confirmatory Factor Analysis (CFA) and Multiple Groups Confirmatory Factor Analysis (MGCFA) through the statistical program (AMOS 22). To judge the suitability of the test factorial structure the following goodness-of-fit indices in Structural equation modeling (SEM) were used: Chi Square ($\chi 2$) test was a statistically significant value of 468.167 and a good comparison index was CFI statistically significant value of 0.934 which is greater than 0.80 and the square of the mean root error approx.

The results showed that there is no Configural invariance according to the location of the school (countryside, city), and the values of the $\chi 2$ were a statistical function and its value is 822.858 and a good comparison index CFI is a statistical function and its value is 0.918 which is greater than 0.80 and the mean square root of the approximate error RMSEA is a statistic function and its value is 0.024 which is less than 0.08 and a root index The mean of the standard residual squares of SRMR is statistically significant and its value is 0.0408 which is less than 0.09.

The results showed that there is no Metric (or weak) invariance according to the location of the school (countryside, city), the values of the $\chi 2$ were a statistical function and its value is 861.321 and a good comparison index CFI is a statistical function and its value is 0.913 which is greater than 0.80 and the average square error of the approximate error RMSEA is a statistical function whose value is 0.024 and is less than 0.08 and an index The root mean standard SRMR squares are 0.0432 and less than 0.09.

The results also showed that there was no Scalar (or strong) invariance according to the location of the school (countryside, city), and the values of the $\chi 2$ were a statistical function and its value was 910.730 and a good comparison index CFI was a statistical function and its value is 0.903 which is greater than 0.80 and the average square of the approximate error root RMSEA statistically significant value and 0.025 which is less than 0.08 and an index The root mean standard SRMR squared is 0.0430 and is less than 0.09.

The results showed that there is no Configural invariance in the supervising authority (Ministry of Education, UNRWA), and the values of the $\chi 2$ were a statistically significant value of 833.660 and a good comparison index CFI was a statistical function with a value of 0.898 which is greater than 0.80 and an average square of the root error of approximate RMSEA statistically significant value of 0.028 which is less than 0.08 The mean index of the standard residual squares SRMR is statistically significant, and its value is 0.0351, which is less than 0.09.

The results showed that there is no Metric (or weak) invariance for the supervising authority (Ministry of Education, UNRWA), and the values of the χ2 were a statistically significant value of 890.903 and a good comparison index CFI was a statistically significant value of 0.883 which is greater than 0.80 and the square of the mean root error of approx. RMSEA statistically significant value of 0.029 which is less than 0.08 and the mean root index of the standard residual squares SRMR are statistically significant and have a value of 0.0620 which is less than 0.09.

The results also showed that there is no Scalar (or strong) invariance of the supervising authority (Ministry of Education, UNRWA), and the values of the χ2 were a statistical function and its value is 993.653 and a good comparison index CFI is a statistical function and its value is 0.854 and it is greater than 0.80 and the square of the average root error of RMSEA is a statistic function and its value is 0.033 which is less than 0.08 and the mean root index of the standard SRMR squares are statistically significant and their value is 0.0634 which is less than 0.09.

Keywords: Measurement Invariance, Factorial Invariance, Mathematics, TIMSS, eighth grade.

الصدق والثبات في القياس لاختبار تمس (TIMSS 2011) في العلوم للصف الثامن عبر موقع المدرسة والجهة المشرفة في الأردن

لطفي فوزي السمكي

وزارة التربية والتعليم || الأردن

الملخص: هدفت هذه الدراسة إلى فحص الصدق والثبات في اختبار TIMSS 2011 في الرياضيات للصف الثامن الأساسي في الأردن عبر كل من متغير موقع المدرسة (مدينة، ريف) ومتغير الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث). ولتحقيق ذلك استخدم الباحث المنهج التحليلي الوصفي في تحليل نتائج اختبار تمس في الرياضيات لطلبة الصف الثامن الذين شاركوا في دراسة تمس في عام 2011. وتكونت العينة من 574 طالبًا وطالبةً على 27 فقرة باستخدام أسلوبي التحليل العاملي التوكيدي CFA والتحليل العاملي التوكيدي متعدد المجموعة MGCFA من خلال البرنامج الإحصائي AMOS 22. وللحكم على مدى ملائمة البناء العاملي للاختبار تم استخدام مؤشرات جودة الملائمة في النماذج الخطية البنائيةSEM وهي: مربع كاي سكوير2χ التي كانت دالة إحصائيا وقيمتها 468.167 ومؤشر حسن المقارنةCFI دالة إحصائيا وقيمتها 0.934 وهي أكبر من 0.80 ومربع متوسط جذر الخطأ التقريبيRMSEA دالة إحصائيا وقيمتها 0.029 وهي أقل من 0.08 ومؤشر جذر متوسط مربعات البواقي المعيارية SRMR دالة إحصائيا وقيمتها 0.0383 وهي أقل من 0.09. وأظهرت النتائج وجود لا تغير تكويني باختلاف موقع المدرسة (ريف، مدينة)، وكانت قيم مربع كاي سكوير2χ دالة إحصائيا وقيمتها 822.858 ومؤشر حسن المقارنةCFI دالة إحصائيا وقيمتها 0.918 وهي أكبر من 0.80 ومربع متوسط جذر الخطأ التقريبيRMSEA دالة إحصائيا وقيمتها 0.024 وهي أقل من 0.08 ومؤشر جذر متوسط مربعات البواقي المعيارية SRMR دالة إحصائيا وقيمتها 0.0408 وهي أقل من 0.09.

وقد أظهرت النتائج وجود لا تغير متري باختلاف موقع المدرسة (ريف، مدينة)، وكانت قيم مربع كاي سكوير2χ دالة إحصائيا وقيمتها 861.321 ومؤشر حسن المقارنةCFI دالة إحصائيا وقيمتها 0.913 وهي أكبر من 0.80 ومربع متوسط جذر الخطأ التقريبيRMSEA دالة إحصائيا وقيمتها 0.024 وهي أقل من 0.08 ومؤشر جذر متوسط مربعات البواقي المعيارية SRMR دالة إحصائيا وقيمتها 0.0432 وهي أقل من 0.09.

كما وأظهرت النتائج وجود لا تغير عددي باختلاف موقع المدرسة (ريف، مدينة)، وكانت قيم مربع كاي سكوير2χ دالة إحصائيا وقيمتها 910.730 ومؤشر حسن المقارنةCFI دالة إحصائيا وقيمتها 0.903 وهي أكبر من 0.80 ومربع متوسط جذر الخطأ التقريبيRMSEA دالة إحصائيا وقيمتها 0.025 وهي أقل من 0.08 ومؤشر جذر متوسط مربعات البواقي المعيارية SRMR دالة إحصائيا وقيمتها 0.0430 وهي أقل من 0.09.

وأظهرت النتائج وجود لا تغير تكويني الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث)، وكانت قيم مربع كاي سكوير2χ دالة إحصائيا وقيمتها 833.660 ومؤشر حسن المقارنةCFI دالة إحصائيا وقيمتها 0.898 وهي أكبر من 0.80 ومربع متوسط جذر الخطأ التقريبيRMSEA

دالة إحصائيا وقيمتها 0.028 وهي أقل من 0.08 ومؤشر جذر متوسط مربعات البواقي المعيارية SRMR دالة إحصائيا وقيمتها 0.0351 وهي أقل من 0.09.

وقد أظهرت النتائج وجود لا تغير متري الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث)، وكانت قيم مربع كاي سكوير$\chi 2$ دالة إحصائيا وقيمتها 890.903 ومؤشر حسن المقارنةCFI دالة إحصائيا وقيمتها 0.883 وهي أكبر من 0.80 ومربع متوسط جذر الخطأ التقريبيRMSEA دالة إحصائيا وقيمتها 0.029 وهي أقل من 0.08 ومؤشر جذر متوسط مربعات البواقي المعيارية SRMR دالة إحصائيا وقيمتها 0.0620 وهي أقل من 0.09.

كما وأظهرت النتائج وجود لا تغير عددي الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث)، وكانت قيم مربع كاي سكوير$\chi 2$ دالة إحصائيا وقيمتها 993.653 ومؤشر حسن المقارنةCFI دالة إحصائيا وقيمتها 0.854 وهي أكبر من 0.80 ومربع متوسط جذر الخطأ التقريبيRMSEA دالة إحصائيا وقيمتها 0.033 وهي أقل من 0.08 ومؤشر جذر متوسط مربعات البواقي المعيارية SRMR دالة إحصائيا وقيمتها 0.0634 وهي أقل من 0.09.

الكلمات المفتاحية: اللاتغير في القياس، اللاتغير العاملي، الرياضيات، اختبار تمس، الصف الثامن.

المقدمة

تسعى كثير من دول العالم إلى إصلاح التعليم وتحسين أدائه لأن للتربية مكانة متميزة وتحتل الصدارة في منظومة الاستراتيجيات التنموية. وترتبط التربية بالعنصر البشري، والذي يعد من أهم العناصر الأساسية في إحداث التنمية، وبجميع مجالات الحياة المختلفة (موسى، 2012). وقد بدأ الاهتمام ومنذ مطلع ستينات القرن الماضي في إجراء الدراسات الدولية التي تهدف بشكل رئيس إلى المقارنة بين اتجاهات وتحصيل طلبة مجموعة من دول العالم، ودراسة العوامل التي تؤثر فيهما. ومن هذه الدراسات برنامج التقييم الدولي للطلبة PISA Program for International Student Assessment والدراسة الدولية للعلوم والرياضيات (TIMSS) Trends In International Mathematics & Science Study وغيرها من الدراسات. وتكمن أهمية الدراسات الدولية في تمكينها للدول المشاركة من فهم أنظمتها التربوية بشكل أفضل، مما يساعد صانعي القرارات وراسمي السياسات التربوية على تحديد معايير حقيقية وواقعية للتحصيل أو الأداء التربوي والتي تعينهم في مراقبة وتقييم نجاحات أو إخفاقات نظمهم التربوية (ابولبدة، حامد، الطويسي، وعبابنة، 2013).

وتشكل الدراسات الدولية إضافة مهمة للدراسات التي تجري على الصعيد الوطني لكل دولة من الدول المشاركة، إذ أنها توسع مدى الخبرة الضرورية لتحسين قياس وتقويم التحصيل التربوي، وتوفر درجة أعلى من الثقة في تعميم الدراسات التي تفسر العوامل المهمة في التحصيل التربوي، وتزيد من احتمال انتشار أفكار جديدة للبحث التربوي في الدول المشاركة، كما توفر دراسة موضوعية وتقييماً غير متحيز للتجديدات التربوية التي تدخلها الدول المشاركة على أنظمتها التربوية، وتُعين في الوقت نفسه على تلافي جوانب القصور والضعف لهذه البرامج واتخاذ الإجراءات ووضع الخطط المستقبلية التي من شأنها أن تحسن من أداء أنظمتها التربوية (أبو لبدة وآخرون، 2013).

وتهدف الدراسات الدولية للرياضيات والعلوم TIMSS إلى التركيز على السياسات والنظم التعليمية، ودراسة فاعلية المناهج المطبقة وطرق تدريسها، والتطبيق العملي لها، وتقييم التحصيل وتوفير المعلومات لتحسين تعليم وتعلم الرياضيات والعلوم. وتتم هذه الدراسة تحت إشراف الهيئة الدولية لتقييم التحصيل التربوي IEA كل أربع سنوات، ولتحقيق العدالة والموضوعية عند مقارنة بيانات الدول المشاركة، يتم إجراء الاختبار في العلوم والرياضيات في نفس الوقت في كل الدول المشاركة في الدراسة. وقد تم التركيز منذ الدورة الثالثة لاختبارات TIMSSعام 2003 على الأسئلة التي تقيس قدرة الطالب على التحليل والتفسير وحل المشكلات. والهدف الرئيسي من دراسة TIMSS هو مقارنة تحصيل الطلبة في العلوم والرياضيات في أنظمة تربوية متباينة في خلفياتها الثقافية والاقتصادية والاجتماعية بهدف

التعرف على مستوى التحصيل في تلك الأنظمة، وقياس مدى تأثير مجموعة من العوامل ذات العلاقة على مستوى التحصيل، وتطوير الإحصائيات الخاصة بأداء الطلاب في المرحلة التأسيسية في مادتي العلوم والرياضيات، وكذلك تدريب الكوادر الوطنية في مجال إجراء الاختبارات القياسية وجمع البيانات الخاصة بأداء العملية التعليمية (Mullis, Ina V. S, Martin, O. M, Ruddock, G. R, Chrisine, Y. O., Alka, A., Ebru, E. 2007)

ويعد تحقق اللاتغير في القياس مؤشرا على الصدق عند المقارنة بين المجموعات المختلفة، ويتم استخدام التحليل العاملي التوكيدي للمجموعات المتعددة Multiple Groups Confirmatory Factor Analysis (MGCFA) لفحص اللاتغير في القياس بين المجموعات (Vandenberg, and Lance, 2000)، وتستخدم كذلك نماذج المعادلة البنائية Structural equation modeling في فحص اللاتغير في القياس، وعادةً يطلق على اللاتغير في القياس باللاتغير العاملي.(Widaman, Ferrer and Conger. 2010) واللاتغير في القياس أو تكافؤ القياس Measurement invariance or equivalence measurement هو خاصية إحصائية تشير إلى نفس البناء أو التركيب الذي يتم قياسه عبر المجموعات. ومن أنواع اللاتغير في القياس اللاتغير التكويني Configural invariance الذي يشير إلى أن للاختبار نفس البناء العاملي في المجموعتين، أي أن المطابقة تتحقق للمجموعتين بصورة متقاربة. ويُعد اللاتغير التكويني أساساً للاتغير المتري Metric invariance الذي يتحقق عندما تتساوى تشبعات العوامل عبر المجموعات وعدم تحققه مؤشر على أن الفقرات تعمل بصورة مختلفة في المجموعات المختلفة. ويعد اللاتغير المتري أساساً للاتغير العددي Scalar invariance الذي يتحقق عندما تتساوى المقاطع intercept والميول slope والبناء العاملي Byrne, Shavelson, and Muthen, (1989).

ويقترح جورجيوس واونيس (Georgios, and, Ioannis. 2013) عند استخدام أنواع اللاتغير في القياس أن تكون استجابة المجموعات المختلفة لجميع الفقرات بنفس الطريقة، ويعتبر اللاتغير في القياس قضية مهمة للدراسات التي تبحث في الفروق بين المجموعات، وقد أكدت كثير من الدراسات عبر ثقافات مختلفة ضرورة فحص اللاتغير في القياس عند إجراء المقارنات بين المجموعات للتأكد من تفسير هذه المقارنات بموثوقية (Vandenberg -, and Lance. 2000). ويتم فحص اللاتغير في القياس باستخدام مؤشرات حسن الملائمة goodness-of-fit indices وهي مؤشرات إحصائية تساعد على تحديد مدى جودة النموذج المقترح عن طريق مقارنته بنموذج آخر (Gadelrab, 2004). ومن أهم هذه المؤشرات مؤشر كاي سكوير $\chi 2$ ومؤشر جذر متوسط مربع الخطأ التقريبي RMSEA ومؤشر المطابقة المقارن Comparative Fit Index (CFI) ومؤشر الجذر المعياري لمتوسط البواقي SRMR (George, and Irini , 2002; Lattin, Carrol, and Green, 2003). ويبين جدول 1 مؤشرات جودة المطابقة المستخدمة في الدراسة الحالية والحدود المقبولة للمطابقة.

جدول (1) مؤشرات جودة المطابقة المستخدمة في الدراسة الحالية وحدودها المقبولة

اسم المؤشر	الرمز	حدود المطابقة	المصدر
كاي سكوير	$\chi 2$	دال إحصائيا	
مؤشر المطابقة المقارن Comparative Fit Index	CFI	يزيد على 0.80	Hair, Anderson, Tatham, and Black. 1995
جذر متوسط مربع الخطأ التقريبي Root Mean Square Error Approximation	RMSEA	يقل عن 0.08	Hu, and Bentler, 1999
الجذر المعياري لمتوسط البواقي Standardized Root Mean Residual	SRMR	يقل عن 0.09	Hu, and Bentler, 1999

وعند فحص اللاتغير المتري والعددي يتم الاعتماد على مؤشرات الفرق في قيمة (CFI)، والفرق في قيمة ($RMSEA$)، والفرق في قيمة ($SRMR$) بين النموذج المستهدف والنموذج الأساس له. وأشار تشين (Chen, 2007) إلى أنه يفضل أن تقل قيمة (ΔCFI) عن 0.01، وأن تقل قيمة ($\Delta RMSEA$) عن 0.015، وأن تقل قيمة ($\Delta SRMR$) عن 0.03. ويرافق استخدام التحليل العاملي التوكيدي استخدام مؤشر التعديل Modification index وهو الحد الأدنى المُقدر للانخفاض المتوقع في مربع كاي الذي يمكن أن ينتج عندما يتم ترك معلمة Parameter معينة حرة دون قيود Unconstrained. واقترح جورسكوج وسوربوم (Joreskog and Sorbom, 1984) أن مؤشر التعديل يجب أن تكون قيمته خمسة على الأقل لتعديل النموذج المقترح من أجل تحسين الملائمة. لذلك يتم التعديل على النموذج أولا بأول بإحداث ترابطات بين أخطاء القياس في البعد نفسه التي لها أكبر قيمة لمؤشر التعديل حتى يتم ملائمة النموذج للبيانات (Model Fit Hox and Bechger, 1998).

مشكلة الدراسة

تبرز مشكلة هذه الدراسة من خلال الاطلاع على النتائج التي تم الحصول عليها لطلاب الصف الثامن الأساسي الأردنيين في اختبار الرياضيات ضمن دراسة التوجهات الدولية في الرياضيات والعلوم TIMSS 2011 التي تنفذها الرابطة الدولية لتقييم التحصيل التربوي The International Association of Educational Achievement (IEA)، التي أظهرت أفضلية لمدارس وكالة الغوث على مدارس الحكومية ولمدارس المدن على مدارس القرى، ويعد الاختلاف في هذه النتائج المحرك الأساسي وراء إجراء هذه الدراسة للوقوف على اللاتغير في القياس عبر مجموعات المقارنة المختلفة. وتأتي الدراسة الحالية كمحاولة لتفسير النتائج في المقاييس الفرعية (مجالات المحتوى للرياضيات) حسب متغيري موقع المدرسة والجهة المشرفة.

أسئلة الدراسة:

وبناء على ما سبق يمكننا تحديد مشكلة الدراسة في السؤالين التاليين:

1- هل نموذج العوامل الأربعة المقترح لمحتوى الاختبار (الأعداد، الجبر، الهندسة، والإحصاء والاحتمالات) يناسب البيانات بصورة مناسبة؟

2- هل يتحقق اللاتغير في البناء العاملي للاختبار عبر كل من موقع المدرسة والجهة المشرفة كما أشار تشين (Chen, 2007) أن تقل قيمة (ΔCFI) عن 0.01، وأن تقل قيمة ($\Delta RMSEA$) عن 0.015، وأن تقل قيمة ($\Delta SRMR$) عن 0.03. ؟

أهداف الدراسة:

تهدف الدراسة إلى:

1. التحقق من ملاءمة البيانات لمقترح الاختبار وهي (الأعداد، الجبر، الهندسة، والإحصاء والاحتمالات).
2. التحقق من اللاتغير في البناء العاملي للاختبار عبر كل من موقع المدرسة والجهة المشرفة حسب الدلالة الإحصائية التي اشار اليها تشين.
3. معرفة مستوى تحصيل الطلبة ضمن المستوى العالمي.
4. إجراء دراسات تتبعيه للطلبة الذين اختبروا في الصف الرابع ثم اختبروا في الصف الثامن ومقارنة نتائج الاختبارات لنفس الطلبة.

أهمية الدراسة

تأتي أهمية الدراسة الحالية من أهمية خاصية اللاتغير في القياس في مجال الاختبارات والمقاييس التربوية والنفسية، حيث لم يجد الباحث في حدود علمه أي دراسة عربية تناولت قياس اللاتغير في القياس لفقرات اختبار TIMSS، ويؤمل الباحث ان تفيد الدراسة الحالية؛ على النحو الآتي:

1. إطارًا نظريًا، يتضمن مجموعة من الحقائق النظرية والإحصائية المتعلقة بخاصية اللاتغير في القياس، من شأنها أن تسهل وتبسط للباحثين والمهتمين التربويين والنفسيين دراسة خاصية اللاتغير في القياس لمقاييس واختبارات أخرى، وهذا كله يعكس الأهمية النظرية للدراسة.
2. أما الأهمية العملية لهذه الدراسة فتنبع من خلال توفير المزيد من مؤشرات الصدق لاختبار تمس في الرياضيات لدى فئات مختلفة من مجتمع طلبة الصف الثامن الأساسي في الأردن.
3. قد تفيد في تزويد المعلمين بمعلومات عن جوانب القوة وجوانب الضعف في أداء الطلبة، وتساعدهم في تقويم طلبتهم ومتابعتهم وتحديد مسار تقدمهم.
4. قد تفيد في دعم الدراسات التربوية المختلفة بهدف مساعدة راسمي السياسة التربوية وطنيا على تطوير مهارات تدريس هذه المباحث، ومقارنة المستوى الوطني عالميا، وتطوير النظام التربوي والارتقاء بنوعية مخرجاته.
5. قد تفيد في تمكين الطلبة من المفاهيم والمهارات الأساسية في العلوم والرياضيات والتي تنعكس على القرارات التي يتخذونها، وإبراز الجوانب التي يجب التركيز عليها في المناهج الدراسية أثناء ممارسة عملية التعليم.

مصطلحات الدراسة

اللاتغير في القياس: Measurement Invariance

تعرفه بايرن وواتكنز (Byrne, and Watkins, 2003) بأنه ثبات العمليات او الوظائف التى من المفترض 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.

التعريف الاجرائي للاتغير في القياس

هو قياس التحليل العاملي التوكيدي CFA والتحليل العاملي التوكيدي للمجموعات المتعددة MGCFA باستخدام مؤشرات حسن المطابقة CFI, RMSEA, SRMR،)) لفقرات اختبار تمس بحيث يكون لها نفس البناء والتركيب بين المجموعات المختلفة.

الدراسة الدولية للرياضيات والعلوم: اختبار تمس (TIMSS)

هي اختبارات عالمية لتقييم تحصيل الطلبة في الرياضيات والعلوم، وتقييم فاعلية تعليم هاتين المادتين في مدارس الدول المشاركة على مستوى العالم. وتتم هذه الدراسة تحت إشراف الرابطة الدولية لتقييم التحصيل التربوي (IEA) ومقرها في أمستردام / هولندا، حيث تنفذ كل أربع سنوات للصفين الرابع والثامن الأساسيين.

التعريف الاجرائي للدراسة الدولية للرياضيات والعلوم لاختبار TIMSS

هي دراسة دولية تهدف الى مقارنة تحصيل الطلاب التربوية والعوامل الديموغرافية المؤثرة بها في أنظمة تربوية مختلفة في خلفياتها الاجتماعية والاقتصادية والثقافية، ويتم إجراء هذه الدراسة كل اربع سنوات.

حدود الدراسة:

اقتصرت هذه الدراسة على اختبار الرياضيات للصف الثامن ضمن دراسة TIMSS وعلى عينة من طلاب الصف الثامن الأساسي في الأردن الذين تقدموا للاختبار في العام 2011، ويتوقف تعميم نتائج الدراسة على مدى تمثيل العينة لمجتمعها.

2- الإطار النظري والدراسات السابقة

أولاً- الإطار النظري:

الاتّجاهات في الدراسة العالمية للرّياضيات والعلوم Trends of the International Mathematics (TIMSS) and Science Studies.

هو مصطلح مختصر لدراسة أجريت عن التوجهات العالمية في العلوم و الرياضيات وهي أداء اختبارات عالمية لتقييم التوجهات في مدى تحصيل الطلاب في العلوم و الرياضيات ويتم تقييم الطلاب في الصفوف الرابع والصفوف الثامن. وهي دراسة عالمية تهدف إلى التركيز على السياسات والنظم التعليمية، ودراسة فعالية المناهج المطبقة وطرق تدريسها، والتطبيق العملي لها، وتقييم التحصيل وتوفير المعلومات لتحسين تعليم وتعلم الرياضيات والعلوم. وتتم هذه الدراسة تحت إشراف الهيئة الدولية لتقييم التحصيل التربوي International Association for The Evaluation of Educational Achievement (IEA).

ولتحقيق العدالة والموضوعية عند مقارنة بيانات الدول المشاركة، يتم إجراء الاختبار في العلوم والرياضيات في نفس الوقت في كل الدول المشاركة في الدراسة. وقد تم التركيز منذ الدورة الثالثة لاختبارات TIMSS سنة 2003م على الأسئلة التي تقيس قدرة الطالب على التحليل والتفسير وحل المشكلات. ولضمان تحقيق أعلى قدر من الجودة والدقة لاختبارات TIMSS فإنه يتم العمل على أن تتطابق جميع إجراءات الاختبار مع المعايير الموضوعة. وتشمل تلك الإجراءات اختيار عيّنة الطلبة، وترجمة الاختبار، وتصميم كراساته والاستبيانات المصاحبة له وإدارته، وتصحيح الإجابات وتحليل النتائج وإعداد التقارير النهائية، وكذلك تنظيم الدورات التدريبية التي تعقد للقائمين على تنفيذ الإجراءات المذكورة.

ومن خلال تطبيق اختبارات TIMSS يتم جمع مصفوفة بيانات عن البيئة التعليمية والمنزلية التي تؤثر في تعليم وتعلم الرياضيات والعلوم وتنعكس آثارها على معدلات تحصيل الطلبة. إن الدراسة الدولية في الرياضيات والعلوم (TIMSS) هي دراسة تُجرى على المستوى الدولي وتُعنى بدراسة معارف ومهارات وقدرات الطلاب في الرياضيات والعلوم. وهي تستقصي إنجازات الطلاب في هاتين المادتين في مجموعة من دول العالم. وقد صُممت الدراسة لتقيس الفروق بين النظم التعليمية الوطنية وتفسير هذه الفروق وذلك للمساعدة في تطوير وتحسين تعليم وتعلم الرياضيات والعلوم في جميع أنحاء العالم. كانت ولا تزال العلوم و الرياضيات منذ فترة طويلة بؤرة الدراسات من قبل IEA الجمعية الدولية لتقييم التحصيل التربوي التي مقرها في أمستردام بهولندا، فقد أشرفت عبر السنوات الماضية على الكثير من الدراسات التي تتعلق بالعلوم و الرياضيات.

ويعود تاريخ إجراء أول دراسة دولية في مادة الرياضيات للعام 1964 وهي الدراسة التي عُرفت باسم (FIMS)، كما تم تقويم أداء الطلاب في مادة العلوم ضمن ست مواد أخرى في عامي 1970-1971. وظلت كل من الرياضيات والعلوم محل اهتمام وتركيز البحوث التربوية الكبرى التي نُفذت في الأعوام 1980 – 1982 و1983– 1984 على التوالي وفي العام 1983 – 1984 قدمت الدراسات العالمية الثانية للعلوم (SISS) بمشاركة 24 دولة.

وفي عام 1990م قرر الاجتماع العام للجمعية الدولية لتقييم التحصيل التربوي القيام بتقويم أداء الطلاب في مادتي الرياضيات والعلوم معاً على نحو دوري كل أربع سنوات. وشكّل ذلك القرار بداية الدراسات الدولية الموسعة لقياس اتجاهات أداء الطلاب، ليبدأ إجراء الدراسة الدولية في الرياضيات والعلوم المعروفة باسم (TIMSS) والتي تم تنفيذها لأول مرة في عام 1995م. وتكرر إجراؤها بعد ذلك كل أربع سنوات منذ عام 1999م.

أدوات الدراسة المستخدمة في الTIMSS :

تتضمن الدراسة عدة أدوات خاصة بالهدف العام وهي على الشكل التالي:

أولاً: كراسات الاختبارات:

وهي عادة ما تكون على شكل كتيبات متكافئة يتراوح عددها بين 7 - 14 كتيب بحيث يشمل كل كتيب عدد من أسئلة الرياضيات و العلوم 70% من هذه الأسئلة من نوع الاختيار من متعدد و 30% الأسئلة ذات الإجابات القصيرة المعتمدة على استنتاج الحل، و توزع هذه الكتيبات على الطلبة الممتحنين بطريقة عشوائية عن طريق البرمجيات الخاصة بهذه الدراسة التي تحدد اسم الطالب و رقم الكتيب الخاص به.

ثانياً- استبانات الدراسة: وتنقسم إلى 4 استبانات:

1- استبانة الطالب: وهي استبانة توفر معلومات حول الخلفية الأسرية والأكاديمية للطلبة، واتجاهاتهم وطموحاتهم والممارسات الصفية لمعلمي الرياضيات والعلوم من وجهة نظر الطلبة.
2- استبانة معلم الرياضيات: و تتعلق فقراتها بالخلفيات العلمية والأكاديمية والممارسات التدريسية واتجاهات معلمي الرياضيات ليجيب عليها معلم الفصل الذي اختير ضمن العينة.
3- استبانة معلم العلوم: و تتعلق فقراتها بالخلفيات العلمية والأكاديمية والممارسات التدريسية واتجاهات معلمي العلوم ليجيب عليها معلم الفصل الذي اختير ضمن العينة.
4- استبانة المدرسة: و تتعلق فقراتها بمعلومات عن البيئة المدرسية والهيئة التدريسية والطلبة والمنهاج والبرامج الدراسية والامكانيات المادية وبرامج تطوير العاملين وعلاقات المدرسة مع المجتمع. ويجيب عليها مديرو المدارس المتوسطة المشاركة في الدراسة. (المصدر: التقرير الوطني الأردني عن الدراسة الدولية للرياضيات والعلوم للعام 2011).

ثالثاً- برمجيات الدراسة:

1- يقوم مدير بيانات المشروع و فريقه بإعداد استمارة نمذجة المدارس (School Sampling) و المحتوية على بيانات جميع طلبة الصف الثاني المتوسط بأي بلد مشترك بما في ذلك المدارس الحكومية و الخاصة و من ثم ارسالها إلى مركز معالجة البيانات DPC في هامبورج بألمانيا ليتم اعتمادها من قبلهم.
2- تقوم اللجنة العالمية بمعالجة استمارة النمذجة و إدخال بياناتها ضمن قاعدة بيانات تابعة لبرنامج WinW3S Windows within School Sampling Softwareو من ثم إرسال هذه الملفات لمديري بيانات المشروع ليتم تفريغ باقي البيانات.
3- بعد إدخال كافة البيانات في قاعدة بيانات برنامجWinW3S ، يقوم مدير بيانات المشروع و بواسطة البرنامج باستخراج العينة العشوائية للصفوف المنتقاة لكل مدرسة لتطبيق الاختبار.
4- يتم استخراج استمارات المتابعة Tracking Forms استعداد لملأها أثناء الاختبار.

5- بعد تطبيق الاختبار يتم إدخال بيانات الاختبار و الاستبانات في قاعدة بيانات أخرى خاصة ببرنامج الـ Windows) (Data Entry Manager، وذلك بعد ربط البرنامجين ببعضهما البعض.

6- بعد إدخال جميع البيانات يتم إرسالها إلى DPC و Statistics Canada ليتم تحليلها بواسطة برنامج التحليل الإحصائي SPSS او البرنامج الإحصائي AMOS. أسئلة اختبار الرياضيات في الدراسة الدولية للعام 2011

بلغ عدد الأسئلة في اختبار الرياضيات للدراسة الدولية للعام 2011 (217) سؤالاً توزعت على المحتويات الاتية وهي الاعداد والجبر والهندسة والبيانات.

وقد غطت الأسئلة المهارات والأداءات المتوقعة في مجالات المعرفة الآتية:

معرفة الحقائق والإجراءات Knowing التطبيق Applying التعليل Reasoning

(المصدر: التقرير الوطني الأردني عن الدراسة الدولية للرياضيات والعلوم للعام 2011).

ثانياً- الدراسات السابقة:

وقد كانت المقارنات بين المجموعات المختلفة على الاختبارات الوطنية والدولية ودراسة البناء العاملي مدار اهتمام العديد من الدراسات. فقد أجرى يلديرم(Yıldırım, 2006) دراسة هدفت إلى تقييم التكافؤ في فقرات الرياضيات لاختبارات TIMSS 1999 و2003 PISA عبر ثقافات ولغات مختلفة، وذلك لمعرفة ما اذا كان الأداء في الرياضيات يعتمد على أي جوانب ثقافية أو لغوية محددة. وكشفت النتائج أن هناك صعوبات في التكيف مع اختبارات TIMSS/ PISA. ومع ذلك فإنه لا يزال من الممكن تحديد الفقرات في الاختبارات الفرعية التي تشكّل الأساس للمقارنة العادلة بين الثقافات.

وأجرى ديمتروف (Dimitrov, 2013) دراسة لفحص الأحادية والبناء العاملي لبيانات اختبار القدرات اللفظية العامة (GAT) في السعودية باستخدام التحليل العاملي التوكيدي (CFA). وأظهرت النتائج حسن مطابقة بيانات اختبار (GAT) لنموذج بثلاثة عوامل مترابطة تمثل الجوانب الكامنة للكفاءة اللفظية.

كما أجرى ايونيس (Ioannis, 2014b) دراسة بهدف فحص البناء العاملي لاختبار القدرات العامة (GAT) في السعودية حسب نوع المدرسة (خاصة، حكومية) وفحص اللاتغير في القياس بأنواعه الثلاث (التكويني، المتري، العددي). وأظهرت النتائج وجود لاتغير تكويني ولاتغير متري ولاتغير عددي في النموذج العاملي من الرتبة الثانية وفي النموذج العاملي الثنائي. وأظهرت النتائج أن أوساط الطلاب في المدارس الخاصة أعلى منها في المدارس الحكومية في كل من القدرات اللفظية والعددية وفي المجالات الفرعية لها باستثناء مجالات الحساب والهندسة.

وفي الأردن تم مقارنة أداء الطلبة على اختبار تمس في الرياضيات بحسب جنس المدرسة وموقعها والجهة المشرفة وغيرها من المتغيرات الديموغرافية. ولكونه لم يتم التحقق من البناء العاملي للاختبار ولم يتم فحص اللاتغير في القياس فقد تكون المقارنات غير صادقة وليست ذات معنى. لذلك جاءت هذه الدراسة بهدف توفير مؤشرات على صدق تلك المقارنات. (التقرير الوطني الأردني عن الدراسة الدولية للرياضيات والعلوم لعام 2011)

3- منهجية الدراسة وإجراءاتها

منهج الدراسة: استخدم الباحث المنهج التحليلي الوصفي في إعداد هذه الدراسة، والذي يتمثل في تحليل نتائج اختبار TIMSS 2011 في الرياضيات لطلبة الصف الثامن الأساسي للعام الدراسي 2010 -2011، وذلك بهدف قياس معرفتهم لمادة الرياضيات.

مجتمع الدراسة وعينتها:

يتكون مجتمع الدراسة الحالية من 7694 طالباً وطالبة من طلبة الصف الثامن الأساسي في الأردن الذين شاركوا في دراسة TIMSS في دورة عام 2011. وقد اختير هؤلاء الطلبة من قبل مركز الإحصاء الكندي وخبراء العينات في الجمعية الدولية IEA بطريقة طبقية عنقودية ذات مرحلتين. في المرحلة الأولى تم اختيار 230 مدرسة من المدارس الأردنية التي تشتمل على ثمانية طلاب على الأقل في الصف الثامن الأساسي بطريقة عشوائية متناسبة الحجم مع الفئات المختلفة للمدارس الأردنية التي تحددها مجموعة من المتغيرات، وفي المرحلة الثانية تم اختيار شعبة واحدة من شعب الصف الثامن من كل مدرسة مختارة. وتكونت عينة الدراسة من 547 طالباً وطالبة هم الذين أجابوا على فقرات الكراستين الثامنة والتاسعة من كراسات اختبار TIMSS. وطريقة تطبيق اختبار TIMSS تضمن تمثيل عينة الطلبة الذين يخضعون لأي كراسة من كراسات الاختبار لمجتمع الدراسة.

أداة الدراسة:

تتكون الأداة الرئيسة للدراسة من اختبار TIMSS في الرياضيات للصف الثامن. ويقوم الاختبار على إطار شامل تم تطويره من قبل خبراء دوليين بالتعاون مع خبراء من الدول المشاركة. ويدور هذا الإطار حول بعدين: البعد الأول يتعلق بالمحتوى ويحدد المجالات والموضوعات الرياضية التي يجب تقييمها، ويتعلق البعد الثاني بالمجالات المعرفية ويحدد العمليات العقلية المتوقعة من الطلبة عندما يتعاملون مع محتوى رياضي. ويشتمل الاختبار بصورته الكلية على 217 فقرة، منها 51% من نوع الاختيار من متعدد، و49% من نوع الاستجابة الحرة. ويصنف محتوى الفقرات على بعد المحتوى في أربعة مجالات هي الأعداد، والجبر، والهندسة، والبيانات والفرص بأوزان نسبية 30%، 30%، 20%، و20% على الترتيب. وتقع العلامات للاختبار على تدريج مداه من 0 إلى 1000 (العلامات في العادة تتراوح بين 300 و700) وبوسط حسابي 500 وانحراف معياري 100. وتستخدم نقطة منتصف التدريج (500) كنقطة مرجعية تبقى ثابتة في كل دورات الدراسة. وتنظم الفقرات في سلسلة من 14 تجمعاً يضم كل تجمع منها ما بين 12 إلى 18 فقرة موزعة على مجالات المحتوى بأوزان نسبية مماثلة للأوزان النسبية المعتمدة في الاختبار الكلي. وللتخفيف على الطلبة المشاركين يتم تكوين 14 كراسة للاختبار بحيث تتضمن الكراسة الواحدة تجمعين من الفقرات. ويكون هناك تجمع مشترك بين كل كراسة والتي تليها. ويخضع كل طالب لكراسة واحدة، أي أنه يتقدم فقط لمجموعة جزئية ممثلة من الفقرات. وتستخدم نظرية الاستجابة للفقرة (IRT) لتقدير الأداء له على كل الفقرات. ولغايات الدراسة الحالية، قام الباحث بتجميع استجابات كراستين (Two Booklet) من كراسات مادة الرياضيات؛ الكراسة الثامنة والكراسة التاسعة، وقد تضمنتا استجابة 547 طالبًا وطالبةً على 34 فقرة.

المعالجة الإحصائية:

بعد الحصول على ملف البيانات من المركز الوطني لتنمية الموارد البشرية، تم التحقق من افتراض التوزيع الطبيعي متعدد المتغيرات (Multivariate Normality) لاستجابات أفراد عينة الدراسة على فقرات الكراستين. وقد تم التحقق من افتراض التوزيع الطبيعي لكل فقرة باستخدام مؤشري الالتواء (Skewness) والتفلطح (kurtosis). وفي هذا السياق يشير وست وفنش وكوران (West, Finch, and Curran, 1995) إلى أنه يفضل أن لا تزيد القيمة المطلقة لمعامل الالتواء على 2 والقيمة المطلقة لمعامل التفلطح على 7. وقد تبين أن هناك 7 فقرات تزيد قيم الالتواء لها على 2 وقيم التفلطح على 7. وهذا يعني أن هذه الفقرات لا تحقق افتراض التوزيع الطبيعي. وقد تم استبعاد هذه الفقرات من التحليل. وبذلك تم تحليل الاستجابات على 27 فقرة، موزعة على أربعة أبعاد هي: بُعد الأعداد ويضم 10 فقرات، وبُعد الجبر ويضم 9 فقرات، وبُعد الهندسة ويضم 4 فقرات، وبُعد البيانات ويضم 4 فقرات.

التوزيع الطبيعي:

تم التحقق من التوزيع الطبيعي متعدد المتغيرات لاستجابات أفراد عينة الدراسة على فقرات الاختبار المكون من 27 فقرة، باستخدام اختبار ماردیا (Mardia's Test) (Smith, 1987)، وتبين قيمة هذا الاختبار (Mardia's Test) عدم تحقق افتراض التوزيع الطبيعي متعدد المتغيرات. ولمعالجة ذلك إحصائيا تم استخدام طريقة (BOOTSTRAP) المتوفرة في برنامج (AMOS) المستخدم في تحليل البيانات.

وللإجابة عن أسئلة الدراسة؛ تم استخدام أسلوب التحليل العاملي التوكيدي CFA وأسلوب التحليل العاملي التوكيدي متعدد المجموعةMGCFA. وللحكم على مدى ملائمة البناء العاملي للبيانات تم استخدام المؤشرات الإحصائية لجودة الملائمة في النماذج الخطية البنائية SEM وهي: إحصائي مربع كاي سكوير (χ^2)، ومؤشر حسن المقارنة CFI، وجذر متوسط مربع الخطأ التقريبي RMSEA، والجذر المعياري لمتوسط البواقي SRMR.

4- عرض نتائج الدراسة ومناقشتها

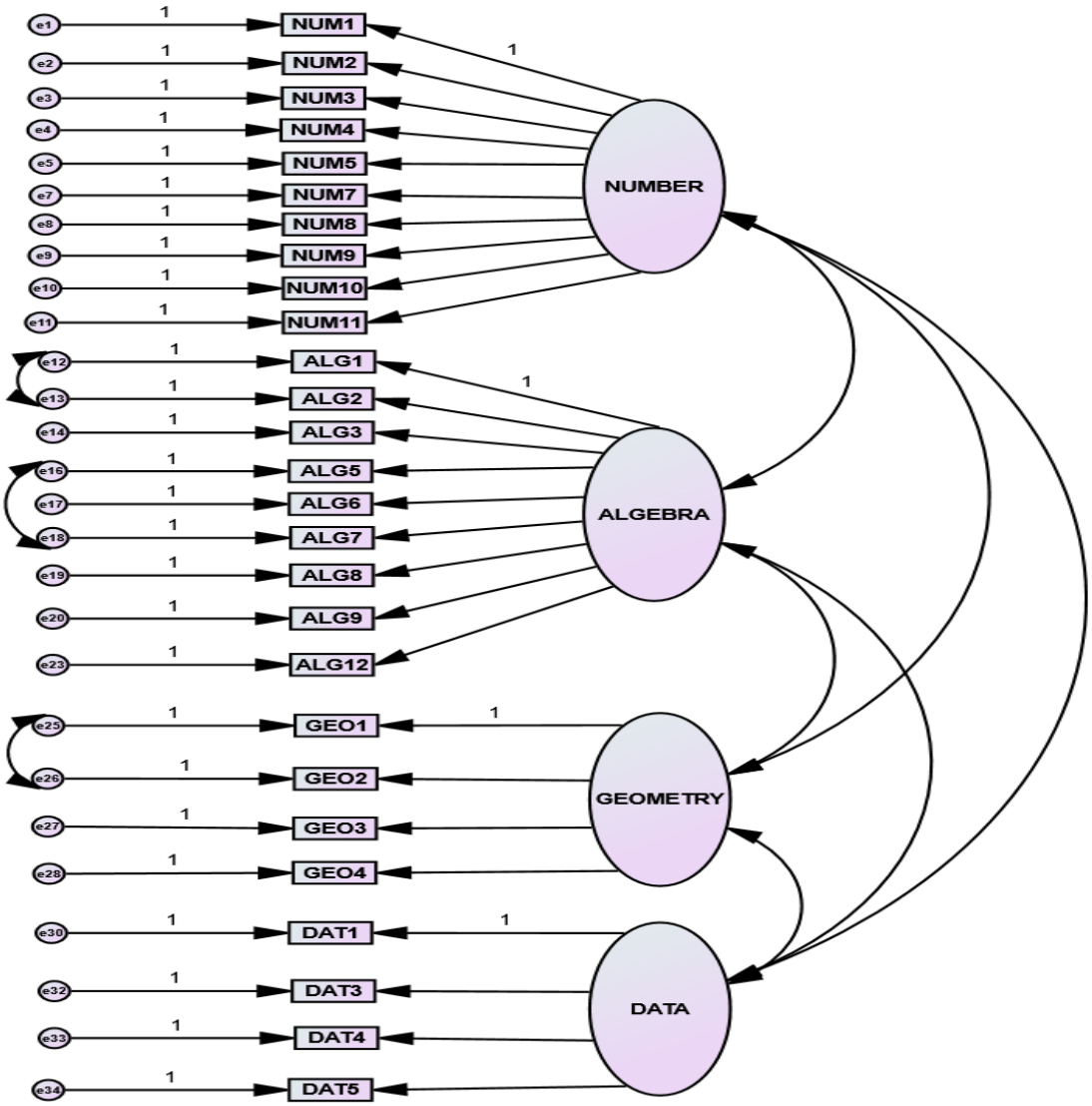
- **النتائج المتعلقة بإجابة السؤال الأول** الذي ينص على: "هل نموذج العوامل الاربعة المقترح لمحتوى الاختبار (الأعداد، الجبر، الهندسة، الإحصاءات والاحتمالات) يناسب البيانات بصورة مناسبة؟"

للتحقق من مناسبة النموذج للبيانات تم استخدام التحليل العاملي التوكيدي المبينة نتائجه في الجدول2.

جدول (2) نتائج التحليل العاملي التوكيدي لفقرات الاختبار (27 فقرة) قبل إجراء التعديلات

المؤشر الإحصائي	χ^2	*Df*	χ^2/df	CFI	RMSEA	SRMR	*sig*
قيمة الاختبار	*468.167	318	1.472	0.934	0.029	0.0383	0.000

يُلاحظ من جدول رقم 2 وجود دلالة إحصائية لقيمة (χ^2)، وكانت جميع قيم مؤشرات (CFI, RMSEA, SRMR) ضمن المدى المقبول لجودة الملائمة. وبالنظر إلى مؤشرات التعديل **Modification Indices** يتبين أنه يمكن تحسين جودة الملائمة من خلال أخذ الارتباطات بين الأخطاء في النموذج المقترح بين (e12) و(e13)، وبين (e16) و(e18)، وبين (e25) و(e26) في الاعتبار. ويبين الشكل **رقم 1** النموذج العاملي المعدل لاختبار الرياضيات.



شكل (1)النموذج العاملي المعدل لاختبار الرياضيات (27 فقرة)

وللتحقق من مناسبة النموذج المعدل تم استخدام التحليل العاملي التوكيدي على الفقرات، ويبين جدول رقم 3 نتائج ملائمة النموذج بعد إجراء التعديلات الممكنة السابقة.

جدول (3)نتائج ملائمة النموذج بعد إجراء التعديلات الممكنة

المؤشر الإحصائي	χ^2	*Df*	χ^2/df	CFI	RMSEA	SRMR	*sig*
قيمة الاختبار	*444.483	315	1.411	0.943	0.027	0.0374	0.000

يُلاحظ من الجدول رقم 3 أن جميع المؤشرات الإحصائية تدل على ملائمة النموذج المستخدم للبيانات، حيث أظهرت النتائج وجود دلالة إحصائية لقيمة (χ^2)، وكانت جميع قيم مؤشرات (CFI, RMSEA, SRMR) ضمن المدى المقبول لجودة الملائمة؛ وكل هذه القيم تدل على أن التركيب العاملي المقترح للاختبار مناسب للبيانات.

- **النتائج المتعلقة بإجابة السؤال الثاني** الذي ينص على: "هل يختلف البناء العاملي المقترح لاختبار الرياضيات باختلاف موقع المدرسة والجهة المشرفة؟"

للإجابة عن السؤال الثاني استخدم التحليل العاملي التوكيدي متعدد المجموعة MGCFA لاختبار اللاتغير في البناء العاملي الرباعي المُقترح لاختبار TIMSS عبر موقع المدرسة والجهة المشرفة، وقد تم اختبار اللاتغير التكويني واللاتغير المتري واللاتغير العددي لكل من موقع المدرسة والجهة المشرفة، وكانت النتائج على النحو الآتي:

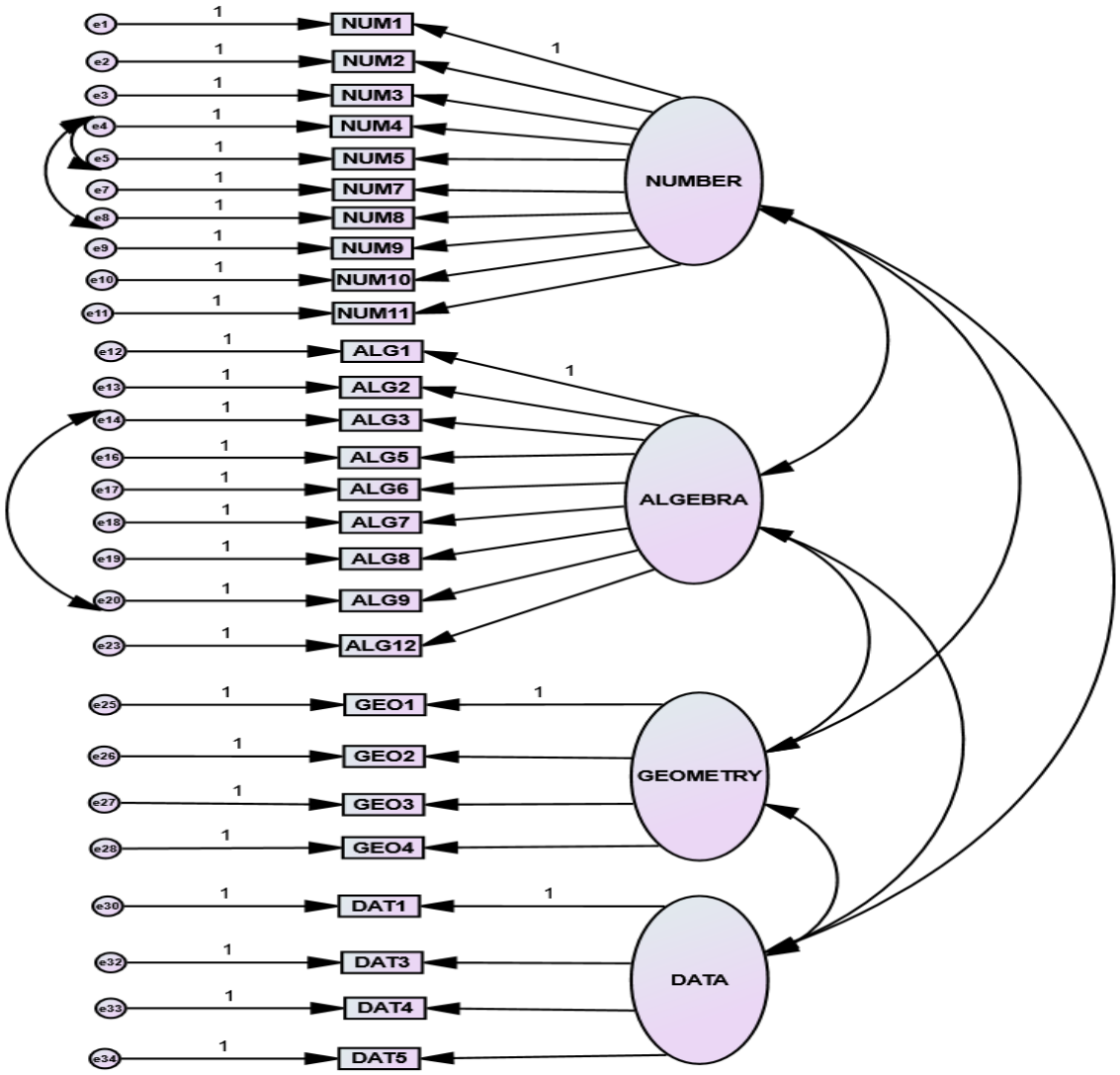
أولاً- اللاتغير التكويني باختلاف موقع المدرسة (مدينة، ريف):

تم أولاً التحقق من درجة ملائمة النموذج لكل من موقع المدرسة (مدينة، ريف) بشكل منفصل. ويكون النموذج المقترح ملائم عند وجود دلالة إحصائية لمؤشر (χ^2)، وقيمة مؤشر (CFI) أكبر من 0.80، وقيمة مؤشر (RMSEA) أقل من 0.08، وقيمة مؤشر (SRMR) أقل من 0.09. ويبين جدول **رقم 4** نتائج ملائمة النموذج المقترح وفقاً لمتغير موقع المدرسة (مدينة، ريف).

جدول (4)نتائج ملائمة النموذج المقترح وفقاً لمتغير موقع المدرسة (مدينة، ريف)

المؤشر الإحصائي		χ^2	Df	χ^2/df	CFI	RMSEA	SRMR	sig
موقع المدرسة	مدينة	*427.902	318	1.346	0.937	0.028	0.0420	0.000
	ريف	*421.824	318	1.326	0.829	0.056	0.0410	0.000

يُلاحظ من الجدول رقم 4 وجود دلالة إحصائية لقيمة (χ^2) لكل من متغير موقع المدرسة (مدينة، ريف) وهذا يؤكد ملائمة النموذج، كما أن قيم مؤشرات (CFI, RMSEA, SRMR) تؤكد ملائمة النموذج لكل من مدارس المدينة ومدارس الريف. وبالنظر لقيم مؤشرات التعديل (Modification Indices) يُلاحظ وجود بعض الارتباطات الواجب أخذها بالاعتبار في البناء العاملي المقترح كما في الشكل **رقم 2**.



شكل (2) النموذج العاملي المعدل لاختبار TIMSS في الرياضيات وفق متغير موقع المدرسة

ويبين الجدول رقم 5 نتائج ملائمة النموذج المعدل بحسب موقع المدرسة (مدينة، ريف) بعد أخذ الارتباطات بالاعتبار من الشكل رقم 2.

جدول (5) نتائج ملائمة النموذج العاملي المعدل وفق موقع المدرسة (مدينة، ريف)

المؤشر الإحصائي		χ^2	*Df*	χ^2/df	CFI	RMSEA	SRMR	sig
موقع المدرسة	مدينة	420.209*	315	1.334	0.939	0.027	0.0420	**0.000**
	ريف	**400.827***	**315**	**1.272**	**0.859**	**0.051**	**0.0748**	**0.000**

يُلاحظ من الجدول رقم 5 وجود دلالة إحصائية لقيمة (χ^2) لكل من متغير موقع المدرسة (مدينة، ريف) وهذا يؤكد ملائمة النموذج المعدل لكلٍ منها، وأن قيم مؤشرات(CFI, RMSEA, SRMR) تؤكد ملائمة النموذج لكل من مدارس المدينة ومدارس الريف. وبعد التأكد من ملائمة النموذج المقترح لفئتي متغير موقع المدرسة (مدينة، ريف)، أُعيد استخدام التحليل العاملي التوكيدي متعدد المجموعة على البيانات كاملة لاختبار اللاتغير التكويني، الذي يتم اجراؤه بدون وجود أية قيود (No constraints). ويبين الجدول رقم 6 نتائج هذا التحليل.

جدول (6) نتائج اختبار اللاتغير التكويني وفقاً لمتغير موقع المدرسة (مدينة، ريف)

المؤشر الإحصائي	χ^2	Df	χ^2/df	CFI	RMSEA	SRMR	sig
موقع المدرسة	*822.858	630	1.306	0.918	0.024	0.0408	0.000

يُلاحظ من الجدول رقم 6 أن جميع المؤشرات الإحصائية تدل على ملائمة النموذج المستخدم للبيانات، حيث أظهرت النتائج وجود دلالة إحصائية لقيمة (χ^2)، وكانت جميع قيم مؤشرات (CFI, RMSEA, SRMR) ضمن المدى المقبول لجودة الملائمة؛ وكل هذه القيم تدل على أن التركيب العاملي للاختبار مناسب للبيانات، وأنه يوجد لاتغير تكويني وفقاً لمتغير موقع المدرسة (مدينة، ريف).

ثانياً- اللاتغير المتري باختلاف موقع المدرسة (مدينة، ريف):

تم ذلك من خلال استخدام (ΔCFI, $\Delta SRMR$, $\Delta RMSEA$) التي تبين الفرق في قيم (CFI, RMSEA, SRMR) بين نموذج (M1) الذي يمثل اللاتغير التكويني ونموذج (M2) الذي يمثل اللاتغير المتري في ظل وضع قيود (constraints) على تشبعات العوامل تتمثل بتساوي التشبعات في المجموعتين (المدينة، الريف). ويبين جدول 7 النتائج الخاصة بالنموذج M2

جدول (7) نتائج نموذج اللاتغير المتري M2

المؤشر الإحصائي	χ^2	Df	χ^2/df	CFI	RMSEA	SRMR	sig
موقع المدرسة	861.321	657	1.311	0.913	0.024	0.0432	0.000

يُلاحظ من الجدول رقم 7 أن النموذج المتري يناسب البيانات. ويبين الجدول التالي نتائج اختبار اللاتغير المتري وفق متغير موقع المدرسة (مدينة، ريف).

جدول (8) نتائج اختبار اللاتغير المتري وفق متغير موقع المدرسة (مدينة، ريف)

النموذج (اللاتغير المتري) Model	الفرق في				
	درجة الحرية	كاي تربيع	CFI	RMSEA	SRMR
اللاتغير المتري في كل الفقرات	27	38.464	-0.005	0.000	0.024

يبين الجدول رقم 8 أن قيم الفرق في مؤشرات (ΔCFI, $\Delta SRMR$, $\Delta RMSEA$) بين النموذج التكويني غير المشروط M1 والنموذج المتري المشروط M2 تدل على وجود لا تغير متري في الاختبار.

ثالثاً: اللاتغير العددي باختلاف موقع المدرسة (مدينة، ريف):

تم ذلك من خلال استخدام (ΔCFI, $\Delta SRMR$, $\Delta RMSEA$) التي تبين الفرق في قيم (CFI, RMSEA, SRMR) بين نموذج (M2) الذي يمثل اللاتغير المتري ونموذج (M3) الذي يمثل اللاتغير العددي في ظل وضع قيود (constraints) على التقاطعات للعوامل (INTERCEPTS) في المجموعتين (المدينة، الريف). ويبين الجدول رقم 9 النتائج الخاصة بالنموذج M3

جدول (9) نتائج نموذج اللاتغير العددي M3

المؤشر الإحصائي	χ^2	Df	χ^2/df	CFI	RMSEA	SRMR	sig
موقع المدرسة	910.730	684	1.331	0.903	0.025	0.0430	0.000

يُلاحظ من الجدول رقم 9 أن النموذج العددي يناسب البيانات، ويبين الجدول التالي نتائج اختبار اللاتغير العددي وفق متغير موقع المدرسة (مدينة، ريف).

جدول (10)نتائج اختبار اللاتغير العددي وفق متغير موقع المدرسة (مدينة، ريف)

النموذج (اللاتغير المتري) Model	الفرق في				
	درجة الحرية	كاي تربيع	CFI	RMSEA	SRMR
اللاتغير المتري في كل الفقرات	27	49.409	01-0.0	10.00	2000.0-

ويبين الجدول رقم 10 أن قيم الفرق في مؤشرات (ΔCFI, $\Delta SRMR$,$\Delta RMSEA$) بين النموذج المتري M2 والنموذج العددي M3 تدل على وجود لا تغير عددي في الاختبار.

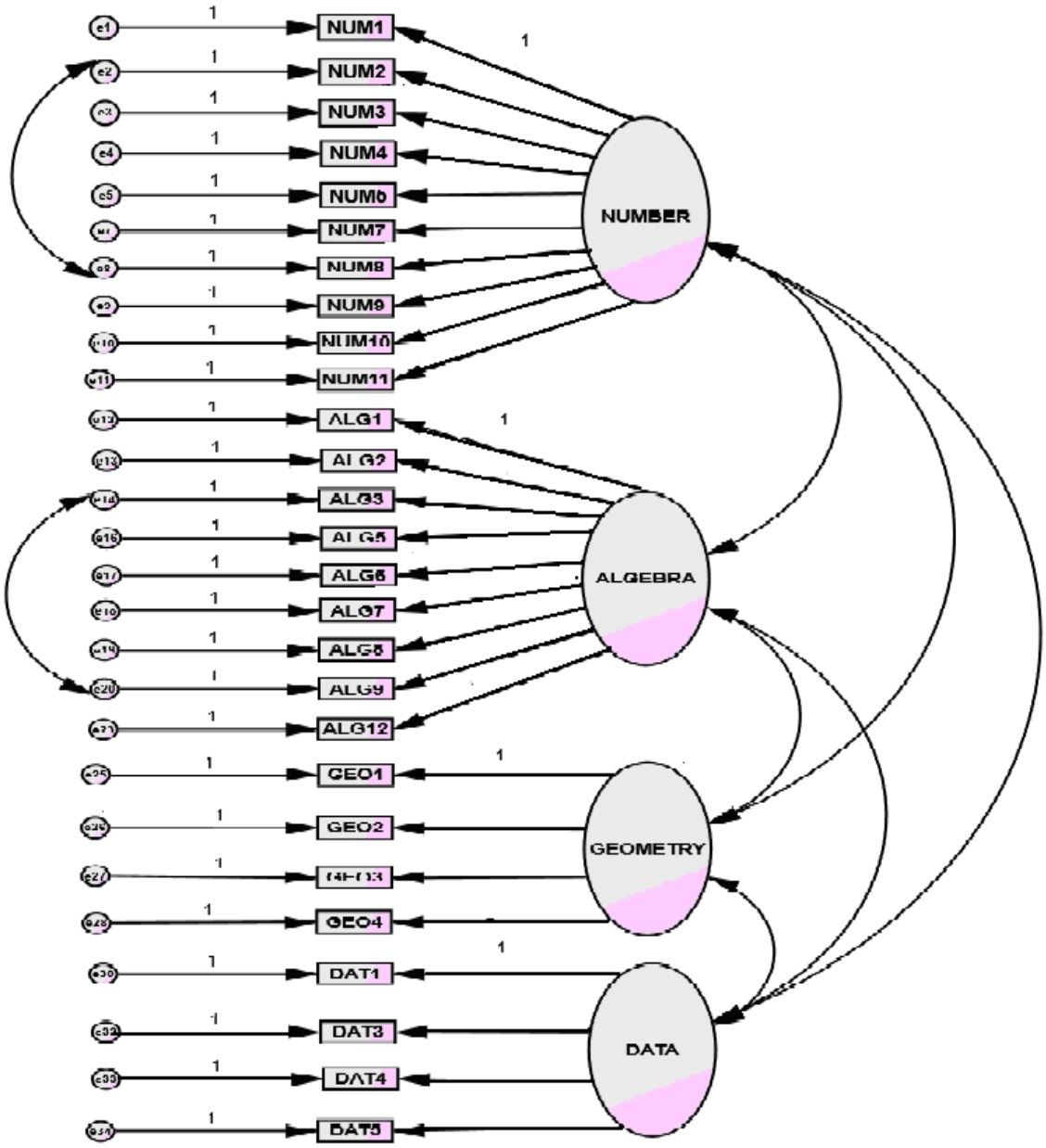
رابعاً- اللاتغير التكويني باختلاف الجهة المشرفة (وزارة التربية، وكالة الغوث):

تم أولاً التحقق من درجة ملائمة النموذج لكل من متغير (وزارة التربية والتعليم، وكالة الغوث) بشكل منفصل. ويكون النموذج المقترح ملائم عند وجود دلالة إحصائية لمؤشر (χ^2)، وقيمة مؤشر (CFI) أكبر من 0.80، وقيمة مؤشر (RMSEA) أقل من 0.08، وقيمة مؤشر (SRMR) أقل من 0.09. ويبين الجدول **رقم 11** نتائج ملائمة النموذج المقترح وفقاً لمتغير الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث).

جدول (11) نتائج ملاءمة النموذج المقترح وفقاً لمتغير الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث)

المؤشر الإحصائي		χ^2	Df	χ^2/df	CFI	RMSEA	SRMR	sig
الجهة المشرفة	وزارة التربية	6537.54	318	1.439	0.890	0.044	0.0557	0.000
	وكالة الغوث	382.103	318	1.202	0.910	0.033	0.0604	0.000

يُلاحظ من جدول رقم 11 وجود دلالة إحصائية لقيمة (χ^2) لكل من (وزارة التربية والتعليم، وكالة الغوث) وهذا يؤكد ملائمة النموذج، كما أن قيم مؤشرات (CFI, RMSEA, SRMR) تؤكد ملائمة النموذج لكل من مدارس وزارة التربية والتعليم ومدارس وكالة الغوث. وبالنظر لقيم مؤشرات التعديل (Modification Indices) يُلاحظ وجود بعض الارتباطات الواجب أخذها بالاعتبار في البناء العاملي المقترح كما في الشكل **رقم 3.**



شكل (3)النموذج العاملي المعدل لاختبار TIMSS في الرياضيات وفق متغير الجهة المشرفة

ويبين الجدول رقم 12 نتائج ملائمة النموذج المعدل بحسب الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث) بعد أخذ تلك الارتباطات بالاعتبار.

جدول (12)نتائج ملائمة النموذج العاملي المعدل وفق الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث)

المؤشر الإحصائي		χ^2	*Df*	χ^2/df	CFI	RMSEA	SRMR	sig
الجهة المشرفة	وزارة التربية والتعليم	454.109	316	1.437	0.891	0.043	0.0555	0.000
	وكالة الغوث	379.535	316	1.201	0.911	0.033	0.0603	0.000

يُلاحظ من الجدول رقم 12 وجود دلالة إحصائية لقيمة (χ^2) لكل من (وزارة التربية والتعليم، وكالة الغوث) وهذا يؤكد ملائمة النموذج المعدل لكلٍ منهم، وأن قيم مؤشرات(CFI, RMSEA, SRMR) تؤكد ملائمة النموذج لكل من مدارس وزارة التربية والتعليم ومدارس وكالة الغوث. وبعد التأكد من ملائمة النموذج المقترح لفئتي متغير

الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث)، أُعيد استخدام التحليل العاملي التوكيدي متعدد المجموعة على البيانات كاملة لاختبار اللاتغير التكويني، الذي يتم إجراؤه بدون وجود أية قيود (No constraints). ويبين الجدول رقم 13 نتائج هذا التحليل.

جدول (13)نتائج اختبار اللاتغير التكويني وفقاً لمتغير الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث)

المؤشر الإحصائي	χ^2	*Df*	χ^2/df	CFI	RMSEA	SRMR	sig
الجهة المشرفة	**833.660**	**632**	**1.319**	**8980.**	**0.028**	**0.0351**	**0.000**

يُلاحظ من الجدول رقم 13 أن جميع المؤشرات الإحصائية تدل على ملائمة النموذج المستخدم للبيانات، حيث أظهرت النتائج وجود دلالة إحصائية لقيمة (χ^2)، وكانت جميع قيم مؤشرات (CFI, RMSEA, SRMR) ضمن المدى المقبول لجودة الملائمة؛ وكل هذه القيم تدل على أن التركيب العاملي للاختبار مناسب للبيانات، وأنه يوجد لاتغير تكويني وفقاً لمتغير الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث).

خامساً: اللاتغير المتري باختلاف الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث):

تم ذلك من خلال استخدام (ΔCFI, $\Delta SRMR$,$\Delta RMSEA$) التي تبين الفرق في قيم (CFI, RMSEA, SRMR) بين نموذج (M1) الذي يمثل اللاتغير التكويني ونموذج (M2) الذي يمثل اللاتغير المتري في ظل وضع قيود (constraints) على تشبعات العوامل تتمثل بتساوي التشبعات في المجموعتين (ذكور، إناث). ويبين الجدول رقم 14 النتائج الخاصة بالنموذج M2

جدول (14)نتائج نموذج اللاتغير المتري M2

المؤشر الإحصائي	χ^2	*Df*	χ^2/df	CFI	RMSEA	SRMR	sig
الجهة المشرفة	**890.903**	**659**	**1.352**	**0.883**	**0.029**	**0.0620**	**0.000**

يُلاحظ من الجدول رقم 14 أن النموذج المتري يناسب البيانات، ويبين الجدول التالي نتائج اختبار اللاتغير المتري وفق متغير الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث).

جدول (15)نتائج اختبار اللاتغير المتري وفق متغير الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث)

النموذج (اللاتغير المتري) Model	الفرق في				
	درجة الحرية	كاي تربيع	CFI	RMSEA	SRMR
اللاتغير المتري في كل الفقرات	27	57.243	-0.015	0.001	0.0065

يبين الجدول رقم 15 أن قيم الفرق في مؤشرات (ΔCFI, $\Delta SRMR$,$\Delta RMSEA$) بين النموذج التكويني غير المشروط M1 والنموذج المتري المشروط M2 تدل على وجود لا تغير متري في الاختبار.

سادساً: اللاتغير العددي باختلاف الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث):

تم ذلك من خلال استخدام (ΔCFI, $\Delta SRMR$,$\Delta RMSEA$) التي تبين الفرق في قيم (CFI, RMSEA, SRMR) بين نموذج (M2) الذي يمثل اللاتغير المتري ونموذج (M3) الذي يمثل اللاتغير العددي في ظل وضع قيود (constraints) على التقاطعات للعوامل (INTERCEPTS) في المجموعتين (ذكور، إناث). ويبين الجدول رقم 16 النتائج الخاصة بالنموذج M3.

جدول(16) نتائج نموذج اللاتغير العددي M3

المؤشر الإحصائي	χ^2	Df	χ^2/df	CFI	RMSEA	SRMR	sig
الجهة المشرفة	993.653	686	1.448	0.845	0.033	0.0634	0.000

يُلاحظ من الجدول رقم 16 أن النموذج العددي يناسب البيانات، ويبين الجدول التالي نتائج اختبار اللاتغير العددي وفق متغير الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث).

جدول (17)نتائج اختبار اللاتغير العددي وفق متغير الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث)

النموذج (اللاتغير المتري) Model	الفرق في				
	درجة الحرية	كاي تربيع	CFI	RMSEA	SRMR
اللاتغير المتري في كل الفقرات	27	102.750	38-0.0	40.00	0140.0-

يبين جدول رقم 17 أن قيم الفرق في مؤشرات (ΔCFI, $\Delta SRMR$,$\Delta RMSEA$) بين النموذج المتري M2 والنموذج العددي M3 تدل على وجود لا تغير عددي في الاختبار.

مناقشة النتائج

أظهرت نتائج الدراسة ملائمة نموذج العوامل الأربعة (الأعداد، الجبر، الهندسة، الإحصاءات والاحتمالات) المقترح لمحتوى اختبار تمس في الرياضيات للصف الثامن الأساسي للبيانات التي تم الحصول عليها في عينة أردنية، وتدل هذه النتيجة على أن هذه المكونات والأبعاد الأربعة تمثل وبشكل واضح محتوى اختبار الرياضيات، وقد يعزى ذلك إلى كون إعداد الاختبار تم وفق أُسس علمية جيدة، قائمة على تحليل المحتوى الرياضي الذي ينبغي معرفته في مثل هذه الفئة العمرية من الطلبة وقد تم تحديد معايير المحتوى والتي تعكس مكونات الاختبار وأبعاده بإشراف هيئة دولية، تضم عدداً كبيراً من الخبراء والمختصين في مجال الرياضيات، وقد تم إعداد الاختبار وتصميم كراساته وترجمته وتطبيقه وتصحيحه وفق ظروف معيارية جعلت منه اختباراً عالمياً موثوقاً. كما أن هذا الاختبار يتم إعداده وفقا لمعايير NCTM في الرياضيات والتي تقسم المعايير إلى مجموعتين؛ تمثل المجموعة الأولى معايير المحتوى، وتتضمن كلا من الأعداد والعمليات والجبر، وفهم الأنماط والعلاقات والاقترانات، والهندسة، والقياس، وتحليل البيانات والاحتمالات. بينما تمثل المجموعة الثانية معايير العمليات: وتشمل المعرفة والتطبيق والتبرير. وهذا يدلل على صدق الاختبار في قياسه لتحصيل الطلبة في الرياضيات. وتتفق هذه النتيجة مع نتيجة دراسة عفانة ونبهان (2004) التي أظهرت جودة مقبولة لاختبار TIMSS في الكشف عن تحصيل الطلبة في الرياضيات، كما اتفقت هذه النتيجة مع نتيجة دراسة كل من ديمتر وديمتروف2013 وايونيس 2014a, 2014b وايونيس 2015 التي دلت على مناسبة التحليل العاملي التوكيدي في الكشف عن ملائمة النموذج العاملي المقترح لبعض الاختبار للبيانات الخاصة بها. ومما يؤكد ملائمة نموذج العوامل الأربعة المقترح لمحتوى اختبار الرياضيات للبيانات؛ وجود ارتباط قوي ودال إحصائيًا بين أبعاده الأربعة المفترضة، حيث تراوحت معاملات الارتباط البينية بينها بين (0.462) و(0.625) كما هو مبين في الجدول رقم 18.

جدول (18)معاملات الارتباط البينية بين الابعاد الاربعة لمحتوى اختبار الرياضيات

الابعاد	الأعداد	الجبر	الهندسة
الجبر	0.625**		
الهندسة	0.466**	0.465**	
البيانات	0.475**	0.501**	0.462**

**دالة إحصائيًا عند مستوى الدلالة الإحصائية ($\alpha = 0.01$)

وأظهرت النتائج وجود لا تغير تكويني باختلاف موقع المدرسة (ريف، مدينة)، وكانت قيم مربع كاي سكوير$\chi 2$ دالة إحصائيا وقيمتها 822.858 ومؤشر حسن المقارنةCFI دالة إحصائيا وقيمتها 0.918 وهي أكبر من 0.80 ومربع متوسط جذر الخطأ التقريبيRMSEA دالة إحصائيا وقيمتها 0.024 وهي أقل من 0.08 ومؤشر جذر متوسط مربعات البواقي المعيارية SRMR دالة إحصائيا وقيمتها 0.0408 وهي أقل من 0.09. وأظهرت النتائج وجود لا تغير متري باختلاف موقع المدرسة (ريف، مدينة)، وكانت قيم مربع كاي سكوير$\chi 2$ دالة إحصائيا وقيمتها 861.321 ومؤشر حسن المقارنةCFI دالة إحصائيا وقيمتها 0.913 وهي أكبر من 0.80 ومربع متوسط جذر الخطأ التقريبيRMSEA دالة إحصائيا وقيمتها 0.024 وهي أقل من 0.08 ومؤشر جذر متوسط مربعات البواقي المعيارية SRMR دالة إحصائيا وقيمتها 0.0432 وهي أقل من 0.09. كما وأظهرت النتائج وجود لا تغير عددي باختلاف موقع المدرسة (ريف، مدينة)، وكانت قيم مربع كاي سكوير$\chi 2$ دالة إحصائيا وقيمتها 910.730 ومؤشر حسن المقارنةCFI دالة إحصائيا وقيمتها 0.903 وهي أكبر من 0.80 ومربع متوسط جذر الخطأ التقريبيRMSEA دالة إحصائيا وقيمتها 0.025 وهي أقل من 0.08 ومؤشر جذر متوسط مربعات البواقي المعيارية SRMR دالة إحصائيا وقيمتها 0.0430 وهي أقل من 0.09.

وأظهرت النتائج وجود لا تغير تكويني الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث)، وكانت قيم مربع كاي سكوير$\chi 2$ دالة إحصائيا وقيمتها 833.660 ومؤشر حسن المقارنةCFI دالة إحصائيا وقيمتها 0.898 وهي أكبر من 0.80 ومربع متوسط جذر الخطأ التقريبيRMSEA دالة إحصائيا وقيمتها 0.028 وهي أقل من 0.08 ومؤشر جذر متوسط مربعات البواقي المعيارية SRMR دالة إحصائيا وقيمتها 0.0351 وهي أقل من 0.09. كما وأظهرت النتائج وجود لا تغير متري الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث)، وكانت قيم مربع كاي سكوير$\chi 2$ دالة إحصائيا وقيمتها 890.903 ومؤشر حسن المقارنةCFI دالة إحصائيا وقيمتها 0.883 وهي أكبر من 0.80 ومربع متوسط جذر الخطأ التقريبيRMSEA دالة إحصائيا وقيمتها 0.029 وهي أقل من 0.08 ومؤشر جذر متوسط مربعات البواقي المعيارية SRMR دالة إحصائيا وقيمتها 0.0620 وهي أقل من 0.09.

وأظهرت النتائج وجود لا تغير عددي الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث)، وكانت قيم مربع كاي سكوير$\chi 2$ دالة إحصائيا وقيمتها 993.653 ومؤشر حسن المقارنةCFI دالة إحصائيا وقيمتها 0.854 وهي أكبر من 0.80 ومربع متوسط جذر الخطأ التقريبيRMSEA دالة إحصائيا وقيمتها 0.033 وهي أقل من 0.08 ومؤشر جذر متوسط مربعات البواقي المعيارية SRMR دالة إحصائيا وقيمتها 0.0634 وهي أقل من 0.09.

وتطابقت نتائج الدراسة مع نتائج وزارة التربية والتعليم التي أكدت على أن مدارس وكالة الغوث كانت الأقل في التراجع في مستوى مادة الرياضيات بالمقارنة مع المدارس الحكومية الاردنية، وأكدت الدراسة تفوق طلبة المدن على الريف في العلوم والرياضيات. هنا يبدو أن مدرسة الريف ما زالت بحاجة إلى رعاية ودعم ومساندة لترق بمستواها إلى مستوى مدرسة المدينة، وهنا يجب التركيز على احتفاظ مدرسة الريف بالمعلم الجيد، وأن لا تكون فقط مدرسة يتدرب فيها المعلم الجديد، وما أن يمتلك المهارات الجيدة يتم نقله إلى مدرسة المدينة، إذ أن الحاجة هنا حقيقية لخلق مجموعة من الحوافز التي تشجع المعلم على البقاء في مدرسة الريف. مما يدل على عدم اختلاف البناء العاملي المقترح والتشبعات العامليه لاختبار الرياضيات بين مدارس وزارة التربية والتعليم ومدارس وكالة الغوث. وهذا يعني أن ما يقيسه الاختبار واحد لدى الفئات المختلفة من الطلبة الأردنيين من حيث الجهة المشرفة (وزارة التربية والتعليم، وكالة الغوث). وقد تعزى مثل هذه النتيجة إلى صدق الاختبار والتزام الهيئات المشرفة على إعداده وتطويره بالمعايير المتفق عليها سواء ما يتعلق بالمحتوى أو العمليات والتزام الجهة المشرفة على تطبيقه في الأردن الدقيق بتعليمات وإجراءات الاختبار. وتعني هذه النتيجة أن هناك مصداقية للمقارنات بين الأوساط الحسابية لأداء المجموعات

المختلفة على الاختبار. فالفروق الموجودة بين تلك المجموعات تعزى للاختلافات فيما بين خصائص تلك المجموعات وقدراتها ولا تعزى لاختلاف خصائص الاختبار من مجموعة إلى أخرى.

وتوفر هذه النتائج دليلاً إضافيا على صدق اختبار تمس في الرياضيات للصف الثامن مما يشجع مدرسي الرياضيات بالرجوع إلى الاختبار والاستفادة منه في التدريس وفي بناء اختباراتهم التحصيلية. وتجدر الإشارة هنا إلى أن النتائج قد تختلف باختلاف الدولة أو لغة الاختبار مما يستدعي إجراء دراسات مستقبلية.

التوصيات والمقترحات:

1- توجيه التدريب إلى تضمين البرامج التدريبية في محتواها إلى أنشطة عن صلب الاختبارات الدولية.
2- إطلاع المشرفين التربويين على نماذج مختلفة من أسئلة الدورات السابقة، ونقلها إلى المعلمين لتناولها في التدريب وفي عملية التقويم.
3- أن تصبح أسئلة الاختبارات الدولية من ضمن قائمة الأسئلة التي يرجع لها المعلم أثناء التدريب والاختبار .
4- محاولة تضمين كتب الرياضيات لعينات من الأسئلة كأنشطة توجيهية أو أسئلة إثرائية او في التقويم.
5- عمل كراسة من الاسئلة من اختبار TIMSSمشفوعة بالحلول ولجزء منها، والجزء الآخر تتم مناقشته في حصص التقوية والإثراء.
6- وضع خطة إعلامية على مستوى كل مدرسة للتعريف بأهمية الدراسات الدولية والوطنية موجهة للطلبة والمعلمين... وأولياء الأمور .
7- اقتراح مجموعة من الحوافز للطلبة والمعلمين ومديرو المدارس والمشرفين الذين يحوزون طلبتهم على نتائج جيدة.

قائمة المراجع

أولاً- المراجع بالعربية:

- ابو لبدة، خطاب وحامد، شيرين والطويسي، احمد وعبابنة، عماد. (2013). التقرير الوطني الأردني، الدراسة الدولية للرياضيات والعلوم لعام2011 (TIMSS 2011)، المركز الوطني لتنمية الموارد البشرية، الأردن.
- عفانة، عزو ونبهان، سعد. (2004). مستوى الجودة في تحصيل الرياضيات باستخدام اختبار TIMSS والاتجاه نحو تعلمها لدى طلبة الصف الثامن بغزة. بحث مقدم إلى مؤتمر التربية في فلسطين وتغيرات العصر. الجامعة الاسلامية، غزة، 23-24 تشرين أول 2004.
- موسى، صالح احمد عطية. (2012). تقويم محتوى كتب العلوم الفلسطينية والإسرائيلية للصف الرابع أساسي في ضوء معايير TIMSS (دراسة مقارنة)، قسم المناهج وطرق التدريس، كلية التربية، الجامعة الإسلامية، غزة.

ثانياً- المراجع بالإنجليزية:

- Byrne, B. M, & Watkins, D. (2003). The issue of measurement invariance revisited. Journal of Cross-Cultural Psychology, 34, 155–175.
- Byrne, B.M. Shavelson, R.J. Muthén, B. (1989). Testing for the equivalence of factor Covariance and mean structures: the issue of partial measurement invariance. Psychol. Bull. 105 (3), 456–466.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. Structural Equation Modeling, 14 (3), 464-504.

- Dimiter, M. D. (2013), GAT-Verbal: Testing for Dimensionality & Validation of Factorial Structure. Administered by the National Center for Assessment in Higher Education (NCA) at Riyadh, Saudi Arabia.
- Gadelrab, H. F. (2004). The effect of model misspecification on goodness-of-fit indices for Structural equation modeling. Unpublished PhD, Wayne State University, Detroit, MI.
- George, A. M. and Irini, M. (2002). Latent Variable and Latent Structure Models, Lawrence Erlbaum Association, Inc, USA.
- Georgios, S. & Ioannis, T. (2013). DIF Analysis for Item and Test on the NCA Tests The General Ability Test (GAT) Art Major. National Center for Assessment in Higher Education (NCA) at Riyadh, Saudi Arabia.
- Hair, J. Anderson, R. Tatham, R. & Black, W. (1995). Multivariate Data Analysis with Readings, Fourth Edition, Prentice-Hall, Inc, New Jersey, USA.
- Hox, J.J. & Bechger, T.M. (1998). An introduction to structural equation modeling, Family Science Review, 11, 354–373 Retrieved from: http://joophox.net/publist/semfamre.
- Hu, L. T., & Bentler, P. M. (1999). Evaluating model fit. In R. H. Hoyle (Ed.), Structural equation modeling: Concepts, issues, and applications (pp. 76-99). Thousand Oaks, CA: Sage.
- Ioannis, T. (2014a). Factorial invariance & latent mean differences of scores on GAT across Gender, National Center for Assessment in Higher Education (NCA) at Riyadh, Saudi Arabia.
- Ioannis, T. (2014b). Factorial Structure of GAT & its Measurement Invariance across School Types. National Center for Assessment in Higher Education (NCA) at Riyadh, Saudi Arabia.
- Ioannis, T. (2015). Factorial invariance & latent mean differences of scores on SAAT across Gender, National Center for Assessment in Higher Education (NCA) at Riyadh, Saudi Arabia.
- Joreskog, K. G., & Sorbom, D. (1984). LISREL-VI: Analysis of linear structural relationships by the method of maximum likelihood [Computer program manual].Mooresville, IN: Scientific Software.
- Kiamanesh, A. R. (2006). Gender Differences in Mathematics Achievement among Iranian Eighth Graders in Two Consecutive International Studies (TIMSS 1999 & TIMSS2003), 2nd IEA International research Conference, http://www.iea.nl.
- Lattin, J. M. Carrol, J. D. & Green, P. E. (2003). Analyzing multivariate data. Brooks/Cole, Thompson Learning, Pacific Grove, CA, US.
- Mullis, Ina V. S, Martin,O. M, Ruddock, G. R. ,Chrisine, Y.,O., Alka, A., Ebru, E. (2007). International Mathematics Report..TIMSS and PIRLS International Study Center. Boston College: USA.
- Smith, R. J. (1987). Testing the normality assumption in multivariate simultaneous limited dependent variable models. Journal of Econometrics, 34 (1-2), 105-123.

- Vandenberg, R. J. Lance, C. E. (2000). A review and synthesis of the measurement invariance Literature: suggestions, practices, and recommendations for organizational research. Organ. Res. Methods 3 (1), 4–69.
- West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with non-normal variables: Problems and remedies. In R. Hoyle (Ed.), Structural equation modeling: Concepts, issues, and applications (pp. 56 –75). Thousand Oaks, CA: Sage.
- Widaman, K. F. Ferrer, E. & Conger, R. D. (2010). Measurement invariance within Longitudinal structural equation models: Evaluating change in the same construct across time. Child Development Perspectives, 4, 10-18.
- Yıldırım, H. H. (2006): The Differential Item Functioning (DIF) Analysis OF Mathematics Item In International Assessment Programs. Middle East Technical University. Chicago.