

أثر معالجة القيم المتطرفة في اختبارات تيمس الدولية (TIMSS, 2011) على تقديرات المعالم للفقرات ودقتها وفقاً لنظرية الاستجابة للفقرة

The Impact of Extreme Values Processing in the TIMSS International Assessment (TIMSS, 2011) on the Parameters and Accuracy of Paragraphs' Parameters in Accordance with the Theory of Response to the Paragraph

إخلاص العطيان

Akhlas Alatyan

وزارة التربية والتعليم، الأردن

بريد الكتروني: eklasflower30@hotmail.com

تاريخ التسليم: (2017/7/24)، تاريخ القبول: (2017/10/24)

ملخص

هدفت هذه الدراسة إلى تحديد نسبة القيم المتطرفة في اختبارات تيمس الدولية (TIMSS, 2011)، والكشف عن أثر معالجة القيم المتطرفة في مادتي العلوم والرياضيات على تقديرات المعالم للفقرات ودقتها وفقاً للنموذج اللوجستي الثلاثي المعلمة، وتكونت العينة من (531) فرد، وتم تحديد القيم المتطرفة بطريقة الرسم الصندوقي ومعالجتها بطريقة متوسط البتر، وتم الحكم على دقة تقدير المعالم باستخدام الخطأ المعياري كمؤشر دقة، حيث أظهرت النتائج قيم متطرفة في اختبار مادة الرياضيات بما نسبته (6.40%)، وفي مادة العلوم بما نسبته (8.29%)، وأظهرت النتائج أن هناك فروق جوهرية بين الأوساط الحسابية للأخطاء المعيارية لتقديرات معالم الفقرات، وأن أسلوب حذف القيم المتطرفة له أثر إيجابي بالمقارنة مع أسلوب الاستبدال والاحتفاظ، وينصح به كأفضل الأساليب بمعالجة القيم المتطرفة في اختبارات تيمس الدولية (2011) في مادة الرياضيات على الوسط الحسابي لتقديرات معالم الفقرات: التمييز والتخمين ودقتها، وفي العلوم على الوسط الحسابي لتقديرات معلمة التمييز ودقتها، كما أظهرت النتائج أن أسلوب الاحتفاظ بالقيم المتطرفة له أثر إيجابي بالمقارنة مع أسلوب الاستبدال والحذف، وينصح به كأفضل وأنجح الأساليب بمعالجة القيم المتطرفة في اختبارات تيمس الدولية (2011) في العلوم على الوسط الحسابي لتقديرات معالم الفقرات: التخمين، والصعوبة ودقتها، وفي الرياضيات على الوسط الحسابي لتقديرات معلمة الصعوبة ودقتها.

الكلمات المفتاحية: اختبار تيمس، معالجة القيم المتطرفة، الخطأ المعياري لدقة التقدير، نظرية الاستجابة للفقرة.

Abstract

This study aimed to identify the percent of Outlier values in TIMSS international tests (TIMSS, 2011) and the impact of the method of treatment in the students' scores at TIMSS test in accuracy of the estimation Items Parameters according to item response theory according used three-parameter logistic model, The sample consists of (531) students', and pointed out the Outlier values by the Box-plot method, it used the trimmed mean in replacing the Outlier values, and it used the standard error for accuracy of estimating, which showed there are Outlier value at Maths data percent of %6.40, Outlier values in Science percent of %8.29, The results showed that substantial differences between the means of standard errors for the Items parameters estimates, the method of deleting the Outlier values has a positive effect comparing with the method of (replacing and keeping) Outlier values, and it is advised that it is the best and the most successful method of Treatment with the Outlier values in TIMSS 2011 in Maths on the mean estimated Items parameters (discrimination and guessing) and its accuracy, and in Science on the mean estimated discrimination parameter and its accuracy. As the results showed that the method of keeping Outlier values has a positive effect comparing with the method of (replacing and deleting), it is advised that it is the best and the most successful method of keeping with the Outlier values in TIMSS at Science on the mean estimated Items parameters (guessing and difficulties) and its accuracy, at Maths on the mean estimated difficulties parameter and its accuracy.

Keywords: TIMSS test, Treatment method of Outlier, Standard Error for Accuracy of Estimating, Item Response Theory.

المقدمة

يعد موضوع القيم المتطرفة في البيانات الإحصائية وأثر وجودها وانعكاساته على نتائج ودقة التحليلات الإحصائية موضع اهتمام العديد من الباحثين في معظم بحوثهم، حيث ورد عن بارنت ولويس (Barnett & Lewis, 1977) أنه يعود الاهتمام بالقيم المتطرفة في مجموعة البيانات الإحصائية إلى ما يزيد على قرنين عندما حاول بوسكوفش Boscovich عام (1750م) تحديد إهليجية الأرض عن طريق إيجاد عشرة قياسات واستبعاد قياسين متطرفين وحساب الوسط لثمانى قياسات فقط، إي أنّ طرائق تشخيص القيم المتطرفة كانت شائعة ومنتشرة منذ

القدم، وزاد الاهتمام بموضوع القيم المتطرفة لأنه عادة ما يتم التعامل معها بالتجاهل، وهذا يؤدي إلى تقديرات ذات كفاءة أقل وتحيز في النتائج وتأثير عليها، لأن وجود القيم المتطرفة التي تختلف بشكل جوهري عن بقية البيانات يؤدي إلى اختلافات كبيرة وتغيرات جوهريّة وتشويه في النتائج الممثلة لبيانات العينة، وعليه يتطلب الأمر البحث والتحقيق في هذه المشاهدات غير العادية والبحث عن طرائق للسيطرة على تأثيرها، وتناول العديد من الباحثين موضوع دراسة القيم المتطرفة في بحوثهم وطرائق الكشف عنها ومعالجتها وأسبابها وأثرها على دقة النتائج، وهناك دراسات عديدة قام بها الباحثون لدراسة موضوع القيم المتطرفة واختلفت آراءهم حول موضوع القيم المتطرفة، وورد عن هادي (Hadi, 2010) إلى أنّ القيم المتطرفة هي مشاهدات غير طبيعية وتوزيعها مختلف عن باقي المشاهدات ضمن العينة الواحدة، ووجودها بالبيانات يؤدي إلى استنتاجات بعيدة عن الواقع بالنسبة للظواهر التي يتم دراستها، ويرى كل من (Shaker, 2009; Hawkins, 1980; Rahman & Al Amri, 2011; Jawad, & Kamal, 2009; Yahya, 2009; Hamoudat, 2009; Dan & Ijeoma, 2013a) أنّ القيم المتطرفة مجموعة من القيم غير الطبيعية، وهي قيم أكبر أو أصغر من مجموعة البيانات الأخرى، وقد يؤثر وجودها على البيانات وتتسبب في نتائج ضعيفة، وهي تأخذ نمط مختلف من التوزيعات ولا تتمثل بنمط معين، ولا تتسجم مع النمط العام للبيانات، وهي قيم غير منطقية وتبدو منحرفة عن باقي مكونات العينة التي وجدت فيها، وقد تنشأ من توزيعات مختلطة، ويجب أنّ يتم الكشف عنها وعلاجها، وكانت بدايات فكرة دراستها عن طريق أفكار بسيطة ومعتمدة على الحدس والتخمين، وقد يؤدي وجود هذه القيم إلى ظهور المشكلات في مجموعة البيانات.

وللأهمية الكبيرة للقيم المتطرفة على نتائج التحليل الإحصائي جاءت ضرورة إجراء هذه الدراسة للكشف عن نسبة القيم المتطرفة وبيان أثر معالجتها في اختبار تيمس (TIMSS, 2011) على تقديرات المعالم للفقرات ودقتها من خلال نظرية الاستجابة للفقرات كما أشار (Mislevy & Bock, 1990) تعد هذه النظرية من التطورات الحديثة في مجالات القياس التربوي والنفسي لما قدمته من طرائق جيدة في تفسير الدرجات على المقاييس التربوية والنفسية عند مقارنتها مع النظرية التقليدية في القياس، وقدمت هذه النظرية العديد من الطرائق السيكمترية الجيدة ببناء المقاييس التربوية والنفسية مثل النماذج ثلاثية وثنائية وأحادية البارامتر، وتم استخدام نموذجها اللوجستي الثلاثي المعلمة بهذه الدراسة لاحتوائه على جميع المعالم للفقرات (الصعوبة، والتمييز، والتخمين)، ويعد أكثر النماذج واقعية، وتم استخدام طريقة الأرجحية العظمى الهامشية في هذه الدراسة لتقدير معالم الفقرات، واستخدام الخطأ المعياري للحكم على دقة تقدير المعالم، وتمت الدراسة بالاستعانة باختبار تيمس الدولي (2011).

الإطار النظري

وفيما يتعلق بأسباب ظهور القيم المتطرفة فهناك عدد من الأسباب تؤدي إلى ظهورها بالبيانات كما ذكر: (Kassem, & Ismail, 2008; Hawkins, 1980; Barnett & Lewis, 1977; Seo, 2006; Osborne & Overbay, 2004; Dan & Ijeoma,

(2013a; Hossain, 1989; Yahya, 2009) أخطاء في عملية اختيار العينات، وفي إمكانية استجابة بعض الأفراد بشكل مختلف عن باقي أفراد العينة، وقيم متطرفة أخرى قد تنتج عن الاختلاف الأصلي (Inherent Variability) أو أخطاء القياس غير المتوقعة من الأفراد، وأخطاء سببها خلل في المقياس المستخدم في عملية القياس، وقد يكون سبب ظهور القيم المتطرفة هو التوزيع الملوث Contaminated Distribution، أو بسبب أخطاء البيانات أثناء جمعها وتسجيلها وإدخالها للحاسوب، أو أخطاء بمنهجية الدراسة، وقد يكون مجتمع القيم المتطرفة مختلف عن مجتمع البحث لباقي المشاهدات، وقد صنف سيو (Seo, 2006) القيم المتطرفة إلى صنفين: الصنف الأول القيم المتطرفة التي تعطي معلومات عن البيانات ولها فائدة كبيرة، والصنف الثاني القيم المتطرفة التي لها أثر سلبي كبير على البيانات.

وهناك عدد من الطرق الإحصائية التي تستخدم للكشف عن القيم المتطرفة في مجموعة البيانات وتحديدها ومن أهمها كما ورد عن كل من: (Ampanthong & Suwattee, 2009; Ismail, 2001; Al- Shmeiri, 2005; Seo, 2006; Kutner, Neter & Wasserman, 2005; Belsley, Kuh & Welsch, 2004; Fox, 1997; Al-Bayati, & Daga, 1996) طريقة مصفوفة القبة (Hat Matrix) وهي إحدى طرق الكشف عن القيم المتطرفة ومن عيوبها الكشف عن القيم المتطرفة في المتغيرات المستقلة فقط، ويمكن اعتبار العنصر القطري (h_i) لمصفوفة القبة مؤشر جيد من مؤشرات الكشف عن القيم المتطرفة في البيانات ويسمى بالرافعة (Leverage)، وأيضاً طريقة البواقي المعيارية Standardized Residuals (ZRESID) وتقوم هذه الطريقة على مبدأ حذف المشاهدة المتطرفة ثم حساب الباقي من نموذج الانحدار، وطريقة اختبار قروب Grubbs Test حيث أن هذا الاختبار يستخدم للكشف عن القيم المتطرفة في البيانات الموزعة توزيع طبيعي وفي حالة المتغير الواحد دون بيان كيفية علاجها، وهو من الطرق المعلمية التي تتطلب معرفة التوزيع الاحتمالي للبيانات، وتحدد القيم المتطرفة في حالة المتغير الواحد، واختبار T^2 Hotelling الذي اقترحه البياتي (Al-Bayati, 1973) من أجل معرفة إذا كان زوج المشاهدات (X_o, Y_o) المشكوك فيه متطرف أم لا، وهو من الطرق المعلمية التي تحدد القيم المتطرفة باتجاهين (Two Dimensions)، وطريقة الانحراف المعياري Standard Deviation إذ تعد من أبسط الطرق التقليدية لفحص القيم المتطرفة في البيانات، والدرجة المعيارية Z-Score وهي طريقة تستخدم لفحص القيم المتطرفة في البيانات باستخدام المتوسط الحسابي والانحراف المعياري، وهناك عدد من المقاييس للكشف عن القيم المتطرفة وتحديدها في البيانات مثل: مقياس مسافة مهلانوبيس Mahalanobis Measure Distance وهو من مقاييس المسافة المهمة التي يتم عن طريقها الكشف عن القيم المتطرفة وأثرها على البيانات، ومقياس المسافة كوك Cook ويستخدم من أجل أن يتم قياس أثر المشاهدة (i) على معاملات نموذج الانحدار المقدر، ومن الطرق الاستكشافية التي تستخدم لتحديد القيم المتطرفة في البيانات وتم استخدامها بهذه الدراسة طريقة Tukey (طريقة الرسم الصندوقي) لأنها حسب رأي الباحثين (Shaker, 2009; Seo, 2006; Hamoudat, 2009; Yahya, 2009) من أفضل الطرق في الكشف عن القيم المتطرفة في البيانات، وهي من الطرق الجيدة

التي تبين القيم ذات التطرف القوي، والقيم ذات التطرف المعتدل، وتعتبر أكثر الطرق حساسية وفعالية وكفاءة بتحديد القيم المتطرفة، وذكر توكي (Tukey, 1977) أن طريقة الصندوق والقطع المخططة مع الملخصات الخمسة من أفضل الطرق البيانية وتشمل: قيمة الوسيط (M)، وقيمة الربع الأدنى (Q_1)، وقيمة الربع الأعلى (Q_3)، وقيمة متطرفة دنيا (E_1)، وقيمة متطرفة عليا (E_2)، وتمثل البيانات من خلال صندوق (Box) ويوجد من الأسفل الربع الأدنى، ومن الأعلى الربع الأعلى، ويكون الوسيط (M) واقع بين هذين الربيعين، ويستخدم للكشف عن القيم المتطرفة، ويتم حساب المدى الربيعي (IQR (Inter Quartile Range باستخدام المعادلة: $IQR = (Q_3 - Q_1)$ حيث (Q_3 قيمة الربع الأعلى Upper Quartile) و(Q_1 قيمة الربع الأدنى Lower Quartile) وحساب الحد الأدنى للرسم الصندوقي باستخدام المعادلة: $(LE = Q_1 - 1.5 IQR)$ وحساب الحد الأعلى للرسم الصندوقي باستخدام المعادلة: $(UE = Q_3 + 1.5 IQR)$.

وهناك عدد من الطرق تستخدم لمعالجة القيم المتطرفة في البيانات أهمها:

1. طريقة الحذف

ويرى قاسم وإسماعيل (Kassem, & Ismail, 2008) والنعيمة (Al-Naimi, 2012) أن ظهور القيم المتطرفة في مجموعة البيانات يؤثر بشكل كبير في تحليلها، وفي معظم الأوقات يجب حذف القيم المتطرفة في البيانات، وذكر رحمان والعُمري (Rahman and Al Amri, 2011) أنه تحذف القيم المتطرفة في البيانات لتحسين دقة معاملات التقدير.

2. الاحتفاظ بالقيم المتطرفة

واقترح المختار (Mukhtar, 1980) وأوليوزي (Olewuezi, 2011) أن هناك عدد من الاحتمالات التي تستخدم لمعالجة القيم المتطرفة في البيانات ومن ضمنها الاحتفاظ بالقيم المتطرفة، لأنه ذا فائدة في نتائج التحليلات الإحصائية وتفسير الظواهر المدروسة.

3. استبدال القيم المتطرفة في البيانات

إذ يتم معالجة القيم المتطرفة في البيانات باستخدام الأساليب الإحصائية المناسبة، واستبدال القيم المتطرفة بطرائق الاستبدال المختلفة التي من أهمها الطرائق الآتية:

أ. طريقة الوسط الحسابي التعويضي Winsorized Mean: فكما ورد عن النور (Al-Nour, 2010) إن هذه الطريقة عبارة عن استخراج وسط حسابي لمجموعة من البيانات تم تقدير القيم المتطرفة فيها عن طريق قيم قريبة منها بدل حذفها.

ب. طريقة الوسط الحسابي المبتور Trimmed Mean: طريقة الوسط الحسابي المبتور التي تم الاعتماد عليها في هذه الدراسة لأهميتها ومرونتها وكفاءتها، فكما ورد عن الجبوري (المشار إليها في Dabdoub, & Younis, 2006) إن طريقة الوسط الحسابي المبتور من طرق علاج القيم المتطرفة في البيانات، وتتسم بالدقة والسهولة وتتم بترتيب القيم في

العمود بشكل تصاعدي ثم حذف أصغر قيمة وأكبر قيمة في بيانات العمود، وإيجاد الوسط الحسابي للبيانات المتبقية، ويكون بمثابة القيمة التقديرية للقيم المتطرفة، أي إيجاد وسط ممتور Trimmed mean لـ (n-2) من القيم، وافترض توكي (Tukey, 1962) أن عدد القيم المراد بترها في الطرفين متساوية، وذلك لصعوبة تحديد موقع القيم المتطرفة، وهذه الطريقة مميزة وتعطي أفضل النتائج عند معالجة القيم المتطرفة.

ويرى كل من: (Al-Zoubi, & Talafah, 2006; Jabouri, 1990; Dabdoub, & Younis, 2006; Al-Fatal, & Interanick, 2009; Hamoudat, 2009; Al-Shmeiri, 2005) أنه يتم علاج القيم المتطرفة بطريقة الوسط الحسابي الممتور من خلال حذف أكبر المشاهدات في العمود، فضلاً عن القيم المتطرفة وإيجاد الوسط الحسابي الذي يمثل تقديراً للقيمة المتطرفة، وهذه الطريقة هي من أفضل وأدق الطرق في علاج القيم المتطرفة، وتعطي أفضل النتائج في استبدال القيم المتطرفة، ولما لهذه الطريقة من ميزات متعددة مثل الدقة والسهولة والمرونة والكفاءة التي قد أجمع عليها معظم الباحثين والمهتمين، فهي تعد من أكثر الأساليب فاعلية في استبدال القيم المتطرفة لعدم تأثرها بها بشكل كبير، ولتميز هذه الطريقة عن الطرائق الأخرى تم الاعتماد عليها بهذه الدراسة من أجل علاج القيم المتطرفة في البيانات، فهي طريقة مميزة وتعطي أفضل النتائج عند استبدال القيم المتطرفة في البيانات.

والقيم المتطرفة تؤثر على دقة تقدير المعالم للفقرات لذا جاءت هذه الدراسة من أجل الكشف عن نسبة القيم المتطرفة وبيان أثر معالجتها في اختبار تيمس (TIMSS, 2011) في مادتي العلوم والرياضيات على تقديرات المعالم للفقرات ودقتها من وجهة نظرية الاستجابة للفقرة نظراً للاهتمام المتزايد بهذه النظرية، فقد بُنيت هذه النظرية على عدد من الافتراضات التي يجب أن يتم تحقيقها في البيانات المستمدة من الاختبار، لكي تؤدي إلى النتائج الموثوق بها كما ذكر كل من (Hambleton & Swaminathan, 1985; Allen & Yen, 1979; Crocker & Algina, 1986) وهي: أحادية البعد Unidimensionality أي هناك سمة واحدة قدرة كامنة واحدة مسؤولة عن أداء الفرد في الاختبار، والاستقلال الموضوعي Local Independence أي استجابة المفحوص (ضمن قدرة معينة) عن زوج من الفقرات مستقلة احصائياً (أن استجابة المفحوص على فقرة معينة لا يؤثر على استجابته على فقرة أخرى)، والتحرر من السرعة Non Speediness قدرة الفرد هي العامل الرئيسي لاستجابته على الاختبار، أي أن عامل السرعة لا يلعب دوراً في الإجابة عن فقرات الاختبار، ومنحنى خصائص الفقرة Item Characteristic Curve (ICC) دالة رياضية تربط بين مستوى القدرة (θ) واحتمالية الإجابة الصحيحة على الفقرة $P(\theta)$ ، إذ تزداد احتمالية إجابة الفرد عن الفقرة بصورة صحيحة بزيادة قدرته، وظهرت نماذج مختلفة بنظرية الاستجابة للفقرة تمكن من تحديد احتمالية إجابة الفرد ذو القدرة (θ) على الفقرة، وأشهر النماذج الأحادية التي تلائم الفقرات ثنائية التدرج كما ورد عن كل من: (Hambleton & Swaminathan, 1985; Hattie, 1985; Baker, 2001; IPL (نموذج راش)

ويفترض أن الفقرات يكون لها نفس مستوى القدرة التمييزية، وأيضاً معامل التخمين للفقرات يساوي صفراً ($0c_i =$)، وتختلف الفقرات أيضاً في معامل الصعوبة $P(\theta) = \frac{1}{1+e^{-1(\theta-b)}}$ (b)، والنموذج اللوجستي ثنائي المعلمة (نموذج بيرنوم) 2PL يفترض أن معامل التخمين لجميع الفقرات يساوي صفراً، وتختلف الفقرات في معامل الصعوبة (b) ومعامل التمييز (a). $P(\theta) = \frac{1}{1+e^{-a(\theta-b)}}$ ، والنموذج اللوجستي ثلاثي المعلمة (نموذج لورد) 3PL يفترض أن الفقرات مختلفة في معامل الصعوبة (b)، معامل التمييز (a)، ومعامل التخمين (c)، وهو النموذج الذي تم الاعتماد عليه في هذه الدراسة $P(\theta) = c + (1 - c) \frac{1}{1+e^{-a(\theta-b)}}$.

ومن خلال نظرية الاستجابة للفقرة يعتمد احتمال الإجابة الصحيحة على الفقرة على معالم الفقرة ومهم بهذه النظرية تقدير معالم الفقرات ودقتها، ويعد ذلك من المكونات الأساسية في نظرية الاستجابة للفقرة، وهناك عدد من الطرق تستخدم من أجل تقدير المعالم اللوجستية للفقرات وأهمها طريقة الأرجحية العظمى الهامشية Maximum Likelihood estimation Marginal وهي التي تطبق على النماذج اللوجستية الأحادية والثنائية والثلاثية كما ذكر كل من: (De Gruijter & Van Der kamp, 2005; Hambleton & Swaminathan, 1985). وقد تم الاعتماد عليها بهذه الدراسة لتقدير المعالم للفقرات، وأما فيما يتعلق بدقة تقدير معالم الفقرات، فقد أشار هامبلتون وجونز (Hambleton & Jones, 1993) أنه بعد عملية تقدير معالم الفقرات وفق نظرية الاستجابة للفقرة فإنه يجب التحقق من مدى دقة التقدير، حيث توجد عدة وسائل للكشف دقة التقدير أبرزها: الخطأ المعياري في التقدير (SEE) واستخدم بهذه الدراسة كمؤشر دقة، ودالة المعلومات للفقرات والاختبار، والكفاءة النسبية للاختبار.

مشكلة الدراسة

تعد عملية تقدير معالم الفقرات وفق نظرية الاستجابة للفقرة من الخطوات الأساسية في تطبيق هذه النظرية، والدراسات والأبحاث التي تتعامل مع العينات لا تخلو من القيم المتطرفة التي يختلف تأثيرها حسب حجمها مقارنة مع البيانات الأصلية وكيفية توزيعها، فغالبا ما يواجه الباحثون مشكلة القيم المتطرفة في البيانات عند تطبيق الاختبارات والمقاييس المتنوعة، ولهذه القيم تأثير على نتائج الاختبارات والتحليل الإحصائي للبيانات، لذا لا بد للباحث أن يعرف كيفية معالجة هذه القيم بالطرائق المناسبة وعدم تجاهلها في البيانات، وهناك عدد من الأبحاث التي تم إجراءها للكشف عن القيم المتطرفة في البيانات، ولكن لا يزال غير واضح طبيعة تأثير القيم المتطرفة على دقة تقديرات معالم الفقرات واتجاه التطرف ونسبته، لذا تحددت مشكلة الدراسة بالكشف عن نسبة القيم المتطرفة في البيانات، وبيان أثر أسلوب معالجتها (احتفاظ، وحذف، واستبدال) على دقة تقدير معالم الفقرات في ضوء نظرية الاستجابة للفقرة، _ وفي حدود علم

الباحثة هناك ندرة في الدراسات والأبحاث التي تناولت هذا الجانب، ولأغراض الدراسة الحالية تم الاستعانة باختبار تيمس (TIMSS, 2011) للكشف عن القيم المتطرفة في أداء الطلبة في هذا الاختبار، ومبررات استخدام هذا الاختبار أهمية الاختبار ونتائجه بالنسبة للأردن، ووجود حجم البيانات الفعلية الكبير الذي لا يمكن توفيره بتطبيق اختبارات مُعدة من قبل الباحثين، ولأن مستوى أداء الطلبة في هذه الدورة كان متراجع وخارج عن السياق المعتاد، حيث كانت نتائجهم متدنية تشكك بالفدرات التي يمتلكونها؛ في حين أن البعض قد عزا هذا التراجع إلى عدم جدية الطلبة في الإجابة عن فقرات الاختبار، إذ ربما تساهم هذه الدراسة في تشخيص الحالة في الكشف عن مدى جدية الطلبة في الإجابة عن فقرات الاختبار، ومعرفة هل وجود القيم المتطرفة في اختبار تيمس يؤثر على معالم الفقرات ودقتها، وتسلط الضوء على نتائج اختبار على درجة من الأهمية، ومعرفة الأسلوب الأفضل للتعامل مع القيم المتطرفة، وبالتحديد إن مشكلة الدراسة تتمثل بالإجابة عن الأسئلة الآتية:

أسئلة الدراسة

1. ما نسبة القيم المتطرفة في درجات طلبة الصف الثامن الأساسي الأردنيين في اختبارات تيمس الدولية المنعقدة في العام 2011 في مبحثي العلوم والرياضيات؟
2. ما أثر معالجة القيم المتطرفة في أداء طلبة الصف الثامن الأساسي الأردنيين في اختبارات تيمس الدولية المنعقدة بالعام 2011 في مبحثي العلوم والرياضيات على تقدير معلمة التمييز ودقتها؟
3. ما أثر معالجة القيم المتطرفة في أداء طلبة الصف الثامن الأساسي الأردنيين في اختبارات تيمس الدولية المنعقدة بالعام 2011 في مبحثي العلوم والرياضيات على تقدير معلمة الصعوبة ودقتها؟
4. ما أثر معالجة القيم المتطرفة في أداء طلبة الصف الثامن الأساسي الأردنيين في اختبارات تيمس الدولية المنعقدة بالعام 2011 في مبحثي العلوم والرياضيات على تقدير معلمة التخمين ودقتها؟

أهداف الدراسة

تهدف الدراسة إلى تحديد نسبة القيم المتطرفة في درجات الطلبة الأردنيين بالصف الثامن الأساسي في اختبارات تيمس الدولية (2011)، ومعرفة أثر معالجة القيم المتطرفة في مادتي العلوم والرياضيات على تقديرات المعالم للفقرات ودقتها وفقاً لنظرية الاستجابة للفقرات وتحديد الطريقة الأفضل لمعالجة القيم المتطرفة في جميع الحالات، وبالأعتماد على الخطأ المعياري كمؤشر للدقة.

أهمية الدراسة

قدمت نظرية الاستجابة للفقرة طرائق فعّالة لاختيار فقرات الاختبار، واختيار الفقرة حسب إسهامها بالمعلومات الكلية التي تمكن من الوصول إلى اختبار يتمتع بالدقة، والنجاح في تطبيق نظرية الاستجابة للفقرة يعتمد على دقة تقدير معالم الفقرات، لذا هذه الدراسة حاولت الكشف عن دقة تقديرات معالم الفقرات وأثر أسلوب معالجة القيم المتطرفة على ذلك، وفي ضوء مراجعة الأدب التربوي المتعلق بالدراسات التي أجريت حول القيم المتطرفة فإن هذه الدراسة تأتي استكمالاً لجهود الباحثين حول القيم المتطرفة، ولكن وفق منحى مختلف لم يتم التطرق له، من خلال دراسة أثر معالجة القيم المتطرفة على دقة تقديرات معالم الفقرات بضوء نظرية الاستجابة للفقرة، وتمثلت أهمية الدراسة بجانبين (نظري وتطبيقي) فمن حيث الأهمية النظرية فإن الدراسة يتوقع أن تساهم في:

1. الكشف عن نسبة القيم المتطرفة باختبار تيمس الدولي (2011) التي يتبين من خلالها جدية المفحوصين بالإجابة عن فقرات الاختبار، وانعكاس ذلك على الترتيب الدولي للأردن بهذا الاختبار.
2. الكشف عن أثر أسلوب معالجة القيم المتطرفة على دقة تقديرات معالم الفقرات في اختبار تيمس الدولي (2011)، وتسهيل مهمة الباحثين بتقديم الأسلوب الأكثر فاعلية بمعالجة القيم المتطرفة في البيانات.
3. دعم القاعدة النظرية للبحوث المتعلقة بتحليل وتفسير النتائج لأثر أسلوب معالجة القيم المتطرفة في البيانات على دقة تقديرات معالم الفقرات وفقاً لنموذج نظرية الاستجابة للفقرة الثلاثي.

أما من الناحية العملية فإن الدراسة يتوقع أن تساهم فيما يلي:

1. هذه الدراسة تعود بفائدة كبيرة على الباحثين والأوساط التربوية والمسؤولين عن الاختبارات وصانعي القرارات، وتعطي تصور واضح حول أثر معالجة القيم المتطرفة عندما تتواجد بيانات النماذج المختلفة للاختبارات المطبقة على الطلبة ضمن برامجها، ولتحقيق ما يسعون إليه من أهداف تربوية والعودة عليهم بالنفع والفائدة.
2. يمكن أن تستفيد من هذه الدراسة الجهات المتخصصة في بناء الاختبارات والذين يهتمون بنظرية الاستجابة للفقرة وموضوع القيم المتطرفة في البيانات لاختيار الأسلوب الأفضل لمعالجتها وتحديثها.
3. لهذه الدراسة فائدة كبيرة على الصعيد التطبيقي وفي واقع الحياة ومواقفها المختلفة؛ مما يوفر إمكانية التفسير والتشخيص واتخاذ قرارات سليمة، ولها أهمية في تحسين الإحصائيات المتعلقة بأداء الطلبة ومستوياتهم، والوصول إلى نتائج أكثر دقة وكفاءة في

التقييمات التربوية، وأيضًا أهميتها لأغراض بحثية ولخدمة أغراض القياس كافة، ولإعداد اختبارات تتمتع بكفاءة وعدم القصور.

تعريف المصطلحات

نظرية الاستجابة للفقرة Item Response Theory: هي اتجاه حديث في القياس التربوي والنفسي، حيث تفترض أنه يمكن التنبؤ بأداء الأفراد وتفسيره في ضوء خاصية مميزة لهذا الأداء تسمى السمات Traits وتفترض وجود سمة واحدة تكمن خلف استجابات الفرد عن فقرات الاختبار، وهذه السمة كامنة لا تلاحظ مباشرة، وتم الاعتماد على هذه النظرية في هذه الدراسة (Hambleton & Swaminathan, 1985).

الخطأ المعياري في التقدير Standard Error of the Estimate: هو القيمة المتوقعة للانحراف المعياري لأخطاء التقدير، وكلما انخفضت قيمته مؤشر على جودة التقدير، وزادت الدقة بدرجات الاختبار، وتم الاعتماد عليه بالدراسة كمؤشر على دقة التقدير (Hambleton & Jones, 1993).

القيمة المتطرفة Outliers: هي مجموعة من القيم بعيدة عن باقي قيم العينة ومختلفة عنها، وهي قيم قصوى ودنيا لعينة ما وتقع خارج سياق نمط البيانات، فهي قيم غير متنسقة وانحرفت عن بقية القيم، وبالدراسة الحالية يتم الكشف عنها أنها تقع خارج حدود الرسم الصندوقي (Rahman and Al Amri, 2011).

اختبار تيمس الدولي & TIMSS: Trends in International Mathematics & Science Study: هو اختبار تشرف عليه الرابطة الدولية لتقييم التحصيل التربوي في أمستردام/هولندا، لتقييم تحصيل الطلاب بمادتي الرياضيات والعلوم بالصف الثامن والرابع الأساسي، وهو اختبار عالمي يهدف إلى دراسة فاعلية المناهج، وطرائق تدريسها ويركز على النظم والسياسات التعليمية وتطويرها، واستخدمت نتائجه نتيجة التطبيق على طلبة الصف الثامن في الأردن للكشف عن القيم المتطرفة الموجودة بنتائجه لتحقيق غرض الدراسة (المركز الوطني لتنمية الموارد البشرية، 2011).

الرسم الصندوقي Box-Plot: طريقة للكشف عن القيم المتطرفة وهي من الطرق الجيدة والأكثر حساسية وفاعلية بتحديد القيم المتطرفة في البيانات (Seo, 2006).

محددات الدراسة

اقتصرت الدراسة الحالية على النموذج اللوجستي الثلاثي المعلمة، وعلى بيانات اختبار TIMSS الدولي لعام 2011، والعينة المكونة من (531) فرد ممن استجابوا على الكراسة الاختبارية ذات الرقم 11 لأنها تحتوي على أكبر عدد من فقرات الاختبار من متعدد لاختباري الرياضيات والعلوم، واستخدام برنامج التحليل الإحصائي SPSS وبرنامج bilog -mg3،

وأسلوب واحد فقط للكشف عن القيم المتطرفة هو الرسم الصندوقي، وأسلوب واحد فقط لعلاجها هو الوسط الحسابي المبتور.

الدراسات السابقة

يتضمن هذا الجزء مراجعة لعدد من الدراسات السابقة العربية والأجنبية ذات العلاقة بموضوع الدراسة التي تم عرضها بعد مراجعة الأدب التربوي، مرتبة وفق التسلسل الزمني ابتداءً من الأقدم إلى الأحدث، وفي الفترة الأخيرة وجه بعض الباحثين اهتمامهم للقيم المتطرفة حيث ظهرت بعض الدراسات ذات العلاقة بموضوع الدراسة، وتناول العديد من الباحثين موضوع التطرف في دراساتهم من خلال التعامل معها من وجهات نظر مختلفة فمنهم من أوصى بحذف القيم المتطرفة، وهناك من أوصى بالاحتفاظ بها أو معالجتها، وفيما يلي عرض موجز لبعض الدراسات المتعلقة بموضوع القيم المتطرفة.

أجرى نوفوتني (Novotny, 2001) دراسة بحثت في أثر القيم المتطرفة المتعددة المتغيرات في مصفوفة التباين للكشف عن تأثير القيم المتطرفة على تقديرات المعالم وعلى تقييم النماذج واختبار نماذج محددة، فبعض القيم المتطرفة لها آثار سيئة على التقديرات الإحصائية، وهناك طرق للكشف عن القيم المتطرفة في البيانات، وبحثت الدراسة في أثر كمية التحيز على تقديرات نموذج المعلمة بحال وجود القيم المتطرفة، وبينت النتائج أن وجود القيم المتطرفة والاحتفاظ بها لا يؤثر في قرارات قبول أو رفض النموذج المستخدم، وأن القيم المتطرفة ليست دائماً بمثابة قيم مؤثرة، وليست جميع المقاييس قادرة على الكشف عن القيم المتطرفة.

وركز زيمرمان (Zimmerman, 2001) في دراسته على البحث في تأثير القيم المتطرفة في الاختبارات المعلمية Parametric tests والاختبارات اللامعلمية Nonparametric tests، حيث بينت الدراسة الخاصة باختبار t المعلمي أن القيم المتطرفة تقلل من احتمالية حدوث زيادة في الأخطاء من النوع الأول وتخفيضها، وتعمل على زيادة احتمالية حدوث الخطأ من النوع الثاني، وبالتالي تخفض من قوة الاختبار، أما فيما يتعلق باختبار مان ويتي وويلكوكسون اللامعلمي، فإن القيم المتطرفة لا تؤثر على احتمالية حدوث زيادة في الخطأ من النوع الأول، وتعمل على زيادة احتمالية حدوث الأخطاء من النوع الثاني، وتقلل من قوة الاختبار، أي أن للدرجات المتطرفة تأثير على اختبار t المعلمي، ولكن الاحتفاظ بها ليس له أي تأثير على الاختبار اللامعلمي.

وأجرى ميشيليدز (Michaelides, 2003) دراسة كانت مكونة من ثلاث مراحل وتهدف إلى فحص تأثير الفقرات المشتركة المتطرفة على عملية المعادلة، وذلك باستخدام نظرية الاستجابة للفقرة، ولتحقيق هدف الدراسة فقد تم استخدام اختبارين لسنتين مختلفتين على التوالي، وأشارت النتائج إلى أن الفقرات المتطرفة عندما تكون سهلة جداً أو صعبة جداً أثرت على إحصاءات المعادلة بشكل أكبر من أن تكون متوسطة الصعوبة، وكانت نتائج المعادلة عند معايرة الفقرات حسب النموذج الأحادي المعلمة IPL أكثر ثباتاً من نتائج المعادلة عندما تمت معايرة الفقرات حسب النموذج الثلاثي المعلمة 3PL.

أما دراسة اوسبورن واوفرباي (Osborne & Overbay, 2004) فقد هدفت إلى الكشف عن أثر القيم المتطرفة على نتائج تحليلات الارتباط واختبار t وتحليل التباين، ولتحقيق ذلك تم دراسة أثر القيم المتطرفة على معدلات الخطأ من النوع الأول ومن النوع الثاني، وأشارت النتائج إلى أن القيم المتطرفة تؤثر بشكل سلبي على معاملات الارتباط، حيث تزداد الدلالة الإحصائية لمعامل الارتباط عند حذف القيم المتطرفة وتقل الأخطاء بصورة دالة إحصائية، ويصبح أكثر دقة وأقرب إلى قيمة معامل الارتباط الحقيقي في المجتمع، كما أشارت النتائج إلى أن وجود قيم متطرفة يؤثر سلباً على دلالة اختبار t وتحليل التباين، وهو يعمل على تقليل الأخطاء، ويحسن من نتائج التحليلات الإحصائية.

وفي دراسة أجرتها ألبرتا (Albera, 2004) كان الهدف منها التحقق في الأساليب الأساسية القائمة على المساواة (المعادلة) في النظرية الحديثة في القياس، وذلك بوجود القيم المتطرفة، والتي لها تأثير كبير وخطير على النتائج الأساسية في النظرية الحديثة في القياس، وقد كان الغرض الأساسي من هذه الدراسة هو التحقق في المقارنة بين عشرة طرائق مساواة أساسية في النظرية الحديثة في القياس، التي يمكن أن تتجاهل القيم المتطرفة، أو تعتبر تأثير القيم المتطرفة متعارض مع تقديرات المعالم، وتبين أنه عندما كانت القيم المتطرفة موجودة في البيانات كان أداء طرائق المعادلة (المساواة) العشرة الأساسية في النظرية الحديثة غير مناسب، وأن هناك عدد من الأساليب القائمة على المساواة في النظرية الحديثة تتأثر بالقيم المتطرفة مثل اختبار المنحنى المميز ووسط سيجما، وعند معالجة القيم المتطرفة كان أداءها بشكل أفضل وأحسن، وعموماً كان أداء الأساليب التي لم تهتم بتأثير القيم المتطرفة غير مثالي ونتائجها ملوثة.

وأجرى سيو (Seo, 2006) دراسة حول مقارنة ومراجعة بين طرائق الكشف عن القيم المتطرفة في مجموعة البيانات أحادية المتغير، وتم تطوير بعض الأساليب للكشف عن القيم المتطرفة بالبيانات، وتبين أن بعض الأساليب تكون حساسة للقيم المتطرفة مثل طريقة الانحراف المعياري، وبعض الطرائق الأخرى تكون مقاومة للقيم المتطرفة كطريقة توكي، ويمكن أن يكون للقيم المتطرفة تأثير سلبي على تحليل البيانات، وقد توفر معلومات مفيدة عن البيانات، وتبين أن القيم المتطرفة تتغير في ضوء الالتواء والكشف عنها وتقصيها، وقد تظهر في حال توزيع متصل السمة لأي معلمة، ولكنها حتماً ستظهر عند التوزيعات الملتوية، فإذا كانت البيانات ملتوية فسوف تشمل قيم متطرفة، وكلما كان الالتواء متطرف أكثر ستزداد نسبة ظهور القيم المتطرفة، وأن أسلوب الرسم الصندوقي يزداد فيه عدد القيم المتطرفة عندما تكون التوزيعات ملتوية، أو الانحراف المعياري صغير للقدرات.

وركزت دراسة ديدوب ويونس (Dabdoub, & Younis, 2006) على أهمية الكشف عن تأثير القيم المتطرفة على نتائج تحليل الانحدار اعتماداً على قيمة معامل التحديد R^2 والقيمة التقديرية لتباين المجتمع/متوسط مربعات الخطأ (MSE: Mean Square Error) وتم الكشف عن القيم المتطرفة كنتيجة لتحليل الانحدار بطريقة الرسم الصندوقي، ومعالجتها بطريقة الوسط الحسابي المبتور، وتم جمع البيانات عن المواليد الخُدج من مستشفى البتول التعليمي للولادة

بالعراق، واستخدام تحليل الانحدار لتحديد القيم المتطرفة، وتبين أن القيم المتطرفة أثرت على تركيبة وبنية النموذج ونتائج تحليل الانحدار، وتبين أنه ليس من الضروري أن يؤثر وجود القيم المتطرفة والاحتفاظ بها سلبيًا على النتائج، وليس بالضرورة أن تؤدي معالجتها لزيادة قيمة معامل التحديد R^2 وانخفاض بقيمة الخطأ MSE، وأوصت الدراسة بالكشف عن القيم المتطرفة عند كل مرحلة من مراحل التحليل.

وتضمنت دراسة قاسم وإسماعيل (Kassem, & Ismail, 2008) البحث في أهمية الكشف عن القيم المتطرفة وتحديدتها في الانحدار الخطي البسيط عن طريق استخدام أسلوب بييز Bayes باستخدام معاينة جيبس Gibbs، وتضمن توزيعين أحدهما ملوث والآخر خالي من القيم المتطرفة باحتمالية محددة، وتم تطبيق النتائج النظرية على بيانات حقيقية من أجل الكشف عن القيم المتطرفة، وتبين أن تحديد القيم المتطرفة بأسلوب بييز يحل العديد من المسائل المعقدة وتكون النتائج ذات كفاءة ودقة، وتبين أن القيم المتطرفة تأتي من البيانات التي تعود إلى توزيعات غير متماثلة أي يكون فيها التواء عال نحو اليسار أو نحو اليمين، وتوصي الدراسة باستخدام معاينة جيبس للكشف عن القيم المتطرفة.

وفي دراسة قامت بها النور (Al-Nour, 2010) كان الهدف منها إيجاد عدد من الطرائق ذات الدقة والموضوعية والكفاءة من أجل تقدير معلمة الصعوبة في حال المتغير الواحد، وذلك في حال التوزيع الطبيعي، وفي حال التوزيع الاحتمالي غير المتماثل، من خلال مقارنة الطرائق المقترحة والطرائق التقليدية والطرائق الحصينة، ومن خلال البحوث تم معرفة أن الطرائق التقليدية ليست ذات كفاءة وذلك بسبب وجود القيم المتطرفة ولأنها تتأثر بالانحراف عن الافتراضات التي تم تحديدها، وتم تحليل النتائج بالاعتماد على قيمة متوسط مربعات الخطأ Mean Square Error، وكان من النتائج أن الوسط الحسابي أفضل مقدر لتقدير معلمة الصعوبة بحال عدم وجود قيم متطرفة بالتوزيع الطبيعي، وبحال التوزيع الملوث يكون الوسط الحسابي المبتور أفضل مقدر لتقدير معلمة الصعوبة.

أما دراسة المحاكاة باستخدام طريقة مونتني كارلو التي قام بها ليوو ويو وزومبو (Liu, Wu & Zumbo, 2010) فقد هدفت الدراسة إلى التحقق من أثر القيم المتطرفة على تقديرات معامل ثبات كرونباخ α لتقدير الثبات لبيانات الاستجابة للفقرة، ولتحقيق أهداف الدراسة تم توليد البيانات باستخدام طريقة مونتني كارلو، حيث تمت إضافة بعض البيانات المتطرفة إلى توزيع طبيعي ليصبح التوزيع ملوث، وأظهرت نتائج الدراسة بأن التحيز والفاعلية في تقديرات معامل ثبات α يزيدان مع زيادة نسبة القيم المتطرفة، وتتضخم قيمة معامل ثبات α ، ولا يتأثر التحيز مع زيادة عدد نقاط المقياس الرتبي، ويقل تضخم الفاعلية مع زيادة معامل الثبات النظري، ولكن يقل التحيز مع زيادة معامل الثبات النظري، كما أظهرت النتائج بأنه في حالة عدم وجود قيم متطرفة فإن التحيز يقل ولا تتأثر الفاعلية مع زيادة عدد نقاط المقياس الرتبي، أي أن تقديرات معامل ثبات كرونباخ α تتضخم وبشكل كبير في ظل وجود القيم المتطرفة التي تؤدي إلى التحيز الإيجابي في قيمة الثبات للمقياس.

وهناك دراسة أخرى أجراها جيمينيز (Jimenez, 2011) كان الغرض منها هو التحقيق في الآثار المترتبة على وجود القيم المتطرفة مع الفقرات المشتركة بتقديرات المعالم في نظرية الاستجابة للفقرة، وأظهرت النتائج بأنه في حالة عدم وجود فقرات متطرفة في الفقرات المشتركة فإن أداء المعادلة جيد عند تقدير معاملات التمييز حسب المجموعات المتكافئة والمجموعات غير المتكافئة، أما بحالات وجود فقرات متطرفة بالفقرات المشتركة كانت الأخطاء متفاوتة على الأسلوبين، وأظهرت النتائج أن وجود الفقرات المتطرفة في الفقرات المشتركة يزيد أخطاء المعادلة، وأشارت النتائج أنه لا يوجد تفاعل بين القيم المتطرفة والمعالم بمجموعة البيانات، أي الاحتفاظ بالقيم المتطرفة ليس له أي أثر سلبي على المعالم.

وأشارت دراسة أوليوزي (Olewuezi, 2011) عن مواصفات النموذج الأحادي المتطرف إلى أن القيم والملاحظات المتطرفة هي مشاهدات غير عادية لها تأثير على البيانات وعلى التحليل الإحصائية والنتائج بشكل عام، ولقد أثبتت الإحصاءات المستمدة من مجموعات البيانات التي تشمل على القيم المتطرفة إلى ضرورة تضمينها وعدم حذفها والاحتفاظ بها في التحليل الإحصائية لمعرفة تأثيراتها المحتملة على نتائج التحليلات الإحصائية، وتبين أنها يمكن أن تكون ملوثة ونتاجها مضللة وسلبية، ويمكن أيضاً أن يكون وجودها مؤشراً على وجود البيانات التي تختلف عن باقي بيانات المجموعة.

وفي دراسة قامت بها النعيمي (Al-Naimi, 2012) من أجل الكشف عن القيم المتطرفة ومعالجتها بالطريقة الحصينة ومقارنتها بطرائق أخرى، كان يتم اللجوء إلى استبعاد القيم المتطرفة عند تحليل البيانات، وتم التعامل مع القيم المتطرفة بأسلوبين الأول في حالة تحليل البيانات الأصلية بعد استبعاد القيم المتطرفة يتم عن طريق زيادة حجم العينة بالنسبة للأعوام الأخرى، والأسلوب الثاني عن طريق استخدام بعض الطرائق الحصينة للكشف عن القيم المتطرفة، وأهم الاستنتاجات أنه يمكن أن نستبعد القيم المتطرفة ونحذفها بحالة أن سببها أخطاء التسجيل للملاحظات، أو بسبب الأجهزة، وأنه عند الكشف عن القيم المتطرفة يتبين أن زيادة البيانات تزداد بزيادة حجم العينة وقد يساعد ذلك في الاعتماد على طرائق التقدير التقليدية في تحليل البيانات لأن هذه الطريقة تمثل المجتمع خير تمثيل، وأن الطرائق الحصينة التي تكشف عن القيم المتطرفة ممكنة وتحتاج إلى التكرار وتستخدم بحال البيانات الكاملة.

وقام يونغ (Yong, 2013) بإجراء دراسة باستخدام أسلوب الانحدار الخطي للكشف عن القيم المتطرفة في النظرية الحديثة في القياس في معادلة ومساواة الفقرات المشتركة، وتم في هذه الدراسة اقتراح طريقة الانحدار الخطي كوسيلة من وسائل الكشف عن القيم المتطرفة في البيانات وتم مقارنة الطريقة المقترحة الحديثة مع الطرق التقليدية، وأكدت النتائج ضرورة الكشف عن الفقرات المشتركة المتطرفة ومعالجتها بالطرائق المناسبة لما لها من تأثيرات سلبية على مجموعة البيانات.

وتناولت دراسة يوسف (Yousef, 2015) تأثير وجود القيم المتطرفة على معالم نموذج تحليل الانحدار الخطي المتعدد والمقارنة بين طرق تشخيص القيم المتطرفة، ومن فرضيات

الدراسة أن القيم المتطرفة لها تأثير معنوي على نتائج نموذج تحليل الانحدار الخطي المتعدد، واعتمدت الدراسة على المنهج الوصفي ومنهج التحليلي الاستدلالي وبعض الطرق الخاصة بتشخيص ومعالجة القيم المتطرفة، وتم الكشف عن القيم المتطرفة بطريقة الرسم الصندوقي وتمت معالجتها بطريقة متوسط البتر، وأهم النتائج أن للقيم المتطرفة تأثير على نموذج الانحدار الخطي المتعدد، وتعمل على تضخم قيمة تباين الأخطاء MSE، وتؤثر معنويًا وبصور كبيرة على عدم تحقق فرضية التوزيع الطبيعي، وأوصت الدراسة بضرورة تشخيص القيم المتطرفة ومعالجتها وعدم حذفها وتقليل أثرها عند تطبيق نموذج تحليل الانحدار المتعدد للحصول على نموذج أفضل ولأهميتها وأثرها على النتائج.

وهدفت دراسة أديبو وجيمس وأكينومي (Adeboye, James & Akinwumi, 2016) إلى النظر في اجراءات الكشف عن القيم المتطرفة في نموذج الانحدار المتعدد، والمقارنة بين طرق الكشف عن القيم المتطرفة والتحقق من كل طريقة، وتمت هذه الدراسة باستخدام بيانات محاكاة، وتم استخدام خمسة طرق للكشف عن القيم المتطرفة في نموذج الانحدار الخطي المتعدد، ثم تمت مقارنة النتائج باستخدام الانحراف المعياري لتقدير المعلمة ومتوسط الانحراف، وتم توليد البيانات وبثلاث نسب مختلفة من القيم المتطرفة فيها، وباستخدام احجام عينات مختلفة، وأظهرت النتائج أن القيم المتطرفة تؤثر على البيانات وبشكل سلبي وتؤثر على معاملات الانحدار وفي جميع الحالات التي تمت دراستها، ويجب الكشف عن القيم المتطرفة وعدم حذفها ومعالجتها لأهميتها على النتائج.

التعقيب على الدراسات السابقة

فيما يتعلق بالدراسات السابقة كان هناك اتفاق عام بين أغلب الباحثين بأن وجود القيم المتطرفة في مجموعة البيانات تساهم بزيادة الأخطاء وتحيز في تقديرات الإحصاءات والمعالم وتؤثر سلبياً على النتائج، ويجب الكشف عنها ومعالجتها بالطرائق المناسبة، وهناك اتفاق بين كل من (Alberta, 2004)، (Jimenez, 2011)، (Al-Nour, 2010)، (Yong, 2013)، (Liu, Wu & 2004)، (Zumbo, 2010)، (Michaelides, 2003)، (Olewuezi, 2011)، (Al-Naimi, 2012)، (Kassem, & Ismail, 2008)، (Yousef, 2015) و (Adeboye, James & Akinwumi, 2016) على ذلك، وأشار (Zimmerman, 2001) إلى أن الخطأ من النوع الأول ينخفض بوجود قيم متطرفة، وتوصل (Novotny, 2001) أن ليس جميع القيم المتطرفة دائماً هي بمثابة قيم مؤثرة، وتوصل (Osborne & Overbay, 2004) أن حذف القيم المتطرفة يعمل على تقليل الأخطاء ويحسن من نتائج التحليلات الإحصائية، وأشار (Dabdoub, & Younis, 2006) أنه ليس ضروري أن تؤثر القيمة المتطرفة سلبياً على النتائج، وأن تؤدي معالجتها إلى زيادة قيمة معامل التحديد R^2 وانخفاض بقيمة الخطأ MSE، وتوصل (Seo, 2006) أن معظم البيانات تحتوي على قيم متطرفة ممكن أن تكون لها تأثير سلبي على تحليل البيانات، وقد توفر معلومات مفيدة عن البيانات.

يتضح من خلال ما تم تناوله من دراسات سابقة ذات علاقة بموضوع الدراسة أن هناك دراسات بيّنت أنّ للقيم المتطرفة أثر إيجابي، ودراسات بيّنت أنّ لها أثر سلبي، ورغم غزارة الإنتاج العلمي بمختلف مفردات الدراسات إلا إن ما يميز هذه الدراسة بأنها جاءت بصورة مختلفة لتلقي الضوء بشكل مباشر على الكشف عن القيم المتطرفة وأثر أسلوب معالجتها على تقدير معالم الفقرات ودقتها، فهي الدراسة الأولى التي قامت بذلك، وأضافت بُعداً جديداً يتمثل باستخدام النموذج اللوجستي الثلاثي المعلمة، وتتميز في أنّها تستخدم بيانات اختبار تيمس الدولي 2011، ومعظم الدراسات السابقة تمت من خلال بيانات محاكاة، وليست باستخدام بيانات فعلية كما في هذه الدراسة، لذلك تعتبر الدراسة استكمالاً للدراسات السابقة التي لم تتناول بشكل مباشر الكشف عن القيم المتطرفة وأثر أسلوب معالجتها على تقدير معالم الفقرات ودقتها من وجهة النظر الحديثة بالقياس، بالاستعانة باختبار تيمس الدولي 2011 للوصول إلى أدق النتائج.

الطريقة والإجراءات

مجتمع الدراسة

تكون مجتمع الدراسة في الاختبار الدولي تيمس من جميع طلبة الصف الثامن الأساسي الأردنيين المشاركين في الاختبار الدولي تيمس عام (2011)، وتألّف العدد الإجمالي من (7694) طالباً وطالبة، من (230) مدرسة اختيرت بالطريقة العشوائية البسيطة من مجتمع مدارس الأردن التي تحتوي على الصف الثامن كواحد من الصفوف التي تدرسها، ويبين الجدول 1 توزيع مجتمع الدراسة حسب السلطة المشرفة، والموقع، وجنس المدرسة، والطبقة.

جدول (1): توزيع مجتمع الدراسة حسب (السلطة المشرفة، والموقع، وجنس المدرسة، والطبقة).

المتغير	مستويات المتغير	الطلبة		المدارس	
		العدد	%	العدد	%
سلطة مشرفة	وزارة التربية والتعليم	6316	82	180	78
	وكالة الغوث	843	11	25	11
	التعليم الخاص	535	7	25	11
	المجموع	7694	100	230	100
موقع	مدينة	6202	81	180	78
	ريف	1492	19	50	22
	المجموع	7694	100	230	100
جنس المدرسة	ذكور	3315	43	101	44
	إناث	2566	33	71	31
	مختلط	1813	24	58	25
	المجموع	7694	100	230	100

...تابع جدول رقم (1)

المدارس		الطلبة		مستويات المتغير	المتغير
%	العدد	%	العدد		
11	25	12	932	استكشافية/وزارة	الطبقة
43	100	41	3145	وزارة التربية والتعليم	
11	25	11	843	وكالة الغوث	
11	25	7	535	التعليم الخاص	
11	25	10	752	مدرستي/وزارة	
13	30	19	1487	دعم التعليم/وزارة	
100	230	100	7694	المجموع	

ويبين الجدول 2 توزيع طلبة مجتمع الدراسة حسب جنس الطالب.

جدول (2): توزيع مجتمع الدراسة حسب جنس الطالب.

النسبة المئوية	عدد الطلبة	الجنس
47	3604	ذكور
53	4090	إناث
100	7694	المجموع

عينة الدراسة

تألفت عينة الدراسة النهائية التي طُبِّقت عليها المعالجات الإحصائية، والتي تم اختيارها بشكل يحقق غرض الدراسة عن طريق تحليل محتوى بنية الاختبار من (531) طالباً وطالبة من طلبة الصف الثامن الأساسي الأردنيين، ممن استجابوا على فقرات الكراسة الإختبارية ذات الرقم 11 التي تحتوي على أكبر عدد من فقرات الاختبار من متعدد لاختباري الرياضيات والعلوم.

أداة الدراسة

تم تطبيق هذه الدراسة بالاعتماد على بيانات الاختبار الدولي تيمس (TIMSS, 2011) لطلبة الصف الثامن الأساسي الأردنيين، ويطبق الاختبار كل أربع سنوات للصفين الثامن والرابع الأساسيين، وتم تطبيقه منذ عام (1995) والدورة الثانية عام (1999) والدورة الثالثة بعام (2003) والدورة الرابعة بعام (2007) والدورة الخامسة في عام (2011)، وتم اختيار بيانات الاختبار في هذه الدورة بالذات بسبب التراجع الكبير بمستوى أداء الطلبة في هذه الدورة عن الدورات السابقة وخروج مستوى أداء الطلبة عن السياق المعتاد، وكان عدد الدول المشاركة في الدورة الخامسة في الصف الثامن الأساسي (45) دولة منها (11) دولة عربية ومن ضمنها الأردن، وتم في الدراسة اختيار الكراسة الإختبارية ذات الرقم (11) المعنية بفقرات الاختبار من

متعدد، وهذه الكراسة الإختبارية هي ناتجة عن تقاطع الكراستين الإختباريتين ذواتي الرقمين 11 و12، وهي تحتوي على أكبر عدد من فقرات الإختيار من متعدد لاختباري الرياضيات والعلوم لذلك تم اختيارها، وبهذا فإن عينة الدراسة النهائية مكونة من (531) طالباً وطالبة ممن استجابوا على فقرات الكراسة الإحصائية ذات الرقم 11.

وفيما يلي نبذة عن هذا الإختبار: كما ورد في التقرير الوطني الاردني عن الدراسة الدولية للعلوم والرياضيات TIMSS 2011 في المركز الوطني لتنمية الموارد البشرية:

الفئة المستهدفة في اختبار تيمس: يتم اختيار الطلبة في نهاية السنوات الاربع الأولى، أي في الصف الرابع الاساسي ومن ثم في نهاية السنوات الاربع التالية أي في الصف الثامن الاساسي من التعليم.

أهمية وأهداف اختبار TIMSS: أن اختبارات تيمس TIMSS تساعد القائمين على التعليم والجهات المعنية بذلك في دعم الدراسات التربوية وتطوير مهارات وطرائق التدريس، وإعادة النظر في المناهج الدراسية، وتطوير النظام التربوي وإصلاحه والارتقاء بمخرجاته ومعرفة مستوى التحصيل عند الطلاب وتقييمهم ومقارنتهم مع الدول الأخرى، وتنمية مهارات التفكير العلمي لديهم، والعمل على تزويد المعلمين بمعلومات عن جوانب القوة والضعف لدى الطلاب من أجل متابعتهم وتقييمهم والوصول إلى تعليم أفضل والعمل على أعداد التربويين وإكسابهم المعرفة والخبرة والمهارات المناسبة، والعمل على تقييم السياسات التربوية وبشكل مستمر.

طريقة بناء أسئلة اختبارات تيمس TIMSS: يتم صياغة السؤال بطريقة تدفع للتفكير بفهم واستيعاب السؤال وكيفية تطبيقه وتحليل عناصره وتركيب جزئياته بطريقة سليمة، ويقدم السؤال كمشكلة أو صورة أو رسم تخطيطي.

أدوات اختبارات تيمس TIMSS: (اختبار في العلوم، واختبار في الرياضيات، واستبانة تعطى إلى مدير المدرسة، واستبانة تعطى إلى الطالب، واستبانة لمعلم العلوم، واستبانة لمعلم الرياضيات).

مراحل اختبار تيمس: (مرحلة الإعداد، ومرحلة التطبيق، ومرحلة التصحيح، ومرحلة التحليل، ومرحلة كتابة التقرير).

صدق أداة الدراسة (اختبار تيمس)

ورد عن الخياط (Al-Khayat, 2012) أنه تم التوصل إلى دلالات صدق اختبار تيمس (TIMSS, 2011) حيث تم التحقق من صدق المحتوى من خلال تحديد الفقرات، وماذا تقيس كل فقرة؟ ومصادر الحصول على الفقرات؟ وطريقة صياغتها؟ وطرائق تحكيمها؟، وتم التحقق من الصدق بدلالة المحك بحساب قيم معاملات الارتباط بين درجات المفحوصين على الاختبار

ودرجات الطلبة في مادة الرياضيات لعينة الدراسة بأكملها، حيث تم إيجاد معامل الارتباط بين درجات الطلبة على الاختبار ودرجاتهم في مادة الرياضيات؛ وبلغت قيمته (0.87).

ثبات أداة الدراسة (اختبار تيمس)

وذكر الخياط (Al-Khayat, 2012) أيضاً أنه تم إيجاد ثبات الاختبار بإيجاد معاملات الارتباط بين الأداء على الفقرة والاختبار ككل، باستخدام معامل كرونباخ α ؛ حيث تم تحليل فقرات الاختبار لجميع أفراد العينة، وبلغت قيمة معامل الثبات للاختبار باستخدام معامل كرونباخ α (0.90).

متغيرات الدراسة

المتغير المستقل: وهو أسلوب معالجة القيم المتطرفة المكتشفة في البيانات الخاصة باستجابات الطلبة على فقرات اختباري الرياضيات والعلوم ويتكون من ثلاثة مستويات؛ وهي: الاحتفاظ بقيم القدرة المتطرفة في مجموعة البيانات، وحذف قيم القدرة المتطرفة بمجموعة البيانات، ومعالجة قيم القدرة المتطرفة بمجموعة البيانات.

المتغيرات التابعة: معالم الفقرات (الصعوبة، التمييز، التخمين)، والخطأ المعياري كمؤشر لدقة التقدير.

المعالجات الإحصائية

للإجابة عن أسئلة الدراسة تم استخدام أسلوب الرسم الصندوقي (Box-Plot) للكشف عن القيم المتطرفة في البيانات الخاصة باستجابات الطلبة على فقرات اختباري الرياضيات والعلوم، وتم حساب الأوساط الحسابية والانحرافات المعيارية، وإجراء تحليل التباين للقياسات المتكررة للأخطاء المعيارية لمعالم (التمييز، الصعوبة، التخمين) لفقرات اختبار تيمس الدولي لدى طلبة عينة الدراسة وفقاً لأسلوب معالجة القيم المتطرفة، واستخدام اختبار (Bonferroni) للمقارنات البعدية المتعددة بين الأوساط الحسابية للخطأ المعياري لمعالم فقرات اختبار تيمس الدولي وفقاً لأسلوب معالجة القيم المتطرفة.

نتائج الدراسة

للإجابة عن سؤال الدراسة الأول؛ الذي ينص على: "ما نسبة القيم المتطرفة في درجات طلبة الصف الثامن الأساسي الأردنيين في اختبارات تيمس الدولية المنعقدة في العام 2011 في مبحتي العلوم والرياضيات؟"

تم استخدام أسلوب الرسم الصندوقي (Box Plot) في الكشف عن القيم المتطرفة في البيانات الخاصة بالدراسة، وتم اختيار عينة الدراسة التي تتكون من (526) طالباً وطالبة ممن أجابوا على الأسئلة الموضوعية في الرزمة الاختبارية ذات الرقم (11) في اختبار الرياضيات بعد حذف (5) من أفراد العينة غير المطابقين وفقاً لاحتمالية الخطأ المعياري في اختبار

الرياضيات، وتبين وجود (34) قيمة متطرفة لدرجات الطلبة في اختبار الرياضيات، وتم اختيار عينة الدراسة التي تتكون من (521) طالباً وطالبة ممن أجابوا على الأسئلة الموضوعية في الرزمة الاختبارية ذات الرقم (11) في اختبار العلوم بعد حذف (10) من أفراد العينة غير المطابقين وفقاً لاحتمالية الخطأ المعياري في اختبار العلوم، حيث تبين وجود (44) قيمة متطرفة لدرجات الطلبة في اختبار العلوم، ويبين الجدول (3) الإحصاءات الوصفية الخاصة بالرسم الصندوقي ونسبة القيم المتطرفة في اختباري الرياضيات والعلوم.

جدول (3): الإحصاءات الوصفية الخاصة بالرسم الصندوقي ونسبة القيم المتطرفة في اختباري الرياضيات والعلوم.

القيمة		الإحصائي
علوم	رياضيات	
2.23	4.00	قيمة القدرة العظمى
لا يوجد	12	عدد القيم المتطرفة فوق الحد الأعلى
	%2.26	نسبة القيم المتطرفة فوق الحد الأعلى
3.0745	3.3814	الحد الأعلى للرسم الصندوقي
0.8469	0.7591	الربيعي الثالث
1.4850	1.7482	المدى الربيعي
-0.6381	-0.9891	الربيعي الأول
-2.8657	-3.6114	الحد الأدنى للرسم الصندوقي
44	22	عدد القيم المتطرفة دون الحد الأدنى
%8.29	%4.14	نسبة القيم المتطرفة دون الحد الأدنى
-4.00	-4.00	قيمة القدرة الصغرى
%8.29	%6.40	نسبة القيم المتطرفة الكلية
%91.71	%93.6	نسبة القيم غير المتطرفة الكلية

وللإجابة عن أسئلة الدراسة الثاني والثالث والرابع ومناقشتهم؛ فقد تمت معالجة القيم المتطرفة بالاستبدال باستخدام طريقة الوسط الحسابي المبتور Trimmed Mean للقيم المتطرفة لقدرات طلبة عينة الدراسة من الصف الثامن الأساسي الأردنيين على اختبار تيمس الدولي في مادة الرياضيات والجدول (4) يبين القيم المتطرفة قبل الاستبدال وبعدها لمادة الرياضيات.

جدول (4): القيم المتطرفة لدرجات الطلبة قبل الاستبدال وبعدها لاختبار الرياضيات.

رقم الطالب	قبل الاستبدال	بعد الاستبدال	رقم الطالب	قبل الاستبدال	بعد الاستبدال
295	-3.626	-0.127	13	4.000	-0.184
299	4.000	-0.182	29	-3.719	-0.141
322	3.974	-0.184	33	-4.000	-0.180
324	-4.000	-0.185	58	-4.000	-0.180
332	-3.626	-0.120	86	4.000	-0.183
353	-4.000	-0.184	135	-4.000	-0.181
363	3.886	-0.185	149	-4.000	-0.182
371	-3.719	-0.134	157	-4.000	-0.182
373	-4.000	-0.177	159	-4.000	-0.183
379	-4.000	-0.169	190	3.802	-0.185
423	-4.000	-0.162	193	3.464	-0.185
453	4.000	-0.182	231	3.394	-0.185
476	4.000	-0.181	240	-4.000	-0.184
506	4.000	-0.180	243	-4.000	-0.184
513	-4.000	-0.155	256	-4.000	-0.185
517	4.000	-0.180	280	-4.000	-0.185
527	-4.000	-0.155	289	-4.000	-0.185

ويبين الجدول (4) أرقام الطلبة ذوي الدرجات المتطرفة، ودرجاتهم قبل الاستبدال وبعده باستخدام الوسط الحسابي المبتور Trimmed Mean حيث تم استبدال 34 قيمة متطرفة في الرياضيات.

ولإجابة أيضاً عن أسئلة الدراسة الثاني والثالث والرابع ومناقشتهم؛ فقد تمت معالجة القيم المتطرفة بالاستبدال باستخدام طريقة الوسط الحسابي المبتور Trimmed Mean للقيم المتطرفة لقدرات طلبة عينة الدراسة من الصف الثامن الأساسي الأردنيين على اختبار تيمس الدولي في مادة العلوم والجدول (5) يبين القيم المتطرفة قبل الاستبدال وبعدها لمادة العلوم.

جدول (5): القيم المتطرفة لدرجات الطلبة قبل الاستبدال وبعدها لاختبار العلوم.

رقم الطالب	قبل الاستبدال	بعد الاستبدال	رقم الطالب	قبل الاستبدال	بعد الاستبدال
294	-4.000	-0.070	29	-4.000	-0.130
320	-3.653	0.098	31	-3.935	0.037
324	-4.000	-0.063	37	-3.624	0.112
332	-3.942	0.029	82	-4.000	-0.123
342	-4.000	-0.055	91	-3.931	0.060
373	-4.000	-0.047	94	-4.000	-0.115
385	-3.640	0.105	124	-3.901	0.091
403	-3.269	0.146	128	-4.000	-0.108
405	-4.000	-0.040	131	-3.270	0.126
407	-3.257	0.159	133	-3.270	0.132
413	-3.934	0.052	144	-3.269	0.139
417	-4.000	-0.032	145	-3.251	0.172
446	-3.911	0.075	149	-3.905	0.083
450	-4.000	-0.024	192	-3.923	0.068
465	-4.000	-0.017	204	-3.232	0.178
467	-3.935	0.045	245	-4.000	-0.100
468	-3.228	0.185	247	-3.257	0.152
469	-4.000	-0.009	256	-4.000	-0.093
497	-4.000	-0.001	263	-4.000	-0.085
522	-4.000	0.006	268	-3.613	0.119
524	-4.000	0.014	283	-4.000	-0.078
531	-4.000	0.022	287	-3.255	0.165

ويبين الجدول (5) أرقام الطلبة ذوي الدرجات المتطرفة، ودرجاتهم قبل الاستبدال وبعده باستخدام الوسط الحسابي المبتور Trimmed Mean حيث تم استبدال 44 قيمة متطرفة بالعلوم. ومن الملاحظ وجود نفس الطلبة في الرياضيات والعلوم ذوي الدرجات المتطرفة مما يدل على تراجع وتدني مستوى أداءهم وخروجه عن السياق المعتاد، وكانت نتائجهم متدنية مما يدل على عدم جدية الطلبة واهتمامهم في الإجابة عن فقرات الاختبار.

وللإجابة أيضًا عن أسئلة الدراسة الثاني والثالث والرابع؛ فقد تم حساب الأوساط الحسابية والانحرافات المعيارية للخطأ المعياري لمعالم فقرات اختبار تيمس لدى طلبة عينة الدراسة بالرياضيات والعلوم وفقًا لأسلوب معالجة القيم المتطرفة كما بالجدول 6 و7.

جدول (6): الإحصاءات الوصفية للخطأ المعياري لمعالم فقرات اختبار تيمس الدولي لدى طلبة عينة الدراسة من الصف الثامن الأساسي الأردنيين في الرياضيات وفقًا لأسلوب معالجة القيم المتطرفة.

الانحراف المعياري	الوسط الحسابي	أسلوب معالجة القيم المتطرفة	الخطأ المعياري لمعلمة
0.018	0.116	احتفاظ	التمييز
0.016	0.100	حذف	
0.166	0.374	استبدال	الصعوبة
0.465	0.536	احتفاظ	
0.816	0.885	حذف	
0.517	0.590	استبدال	التخمين
0.009	0.034	احتفاظ	
0.011	0.032	حذف	
0.026	0.036	استبدال	

يلاحظ من الجدول 6 أنَّ الوسط الحسابي للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس بالرياضيات وفقًا لأسلوب الحذف بمعالجة القيم المتطرفة أصغر مما هو عليه وفقًا لأسلوبي الاحتفاظ والاستبدال، وأنَّ الانحراف المعياري للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس بالرياضيات وفقًا لأسلوب الحذف بمعالجة القيم المتطرفة أصغر مما هو عليه وفقًا لأسلوبي الاحتفاظ والاستبدال، وأنَّ الوسط الحسابي للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس بالرياضيات وفقًا لأسلوب الاحتفاظ بمعالجة القيم المتطرفة أصغر مما هو عليه وفقًا لأسلوبي الاستبدال والحذف، والانحراف المعياري للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس بالرياضيات وفقًا لأسلوب الاحتفاظ بمعالجة القيم المتطرفة أصغر مما هو عليه وفقًا لأسلوبي الاستبدال والحذف، والوسط الحسابي للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس بالرياضيات وفقًا لأسلوب الحذف بمعالجة القيم المتطرفة أصغر مما هو عليه وفقًا لأسلوبي الاحتفاظ والاستبدال، وأنَّ الانحراف المعياري للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس بالرياضيات وفقًا لأسلوب الاحتفاظ بمعالجة القيم المتطرفة أصغر مما هو عليه وفقًا لأسلوبي الحذف والاستبدال.

جدول (7): الإحصاءات الوصفية للخطأ المعياري لمعالم فقرات اختبار تيمس الدولي لدى طلبة عينة الدراسة من الصف الثامن الأساسي الأردنيين في العلوم وفقاً لأسلوب معالجة القيم المتطرفة.

الانحراف المعياري	الوسط الحسابي	أسلوب معالجة القيم المتطرفة	الخطأ المعياري لمعلمة
0.194	0.406	احتفاظ	التمييز
0.215	0.347	حذف	
0.228	0.423	استبدال	
0.122	0.232	احتفاظ	الصعوبة
1.438	0.685	حذف	
0.770	0.470	استبدال	
0.025	0.049	احتفاظ	التخمين
0.024	0.056	حذف	
0.029	0.053	استبدال	

يلاحظ من الجدول 7 أنّ الوسط الحسابي للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس بالعلوم وفقاً لأسلوب الحذف بمعالجة القيم المتطرفة هو أصغر مما هو عليه وفقاً لأسلوب الاحتفاظ والاستبدال بمعالجتها، وأنّ الانحراف المعياري للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي بالعلوم وفقاً لأسلوب الاحتفاظ بمعالجة القيم المتطرفة أصغر مما هو عليه وفقاً لأسلوب الحذف والاستبدال بمعالجتها، وأنّ الوسط الحسابي للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس بالعلوم وفقاً لأسلوب الاحتفاظ بمعالجة القيم المتطرفة هو أصغر مما هو عليه وفقاً لأسلوب الحذف والاستبدال بمعالجتها، وأنّ الانحراف المعياري للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس بالعلوم وفقاً لأسلوب الاحتفاظ بمعالجة القيم المتطرفة أصغر مما هو عليه وفقاً لأسلوب الحذف والاستبدال بمعالجتها، ويلاحظ من الجدول 7 الوسط الحسابي للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس بالعلوم وفقاً لأسلوب الاحتفاظ بمعالجة القيم المتطرفة أصغر مما هو عليه وفقاً لأسلوب الحذف والاستبدال بمعالجتها، وأنّ الانحراف المعياري للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس بالعلوم وفقاً لأسلوب الاحتفاظ بمعالجة القيم المتطرفة هو أصغر مما هو عليه وفقاً لأسلوب الاحتفاظ والاستبدال بمعالجة القيم المتطرفة.

للإجابة عن سؤال الدراسة الثاني؛ الذي ينص على: ما أثر معالجة القيم المتطرفة في أداء طلبة الصف الثامن الأساسي الأردنيين في اختبارات تيمس الدولية المنعقدة في العام 2011 في مبحثي العلوم والرياضيات على تقدير معلمة التمييز ودقتها؟

يلاحظ من الجدول (6) وجود فروق ظاهرية بين الأوساط الحسابية للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي بالرياضيات ناتجة عن اختلاف مستويات أسلوب معالجة القيم المتطرفة؛ ويهدف التحقق من جوهرية الفروق الظاهرية؛ تم إجراء تحليل التباين للقياسات المتكررة للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس بالرياضيات وفقاً لأسلوب معالجة القيم المتطرفة، كما في الجدول 8.

جدول (8): نتائج تحليل التباين للقياسات المتكررة للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس بالرياضيات وفقاً لأسلوب معالجة القيم المتطرفة.

آثار الاختبارات لـ:	مصدر التباين	مجموع المربعات	درجة الحرية	متوسط مجموع المربعات	قيمة ف المحسوبة	الدلالة الإحصائية
داخل أسلوب معالجة القيم المتطرفة	الخطأ المعياري للتمييز	0.66	2	0.33	34.46	0.00
	(الخطأ المعياري للتمييز)	0.25	26	0.01		
	الكلية	0.91	28	0.03		
بين الفقرات	الخطأ	0.12	13	0.01		
الكلية		1.03	41	0.03		

يتضح من الجدول 8 وجود فروق دالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين الأوساط الحسابية للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي في الرياضيات يُعزى لأسلوب معالجة القيم المتطرفة؛ وللكشف عن مواقع الفروق تم استخدام اختبار Bonferroni للمقارنات البعدية المتعددة بين الأوساط الحسابية للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي في الرياضيات وفقاً لأسلوب معالجة القيم المتطرفة، وذلك كما في الجدول 9.

جدول (9): نتائج اختبار Bonferroni للمقارنات البعدية المتعددة للأوساط الحسابية للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي في الرياضيات وفقاً لأسلوب معالجة القيم المتطرفة.

الخطأ المعياري للتمييز	حذف	احتفاظ	استبدال
Bonferroni	0.10	0.12	0.37
حذف	0.10		
احتفاظ	0.12	0.02	
استبدال	0.37	0.27	0.26

يتضح من الجدول (9) أن الفرق بين الوسطين الحسابيين للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي في الرياضيات قد كان أقل بفارق جوهري وفقاً لأسلوب الحذف بمعالجة القيم المتطرفة مقارنة بأسلوب الاستبدال بمعالجة القيم المتطرفة، كما يتضح من الجدول (9) أن الفرق بين الوسطين الحسابيين للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس بالرياضيات قد كان أقل بفارق جوهري وفقاً لأسلوب الاحتفاظ بمعالجة القيم المتطرفة مقارنة بأسلوب الاستبدال بمعالجة القيم المتطرفة، وكذلك يتضح من الجدول (9) أن الفرق بين الوسطين الحسابيين للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي في الرياضيات كان أقل بفارق جوهري وفقاً لأسلوب الحذف بمعالجة القيم المتطرفة مقارنة بأسلوب الاحتفاظ بمعالجة القيم المتطرفة.

وللإجابة عن سؤال الدراسة الثاني؛ يلاحظ من الجدول (7) وجود فروق ظاهرية بين الأوساط الحسابية للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي في العلوم ناتجة عن اختلاف مستويات أسلوب معالجة القيم المتطرفة؛ وبهدف التحقق من جوهرية الفروق الظاهرية؛ تم إجراء تحليل التباين للقياسات المتكررة للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي بالعلوم وفقاً لأسلوب معالجة القيم المتطرفة، وذلك كما في الجدول 10.

جدول (10): نتائج تحليل التباين للقياسات المتكررة للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي في العلوم وفقاً لأسلوب معالجة القيم المتطرفة.

آثار الاختبارات لـ:	مصدر التباين	مجموع المربعات	درجة الحرية	متوسط مجموع المربعات	قيمة ف المحسوبة	الدلالة الإحصائية
داخل أسلوب معالجة القيم المتطرفة	معلمة الخطأ المعياري للتمييز	0.06	2	0.03	6.51	0.00
	الخطأ (معلمة الخطأ المعياري للتمييز)	0.15	34	0.00		
	الكلية	0.21	36	0.01		
بين الفقرات	الخطأ	2.16	17	0.13		
	الكلية	2.37	53	0.04		

يتضح من الجدول 10 وجود فروق دالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين الأوساط الحسابية للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي بالعلوم يُعزى لأسلوب معالجة القيم المتطرفة وللكشف عن مواقع الفروق تم استخدام اختبار Bonferroni

للمقارنات البعدية المتعددة بين الأوساط الحسابية للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي بالعلوم وفقاً لأسلوب معالجة القيم المتطرفة، بالجدول 11.

جدول (11): نتائج اختبار Bonferroni للمقارنات البعدية المتعددة للأوساط الحسابية للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي في العلوم وفقاً لأسلوب معالجة القيم المتطرفة.

الخطأ المعياري للتمييز		حذف	احتفاظ	استبدال
Bonferroni	الوسط الحسابي	0.35	0.41	0.42
حذف	0.35			
احتفاظ	0.41	0.06		
استبدال	0.42	0.08	0.02	

يتضح من الجدول 11 أنّ الفرق بين الوسطين الحسابيين للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس بالعلوم أقل بفارق جوهري وفقاً لأسلوب الحذف بمعالجة القيم المتطرفة مقارنة بأسلوب الاستبدال، والفرق بين الوسطين الحسابيين للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس بالعلوم أقل بفارق جوهري وفقاً لأسلوب الحذف بمعالجتها مقارنةً بأسلوب الاحتفاظ.

وللإجابة عن سؤال الدراسة الثالث؛ الذي ينص على: ما أثر معالجة القيم المتطرفة في أداء طلبة الصف الثامن الأساسي الأردنيين في اختبارات تيمس الدولية المنعقدة في العام (2011) في مبحثي العلوم والرياضيات على تقدير معلمة الصعوبة ودقتها؟

يلاحظ من الجدول (6) وجود فروق ظاهرية بين الأوساط الحسابية للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس الدولي بالرياضيات ناتجة عن اختلاف مستويات أسلوب معالجة القيم المتطرفة؛ وللتحقق من جوهرية الفروق الظاهرية؛ تم إجراء تحليل التباين للقياسات المتكررة للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس بالرياضيات وفقاً لأسلوب معالجة القيم المتطرفة، كما بالجدول (12).

جدول (12): نتائج تحليل التباين للقياسات المتكررة للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس الدولي في الرياضيات وفقاً لأسلوب معالجة القيم المتطرفة.

الدلالة الإحصائية	قيمة ف المحسوبة	متوسط مجموع المربعات	درجة الحرية	مجموع المربعات	مصدر التباين	أثار الاختبارات لـ:
0.00	10.80	0.49	2	0.99	الخطأ المعياري للصعوبة	داخل أسلوب معالجة القيم المتطرفة
		0.05	26	1.19	الخطأ (الخطأ المعياري للصعوبة)	
		0.08	28	2.18	الكلية	
		1.06	13	13.74	الخطأ	بين الفقرات
		0.39	41	15.92		الكلية

يتضح من الجدول 12 وجود فروق دالة إحصائية عند مستوى الدلالة $\alpha=0.05$ بين الأوساط الحسابية للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس بالرياضيات يُعزى لأسلوب معالجة القيم المتطرفة؛ ولكشف مواقع الفروق تم استخدام اختبار Bonferroni للمقارنات البعدية المتعددة بين الأوساط الحسابية للخطأ المعياري لمعلمة الصعوبة لفقرات الاختبار كما بالجدول 13:

جدول (13): نتائج اختبار Bonferroni للمقارنات البعدية المتعددة للأوساط الحسابية للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس الدولي في الرياضيات وفقاً لأسلوب معالجة القيم المتطرفة.

حذف	استبدال	احتفاظ	الخطأ المعياري للصعوبة	
0.88	0.59	0.54	الوسط الحسابي	Bonferroni
			0.54	احتفاظ
		0.05	0.59	استبدال
	0.29	0.35	0.88	حذف

يتضح من الجدول 13 أنّ الفرق بين الوسطين الحسابيين للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس بالرياضيات أقل بفارق جوهري وفقاً لأسلوب الاحتفاظ بمعالجة القيم المتطرفة مقارنة بأسلوب الحذف، وأنّ الفرق بين الوسطين الحسابيين للخطأ المعياري لمعلمة

الصعوبة لفقرات اختبار تيمس بالرياضيات أقل بفارق جوهري وفقاً لأسلوب الاستبدال بمعالجة القيم المتطرفة مقارنةً بأسلوب الحذف.

وللإجابة عن سؤال الدراسة الثالث؛ يلاحظ من الجدول 7 وجود فروق ظاهرية بين الأوساط الحسابية للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس الدولي بالعلوم ناتجة عن اختلاف مستويات أسلوب معالجة القيم المتطرفة؛ وبهدف التحقق من جوهريّة الفروق الظاهرية؛ تم إجراء تحليل التباين للقياسات المتكررة للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس الدولي في العلوم وفقاً لأسلوب معالجة القيم المتطرفة، بالجدول 14.

جدول (14): نتائج تحليل التباين للقياسات المتكررة للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس الدولي في العلوم وفقاً لأسلوب معالجة القيم المتطرفة.

آثار الاختبارات لـ:	مصدر التباين	مجموع المربعات	درجة الحرية	متوسط مجموع المربعات	قيمة ف المحسوبة	الدلالة الإحصائية
داخل أسلوب معالجة القيم المتطرفة	معلمة الخطأ المعياري للسعوبة	1.85	2	0.92	1.98	0.15
	الخطأ (معلمة الخطأ المعياري للسعوبة)	15.83	34	0.47		
	الكلية	17.67	36	0.49		
بين الفقرات	الخطأ	29.65	17	1.74		
الكلية		47.33	53	0.89		

يتضح من الجدول 14 عدم وجود فروق دالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين الأوساط الحسابية للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس بالعلوم يُعزى لأسلوب معالجة القيم المتطرفة.

وللإجابة عن سؤال الدراسة الرابع؛ الذي ينص على: ما أثر معالجة القيم المتطرفة في أداء طلبة الصف الثامن الأساسي الأردنيين في اختبارات تيمس الدولية المنعقدة في العام 2011 في مبحثي العلوم والرياضيات على تقدير معلمة التخمين ودقتها؟

يلاحظ من الجدول (6) وجود فروق ظاهرية بين الأوساط الحسابية للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس الدولي بالرياضيات ناتجة عن اختلاف مستويات أسلوب معالجة

القيم المتطرفة؛ وللتحقق من جوهرية الفروق الظاهرية تم إجراء تحليل التباين للقياسات المتكررة للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس الدولي بالرياضيات وفقاً لأسلوب معالجة القيم المتطرفة، كما بالجدول (15).

جدول (15): نتائج تحليل التباين للقياسات المتكررة للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس الدولي في الرياضيات وفقاً لأسلوب معالجة القيم المتطرفة.

الدلالة الإحصائية	قيمة ف المحسوبة	متوسط مجموع المربعات	درجة الحرية	مجموع المربعات	مصدر التباين	أثار الاختبارات لـ:
0.65	0.43	0.00007	2	0.00014	الخطأ المعياري للتخمين	داخل أسلوب معالجة القيم المتطرفة
		0.00016	26	0.00415	الخطأ (الخطأ) المعياري للتخمين	
		0.00015	28	0.00429	الكلية	
		0.00057	13	0.00744	الخطأ	بين الفقرات
		0.00029	41	0.01173	الكلية	

يتضح من الجدول (15) عدم وجود فروق دالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين الأوساط الحسابية للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس الدولي بالرياضيات يُعزى لأسلوب معالجة القيم المتطرفة.

وللإجابة عن سؤال الدراسة الرابع؛ يلاحظ من الجدول 7 وجود فروق ظاهرية بين الأوساط الحسابية للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس الدولي في العلوم ناتجة عن اختلاف مستويات أسلوب معالجة القيم المتطرفة؛ وبهدف التحقق من جوهرية الفروق الظاهرية؛ تم إجراء تحليل التباين للقياسات المتكررة للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس الدولي بالعلوم وفقاً لأسلوب معالجة القيم المتطرفة، كما بالجدول 16.

جدول (16): نتائج تحليل التباين للقياسات المتكررة للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس الدولي في العلوم وفقاً لأسلوب معالجة القيم المتطرفة.

آثار الاختبارات لـ:	مصدر التباين	مجموع المربعات	درجة الحرية	متوسط مجموع المربعات	قيمة ف المحسوبة	الدلالة الإحصائية
داخل أسلوب معالجة القيم المتطرفة	معلمة الخطأ المعياري للتخمين	0.00056	2	0.00028	6.47	0.00
	الخطأ (معلمة الخطأ المعياري للتخمين)	0.00148	34	0.00004		
	الكلية	0.00204	36	0.00006		
بين الفقرات	الخطأ	0.03312	17	0.00195		
	الكلية	0.03516	53	0.00066		

يتضح من الجدول 16 وجود فروق دالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين الأوساط الحسابية للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس الدولي بالعلوم يُعزى لأسلوب معالجة القيم المتطرفة، وللكشف عن مواقع الفروق تم استخدام اختبار Bonferroni للمقارنات البعدية المتعددة بين الأوساط الحسابية للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس الدولي بالعلوم وفقاً لأسلوب معالجة القيم المتطرفة، وذلك كما بالجدول 17.

جدول (17): نتائج اختبار Bonferroni للمقارنات البعدية المتعددة للأوساط الحسابية للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس الدولي في العلوم وفقاً لأسلوب معالجة القيم المتطرفة.

حذف	استبدال	احتفاظ	الخطأ المعياري للتخمين	
0.056	0.053	0.049	الوسط الحسابي	Bonferroni
			0.049	احتفاظ
		0.0043	0.053	استبدال
	0.0036	0.0079	0.056	حذف

يتضح من الجدول 17 أنَّ الفرق بين الوسطين الحسابيين للخطأ المعياري لمعلمة التخمين

لفقرات اختبار تيمس بالعلوم أقل بفارق جوهري وفقاً لأسلوب الاحتفاظ بمعالجة القيم المتطرفة مقارنة بأسلوب الحذف والاستبدال.

مناقشة نتائج الدراسة

مناقشة النتائج المتعلقة بسؤال الدراسة؛ الذي ينص على: "ما نسبة القيم المتطرفة في درجات طلبة الصف الثامن الأساسي الأردنيين في اختبارات تيمس الدولية، TIMSS, (2011) في مادتي العلوم والرياضيات؟"

أظهرت النتائج المتعلقة بهذا السؤال وجود نسبة من القيم المتطرفة في اختبارات تيمس الدولية (TIMSS, 2011) في مادتي العلوم والرياضيات، حيث بينت نتائج السؤال وجود 34 قيمة متطرفة لقدرات الطلبة في اختبار الرياضيات بما نسبته تقريباً (6.40%) من الطلبة، منها 22 قيمة متطرفة أقل من الحد الأدنى للرسم الصندوقي بما نسبته (4.14%)، و 12 قيمة متطرفة أكبر من الحد الأعلى للرسم الصندوقي بما نسبته (2.26%)، وتبين من نتائج السؤال وجود 44 قيمة متطرفة لقدرات الطلبة في اختبار العلوم بما نسبته (8.29%) من الطلبة، منها 44 قيمة متطرفة أقل من الحد الأدنى للرسم الصندوقي بما نسبته (8.29%)، وصفر قيمة متطرفة أكبر من الحد الأعلى للرسم الصندوقي، ويلاحظ من النتائج أن هناك زيادة في نسبة القيم المتطرفة لقدرات الطلبة في اختبار العلوم عنه في اختبار الرياضيات، وتعزو الباحثة السبب في ذلك أن المدى الربيعي لمادة العلوم (1.4850) أقل من المدى الربيعي لمادة الرياضيات (1.7482)، مما يقود رياضياً إلى إمكانية وجود فرصة لظهور نسبة أكبر من القيم المتطرفة في اختبار العلوم مما هي عليه في اختبار الرياضيات، لأنه كلما كان المدى الربيعي أضيق يعطي فرصة أكبر لظهور قيم متطرفة، وأيضاً تطرف القيم لقدرات الطلبة في اختبار الرياضيات كانت موزعة على طرفي التوزيع التكراري، بسبب تجمع (93.60%) من قدرات الطلبة في مركز التوزيع التكراري، مما جعل التوزيع التكراري نوعاً ما متزن وأقرب للتوزيع الطبيعي، وبالتالي ساهم بظهور التطرف على كلا الطرفين، وقد كان الالتواء أكثر سلبية في حال اختبار العلوم مما سبب تطرف القيم لقدرات الطلبة في اختبار العلوم أقل من الحد الأدنى للرسم الصندوقي Box-Plot وإلى تجمع (91.71%) من قدرات الطلبة في الجهة اليمنى من التوزيع التكراري، مما جعل التوزيع التكراري ملتويًا نحو اليسار أكثر، الأمر الذي أدى إلى ظهور القيم المتطرفة عليه أكثر وكانت أقل من الحد الأدنى للرسم الصندوقي، لذلك فإن التطرف غير متماثل على قيم قدرات الطلبة في اختبار العلوم، وبالتالي يؤثر على الوسط الحسابي سلبيًا لكونه من الأسفل، وبذلك فإن تجمع قدرات الطلبة في الجهة اليمنى من التوزيع التكراري جعل التوزيع التكراري ملتويًا نحو اليسار، وهذا يتفق مع ما أشار إليه سيو (Seo, 2006) إلى أن أسلوب الرسم الصندوقي يزداد به عدد القيم المتطرفة عندما تكون التوزيعات ملتوية، وتعزو الباحثة أيضًا وجود نسبة من القيم المتطرفة في الطرف السفلي للتوزيع وأن نسب القيم المتطرفة متدنية إلى أن خصائص الطلبة مختلفة، ودليل على عدم جديتهم بالإجابة عن فقرات الاختبار والالتزام بتعليماته، ويدل على عدم اهتمام الطلبة وتدني تحصيلهم وتراجع مستوى أداءهم، مما أثر على مستوى أداء الطلبة العام،

وبالتالي تدني مستوى أداء الطلبة هو من أحد الأسباب الرئيسية في تراجع ترتيب الاردن دولياً في هذا الاختبار، حيث أكدت نتائج اختبار تيمس الدولي 2011 على ذلك.

مناقشة سؤال الدراسة الثاني؛ الذي ينص على: ما أثر معالجة القيم المتطرفة في أداء طلبة الصف الثامن الأساسي الأردنيين في اختبارات تيمس الدولية المنعقدة في العام 2011 في مبحتي العلوم والرياضيات على تقدير معلمة التمييز ودقتها؟

أظهرت نتائج السؤال؛ أنّ الوسط الحسابي للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي في مادة الرياضيات (0.100) وفقاً لأسلوب الحذف بمعالجة القيم المتطرفة، مما يُعنى أنّ الوسط الحسابي بحالة الحذف أصغر مما هو عليه وفقاً لأسلوبي الاحتفاظ والاستبدال، وأنّ أسلوب حذف القيم المتطرفة عاد بشكل إيجابي على الوسط الحسابي هنا، وينصح به كأفضل وأنجح أسلوب لمعالجة القيم المتطرفة في هذه الحالة، وكانت قيمة الانحراف المعياري للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي في مادة الرياضيات وفقاً لأسلوب الحذف بمعالجة القيم المتطرفة (0.016)، وهي أصغر مما هي عليه وفقاً لأسلوبي الاحتفاظ والاستبدال والاحتفاظ أي موثوقية الوسط الحسابي وفقاً لأسلوب الحذف عالية.

وأظهرت نتائج السؤال أنّ الوسط الحسابي للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي في مادة العلوم (0.347) وفقاً لأسلوب الحذف بمعالجة القيم المتطرفة، والنتيجة هنا لصالح أسلوب حذف القيم المتطرفة كأفضل أسلوب لمعالجة القيم المتطرفة بهذه الحالة، حيث كان الوسط الحسابي أقل ما يمكن بهذه الحالة، وكانت قيمة الانحراف المعياري للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي في مادة العلوم وفقاً لأسلوب الاحتفاظ بمعالجة القيم المتطرفة (0.194)، وهي أصغر مما هي عليه وفقاً لأسلوبي الاستبدال والحذف، أي أنّ موثوقية الوسط الحسابي وفقاً لأسلوب الاحتفاظ أعلى من موثوقية الوسط الحسابي للحذف.

وأظهرت أيضاً النتائج التي تم التوصل إليها بعد إجراء تحليل التباين للقياسات المتكررة واستخدام اختبار Bonferroni للمقارنات البعدية المتعددة بين الأوساط الحسابية للخطأ المعياري لمعالم فقرات اختبار تيمس الدولي في مادتي الرياضيات والعلوم وفقاً لأسلوب معالجة القيم المتطرفة، وجود فروق دالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين الأوساط الحسابية للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي في مادتي الرياضيات والعلوم، ويتضح أنّ الفرق بين الوسطين الحسابيين للخطأ المعياري لمعلمة التمييز لفقرات اختبار تيمس الدولي بمادتي الرياضيات والعلوم أقل وفقاً لأسلوب الحذف بمعالجة القيم المتطرفة مقارنة بأسلوبي الاستبدال والاحتفاظ بمعالجتها، وتتفق هذه النتائج تماماً مع نتائج الاحصاءات الوصفية السابقة الذكر.

وتعني هذه النتائج جميعها أنّ أسلوب حذف القيم المتطرفة له أثر إيجابي وأكثر دقة وثقة بالمقارنة مع أسلوبي (الاستبدال والاحتفاظ)، فهو مؤشر إيجابي وأفضل وأنجح الأساليب بمعالجة القيم المتطرفة في اختبارات تيمس الدولية (TIMSS, 2011) في مادتي الرياضيات والعلوم على الوسط الحسابي لتقديرات معلمة التمييز لفقرات ودقتها، وينصح بحذف القيم

المتطرفة كأفضل وأنجح أسلوب بديل لمعالجتها في هذه الحالة، وتعزو الباحثة السبب في أفضلية أسلوب الحذف بمعالجة القيم المتطرفة، إلى أنّ المرجعية في أفضلية أسلوب معالجة القيم المتطرفة هو شكل التوزيع العيني النظري لأي معلمة من المعالم الثلاث، فكما هو معلوم فإنّ التوزيع النظري لمعلمة التمييز هو Log Normal، وهو توزيع ملتوي، فإذا كان شكل التوزيع العيني لتلك المعالم ملتويًا فإن أنسب أسلوب لمعالجة القيم المتطرفة هو الحذف، لأنّه يقلص من ميل المعالم إلى الالتواء، ويمكن من خلاله الحصول على نتائج أفضل، وفي حالة معلمة التمييز في مادتي الرياضيات والعلوم فإن توزيع المعلمة ملتويًا لذا فإن الأفضلية تكون لأسلوب حذف القيم المتطرفة.

وتتفق هذه النتيجة في ذلك مع ما أشار إليه (Seo, 2006) أنّ القيم المتطرفة تتغير وفق أسلوب الكشف عنها وتفصيلها، وتتغير في ضوء الالتواء وقد تظهر في حال توزيع متصل السمة لأي معلمة، ولكنها حتمًا ستظهر عند التوزيعات الملتوية، فإذا كانت البيانات ملتوية فسوف تشمل قيم متطرفة، وكلما كان الالتواء متطرف أكثر ستزداد نسبة ظهور القيم المتطرفة، وتتفق هذه النتيجة مع ما أشار إليه قاسم وإسماعيل (Kassem, & Ismail, 2008) من حيث أنّ القيم المتطرفة تأتي أكثر من البيانات التي تعود إلى توزيعات غير متمائلة، أي يكون فيها التواء عال نحو اليسار أو نحو اليمين، وتتفق النتيجة مع إحدى نتائج دراسة النعيمي (Al-Naimi, 2012) التي أشارت إلى أنّه يمكن أن نستبعد القيم المتطرفة ونحذفها في بعض الأحيان خصوصًا إذا تم التعرف على أسبابها.

وتختلف هذه النتيجة مع نتائج دراسات: ألبرتا (Alberta, 2004)، ويونغ (Yong, 2013)، والنعيمي (Al-Naimi, 2012)، وقاسم وإسماعيل (Kassem, & Ismail, 2008)، ويوسف (Yousef, 2015)، حيث أوصت هذه الدراسات بالكشف عن القيم المتطرفة ومعالجتها، وعدم حذفها لأهميتها وأثرها على النتائج.

مناقشة سؤال الدراسة الثالث؛ الذي ينص على: ما أثر معالجة القيم المتطرفة في أداء طلبة الصف الثامن الأساسي الأردنيين في اختبارات تيمس الدولية المنعقدة في العام 2011 في مبحثي العلوم والرياضيات على تقدير معلمة الصعوبة ودقتها؟

أظهرت النتائج الخاصة بالوسط الحسابي للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس الدولي في مادة الرياضيات كان (0.536) وفقًا لأسلوب الاحتفاظ بمعالجة القيم المتطرفة، والنتيجة لصالح الاحتفاظ بالقيم المتطرفة كأفضل أسلوب لمعالجة القيم المتطرفة بهذه الحالة، حيث الوسط الحسابي أقل ما يمكن بهذه الحالة، وقيمة الانحراف المعياري للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس الدولي في مادة الرياضيات وفقًا لأسلوب الاحتفاظ بمعالجة القيم المتطرفة (0.465)، وهي أصغر مما هي عليه وفقًا لأسلوب الاستبدال والحذف، أي موثوقة الوسط الحسابي وفقًا لأسلوب الاحتفاظ عالية.

وأظهرت النتائج الخاصة بالوسط الحسابي للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس الدولي في مادة العلوم أنّه كان (0.232) وفقًا لأسلوب الاحتفاظ بمعالجة القيم المتطرفة، والنتيجة هنا لصالح الاحتفاظ بالقيم المتطرفة كأفضل أسلوب للتعامل مع القيم المتطرفة بهذه

الحالة، حيث الوسط الحسابي أقل ما يمكن بهذه الحالة، وقيمة الانحراف المعياري للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس الدولي في العلوم وفقاً لأسلوب الاحتفاظ بمعالجة القيم المتطرفة (0.122)، وهي أصغر مما هي عليه وفقاً لأسلوب الاستبدال والحذف، أي أنّ موثوقية الوسط الحسابي وفقاً لأسلوب الاحتفاظ عالية.

وأظهرت أيضاً النتائج التي تم التوصل إليها بعد إجراء تحليل التباين للقياسات المتكررة واستخدام اختبار Bonferroni للمقارنات البعدية المتعددة بين الأوساط الحسابية للخطأ المعياري لمعالم فقرات اختبار تيمس الدولي في مادتي الرياضيات والعلوم وفقاً لأسلوب معالجة القيم المتطرفة، وجود فروق دالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين الأوساط الحسابية للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس الدولي في مادة الرياضيات، ويتضح أنّ الفرق بين الوسطين الحسابيين للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس الدولي بمادة الرياضيات أقل وفقاً لأسلوب الاحتفاظ بمعالجة القيم المتطرفة مقارنة بأسلوب الاستبدال والحذف بمعالجتها، ويلاحظ من النتائج عدم وجود فروق دالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين الأوساط الحسابية للخطأ المعياري لمعلمة الصعوبة لفقرات اختبار تيمس الدولي بمادة العلوم يُعزى لأسلوب معالجة القيم المتطرفة، أي أسلوب معالجة القيم المتطرفة لا يؤثر في النتائج بهذه الحالة، وتتفق هذه النتائج تماماً مع نتائج الاحصاءات الوصفية السابقة الذكر.

وتعني هذه النتائج أنّ أسلوب الاحتفاظ بالقيم المتطرفة له أثر إيجابي وأكثر دقة وثقة بالمقارنة مع أسلوب (الاستبدال والحذف)، فهو مؤشر إيجابي وأفضل وأنجح الأساليب بمعالجة القيم المتطرفة في اختبارات تيمس الدولية (TIMSS, 2011) في مادة العلوم والرياضيات على الوسط الحسابي لتقديرات معلمة الصعوبة ودقتها، وينصح بالاحتفاظ بالقيم المتطرفة كأفضل وأنجح أسلوب بديل لمعالجتها في هذه الحالة.

وتعزو الباحثة السبب في أفضلية أسلوب الاحتفاظ بمعالجة القيم المتطرفة، أنّ المرجعية في أفضلية أسلوب معالجة القيم المتطرفة هو شكل التوزيع العيني لأي معلمة من المعالم الثلاث كما ذكر سابقاً، فإذا كان شكل التوزيع العيني لتلك المعالم الاعتدالية وأقرب للتوزيع الطبيعي كما في معلمة الصعوبة (حيث التوزيع النظري لمعلمة الصعوبة Normal)، فإن أنسب أسلوب لمعالجة القيم المتطرفة الاحتفاظ بها، لأنها ليست مفرطة في الالتواء والاحتفاظ بها يحافظ على الاعتدالية وعدم الإخلال بالتوزيع الطبيعي، ويمكن الحصول على نتائج أفضل، وفي حالة معلمة الصعوبة في مادتي الرياضيات والعلوم فإن توزيع المعالم كان غير ملتوي وأكثر اعتدالية لذا فإن الأفضلية لأسلوب الاحتفاظ بالقيم المتطرفة.

وتتفق هذه النتيجة مع نتيجة دراسة أوليوزي (Olewuezi, 2011) الذي أشار إلى أنّ الإحصاءات المستمدة من مجموعات البيانات التي تشمل على وجود القيم المتطرفة أثبتت ضرورة تضمينها والاحتفاظ بها في التحاليل الإحصائية لمعرفة تأثيراتها المحتملة على نتائج التحليلات الإحصائية، ومع إحدى نتائج دراسة زيمرمان (Zimmerman, 2001) الذي أكد على أنّ الاحتفاظ بالقيم المتطرفة في مجموعة البيانات ليس له أي تأثير على الاختبار

اللامعلمي، وتتفق مع إحدى نتائج دراسة جيمينز (Jimenez, 2011) الذي أشار الى انه لا يوجد تفاعل بين القيم المتطرفة والمعلم في مجموعة البيانات، أي الاحتفاظ بالقيم المتطرفة ليس له أثر سلبي على المعلم، وتتفق أيضًا مع نتائج دراسة نوفوتني (Novotny, 2001) الذي توصل إلى أن وجود القيم المتطرفة والاحتفاظ بها لم يكن لها تأثير في قرارات قبول أو رفض النموذج المستخدم، مما يدل على أنه ليس جميع القيم المتطرفة دائمًا هي بمثابة قيم مؤثرة، بينما تختلف هذه النتيجة مع نتائج دراسات ألبرتا (Albera, 2004)، ويونغ (Yong, 2013) وليو وآخرون (Liu, Wu & Zumbo, 2010) وميشيليدز (Michaelides, 2003)، حيث أكدت هذه الدراسات على ضرورة الكشف عن القيم المتطرفة ومعالجتها لما لها من تأثيرات غير مثالية ونتائج ملوثة، ولأن وجود القيم المتطرفة بمجموعة البيانات والاحتفاظ بها تساهم بزيادة الأخطاء ولها آثار سلبية على النتائج.

مناقشة سؤال الدراسة الرابع؛ الذي ينص على: ما أثر معالجة القيم المتطرفة في أداء طلبة الصف الثامن الأساسي الأردنيين في اختبارات تيمس الدولية المنعقدة في العام 2011 في مبني العلوم والرياضيات على تقدير معلمة التخمين ودقتها؟

أظهرت النتائج الخاصة بالوسط الحسابي للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس الدولي في مادة الرياضيات أنه كان (0.032) وفقاً لأسلوب الحذف بمعالجة القيم المتطرفة، أي أن أسلوب الحذف أفضل للتعامل مع القيم المتطرفة بهذه الحالة، وكانت قيمة الانحراف المعياري للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس الدولي في مادة الرياضيات وفقاً لأسلوب الاحتفاظ بمعالجة القيم المتطرفة (0.009) وهي أصغر مما هي عليه وفقاً لأسلوب الاستبدال والحذف، أي أن موثوقية الوسط الحسابي وفقاً لأسلوب الاحتفاظ كانت أعلى من موثوقية الوسط الحسابي للحذف.

وأظهرت النتائج التي تم التوصل إليها بعد إجراء تحليل التباين للقياسات المتكررة واستخدام اختبار Bonferroni للمقارنات البعدية المتعددة بين الأوساط الحسابية للخطأ المعياري لمعلم فقرات اختبار تيمس الدولي في مادة الرياضيات وفقاً لأسلوب معالجة القيم المتطرفة، عدم وجود فروق دالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين الأوساط الحسابية للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس بمادة الرياضيات يُعزى لأسلوب معالجة القيم المتطرفة، أي أسلوب معالجة القيم المتطرفة لا يؤثر بالنتائج بهذه الحالة، وتتفق هذه النتائج تماماً مع نتائج الاحصاءات الوصفية السابقة الذكر.

وتعني هذه النتيجة أن أسلوب حذف القيم المتطرفة له أثر إيجابي وأكثر دقة وثقة بالمقارنة مع أسلوب (الاستبدال والاحتفاظ)، فهو مؤشر إيجابي وأفضل وأنجح الأساليب بمعالجة القيم المتطرفة في اختبارات تيمس الدولية (TIMSS, 2011) في مادة الرياضيات على الوسط الحسابي لتقديرات معلمة التخمين للفقرات ودقتها، وتعزو الباحثة أن السبب في أفضلية أسلوب الحذف بمعالجة القيم المتطرفة، إلى أن المرجعية في أفضلية أسلوب معالجة القيم المتطرفة هو شكل التوزيع العيني النظري لأي معلمة من المعلم الثلاث، فكما هو معلوم فإن التوزيع النظري

لمعلمة التخمين هو Beta، وهو توزيع ملتوي، فإذا كان شكل التوزيع العيني ملتويًا فإن أنسب أسلوب لمعالجة القيم المتطرفة هو الحذف، لأنه يقلص من ميل المعالم إلى الالتواء، ويمكن من خلاله الحصول على نتائج أفضل، وفي حالة معلمة التخمين في مادة الرياضيات، فإن توزيع المعلمة ملتويًا لذا فإن الأفضلية تكون لأسلوب حذف القيم المتطرفة.

وتتفق هذه النتيجة مع نتائج دراسة اوسبورني واوفرباي (Osborne & Overbay, 2004) الذي أشار إلى أن حذف القيم المتطرفة يعمل على تقليل الأخطاء بصورة دالة إحصائية ويزيد من الدقة بصورة كبيرة، وتختلف هذه النتيجة مع نتائج دراسة وأديبو وجيمس وأكينومي (Adeboye, James & Akinwumi, 2016) حيث أوصت هذه الدراسة بضرورة الكشف عن القيم المتطرفة ومعالجتها، وعدم حذفها لأهميتها وأثرها على النتائج.

وأظهرت نتائج الوسط الحسابي للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس الدولي في العلوم أنه كان (0.049) وفقاً لأسلوب الاحتفاظ بمعالجة القيم المتطرفة، والنتيجة هنا لصالح أسلوب الاحتفاظ بالقيم المتطرفة كأفضل وأنجح أسلوب لمعالجة القيم المتطرفة بهذه الحالة، حيث الوسط الحسابي أقل ما يمكن بهذه الحالة، وقيمة الانحراف المعياري للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس الدولي في العلوم وفقاً لأسلوب الحذف بمعالجة القيم المتطرفة (0.024) وهي أصغر مما هي عليه وفقاً لأسلوب الاستبدال والاحتفاظ، أي أن موثوقية الوسط الحسابي وفقاً لأسلوب الحذف عالية.

وأظهرت أيضاً النتائج التي تم التوصل إليها بعد إجراء تحليل التباين للقياسات المتكررة واستخدام اختبار Bonferroni للمقارنات البعدية المتعددة بين الأوساط الحسابية للخطأ المعياري لمعالم فقرات اختبار تيمس الدولي في مادة العلوم وفقاً لأسلوب معالجة القيم المتطرفة، وجود فروق دالة إحصائية عند مستوى الدلالة ($\alpha=0.05$) بين الأوساط الحسابية للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس الدولي في مادة العلوم، ويتضح أن الفرق بين الوسطين الحسابيين للخطأ المعياري لمعلمة التخمين لفقرات اختبار تيمس الدولي بمادة العلوم أقل وفقاً لأسلوب الاحتفاظ بمعالجة القيم المتطرفة مقارنة بأسلوب الاستبدال والحذف بمعالجتها، وتتفق هذه النتائج مع نتائج الاحصاءات الوصفية السابقة الذكر.

وتعزو الباحثة سبب أفضلية أسلوب الاحتفاظ بمعالجة القيم المتطرفة بحالة معلمة التخمين بمادة العلوم لتقلص قيم المؤشرات الإحصائية في حالة الاحتفاظ مقارنة بكل من حالتي الحذف والاستبدال لمعلمة التخمين بالعلوم، لذلك تراجع أثر قيم القدرة المتطرفة على معلمة التخمين بهذه الحالة، وكان أسلوب الاحتفاظ بالقيم المتطرفة بهذه الحالة أنجح أسلوب ويعطي نتائج أفضل.

وتتفق هذه النتيجة مع إحدى نتائج دراسة دبدوب ويونس (Dabdoub, & Younis, 2006) إلى أنه ليس من الضروري أن يؤثر وجود القيمة المتطرفة والاحتفاظ بها سلباً على النتائج، بينما تختلف هذه النتيجة مع نتائج دراسات جيمينز (Jimenez, 2011)، والنور (Al-Nour, 2010)، حيث أكدت هذه الدراسات على ضرورة الكشف عن القيم المتطرفة

ومعالجتها لأن الاحتفاظ بالقيم المتطرفة بمجموعة البيانات يساهم بزيادة الأخطاء ولها آثار سلبية على النتائج.

وتلخص نتائج الدراسة أنّ أسلوب حذف القيم المتطرفة له أثر إيجابي وأنجح الأساليب بمعالجة القيم المتطرفة باختبارات تيمس الدولية (2011) بالرياضيات على الوسط الحسابي لتقديرات معالم الفقرات (التمييز والتخمين) ودقتها، وبالعلوم على الوسط الحسابي لتقديرات معلمة التمييز ودقتها، وأنّ أسلوب الاحتفاظ بالقيم المتطرفة له أثر إيجابي وأنجح الأساليب بمعالجة القيم المتطرفة في اختبارات تيمس الدولية (2011) بالعلوم على الوسط الحسابي لتقديرات معالم الفقرات (التخمين، والصعوبة) ودقتها، وبالرياضيات على الوسط الحسابي لتقديرات معلمة الصعوبة ودقتها.

التوصيات

في ضوء ما توصلت إليه الدراسة من نتائج توصي الباحثة بما يلي:

1. الاهتمام بعملية تقدير القيم المتطرفة والكشف عنها بمجموعة البيانات الإحصائية في البحوث والدراسات المستقبلية دون بتر هذه القيم.
2. قيام الباحثين بمزيد من التطبيقات العملية على البيانات بموضوع القيم المتطرفة.
3. إجراء دراسات مستقبلية أخرى عن أثر معالجة القيم المتطرفة في اختبارات ومتغيرات أخرى، وعلى مقاييس الاتجاهات، وباستخدام بيانات مولدة، ونسب قيم متطرفة مختلفة، واستخدام النظرية الكلاسيكية في القياس والنماذج الأخرى لنظرية الاستجابة للفقرة.
4. دعوة الباحثين والأوساط التربوية والمسؤولين عن الاختبارات وصانعي القرارات والجهات المسؤولة إلى الاهتمام أكثر بالاختبارات الوطنية والدولية وبطرائق الكشف عن القيم المتطرفة فيها وطرق معالجتها.
5. تحفيز الطلبة والمعلمين وإدارات المدارس على الاهتمام بالأداء بالاختبارات الوطنية والدولية، من خلال اعتبار نتائجها جزءاً له اعتباره بنظام تقييم أداء الطلبة.

References (Arabic & English)

- Adeboye, A. James, N. & Akinwumi, O. (2016). Assessment of Outlier Detection Procedures in Analysis of Regression Model. Department of Biostatistics and Epidemiology, *Pakistan Journal of Social Sciences*, University of Fort Hare, 13(3), 25-31.
- Al-Bayati, M. & Daga, D. (1996). *Determination of Extreme Values using Exploration Methods and Comparison with Informal Methods*. Unpublished Master Thesis. University of Baghdad, Iraq.
- Alberta, E. (2004). *Investigation of IRT-Based Equating Methods in the Presence of Outliers*. by Huiqin hu, University of Alberta.
- Al-Fatal, H. & Interanick, E. (2009). Explore and estimate extreme values in some nonlinear models. *Journal of Management and Economics*, Mustansiriya University, 77, 214-224.
- Al-khayat, M. (2012). The degree of achievement of the achievement test according to a single parameter rash model in the detection of the level of scientific knowledge in the mathematical skills of the eighth grade students. *Al-Aqsa University Journal*, (1) 16, 87-111.
- Allen, M. & Yen, W. (1979). *Introduction to measurement theory*. California: Brooks Cole publishing company. Monterey.
- Al-Naimi, A. (2012). Detect and manipulate extreme values in a robust manner and compare them in other ways. *Tikrit Journal of Administrative Sciences and Economics*, University of Mosul, 8 (25), 124-141.
- Al-Nour, N. (2010). Comparison of some robust methods to estimate the location parameter of some probability distributions. *Kufa Journal of Mathematics and Computers*, University of Mustansiriya, 1(1), 1-22.
- Al-Shmeiri, K. (2005). *Remove the effect of extreme values in some linear and nonlinear models*. Unpublished Master Thesis. University of Aden, Yemen.

- Al-Zoubi, M. & Talafah, A. (2006). *Statistical System SPSS v23: Understanding and Analyzing Statistical Data*. Amman: Dar Wael for printing, publishing and distribution.
- Ampanthong, P. & Suwattee, P. (2009). *A comparative Study of Outlier Detection Proceedings in Multiple Linear Regression. Proceedings of the International Multiconference of Engineers and Computer Scientists, Vol.1, IMECS, Hong Kong.*
- Baker, F.B. (2001). *The Basics of Item Response Theory. (2nd ED). College Park, MD: ERIC Clearing house on Assessment and Evaluation.* (ERIC Document Reproduction Service No. ED 458 219).
- Barnett, V. & Lewis, T. (1977). *Outlier in Statistical Data. and Sons, Third Edition.* New York: John Wiley.
- Belsley, D. Kuh, E. & Welsh, R. (2004). *Regression Diagnostics Identifying Influential Data and Sources of Collinearity.* New Jersey: Willy Inter science, Hoboken.
- Crocker, L. & Algina, J. (1986). *Introduction to classical and modern Test theory.* New York: Holt, Rinehart & Winston Inc.
- Dabdoub, M. & Younis, F. (2006). The effect of extreme values on the results of regression analysis with application to preterm infants. *Journal of Mesopotamia, University of Mosul, 17 (1), 62-81.*
- Dan, E. D. & Ijeoma, O. A. (2013a). Statistical Analysis/Methods of Detecting Outliers in A Multivariate Data in A Regression Analysis Model. *Journal of International Academic Research for Multidisciplinary, 1(3), 302-337.*
- De Gruijter, D.M. & Van Der Kamp, L.J. (2005). *Statistical test theory for education and Psychology.* New York.
- Fox, J. (1997). *Applied regression analysis, linear models, and related Methods.* McMaster University, Hamilton, Ontario, Canada, Second Edition.

- Hadi, F. (2010). Study of the effect of extreme values on the Box-Jenkins and Furer methods. *Qadisiyah Journal of Administrative and Economic Sciences*, Al-Qadisiyah University, 12(3), 121- 134.
- Hambleton, R.K. & Jonse, R.W. (1993). *Comparison of Classical Test Theory and Item Response Theory and Their Application to Test Development*. NCME items, The Instructional topics in educational measurement series, Educational Measurement, Issues and Practice, 38-47.
- Hambleton, R.K. & Swaminathan, H. (1985). *Item Response Theory: principles and applications*. Boston, MA: Kluwer -Nijhoff.
- Hamoudat, A. (2009). *The effect of extreme values and origin points on regression analysis results*. Tikrit University of Pure Sciences, University of Mosul, Iraq, 15 (1), 310- 319.
- Hattie, J. (1985). *Methodology Review: Assessing Unidimensionality of Tests and Items*. *Applied Psychological Measurement*, University of New England, Australia, 9, 139–164.
- Hawkins, D.M. (1980). *Identification of outlier*. London: Chapman and Hall.
- Hossain, D. (1989). *Detection of outliers and influential observations in regression models*. Old Dominion University, UMI 300 N. Zeeb Rd. Ann Aibor, MI 48106.
- Ismail, M. (2001). *Linear regression analysis*. Riyadh: Research Center Institute of Public Administration
- Jabouri, M. (1990). *Partial detection of extreme views and estimation methods in a multivariate state*. Unpublished Master Thesis. University of Mustansiriya, Iraq.
- Jawad, N. & Kamal, G. (2009). Comparison of methods of estimating the parameters of the regression model in the case of the problem of linear multiplicity and extreme values. *Journal of Administrative and Economic Sciences*, University of Baghdad, Iraq, 55, 166-153.

- Jimenez, F. A. (2011). *Effect of Outlier Item Parameters on IRT Characteristics Curve Linking Methods Under the Common Item Nonequivalent Groups Design*. Unpublished Doctoral Dissertation, University of Florida, Florida: USA.
- Liu, Y. & Wu, A. D. & Zumbo, B. D. (2010). *The impact of outliers on Cronbach's coefficient alpha estimate of reliability: Ordinal/Rating Scale Item Response*. *Education and Psychological Measurement*, 70(1), 5-21.
- Michaelides, M. P. (2003). *Effects of Common-item Selection on the Accuracy of Item Response Theory Test Equating with Nonequivalent Groups*. Unpublished Doctoral Dissertation, Stanford University, Stanford, California.
- Mislevy, R.J. & Bock, R.D. (1990). *Item analysis and test scoring with binary logistic models*. (2nd edition). Scientific software, Inc.
- Mukhtar, S. (1980). *Extreme values and their impact on statistical data analysis. Master's thesis unpublished in statistics*. University of Baghdad, Iraq.
- Kassem, M. & Ismail, Y. (2008). Expose extreme values in Bez style using a Gypsum preview. *Iraqi Journal of Statistical Sciences*, University of Mosul, 14, 88-68.
- Kutner, M. Neter, J. & Wasserman, W. (2005). *Applies Linear Statistical Models*, (2nd edition). New York: McGraw-Hill, U.S.A. 250-482.
- National Center for Human Resources Development, Jordan National Report on International Mathematics and Science Study 2011. (TIMSS, 2011). Amman: Monitoring and Planning Unit, Center Publication Series, 2013.
- Novotny, J. (2001). *Detection and Evaluation of Multivariate Outliers*. A Dissertation Presented to The Faculty of the Department of Psychology University of Denver.

- Olewuezi, N. P. (2011). Specification of the Outlier Univariate Model Techniques. *Journal of Mathematics and Statistics*, 7(4), 348-355.
- Osborne, J. W. & Overbay, A. (2004). *The power of outliers (and why researchers should always check for them)*. Practical Assessment, Research & Evaluation, 9(6), 1-8.
- Rahman, M. & Al Amri, K. (2011). Effect of Outlier on Coefficient of Determination. *International Journal of Education Research*, Academic Journal, 1(6). P 9.
- Seo, M.S. (2006). *A Review and Comparison of Methods for Detecting Outliers in University Data Sets*. Unpublished Master Thesis, University of Pittsburgh.
- Shaker, S. (2009). Improvement of the M-hippocampal method in estimating linear regression model parameters. *Iraqi Journal of Statistical Sciences*, University of Mosul, 16, 242-219.
- Tukey, J.W. (1962). *The Future of Data Analysis*. Ann, Math, Statist.
- Tukey, J.W. (1977). *Exploratory Data Analysis*. Addison-Wesley Redding, MA, USA.
- Yahya, M. (2009). Unusual observations in multiple linear regression and some methods of diagnosis with application. *Iraqi Journal of Statistical Sciences*, University of Mosul, 9(15), 161. -192.
- Yong, H. (2013). *Using a Linear Regression Method to Detect Outliers in IRT Common Item Equating*. University of Missouri, Columbia, USA.
- Yousef, A. (2015). *The effect of Outliers in multiple linear regression regression model parameters. A magister message that is not published*. Sudan University of Science and Technology, Khartoum.

- Zimmerman, D. W. (2001). A Note on the influence of Outliers on Parametric and Non parametric Tests. *The Journal of General Psychology*, 121(4), 391-401.