دراسات، العلوم التربوية، مجلد 44، عدد 4، ملحق 4، 2017

أثر طريقة التعويض عن الاستجابات المفقودة على دقة معادلة صور مختلفة لاختبار باستخدام نماذج نظرية الاستجابة للفقرة *

نضال كمال الشريفين**

ملخص

هدفت هذه الدراسة إلى تقصي أثر طرق تعويض القيم المفقودة على دقة المعادلة الأفقية للاختبار باستخدام نظرية الاستجابة للفقرة. ولتحقيق هدف الدراسة، تم استخدام بيانات مُولَّدة، باستخدام برنامج (WINGEN) لتوليد استجابات (1000) مُستجيب على اختبارين مكوّنَين من (50) فقرة ثُنائية الاستجابة، ومطابقة للنموذج اللوجستيّ الأُحادي المَعْلَمَة والثُلاثي المَعْلَمَة، حيث تراوحت قيم مَعْلَمَة التمييز لفقرات الاختبار بين (0.1)، و(2.0)، والصعوبة بين (3.50−)، و(3.50)، بافتراض أن قُدرات المُستجيبيّن تتوزع توزيعاً طبيعياً بوسط حسابي مقداره صفر، وانحراف معياري مقداره واحد N(0, 1). وباستخدام برنامجي (SPSS)، و(EXCEL)، تم الحصول على بيانات تتضمن استجاباتٍ مفقودة بنسبة (20%)، وتمَّت معالجة هذه الاستجابات بطرق المعالجة للقيم المفقودة وهي:(تعظيم التوقعات، والقيم التعويضية المتعددة، ودالة استجابة الفقرة). وبعد التأكد من أُحادية البُعد للبيانات باستخدام التحليل العاملي، تمت مطابقة الفقرات، والأفراد للنموذج المُستَخدَم، وقد تمَّ استبعاد عددٍ من الفقرات، والأفراد (6 فقرات، و 14 فرد) للنموذج الأحادي المعلمة، (فقرتان، و9 أفراد) للنموذج ثلاثي المعلمة، ومن ثم تمَّ تقدير معالم الفقرات والأفراد، والأخطاء المعيارية لكل منها بطريقة الأرجحية العظمى، وذلك باستخدام برنامج التحليل (BILOG-MG3)، وللكشف عن دقة المعادلة الأفقية للاختبار باستخدام نظرية الاستجابة للفقرة، باختلاف طرق المعالجة للقيم المفقودة، تم استخدام عدة محكات هي: محك الخطأ المعياري للمعادلة، ومحك الصدق التقاطعي. حيث أظهرت النتائج وجود اختلاف في تقدير قُدرات الأفراد باختلاف طرق تعويض القيم المفقودة، وأن دقة المعادلة الخطية تبلغ أكبر قيمة لها بالاقتراب من الدرجة المعيارية (صفر)، ووجود فروق دالة إحصائياً بين المتوسطات الحسابية لأخطاء المعادلة تُعزى لمتغير (طريقة تعويض القيم المفقودة) في النموذجين الأحادي والثلاثي المعلمة لصالح طريقة تعويض القيم المفقودة (EM) مقارنة بطريقة تعويض القيم المفقودة (MI)، ولصالح طريقة تعويض القيم المفقودة (RF) مقارنة بطريقة تعويض القيم المفقودة (EM). ولصالح طريقة تعويض القيم المفقودة (RF) مقارنة بطريقة تعويض القيم المفقودة (MI). وأخيراً أكدت نتائج الصدق التقاطعي، ومعاملات درابا أفضلية نموذج نظرية الاستجابة للفقرة ثُلاثي المَعْلَمَة بعد تعويض القيم المفقودة بطريقة (RF) في دقة المعادلة الخطية الأفقية.

الكلمات الدالة: الاستجابات المفقودة، تعظيم التوقعات، القيم التعويضية المتعددة، دالة الاستجابة للفقرة، دقة المعادلة الأفقية.

المقدمة

تُعدُّ الاختبارات من أهم الأدوات المستخدمة في تقدير ما يمتلكه الفرد من سمة أو قُدرة معينة وذلك عن طريق الاستجابة لمجموعة من الفقرات يُفترض أنها تمثل تلك السمة أو القُدرة، وقد يُحدد الاختبار مجالات العمل ويُوجّه اهتمامات الأفراد، كما يُحدّد مسار حياتهم اعتماداً على القُدرات التي يمتلكونها، لذلك فإنه من حق الأفراد معرفة كيفية تقويم أدائهم وقدراتهم، والحكم عليهم، ومن الواجب طمأنتهم بأن أوراقهم تُصحح دون تَحيُّز، وأنهم يخضعون لاختبارات متكافئة عندما يكون للاختبار صور متعددة، وأنه سيكون من الضروري معادلة صور الاختبارات المختلفة في صعوبة فقراتها (Green, 1981).

والمقصود بالمعادلة هي: تحويل نظام وحدات القياس الخاص بإحدى الصورتين إلى نظام وحدات القياس الخاص بالصورة الأخرى، بحيث تصبح القياسات المستمدة من علامات كل من الصورتين متكافئة بعد إجراء هذا التحويل (علام، Lord, 1980, 2005)، وعرفها كولن وبرينان (Kolen & Brennan, 2004) بأنها إجراء إحصائي لإزالة الآثار التي

**البحث مستل من أطروحة دكتوراه مقدمة من الطالب محمود البراسنه في الفصل الأول من العام الدراسي 2015/2014 بإشراف الدكتور نضال الشريفين.

* قسم علم النفس الارشادي والتربوي، جامعة اليرموك. تاريخ استلام البحث 2015/6/24، وتاريخ قبوله 2015/8/26.

©2017 عمادة البحث العلمي وضمان الجودة/ الجامعة الأردنية. جميع الحقوق محفوظة.

يتركها الفرق بين معدل مستويات الصعوبة لفقرات الاختبار، وكذلك الفروق في معدل مستويات التمييز لفقرات الاختبار، على درجات الاختبار، وذلك بين صورتين من الاختبار نفسه لجعلهما متكافئين، ويعرف هيلز وصبحية وهيرش (Hills, Subhyiah & Hirsch, 1988) معادلة الدرجات بأنها: الحصول على درجات متكافئة لأداتين تقيسان السمة نفسها، وأن الدرجات المتكافئة هي تلك الدرجات التي تقيس السمة نفسها وبمقدار متساوٍ من الثبات، ويرى هولاند وروبن (Holland & Rubin, 1982) أن عملية معادلة الاختبارات تعني: إجراء تعديلات عددية على علامات كل صورة من صور الاختبار للتعويض عن اختلاف الصعوبة من صورة لأخرى، والهدف من المعادلة هو: تعديل الفروق الناتجة عن الصعوبة بحيث تسمح باستخدام صور الاختبار بشكل متبادل. وبعد إجراء المعادلة يمكن الحصول على تقدير واحد لقدرة المفحوص بغض النظر عن صورة الاختبار التي يتقدم لها، وهناك نوعان من أنواع معادلة الاختبارات هما: المعادلة الأفقية Horizontal Equating حيث تكون الصور المختلفة للاختبار متماثلة في المحتوى، ومتقاربة في مستوى صعوبة فقراتها، وتقيس السمة نفسها في المجتمع نفسه، ويتلخص هدف المعادلة في تعديل الفروق الناجمة عن اختلاف صعوبة الفقرات في الصور المختلفة للاختبار، والمعادلة العمودية Vertical Equating، وتكون صور الاختبار مختلفة في مستوى صعوبتها، ولكنها تقيس نفس النطاق من المحتوى أو المهارات، وتكون عينات الأفراد مختلفة في مستوى قدرتها، والافتراض في مثل هذه الحالة أن القدرة، أو السمة واقعة على متصل واسع المدى أو متعدد المستويات كما هو الحال في السمات التطورية، حيث يصعب قياس المدى باختبار واحد، لذا تنشأ الحاجة إلى عدة اختبارات تقيس السمة نفسها، ولكل اختبار مستوى أو جزءٌ من المدى للمتصل، وكون هذه الاختبارات تقيس السمة نفسها، ولكن باختلاف نظام الوحدات في الاختبارات لذا تنشأ الحاجة الى معادلة هذه الاختبارات (Lissitz & Yang, 1999)، والهدف هنا هو تحويل علامات صيغ الاختبارات المصمّمة لمجتمعات الأفراد في مستويات تعليمية مختلفة إلى ميزان واحد.

وفيما يتعلق بطرق معادلة الاختبارات (Test Equating Methods)، فقد ميّزت أدبيات القياس التربوي والنفسي بين نوعين من الطرق في معادلة الاختبارات وهما: طرق المعادلة في النظرية الكلاسيكية، وطرق المعادلة بواسطة نظرية الاستجابة للفقرة، حيث أشار كروكر وألجينا (Crocker & Algina,1986) أن هناك طريقتين رئيسيتين لمعادلة الاختبارات وفق النظرية الكلاسيكية هما: المعادلة الخطية Linear Equating حيث يتم تحويل علامات إحدى صورتي الاختبار إلى تدريج علامات الصورة الأخرى، وذلك بجعل الدرجات المعيارية المناظرة لكل منها متساوية، حيث تكون العلامتان على صورتي الاختبار متكافئتين إذا كانتا تناظران العدد نفسه من الانحرافات المعيارية عن الوسط الحسابي في مجموعة معينة من المفحوصين (Angoff & Cowell, 1986). والمعادلة المئينية Equipercentile Equating، التي تعتبر أن الدرجتين اللتين يحصل عليهما المفحوص في صورتي الاختبار (Y)، (X) يمكن اعتبارهما مُتكافئتيْن إذا كانت الرتبتان المئينيتان المتناظرتان لهما بالنسبة لمجموعة معينة من الأفراد متساويتين (أي أن الدرجتين لهما نفس الرتبة المئينية).

إن معادلة الاختبارات المعتمدة على الدرجات الخام ضمن النظرية الكلاسيكية في القياس قد لا يكون مرغوباً فيه بسبب إخفاقها في تحقيق جميع شروط معادلة الاختبارات كالعدل، والمساواة، والتماثل واللاتغير (Hambleton & Swaminthan, 1985). لذلك فإن معادلة الاختبارات بواسطة نظرية الاستجابة للفقرة تحل الكثير من المشكلات التي عجزت عنها النظرية الكلاسيكية في القياس بشرط مطابقة البيانات للنموذج المستخدم (Kolen, 1981). فنظرية الاستجابة للفقرة تقوم على افتراض أن احتمال إجابة المفحوص على الفقرة إجابة صحيحة يتحدد بقُدرة المفحوص، بحيث يمكن التنبؤ بأداء الفرد في ضوء خاصية، أو خصائص مميزة لهذا الأداء تسمى بالسمات (Traits). فالاختبارات النفسية، والتربوية بعامة تفترض أن هناك سمات، أو خصائص معينة يشترك فيها جميع الأفراد، ولكنهم يختلفون في مقدارها وبالرغم من أن هذه السمات غير منظورة، إلا أنه يمكن الاستدلال على مقدارها من السلوك الملاحظ للفرد المتمثل في استجاباته على فقرات الاختبار، وهذا ما يبرر تسميتها بالسمات الكامنة (علام، 1987).

وتميّزت نظرية الاستجابة للفقرة عن نظرية القياس الكلاسيكية باستقلالية تقدير قُدرة الفرد عن عينة الفقرات التي تطبق عليه، أي أن تقدير قُدرات الأفراد متحرّر من الفقرات Item Free. واستقلالية تقديرات معالم الفقرات (الصعوبة، والتمييز) عن عينة الأفراد المستخدمة في تقدير معالم هذه الفقرات، أي أن معالم الفقرات لا تتأثر بالعينة المختارة من المجتمع Person Free. وأخيراً أنه يمكن الحصول على إحصائي الخطأ المعياري في التقدير لتقدير درجة الدقة في قياس القُدرة لكل فرد، وربما يختلف هذا الإحصائي من فردٍ إلى آخر.

وتقوم نظرية الاستجابة للفقرة على أربعة افتراضات أساسية هي: أُحادية البعد Unidimensionality: بمعنى وجود سمة

واحدة تفسر أداء الفرد في الاختبار. والاستقلال الموضعي Local Independence: بمعنى أن استجابات المفحوص على الفقرات المختلفة في الاختبار مستقلة إحصائياً عند مستوى قُدرة معين، أي لا تؤثر استجابة المفحوص عن فقرة ما في الاختبار على استجابته عن الفقرات الأخرى. ووتيرية العلاقة بين السمة، أو قدرة المفحوص واحتمال الإجابة عن الفقرة اجابة صحيحة أو ما يسمى بمنحنى الخصائص للفقرة Item Characteristic Curve (ICC) ، حيث يمكن وصف العلاقة بين تحصيل المفحوص على الفقرة، وقدرته المُقاسة على الاختبار من خلال دالة تراكمية صاعدة تعرف باسم منحنى خصائص الفقرة (ICC)، وهذا المنحنى يشير إلى أن احتمال إجابة الفقرة إجابة صحيحة يزداد بازدياد قُدرة المفحوص(Hambleton & Swaminithan, 1985)، وتوصف هذه المنحنيات عادةً بدلالة معلمة واحدة (صعوبة الفقرة) حيث يكون الاختلاف فقط في معامل الصعوبة الذي يمثل موقع الفقرة Item Location على متصل القُدرة، أو معلمتين (صعوبة الفقرة، تمييز الفقرة) حيث يكون الاختلاف في كل من الصعوبة والتمييز الذي يمثل ميل المنحنى، أو ثلاثة معالم (صعوبة الفقرة، وتمييز الفقرة، والتَّخمين)، حيث يكون الاختلاف في كل من الصعوبة، والتمييز، وكذلك في التّخمين الذي يمثل خط التقارب السفلى للمنحنى، والتحرر من السرعة **'Speediness'**، أي أن الاختبار لا يُطبّق تحت شرط السرعة في الأداء، وهذا يعني أن عامل السرعة لا يلعب دوراً في الإجابة عن فقرات الاختبار.

طرق المعادلة وفق نظرية الاستجابة للفقرة (IRT Equating)

ذكر هامبلتون وسوامنثيان وروجرز (Hambleton, Swaminthan & Rogers, 1991) ثلاث طرق لمعادلة الاختبار وفق هذه النظرية، وهي: أولاً: المعادلة باستخدام الدرجات الحقيقية (True-Score Equating)، حيث تكون الدرجة الحقيقية للمفحوص هي مجموع احتمالات الإجابة الصحيحة على جميع الفقرات التي يستجيب عليها المفحوص ($\sum_{i=1}^{n} P_i(\theta)$)، ومن عيوب هذه الطريقة أن الدرجات الحقيقية المُقدرة لا تُناظر الدرجات الخام مناظرة واحدٍ لواحدٍ، كما أن الدرجات الحقيقية المُقدرة توضع على التدريج نفسه مع الدرجات الخام (Peterson, Kolen & Hoover, 1989). ثانياً: المعادلة باستخدام علامات القُدرة (Ability Score Equating)، وهنا يتم الافتراض أن الاختبارين يقيسان القُدرة نفسها، وعند قياس معالم القُدرة، ومعالم الفقرات في الوقت نفسه لكلا الاختبارين تكون هذه المعالم على نفس التدريج. وعند تقدير معالم القُدرة، والصعوبة، والتمييز لكل اختبار على شكل منفرد فإنه ينتج علاقة خطية بين تدريج القُدرة للاختبارين. وتتم عملية المعادلة بين الاختبارين بتحديد العلاقة الخطية بين علامات القُدرة المقُدرة. ثالثاً: المعادلة باستخدام الدرجات المشاهدة (Observed Score Equating)، وعند استخدام طريقة الدرجات الحقيقية في معادلة الاختبارات فإنه لا يُمكن معادلة علامة أدنى من درجة التخمين، حيث تكون أدنى علامة في الاختبار هي الصفر، بينما أدنى علامة حقيقية هي ($\xi = \sum_{i}^{n} c_i$). حيث c_i : درجة التخمين للفقرة **i** ، وتقوم هذه الطريقة على فكرة التنبؤ بتوزيع نظري للعلامات الخام للاختبار ببناء التوزيع التكراري للعلامات الخام للاختبار للمفحوصين ذوي القُدرة **(θ)**، وعند استخراج التوزيعات النظرية للعلامات الخام المشاهدة للاختبارين يمكن إجراء المعادلة المئينية.

أما فيما يتعلق بتقييم دقة المعادلة فقد أشار كولن (Kolen, 1981) أنه لا يوجد محك واضح يُبيّن أنه الأفضل فيما يتعلق بالمقارنة بين دقة الطرق المختلفة وفاعليتها ويمكن ذكر نوعين من هذه المعايير هما: أولا: الخطأ المعياري للمعادلة (Standard Error of Equating) حيث يُعرّف بأنه: الانحراف المعياري للعلامات المحوّلة لإحدى الصورتين (Y) المناظرة لعلامة معينة على الصورة الأخرى (X) تُؤخذ من تحويل مبني على عينات مستقلة اُختيرت من مجتمعها، وتستخدم الأخطاء المعيارية كوسيلة للتعبير عن دقة المعادلة، كما تستخدم في تقدير حجم العينة المطلوب ليُحقّق مستوى معين من دقّة المعادلة، وتُستعمل للمقارنة بين طرق المعادلة وتصاميمها (Peterson, Kolen, and Hoover, 1989)، ثانيا: محك الصدق التقاطعي (Cross- Validation Criterion) هو: متوسط مربعات الانحرافات عندما يُطبّق على عيّنات مختلفة لمعرفة مدى استقرار النتائج.

وعادةً ما تتم معادلة الاختبارات من خلال افتراض أن جميع أفراد العينة قيد الدراسة سيقومون بالاستجابة عن جميع فقرات الاختبار، ولكنَّ ذلك نادرُ الحدوث، إذ إن بعض الأفراد في العينة قد لا يستجيبون على بعض الفقرات لسببٍ أو لآخر؛ مما يتسبب بظهور مشكلة ما يسمى بالقيم المفقودة Missing Values، وهي من المشكلات التي تواجه الباحثين أو القائمين على اتخاذ القرارات، إذ إن وجود القيم المفقودة في البيانات يُشكّل تهديداً لصدق النتائج، وخاصة اذا كانت القيم المفقودة بشكل يجعل العينة غير ممثّلة للمجتمع الذي أُخذت منه (Patrick, 2007) بمعنى أنه: ينتج مجموعة من البيانات لدراسة متغير الدراسة بحجم أقل من الحجم الذي خُطط له، وهذا يؤدي إلى زيادة في تقدير التباينات والتقديرات المتَحيّزة

(Little & Rubin, 2002). وهناك عدة طرق تستخدم في معالجة البيانات المفقودة ، ولكن لم يتبين من الدراسات السابقة ما هي طريقة التعويض المناسبة في معالجة القيم المفقودة عندما يكون هناك صور متعددة للاختبار، ويرغب الباحث بمعادلتها ؛ ونظراً لأهمية القرارات المترتبة على نتائج تطبيق الصور المختلفة لتلك الاختبارات كان لا بد من البحث عن الطريقة الأكثر دقة في عملية التعويض للبيانات المفقودة ، والتي تعطي أفضل النتائج في معادلة الاختبارات.

دأب الباحثون في التعامل مع القيم المفقودة على أنها خاطئة، خاصةً عندما يكون الاهتمام بتقدير قدرة المفحوص بشكل دقيق، وبالتحديد في الاختبارات التحصيلية، وربما أحياناً يتم التعامل مع تلك الفقرات بحذفها، حيث يعد ذلك أكبر تحيز في تقدير القدرة عندما تحذف الفقرة باعتبارها خاطئة (De Ayala, Plake, Impara, 2006)، ولمساعدة الباحثين في اختيار الطريقة المناسبة للتعامل مع القيم المفقودة، يجب عليه معرفة نمط Patterns وآليّة Mechanism الفَقد في البيانات. فكيفية ظهور القيم المفقودة في مجموعة البيانات يُقصد به نمط فقد القيم، حيث أشار اندرس (Enders, 2010) إلى وجود ستة أنواع من أنماط فقد القيم، وهي: النمط الاعتباطي Arbitrary Pattern وفيه تكون القيم المفقودة منتشرة بشكل عشوائي (بدون شكل معين)، والنمط وحيد المُتغير Univariate Pattern وفيه تكون القيم المفقودة متعلقة بفقرة واحدة فقط من فقرات المقياس، أو الاختبار (مُتغير واحد)، والنمط الوتيري Monotone Pattern وفيه تظهر القيم المفقودة على شكل درج بحيث إن القيم المفقودة تزداد مع ازدياد المرحلة أو مع ازدياد صعوبة الفقرة، كما يَظهر في هذا النمط أثر الهدر Attrition لأفراد العينة، ونمط وحدة عدم الاستجابة Unit Nonresponse Pattern ويحدث غالباً في البحوث المسحية، بمعنى أنه لو كان هناك ثلاثة مُتغيرات: اثنان منها بياناتهما متوفرة لجميع المُستجيبين، والمُتغير الآخر يرفض جميع المُستجيبين الإجابة عليه، ونمط البيانات المفقودة المُخطط لها The planned Missing Data Pattern وفيه يتم التخطيط لهذا الفَقد من قبل الباحثين عند عملية جمع البيانات، وتجهيز أدوات الدراسة، ونمط المُتغير الكامن Latent Variable Pattern وفيه تُفقد البيانات بسبب مُتغير كامن (خاصية) ولجميع المُستجيبين.

أما آلية فقد البيانات فيُقصد بها: الكيفيّة التي تمت من خلالها عملية فقد بعض القيم من بعض متغيرات الدراسة، حيث تُحدّد آليات فقد القيم العلاقات الاحتمالية بين الفَقد، والقيم المشاهدة في مجموعة البيانات. على افتراض أن مجتمع إحصائي حجمه (N) من الأفراد، وكان المتغير (Y) هو متغير الدراسة، و (Yi) يمثل قيمة متغير الدراسة (Y) للفرد $(i \in U)$ حيث $U = \{1,2,.....,N\}$، وعلى فرض أنه تم سحب عينة احتمالية (S) حجمها (n) من المجتمع (U)، وأن جميع قيم المتغير المساعد (X): ($X_1, X_2, ... X_n$) تكون معلومة لجميع وحدات المعاينة في المجتمع، وبوجود القيم المفقودة؛ عُرّف مؤشر القيم المفقودة (R) – Missing Values Indicator كما يلي:

$$R = \begin{cases} R_i = 1 \ , & Y_i \text{ ليست مفقودة} \\ R_i = 0 \ , & Y_i \text{ مفقودة} \end{cases}$$

لذلك يكون $R=(R_1, R_2, ..., R_n)$ يمثل نمط القيم المفقودة. وتمثل $r = \{i / i \in S, R_i = 1\}$ مجموع الاستجابات للمتغير Y.

وقد صنّف إندريز (Endres, 2010) آ(**MAR**)، القيم إلى ثلاث آليات هي: الفَقد العشوائي الكامل Missing Completely at Random (MCAR) ، حيث تكون القيم مفقودة عشوائياً بالكامل عندما تكون عملية الفَقد بسبب العشوائية المحضة، حيث إن البيانات المفقودة لا تتأثر بخصائصها أو بخصائص الأفراد، وتكون القيم المفقودة تمثل عينة عشوائية من كامل البيانات. والفَقد العشوائي Missing at Random (MAR)، وفي هذا النوع من الفَقد تتأثر البيانات المفقودة بخصائص الأفراد، ولا تتأثر بخصائص البيانات المفقودة، وهنا المستجيبون لا يمثلون عينة عشوائية. والفَقد غير العشوائي Missing Not At Random (MNAR)، حيث يكون احتمال الاستجابة يعتمد على القيم المفقودة، ولا يمكن تفسيره من خلال البيانات التي تم قياسها لجميع عناصر العينة، وتكون القيم المفقودة هنا ليس لها أي علاقة بالعشوائية، أي أن الفَقد غير عشوائي، وغير قابل للتنبؤ به من المتغيرات الأخرى.

وبعد معرفة الباحث لنمط وآليّة الفَقد في البيانات يمكنه استخدام الطريقة المناسبة لتعويض Imputation أو معالجة Handling القيم المفقودة، ونتيجة للتقدم الملموس في الأساليب، والإجراءات الإحصائية في معالجة القيم المفقودة، وتزايد الاهتمام من قبل الباحثين، والإحصائيين بأهمية معالجة القيم المفقودة بالطرق التي تُقلّل ما أمكن من التَحيُّز في تقدير معالم المجتمع بوجود القيم المفقودة، فقد تنوعت الأساليب والطرق التي تعالج القيم المفقودة، والتي تتم قبل البدء بعملية تحليل البيانات والتي يلجأ إليها الباحث للتعامل مع القيم

المفقودة لتظهر البيانات على أنها مكتملة وليس هناك فقد للبيانات (Little & Rubin, 2002).

طرق معالجة القيم المفقودة

يوضح جدول (1) الطرق المختلفة في معالجة القيم المفقودة، والتي يمكن تصنيفها إلى مجموعتين رئيسيتين: الأولى تقوم على الحذف Deletion (حذف القيم المفقودة)، والثانية تقوم على التعويض Imputation (احتساب قيم بديلة للقيم المفقودة).

جدول (1): الطرق المختلفة في معالجة القيم المفقودة

الطريقة	أشكال الطريقة	أنواع الطريقة
الطرق القائمة على الحذف	حذف الحالة	حذف الحالة الكاملة. حذف الحالة المتوفرة.
	طريقة الخطأ طريقة غير الموجودة	
الطرق القائمة على احتساب قيمة تعويضية	الطرق الصريحة	التعويض من خلال الوسط الحسابي. التعويض من خلال الانحدار.
	الطرق الضمنية	التعويض بطريقة دالة الاستجابة. خوارزمية تعظيم التوقعات. حساب قيم تعويضية متعددة. حساب قيمة تعويضية للوسط المصحح للفقرة. حساب قيمة تعويضية من توزيع مشروط. حساب قيمة تعويضية من توزيع غير مشروط.

ففيما يتعلق بالطرق القائمة على الحذف Methods Depends on Deletion، التي يلجأ فيها الباحث إلى إظهار البيانات التي تحتوي على قيم مفقودة على شكل بيانات كاملة، وذلك عن طريق حذف الحالات التي تحتوي قيم مفقودة، ولكن هذه الطرق بالرغم من بساطتها إلا أنها غالباً ما تؤدي إلى نتائج متَحيِّزة وغير فعّالة (Little & Rubin, 2002). ويندرج ضمنها ثلاث طرق أولها: طريقة حذف الحالة Case Deletion، حيث يقوم عمل هذه الطريقة على التخلص من الأفراد الذين لديهم قيم مفقودة على بعض فقرات المقياس، وإن استخدام هذه الطريقة يؤدي إلى تقليل حجم العينة، مما يؤدي إلى تعظيم الخطأ المعياري ويقلل مستوى الدلالة (Acock, 2005)، وتصنف طرق حذف الحالة إلى أسلوبين في الحذف هما: حذف الحالة الكاملة، وسُميت بهذا الاسم لأنها تُحلل بيانات المفحوصين الذين لديهم استجابات كاملة على جميع الفقرات، ولهذا الأسلوب مُحددَيْن رئيسيان، يتمثل الأول: بافتراض أن تكون البيانات مفقودة عشوائيا بالكامل (MCAR)، وإلا فإن هذا الأسلوب يؤدي إلى نتائج متَحيِّزة بسبب أن باقي العينة لا تمثل المجتمع الأصلي. والمُحدد الثاني: يتمثل في أن هذا الأسلوب يُقلل من حجم العينة المتوفرة، (Gemici, Bednarz, & Lim, 2012)، أما الأسلوب الثاني في حذف الحالة هو: حذف الحالة المتوفرة، إذ يهتم هذا الأسلوب بالأفراد الذين لديهم استجابات كاملة على أي من المتغيرين، ويتم استخدام جميع الحالات الكاملة، والتي لا تحتوي على قيم مفقودة لكلا المتغيرين (Acock, 2005).

وثاني الطرق القائمة على الحذف، هي طريقة الخطأ (Incorrect: IN)، حيث يتم التعامل مع الفقرة المفقودة (غير المجابة) على أنها إجابة خطأ، وتُدخل خطأ في نظام التصحيح (أي تأخذ القيمة صفر)، وقد أشار أيالا وبلاك وإمبارا (De-Ayala, Plak, & Impara, 2001) إلى أن استخدام هذه الطريقة في معالجة القيم المفقودة، تعطي نتائج غير دقيقة، وتقديرات متَحيِّزة لقُدرات الأفراد.

أما الطريقة الثالثة من الطرق القائمة على الحذف، فهي طريقة غير الموجودة. حيث تُعامل الفقرة المفقودة كأنها غير ممثلة، أي أن الفقرة المفقودة للفرد يتم إهمالها وعدم إدخالها ضمن تقديرات معالم الفقرات، وكذلك معالم المفحوصين (Pigott, 2001).

وفيما يتعلق بالطرق القائمة على احتساب قيمة تعويضية، حيث تعتمد هذه الطرق على تقدير قيم معينة وتعويضها بدلاً من القيم المفقودة، وذلك قبل إجراء التحليل الإحصائي للبيانات، ومن هذه الطرق: حساب قيمة تعويضية واحدة Single Imputation، حيث يُقصد بالتعويض الأُحادي للقيم المفقودة تعبئة القيمة المفقودة بقيمة بديلة عنها أُحادية، ويقوم التعويض هنا على عدم إهدار أي معلومات من أفراد العينة، والاحتفاظ بحجم العينة كما هو (Schafer & Graham, 2002). ويُمكن تصنيف الطرق القائمة على حساب قيمة

تعويضية واحدة إلى فئتين هما:

أ) الطرق الصريحة Explicit Methods. حيث يتم استخدام نظام إحصائي يُمكّن الباحث من استبدال القيم المفقودة بقيم مقدرة، ويتم ذلك بطريقتين هما:

1- حساب قيمة تعويضية من خلال الوسط Mean Imputation: وتقوم هذه الطريقة على استبدال البيانات المفقودة على الفقرة بمتوسط الدرجات المتوفرة على الفقرة، وتعويض هذا الوسط لجميع البيانات المفقودة على الفقرة، أو عن طريق حساب المتوسط الحسابي للمفحوص الواحد من خلال استجاباته على جميع فقرات الاختبار ثم تعويض هذا المتوسط بدلاً من جميع الفقرات المفقودة لدى ذلك المفحوص، وفي هذه الطريقة يتم الحفاظ على متوسط التغيّر، ولكن ذلك يؤثر على شكل التوزيع، كما أنها تخفض من التباين وكذلك تفسد الارتباطات والتباينات المشتركة مع المتغيرات الأخرى (Enders, 2010).

2- حساب قيمة تعويضية من خلال الانحدار Regression Imputation: وتعتمد هذه الطريقة على تعويض القيم المفقودة بقيم مشتقة من معادلة الانحدار بالاعتماد على متغيرات من مجموعة البيانات التي لا تحتوي على قيم مفقودة، ثم نستخدم معادلة الانحدار الناتجة للحصول على تقديرات للقيم المفقودة (Little & Rubin, 2002). ومن العيوب التي تُؤخذ على طريقة تعويض الانحدار أنها تضخم الارتباطات، والارتباطات المشتركة بين المتغيرات؛ حيث يتم أخذ القيم التي تقع تماماً على خطِّ الانحدار كتعويض للقيم المفقودة كما أنها بحاجة إلى تحديد نموذج الانحدار، والقيم التي تم التنبؤ بها من الممكن أن تزيد عن قيم المدى المنطقي للبيانات المفقودة (Enders, 2010).

ب) الطـرق الضـمنية Implicit Methods: حيث يتم الاعتماد على أفراد العينة في احتساب قيمة تعويضية للقيم المفقودة من خلالها، وتشمل هذه الطرق الأنواع الآتية:

1- حساب القيمة التعويضية بطريقة دالة الاستجابة Response Function Imputation (RF): حيث يتم تعويض القيم المفقودة للمفحوص من نفس استجاباته، ولا تتأثر استجابات المفحوص باستجابات مفحوص آخر، وفي حال تغيير أي من الفقرات لمفحوص معين فإن ذلك لا يؤثر على المفحوص الآخر، إذ إنه لا يعتمد على متوسطات الفقرات، ولا يفترض أي شيء حول معالم الفقرة، ولا يحاول أن يقدرها ولكنه يفترض وجود معلمة قُدرة (θ) (Little & Rubin, 2002). ولكي يتم إسناد قيمة للفقرة **(j)** الخاصة بالمفحوص **(i)** وذلك اعتماداً على استجابات الفقرات غير المفقودة للمفحوص وعدد فقرات الاختبار الكلي وذلك من خلال المعادلة:

$$\hat{R}(-j)_i = PM_i(j-1)$$

حيث: $\hat{R}(-j)_i$: القيمة التي سيتم وضعها مكان القيمة المفقودة للفقرة j الخاصة بالمفحوص i (وتعرف بالقيمة المتبقية Rest Score).

J: العدد الكلي للفقرات. PM_i : الوسط الحسابي للمفحوص.

القيم الناتجة (القيم المتوفرة) التي يتم الحصول عليها تختلف عن نمط الاستجابة **(0, 1)** لذلك هناك قاعدة لتعويضها مكان القيم المفقودة كما يلي: إيجاد المتوسط الحسابي لجميع القيم المتبقية للمستجيبين ثم مقارنة كل قيمة من القيم المتوفرة بالمتوسط الحسابي ، فإذا كانت القيمة المتبقية للفرد أكبر أو تساوي المتوسط الحسابي تعطى القيمة المتبقية نمط الاستجابة (1)، وإذا كانت القيمة المتبقية للفرد أقل من المتوسط الحسابي لجميع القيم المتوفرة تعطى نمط الاستجابة (0). وهذه الطريقة من الطرق التي سيتم التعامل معها في هذه الدراسة.

2- خوارزمية تعظيم التوقعات (Expectation–Maximization Algorithm: EM) وتتكون هذه الطريقة من خطوتين رئيسيتين هما: خطوة التوقُّع E-step وخطوة التعظيم M-step. ففي خطوة التوقُّع E-step يتم حساب التوقُّع للإحصائيات المحسوبة من البيانات المكتملة، وذلك للتوصل إلى التقديرات **(θ)** الحالية للمعالم، أما في خطوة التعظيم M-step يتم فيها تحديث التقديرات للمعالم من خلال استخدام أسلوب الأرجحية العظمى (MLE)، والذي يعتمد على القيم الحالية للإحصائيات المكتملة (Dempster, Laird, & Rubin, 1977).

وفي الخطوة الأولى (خطوة التوقُّع) نحتاج إلى تقدير ابتدائي لمصفوفة التباين المشترك، إذ يمكن استخدام طرق الحذف لإيجاد التقدير الابتدائي ثم تُحسب القيمة التعويضية من خلال الانحدار لتقدير القيم المفقودة. أما في الخطوة الثانية (خطوة التعظيم) فهي عبارة عن تحديث على مصفوفة التباينات المشتركة باستخدام البيانات المحسوبة من الخطوة السابقة (خطوة التوقُّع)، لذا تتولَّد تقديرات جديدة للقيم المفقودة.

تُعاد خطوتي التوقُّع والتعظيم بشكل متكرر حتى نحصل أخيراً على مصفوفتين متقاربتين (الفرق بينهما قليل)، وتكون المصفوفة الأخيرة هي مصفوفة تعظيم التوقعات حيث يمكن تحويلها إلى مصفوفة ارتباط لملاحظة الخطأ المعياري (Little& Rubin, 2002).

3- حساب قيم تعويضية متعددة Multiple

(MI) Imputation: وهي طريقة تم تطويرها كبديل للأساليب التقليدية في معالجة القيم المفقودة، حيث يتم فيها استبدال كلِّ قيمة مفقودة بمتوسط مجموعة من القيم المختارة عشوائياً، وتعدّ هذه الطريقة من الطرق الجذابة لمعالجة القيم المفقودة، إذ إنها تُعالج القيم المفقودة قبل التحليل الفعلي لها من خلال عدة قيم منسوبة حيث يتم اختيار قيمة واحدة فقط من هذه القيم، وهي تقدم درجة من العشوائية لحساب القيمة التعويضية للحصول على أخطاء معيارية غير متَحيِّزة في التحليلات الإحصائية، وذلك على عكس حساب قيمة تعويضية واحدة (Enders, 2010).

إن استخدام التعويض المتعدد (MI) يعطي نتائج معالجة أفضل من التعويض المفرد، أو الأُحادي على الرغم من أنه يتطلب مجهودا أكبر لمعالجة البيانات، وحساب التقديرات (Little & Rubin, 2002). كما أنها مناسبة عندما تكون آلية فقد البيانات تتبع آلية الفَقد العشوائي (MAR)، وهذه الطريقة هي من الطرق التي سيتم تناولها في هذه الدراسة.

4- حساب قيمة تعويضية للوسط المصحح للفقرة: حيث يتم تعويض القيم المفقودة للمفحوص من خلال استجاباته أو استجابات المفحوصين الآخرين على نفس الاختبار.

5- طريقة حساب قيمة تعويضية من توزيع مشروط: وفي هذه الطريقة يتم تكوين معادلة انحدار لكل فقرة، أو تكوين عدد من المعادلات بطرق مختلفة لنفس الفقرة، ثم اختيار أحد هذه المعادلات عشوائياً للحصول على قيمة تقديرية للقيمة المفقودة، حيث إن هذه الطريقة تمزج بين طريقة الانحدار والاختيار العشوائي.

6- طريقة حساب قيمة تعويضية من توزيع غير مشروط: حيث يتم احتساب قيمة تعويضية للقيم المفقودة للمفحوص من خلال الاختيار العشوائي لإحدى القيم من بين الاستجابات الموجودة على الفقرة للمفحوصين.

الدراسات السابقة

يُمكن تصنيف الدراسات السابقة المرتبطة بهذه الدراسة في فئتين رئيسيين هما: **أولا: الدراسات التي تناولت طرق تعويض القيم المفقودة،** ومنها دراسة وايمان (Wayman, 2003) التي هدفت لمقارنة ثلاث من طرق التعامل مع القيم المفقودة، وبينت نتائج الدراسة إلى أن طريقة حساب قيم تعويضية متعددة (MI) أعطت نتائج ذات تَحيُّز أقل من حيث الأخطاء المعيارية مقارنة مع بقية الطرق المستخدمة في الدراسة، وفي دراسة لهاوثورن وإليوت (Hawthorn & Elliott, 2004) بهدف تحديد الطريقة الأمثل للتعامل مع القيم المفقودة اعتماداً على نسبة القيم المفقودة، ولتحقيق أهداف الدراسة تم استخدام خمس طرق لمعالجة القيم المفقودة، وأظهرت نتائج الدراسة أن طريقة المتوسط الحسابي للشخص وطريقة حساب قيم تعويضية من توزيعات غير مشروطة هما الأفضل للتعامل مع القيم المفقودة، وأجرى إندرز (Enders,2004) دراسة هدفت لمعرفة أثر طريقة التعامل مع القيم المفقودة في تقدير الثبات لبيانات مقياس ليكرت، ولتحقيق أهداف الدراسة استخدم الباحث أربع طرق للتعامل مع البيانات المفقودة، وأشارت نتائج الدراسة إلى أن طريقة "خوارزمية تعظيم التوقعات" أعطت أقل تَحيُّز باستخدام محك الجذر التربيعي لمتوسط مربعات الأخطاء (RMSE)، الذي يعتمد على الفروق بين المعالم الحقيقية، والمقُدرة واقترحت الدراسة على الباحثين أن يدركوا أثر البيانات المفقودة على تقدير الثبات، واقترحت أيضا أن يتم استخدام طريقة خوارزمية تعظيم التوقُّعات (EM) لحساب قيمة تعويضية للفقرات المفقودة.

وفي دراسة ديالا وبلاك وإمبارا (De-Ayala, Plake & Impara, 2001) التي هدفت لمعرفة تأثير البيانات المفقودة على دقة تقديرات قُدرات المفحوصين في نظرية الاستجابة للفقرة، من خلال استخدام نموذج ثُلاثي المَعْلَمَة (3PL)، حيث تم استخدام طريقتين للتعامل مع البيانات المفقودة هما: الطريقة غير الموجودة (Not Present)، وطريقة "الخطأ" (Incorrect). وأظهرت النتائج الآتي: إن أكبر تَحيُّز في تقدير القُدرة يحدث عندما يتم حذف الفقرة باعتبارها "خطأ"، وإن تقديرات الطريقة "غير الموجودة" في معالجة القيم المفقودة أدق من تقديرات الطريقة "الخطأ"، وأجرى أليسون (Allison, 2006) دراسة هدفت معرفة تأثير استخدام طريقة حساب القيم التعويضية المختلفة للتعامل مع البيانات المفقودة على البيانات التصنيفية، حيث استخدم الباحث آليّة الفَقد العشوائي الكامل (MCAR)، وآليّة الفَقد العشوائي (MAR)، كما استخدم الباحث خمس طرق للتعامل مع القيم المفقودة، وأشارت النتائج إلى أنه يفضل استخدام آلية الفَقد العشوائي (MAR) بدلاً من آلية الفَقد العشوائي الكامل (MCAR) حيث أظهرت بأن طرق حساب القيم التعويضية في آلية الفَقد (MAR) تمتلك أخطاء معيارية أقل.

وأجرى فينج (Finch, 2008) دراسة هدفت إلى الكشف عن كفاءة الطرق المختلفة في معالجة القيم المفقودة لتقدير معالم الفقرة في نظرية الاستجابة للفقرة (IRT) باستخدام النموذج ثلاثي المعلمة، وقد تم استخدام نموذج نظرية الاستجابة للفقرة ثُلاثي المَعْلَمَة (3PL) في هذه الدراسة، وأظهرت النتائج الآتي: بالنسبة لمعلمة التمييز: أشارت النتائج إلى أن أداء طريقة حساب قيم تعويضية متعددة (MI) كانت

أفضل من بقية الطرق الأخرى حيث ارتبطت بنسب تَحيُّز أقل من غيرها، وأن الأخطاء المعيارية في آلية الفَقد غير العشوائي (MNAR) كانت أقلَّ منها في آلية الفَقد العشوائي (MAR)، وبالنسبة لمعلمة الصعوبة: أشارت النتائج إلى أن أداء طريقة "حساب قيم تعويضية متعددة" (MI) كانت أفضل من بقية الطرق الأخرى، وأن الأخطاء المعيارية في آلية الفَقد غير العشوائي (MNAR) كانت أقل منها في آلية الفَقد العشوائي (MAR). أما بالنسبة لمعلمة التخمين: أشارت النتائج المتعلقة إلى أن جميع طرق التعويض كانت متقاربة من حيث الأخطاء المعيارية باستثناء طريقة (الخطأ) حيث كانت أقل من أي طريقة أخرى.

وفي دراسة لبني عواد (2010)، هدفت لمقارنة طرق التعامل مع البيانات المفقودة في تقدير معالم الفقرات، وقُدرات الأفراد، حيث تمت معالجة البيانات المفقودة باستخدام ثماني طرق مختلفة من طرق معالجة القيم المفقودة، وأشارت النتائج إلى وجود فروق ذات دلالة إحصائية بين المتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة القُدرة للأفراد، وكذلك في معالم الفقرات (الصعوبة، والتمييز، والتخمين) تعزى لطريقة المعالجة، ولطول الاختبار، وللتفاعل بينهما. كما أظهرت النتائج أيضا وجود فروق ذات دلالة إحصائية بين المتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة التمييز تعزى لمتغير الطريقة والتفاعل بينهما، وعدم وجود فروق تعزى لمتغير طول الاختبار، كما أظهرت النتائج وجود فروق ذات دلالة إحصائية بين المتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة الصعوبة تعزى لمتغير الطريقة، ولمتغير طول الاختبار، والتفاعل بينهما.

وفي دراسة كوكلك وكيري (Cokluk & Kayri, 2011)، بعنوان أثر طرق تعويض القيم المفقودة على صدق وثبات المقاييس، هدفت إلى فحص ومقارنة معاملات الثبات الناتجة عن تطبيق خمس طرق من طرق التعامل مع القيم المفقودة، فقد أظهرت نتائج الدراسة أن الطرق المختلفة في معالجة القيم المفقودة تتسبب في تقليل نسب التباين المُفَسَّر للطرق المستخدمة في الدراسة. أما بالنسبة للجذور الكامنة (Eigenvalues)، ومعاملات ثبات كرونباخ ألفا للاتساق الداخلي فقد أظهرت النتائج التي تتضمن قيم مفقودة أن تعويض القيم المفقودة تسبب في تخفيض معامل كرونباخ ألفا.

وفي دراسة لجيمسي وبندنارز وليم (Gemici, Bendnarz & Lim, 2011) حول طرق التعامل مع القيم المفقودة لبيانات التعليم، والتدريب للتعليم المهني (Vocational Education and Training VET)، حيث أشارت النتائج إلى الأداء الضعيف للطرق البسيطة مثل: طريقة الحذف المزدوج Listwise Deletion، والتعويض بقيمة ثابتة، إذ إن هذه الطرق قد تؤدي إلى أحكام خاطئة، أما الطرق الحديثة في التعامل مع القيم المفقودة مثل التعويض المتعدد (MI) لا يمكن اعتبارها علاج لكل المشاكل الناتجة عن القيم المفقودة، ولكنها غالباً ما تساعد في التقليل من خطر التَحيُّز، وأوصت الدراسة باستخدام الطرق الحديثة في التعامل مع القيم المفقودة كلما أمكن ذلك.

وأجرى الدرابسة (2012) دراسة هدفت إلى بيان أثر طريقة معالجة القيم المفقودة، وطريقة تقدير قُدرات الأفراد على دقة تقدير معالم الفقرات والأفراد، وتمت معالجة القيم المفقودة بثلاث طرق، حيث أظهرت النتائج عدم وجود اختلاف في دقة تقدير معلمة الصعوبة أو معلمة التخمين يعزى لاختلاف طريقة المعالجة للقيم المفقودة، أو طريقة التقدير، أو للتفاعل بينهما، وكذلك وجود اختلاف في دقة تقدير قُدرات الأفراد يعزى لطريقة معالجة القيم المفقودة، وتشير النتائج إلى وجود اختلاف في دقة تقديرات قُدرات الأفراد يعزى للتفاعل بين متغيري طريقة المعالجة للقيم المفقودة، وطريقة التقدير.

وفي دراسة للزعبي (2013) هدفت إلى بيان أثر نسبة البيانات المفقودة، وطريقة التعويض عنها في دقة تقدير معالم الفقرات والأفراد، للنموذجين: الأُحادي المَعْلَمَة، والثنائي المَعْلَمَة. ومعالجة هذه الاستجابات بطريقتي المعالجة للقيم المفقودة: تعظيم التوقعات، والقيم التعويضية المتعددة، وأظهرت النتائج وجود أثر دالٍّ إحصائياً في تقدير معلمة الصعوبة للفقرات يعزى لكل من طريقة التعويض في النموذج الأُحادي المَعْلَمَة، ولصالح طريقة حساب قيم تعويضية متعددة، عند نسبة فَقد (5%)، كما أظهرت النتائج وجود أثر دال في دقة تقدير معلمة تمييز الفقرات في النموذج الثنائي المَعْلَمَة يعزى لطريقة التعويض ولصالح طريقة حساب قيم تعويضية متعددة، عند نسبة الفَقد (5%). في حين أظهرت النتائج وجود أثر دال في دقة تقدير قُدرات الأفراد في النموذجين: الأُحادي، والثنائي يعزى لطريقة التعويض، ولصالح طريقة حساب قيم تعويضية متعددة.

ثانيا: الدراسات التي تناولت معادلة علامات الاختبارات، ومنها: دراسة أيوب **(1994)** التي هدفت إلى المقارنة بين أربع طرق لمعادلة الاختبارات وهي: الطريقة الخطية، والطريقة المئينية (وفقا للنظرية الكلاسيكية)، وباستخدام النموذج أُحادي المَعْلَمَة وثنائي المَعْلَمَة (وفقا للنظرية الحديثة)، واستخدم معامل الصدق التقاطعي للحكم على دقة المعادلة. وأشارت نتائج الدراسة أن نماذج النظرية الحديثة في القياس كانت أكثر فاعلية من طريقتي

المعادلة الخطية والمئينية، أما فيما يتعلق بالمعادلة العمودية، والتي أجريت على المجموعات غير المتكافئة فقد أشارت النتائج إلى أن الطريقة المئينية كانت أكثر فاعلية من الطرق الأخرى المستخدمة في الدراسة، ويليها النموذج ثنائي المَعْلَمَة، ثم أُحادي المَعْلَمَة، وأخيراً الطريقة الخطية.

وأجرى وايلي (Wiley, 1999) دراسة حول معادلة نموذجين من الاختبارات متعددة الفقرات باستخدام التصميم القائم على أساس وجود جذع مشترك، ومجموعات غير متكافئة، وأشارت النتائج أن خطأ المعادلة يتأثر بشكل أكبر بقُدرات المفحوصين، وليس بتصميم الدراسة، أو إجراءات المعادلة.

وفي دراسة الشريفين (2003) التي هدفت إلى الكشف على مدى تحقق معايير الفاعلية في معادلة اختبارين، أحدهما ثنائي التدريج، والآخر متعدد التدريج وفق صور النظرية الكلاسيكية والنظرية الحديثة في القياس. أشارت نتائج المعادلة الأفقية أن نموذج نظرية الاستجابة للفقرة أُحادي المَعْلَمَة كان الأكثر فاعلية من طريقتي المعادلة الخطية والمئينية، إذ بلغت قيم معامل الصدق التقاطعي أقل ما يمكن للنموذج أُحادي المَعْلَمَة، ثم ازدادت تدريجياً في المعادلة الخطية، والمعادلة المئينية، أما وفق معيار الخطأ المعياري للمعادلة فقد كانت المعادلة الخطية أكثر فاعلية من المعادلة المئينية.

وقام فاندر (Vander, 2006) بدراسة هدفت إلى التحقّق من معادلة الاختبارات التكيفية حاسوبياً، وذلك من خلال استخدام أربع طرق لمعادلة الاختبارات (المئينية، منحنى خصائص الفقرة، الدرجات الحقيقية والدرجات المتوقعة)، وقد اعتمد الباحث على تصميم الحالة المفردة، وأشارت النتائج إلى أنّ طريقة منحنى خصائص الفقرة والمئينية تعطيان أقلَّ الأخطاء في معادلة الاختبارات، وفي نفس الوقت فان طريقة الدرجة الحقيقية، وطريقة الدرجة المتوقعة تعطيان دقة اقلَّ في المعادلة.

وفي دراسة أخرى للشريفين **(2009)** هدفت إلى معادلة درجات سبعة نماذج مختلفة لاختبار الكفاءة في اللغة الانجليزية الذي تعقده جامعة اليرموك، حيث أشارت نتائج الدراسة أن النماذج السبعة لم تكن متكافئة في أوساطها الحسابية وتبايناتها، ولمعادلتها فقد تم استخدام نموذج راش **(Rasch)**، والمعادلة الخطية، والمعادلة المئينية، وبينت نتائج الدراسة أن القدرات المتقابلة والدرجات الخام على النماذج متفاوتة، وبناءً على النتائج أوصى الباحث باستخدام قيم القدرة المتعادلة، أو الدرجات الخام المتعادلة، أو الدرجات الخام المقابلة للرتب المئينية نفسها كأساس لاتخاذ قرار النجاح، أو الرسوب.

وأجرى الكوفحي (2013) دراسة هدفت إلى أثر طريقة تصفية المموهات (التصفية العشوائية للمموهات، والتصفية التجريبية، والتصفية بالتحكيم، والتصفية المنطقية) في فقرات الاختبار من متعدد على معادلة الاختبار، وأظهرت النتائج تأثّر دقة وفاعلية المعادلة باختلاف طريقة تصفية المموهات، وعدم تفوق أي طريقة من طرق تصفية المموهات على الطرق الأخرى من حيث دقة، وفاعلية المعادلة، إذ إن تفوق أي طريقة على الأخرى اختلف باختلاف طريقة المعادلة المستخدمة، وكشفت النتائج تفوق طريقة المعادلة وفق نموذج راش على طريقتي المعادلة الخطية، والمعادلة المئينية وفق النظرية الكلاسيكية في القياس.

وأجرى الرحيل (2013) دراسة هدفت إلى الكشف عن أثر وجود أداء تفاضلي للجنس في الفقرات ذات الجدع المشترك **(anchor items)** على دقة المعادلة العمودية لاختبار اوتيس لينون للقدرة العقلية بمستوييه المتوسط، والمتقدم الصورة " **K** "، والمعدلتين للبيئة الأردنية، وقد أشارت النتائج إلى أن إجراء المعادلة في ظل حذف الفقرات ذات الجدع المشترك، وذات الأداء التفاضلي زاد من دقة المعادلة العمودية بالاستناد إلى قيم معاملات الصدق التقاطعي.

ومن خلال مراجعة الدراسات السابقة، تبين أنه لا توجد دراسات تناولت طرق معادلة الاختبار وتأثرها بطرق تعويض القيم المفقودة. لذلك فإنه من المؤمل من هذه الدراسة أن تكشف عن الطريقة الأكثر فاعلية في تعويض القيم المفقودة واستخدامها في معادلة الاختبار.

مشكلة الدراسة وأسئلتها

غالباً ما يُواجه الباحثون مشكلة البيانات المفقودة الناتجة عن تطبيق الاختبارات، والمقاييس المتنوعة، وقلّما يجد الباحث دراسة تتعامل مع عينات في التصاميم البحثية، وخاصة الاختبارات التي تطبق على عدد كبير من المفحوصين، وتستخدم فيها صور مختلفة للاختبار مثل: الاختبار الوطني في اللغة الانجليزية، واختبارات القدرات، واختبار التوفلTOEFL، واختبار الأيلتس IELTS وغيرها، إلاّ وفيها نسبة من البيانات المفقودة، ويُتوقع أن يختلف تأثير فقد البيانات على حجم المفقود بالنسبة لحجم البيانات الأصلية، وكيفيّة توزيع المفقود على مصفوفة البيانات، إلا أنَّ تجاهل التعامل مع البيانات المفقودة سينعكس سلباً على النتائج ومنها معادلة الاختبارات، لذا لا بُدّ للباحث من أن يتعامل مع تلك البيانات بطريقة موضوعية.

ولما كانت الطريقة التي يتم التعامل فيها مع البيانات المفقودة تؤثر في خصائص فقرات اختبار الاختيار من متعدد (الصعوبة، والتمييز، والتخمين)، وبالتالي تؤثر على درجات الأفراد في الاختبار، وكذلك في تقديرات القُدرة، حيث إن هناك

العديد من الاختبارات التي تُستخدم فيها صور مختلفة، فقد كان من الضروري تعديل الفروق الناجمة عن اختلاف خصائص الفقرات وقُدرات الأفراد في صور اختبار الاختيار من متعدد ومعادلتها، وكيفية التعامل مع القيم المفقودة بالطرق المناسبة، ولذلك أيضاً تم استخدام النموذج الثلاثي المعلمة الذي يعد من أكثر النماذج واقعية، والطريقة الفضلى في تقدير القدرة؛ نظراً لاحتوائه على معلمة التخمين الخاصة بالأفراد ذوي القدرة المتدنية، كما أن هذا النموذج يشتمل إضافة إلى المعالم التي يشتملها النموذج ذو المعلمتين على معلمة التخمين للفقرة، أما النموذج الأحادي المعلمة فهو أكثر النماذج استخداما لسهولة التعامل معه بالمقارنة مع النماذج الأخرى، ومن المميزات التي أعطت هذا النموذج هذه الأهمية هي أنه عندما تتطابق البيانات مع هذا النموذج فإن معالم الفقرات التي تتمثل في درجة صعوبتها يمكن تقديرها بشكل مستقل عن العينة المستخدمة، كما أن قدرات الأفراد يمكن تقديرها بشكل مستقل عن درجة صعوبة الفقرات، وبالتالي فاستخدامهما في هذه الدراسة ربما يساهم في المقارنة بينهما، وبيان أي طرق تعويض البيانات تكون أكثر فاعلية في كلا النموذجين.

لذا جاءت هذه الدراسة لفحص أثر طريقة تعويض البيانات المفقودة على دقّة معادلة الاختبار في ضوء نظرية الاستجابة للفقرة، وباستخدام بيانات مولّدة، حيث استخدم الباحثان لأغراض هذه الدراسة الطرق الآتية: حساب قيمة تعويضية باستخدام خوارزمية تعظيم التوقعات EM Algorithm، وحساب قيمة تعويضية بطريقة دالة الاستجابة Response Function Imputation (RF)، وحساب قيم تعويضية متعددة Multiple Imputation (MI).

واعتمد الباحثان على طرق نظرية الاستجابة للفقرة لمعادلة صور الاختبار باختلاف طرق التعامل مع البيانات المفقودة، وبالتحديد فإن الدراسة سعت للإجابة عن الأسئلة الآتية:

1. ما قيم القُدرة المُتعادَلة لصور الاختبار المُولَّد وفق نموذج نظرية الاستجابة للفقرة أُحادي المَعْلَمَة وثُلاثي المَعْلَمَة باختلاف طريقة تعويض القيم المفقودة فيه؟.

2. هل تختلف دقة معادلة الاختبار وفاعليتها في نموذج نظرية الاستجابة للفقرة الأُحادي المَعْلَمَة باختلاف طريقة تعويض القيم المفقودة ؟.

3. هل تختلف دقة معادلة الاختبار وفاعليتها في نموذج نظرية الاستجابة للفقرة الثُلاثي المَعْلَمَة باختلاف طريقة تعويض القيم المفقودة ؟.

أهمية الدراسة

تنبثق أهمية هذه الدراسة في جانبين: الأول نظريٌ، والثاني عملي فمن حيث الأهمية النظرية فإن الدراسة يتوقع أن تساهم في الكشف عن الطريقة الفضلى للتعامل مع القيم المفقودة باستخدام نموذجي نظرية الاستجابة للفقرة: أُحادي المَعْلَمَة، وثُلاثي المَعْلَمَة، وخاصة عندما تكون هناك صور متعددة للاختبار، عن طريق دراسة أثر طرق تعويض القيم المفقودة على دقة المعادلة الأفقية للاختبار باستخدام النموذجين: الأُحادي المَعْلَمَة، والثُلاثي المَعْلَمَة، باستخدام بيانات مولّدة تحتوي على نسبة من القيم المفقودة يتم التعويض عنها باستخدام طرق مختلفة للتعامل مع القيم المفقودة (MI, EM, RF)، حيث تعد هذه الطرق من الطرق البديلة للطرق التقليدية في التعامل مع البيانات المفقودة مثل الطرق القائمة على الحذف، وطرق التعويض بالوسط الحسابي، أو بالانحدار، أما من حيث الأهمية العملية (التطبيقية) فإن هذه الدراسة يمكن أن يستفيد منها مطورو الاختبارات الذين يلجأوون إلى نظرية الاستجابة للفقرة لحل المشكلات المتعلقة بمعادلة الاختبارات لاختيار الطريقة الأنسب في التعامل مع القيم المفقودة وفق النموذجين: الأحادي المعلمة، والثلاثي المعلمة، إذ إن وجود القيم المفقودة في البيانات يُشكّل تهديدا لصدق النتائج، وتعد معادلة صور الاختبار من المشكلات التي تواجه الباحثون أو القائمون على اتخاذ القرارات، وخاصة إذا كانت القيم المفقودة بشكل يجعل العينة غير ممثّلة للمجتمع الذي أُخذت منه؛ ونظراً لأهمية القرارات المترتبة على نتائج تطبيق الصور المختلفة للاختبارات كان لا بد من البحث عن الطريقة الأكثر دقة في عملية التعويض للبيانات المفقودة ، والتي تعطي أفضل النتائج عند معادلة الاختبارات.

مصطلحات الدراسة

الاستجابات المفقودة: وهي الاستجابات التي يتم الحصول عليها نتيجة عدم الاستجابة على بعض فقرات الاختبار، أو المقياس من قبل المفحوص، وترك هذه الإجابات فارغة، ويطلق عليها أيضاً القيم المفقودة.

القيمة التعويضية: وهي تلك القيمة التي تستبدل بها القيمة المفقودة بعد أن يتم معالجتها بإحدى طرق معالجة القيم المفقودة.

معادلة الاختبار: وهي تحويل نظام وحدات القياس الخاص بإحدى صورتي اختبار إلى نظام وحدات القياس الخاص بالصورة الأخرى، بحيث تُصبح القياسات المستمدة من علامات كل من الصورتين متكافئة بعد إجراء التحويل.

معالم الاختبار: وهي قيم إحصائية يتم تقديرها باستخدام معادلات رياضية، وتشمل كلا من معالم الأفراد (القُدرة)، ومعالم الفقرات (الصعوبة، والتمييز، والتخمين).

البيانات المُولَّدة: وهي تلك البيانات التي يتم استخلاصها بواسطة برنامج خاص لتوليد البيانات (WINGEN)، وذلك

وفق خصائص معينة يتم اختيارها من خيارات البرنامج.

المعادلة الأفقية: إحدى طرق المعادلة التي تكون فيها صور الاختبار الواحد معدة لمنطقة سلوك واحدة، وبالتالي لمجتمع محدد من الأفراد يتمثل فيه توزيع القدرة نفسه، والهدف من إجراء المعادلة هو: تعديل الفروق الناتجة عن اختلاف مستوى الصعوبة عندما تكون الصور متقاربة في مستوى صعوبتها.

معامل الصدق التقاطعي: مؤشر إحصائي يستخدم للحكم على فاعلية المعادلة، ويتم حسابه عن طريق حساب متوسط مربعات الانحرافات للعلامات المُعادَلة عندما يطبق الاختبار على عينات مختلفة لمعرفة مدى استقرار النتائج.

الخطأ المعياري للمعادلة: مؤشر إحصائي يستخدم للحكم على فاعلية المعادلة ويتم حسابه من خلال حساب الانحراف المعياري لدرجات محولة لإحدى الصورتين (Y) المناظرة لدرجة معينة على الصورة الأخرى (X)، وتؤخذ من تحويل مبني على عينات مستقلة اختيرت من مجتمعها.

محددات الدراسة

اقتصرت هذه الدراسة على:

– استخدام ثلاث طرق لتعويض القيم المفقودة هي: خوارزمية تعظيم التوقعات (RF)، والتعويض المتعدد (MI)، ودالة الاستجابة (RF)، وبالتالي لا يمكن تعميم النتائج على كل الطرق.

– استخدام نموذجي نظرية الاستجابة للفقرة أُحادي المَعْلَمَة، وثُلاثي المَعْلَمَة، وبالتالي لا يمكن تعميم النتائج على كل النماذج.

– نسبة واحدة للفقد، وهي 20%، وعلى إحدى صورتي الاختبار فقط، إذ ربما اختلاف نسب الفقد يتفاعل مع طريقة التعويض ومع النموذج.

– استخدام طريقة المعادلة الأفقية، وباستخدام نظرية الاستجابة للفقرة، إذ ربما أن اختلاف طريقة المعادلة يكون لها دور أيضاً.

مُتغيرات الدراسة

– المتغيرات المستقلة: واشتملت على طرق تعويض القيم المفقودة (EM, MI, RF)،

– المتغيرات التابعة: واشتملت على قُدرات الأفراد، والخطأ المعياري للمعادلة، ومعامل الصدق التقاطعي.

منهجية الدراسة وإجراءاتها

هدفت هذه الدراسة إلى معرفة أثر طرق تعويض القيم المفقودة على دقة المعادلة الأفقية للاختبار باستخدام نظرية الاستجابة للفقرة، ولتحقيق أهداف الدراسة تم استخدام تصميم المجموعات المُتكافئة (Equivalent Group Design EGD) الذي يتطلب وجود استجابات مجموعتين مُتكافئتيْن من الأفراد على صورتي اختبار مُتكافئتيْن بعدد معين من الفقرات الثنائية الاستجابة، وقام الباحثان بمجموعة من الإجراءات القائمة على البيانات المُولَّدة Generating Data (دراسة محاكاة)، من خلال برنامج التوليد (WINGEN3)، وقد سارت خطوات عملية توليد البيانات على النحو الآتي:

أ) باستخدام برنامج التوليد (WINGEN3)، تم توليد قُدرات مجموعتين من الأفراد (A, B) – وفق نموذج نظرية الاستجابة للفقرة أُحادي المَعْلَمَة، وثلاثي المعلمة – بلغ عددهم 1000 فرد (مُستجيب) في كل مجموعة، موزعةً توزيعاً طبيعياً (1، 0) N~θ وتحقق الباحثان من أن القُدرة تتوزع توزيعاً طبيعياً بمتوسط حسابي (صفر)، وانحراف معياري (واحد) للمجموعتين، من خلال تطبيق اختبار كولمجروف - سميرنوف Kolmogorov-Smirnov test، ويُبيّن جدول (2) ملخصاً لنتائج هذا الاختبار.

جدول (2): الاحصاءات الوصفية واختبار كولمجروف - سميرنوف (Kolmogorov-Smirnov) لقُدرات الأفراد المُولَّدة ولمجموعتي الدراسة (A, B) للنموذجين أحادي وثلاثي المعلمة

النموذج	قُدرات المجموعة	الاحصائي الوصفي					القيمة المحسوبة لاختبار (Kolmogorov-Smirnov)	الدلالة الاحصائية
		العدد	المتوسط الحسابي	الانحراف المعياري	القيمة الصُغرى	القيمة العُظمى		
أحادي المعلمة	الأولى(A)	1000	-.0359	.9807	-3.3131	3.5500	1.320	.061
	الثانية(B)	1000	.0430	1.0409	-3.5506	3.5500	1.309	.065
ثلاثي المعلمة	الأولى(A)	1000	-.215	986.9	-3.5412	32143.	2021.	18.0
	الثانية(B)	1000	321.0	2341.0	-3.6124	04503.	35151.	63.0

يُلاحظ من جدول (2) أن قيمة اختبار كولمجروف - سميرنوف (Kolmogorov-Smirnov) لمجموعتي الدراسة (A,B) في النموذج الأحادي المعلمة بَلَغتَا (1.320، 1.309) على الترتيب، وبدلالة إحصائية (0.061، 0.065)، وللنموذج ثلاثي المعلمة بلغتا (1.220، 1.3515)، على الترتيب وبدلالة إحصائية(0.081، 0.063) وهما أكبر من مستوى الدلالة الإحصائية (α = 0.05)، مما يدل على أن توزيع القُدرة لا يختلف جوهرياً عن التوزيع الطبيعي بمتوسط حسابي يساوي (صفر)، وانحراف معياري يساوي (واحد).

ب) تمَّ توليد معالم فقرات صورتي اختبار مُتكافئتيْن في كل منها (50) فقرة ثُنائية الاستجابة، وفق نموذجي الاختبار اللوجستيّ أُحادي، وثلاثي المَعْلَمَة.

ت) تم توليد استجابات 1000 فردًا على 50 فقرة اختبار وفق النموذج اللوجستيّ الأُحادي، والثلاثي المَعْلَمَة ولكل مجموعة، بعد أن تمَّ توليد قُدرات الأفراد، وكذلك معالم فقرات الاختباريْن، ليتم حفظ هذه الاستجابات في ملفين منفصلين، الملف الأول: صورة الاختبار الأول (A3, A1) يحتوي على استجابات 1000 فرد على 50 فقرة ملائمة للنموذج اللوجستيّ الأُحادي، والثلاثي المَعْلَمَة، والملف الثاني: صورة الاختبار الثاني (B) يحتوي على استجابات 1000 فرد على 50 فقرة ملائمة للنموذج اللوجستيّ الأُحادي، والثلاثي المَعْلَمَة.

ث) تم أخذ استجابات أفراد مجموعتي الدراسة (A, B)، والذي يمثل كل منهما استجابة (1000) فرد على صورتي اختباريْن متكافئيْن في كل منهما (50) فقرة، بعد إعادة ترتيب الفقرات المتكافئة في الصورتيْن، ليتم وضعهما في ملفين منفصلين، وإعادة تحليلهما باستخدام برنامج (BILOG-MG3)، لتقدير قُدرات الأفراد وصعوبة الفقرات، بعد مطابقة الاستجابات عن فقرات كل صورة من صورتي الاختباريْن لنموذج نظرية الاستجابة للفقرة أُحادي وثلاثي المَعْلَمَة.

ج) تم معالجة فَقْد القيم باستخدام برنامج التحليل (SPSS)، وبرنامج الجداول الحسابية (EXCEL)، تمت عملية فَقْد الاستجابات وبنسبة فَقْد (20%) فقط لفقرات اختبار (B). ليتم بعد ذلك إعداد ستة ملفات تحتوي على القيم المفقودة، الملف الأول: سُمّي (EM1PL): وتُعالج فيه القيم المفقودة بطريقة خوارزمية تعظيم التوقعات (EM) للبيانات الخاصة بالنموذج أُحادي المَعْلَمَة. والملف الثاني: سُمّي (MI1PL): وتُعالج فيه القيم المفقودة بطريقة حساب قيم تعويضية متعددة (MI) الموجودة أيضاً في برنامج للبيانات الخاصة بالنموذج أُحادي المَعْلَمَة. والملف الثالث: سُمّي (RF1PL): وتُعالج فيه القيم المفقودة بطريقة حساب قيم تعويضية باستخدام دالة الاستجابة (RF) للبيانات الخاصة بالنموذج أُحادي المَعْلَمَة: والملف الرابع: سُمّي (EM3PL): وتُعالج فيه القيم المفقودة بطريقة خوارزمية تعظيم التوقعات (EM) للبيانات الخاصة بالنموذج ثُلاثي المَعْلَمَة، والملف الخامس: سُمّي (MI3PL): وتُعالج فيه القيم المفقودة بطريقة حساب قيم تعويضية متعددة (MI) للبيانات الخاصة بالنموذج ثُلاثي المَعْلَمَة. والملف السادس: سُمّي (RF3PL): وتُعالج فيه القيم المفقودة بطريقة حساب قيم تعويضية باستخدام دالة الاستجابة (RF) للبيانات الخاصة بالنموذج ثُلاثي المَعْلَمَة؛ لينتج – بعد عملية الفَقد، والمعالجة بالطرق الثلاث – ستة ملفات كاملة الاستجابات دون وجود أيٍّ من القيم المفقودة فيها. وبذلك تكون هذه الملفات الست قد أصبحت جاهزة ليتم قراءتها بواسطة برنامج (BILOG- MG) في مرحلةٍ لاحقة.

ح) تم التحقق من افتراضات نموذج استجابة الفقرة الأُحادي، والثُلاثي المَعْلَمَة لبيانات الملفات الست (EM1PL, MI1PL, RF1PL, EM3PL, MI3PL, RF3PL) والملفين (A1, A3): وقد تحقق الباحثان من الافتراضات التي تتطلبها نظرية الاستجابة للفقرة في طرق المعالجة الثلاث كما يلي:

أولاً: التحقق من افتراض أُحادية البُعد Unidimensionality: تم التحقق من افتراض أُحادية البُعد – الذي يدلّ على وجود قُدرة واحدة تُفسّر أداء المُستجيب في الاختبار – من خلال عدة مؤشرات: **أولها** مؤشر معامل الارتباط بوينت– بايسيريال الذي يمثّل معامل ارتباط بين الأداء على الفقرة والأداء على الاختبار الكلي بعد حذف الفقرة corrected item total correlation، حيث أشار هيتي (Hattie, 1985) أن الاختبار الذي تكون معظم فقراته ذات معاملات ارتباط بوينت بايسيريال ا تزيد عن (0.20) يعَدُّ مُؤشراً على أُحادية البُعد، حيث تم حذف أربع فقرات معامل ارتباط بوينت بايسريال لها أقل من (0.20) من النموذج أحادي المعلمة وهي ذات الأرقام (8، 21، 39، 47) من الملفات الثلاثة ومن الملف (A1)–، وفقرتين من النموذج ثلاثي المعلمة ذات الأرقام (39، 44) من الملفات الثلاثة الأخرى ومن الملف (A3). **وثانيها:** معامل ثبات الاتساق الداخلي (كرونباخ ألفا)، حيث اعتبر هيت (Hattie, 1985)، بأنه أكثر مؤشرات أُحادية البُعد استخداماً. وتم تقدير معاملات الثبات باستخدام معادلة (كرونباخ ألفا α) لفقرات الاختبار الثاني (B) في الملفات الست، ولفقرات صورة الاختبار الأول (A)، وجدول (7) يبين ذلك.

جدول (7): معاملات ثبات الاتساق الداخلي (كرونباخ ألفا) لفقرات الاختبار الثاني صورة الاختبار الثاني (B) في الملفات الست (EM1PL, MI1PL, RF1PL, EM3PL, MI3PL, RF3PL) ولفقرات صورة الاختبار الأول (A)

النموذج	اسم الملف	معامل ثبات الاتساق الداخلي (كرونباخ ألفا)
	EM1PL	0.90
	MI1PL	0.79
أحادي المعلمة	**RF1PL**	0.93
	A1	0.84
	EM3PL	0.92
ثلاثي المعلمة	**MI3PL**	083
	RF3PL	0.95
	A3	0.86

يُلاحظ من جدول (7) أن قيم معاملات ثبات الاتساق الداخلي تراوحت بين (0.79)، و (0.95) وجميعها قيم عالية، مما يدلُّ على تحقُّق أُحادية البُعد. **وثالثها:** مؤشرات تعتمد على أسلوب التحليل العاملي (Factor Analysis): حيث تم استخدام التحليل العاملي Factor Analysis، بأسلوب المُكوِّنات الأساسية، وذلك للاستجابات على فقرات صورة الاختبار الثاني (B) في الملفات الست، ولفقرات صورة الاختبار الأول (A1) للنموذج أحادي المعلمة، و(A3) للنموذج ثلاثي المعلمة، وقد تم إيجاد قيم الجذور الكامنة Eigenvalues، ونِسَب التباين المُفَسَّر لكل عامل من العوامل، وجرى تدوير تلك العوامل المُستخلصة – التي كانت قيمة الجذر الكامن لها أكبر من (واحد)– بطريقة التدوير المتعامد Varimax-Rotation، والملحق (1) يُبيّن نتائج التحليل، حيث تبين أن أُحادية البعد قد تحققت في صورة الاختبار الثاني (B) في الملفات الست (EM1PL, MI1PL, RF1PL, EM3PL, MI3PL, RF3PL) ولفقرات صورة الاختبار الأول (A1, A3) في النموذجين الأحادي المعلمة، والثلاثي المعلمة؛ وذلك اعتماداً على محك الجذر الكامن كمؤشر على أُحادية البعد، فكما هو واضح من الملحق (1) أنه بالرغم من أن نسبة التباين المُفَسَّر للعامل الأول ليست مرتفعة، وخاصة في النموذج الأحادي المعلمة – أقل من 20%–، إلا أن الجذر الكامن للعامل الأول يطغى بشكل جليّ على الجذور الكامنة لبقية العوامل، حيث إن نسبة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني تزيد عن (2). كما ويُلاحظ أن نسبة التباين المُفَسَّر لكل من العوامل المتبقية متقاربة جداً، أي أن هناك شبه استقرار في نسب التباين المُفَسَّرة للعوامل جميعها، وهذا يعني تحقق أُحادية البعد في بيانات هذا الاختبار.

ثانياً: التحقق من مطابقة الأفراد، والفقرات لنموذج نظرية الاستجابة للفقرة الأُحادي المَعْلَمَة، والثُلاثي المَعْلَمَة، حيث تم تحليل استجابات أفراد عينة الدراسة الذين أجابوا عن فقرات صورة الاختبار الثاني (B) في الملفات الست باستخدام برنامج (BILOG-MG3) المصمم لتحليل البيانات وفقاً لنموذج نظرية الاستجابة للفقرة أُحادي، وثُلاثي المَعْلَمَة. حيث أظهرت نتائج التحليل عدم وجود أي فقرة غير مطابقة في صورة الاختبار الثاني (B) في الملفات الست، ووجود (14، 9) أفراد غير مطابقين لنموذجي نظرية الاستجابة للفقرة أُحادي، وثُلاثي المَعْلَمَة على الترتيب، ليتم حذفهم أيضاً من صورة الاختبار الأول للنموذج الأحادي المعلمة (A1)، ومن صورة الاختبار الأول للنموذج الثلاثي المعلمة (A3). وبذلك أصبح الاختبار الأول الصورة (A1)، والاختبار الثاني الصورة (B) مكونين من (46) فقرة لكل منهما يجيب على كل منهما (986) فرداً وذلك في النموذج الأحادي المعلمة، و (48) فقرة يجيب عليها (991) فرداً وذلك في النموذج الثلاثي المعلمة.استخدام برنامج (BILOG-MG3) لتقدير قُدرات الأفراد في الملفات الست بالإضافة إلى ملف صورة الاختبار الأول (A1, A3)، وتقدير معالم الفقرات، والأفراد في الملفات الست وفق نموذجي نظرية الاستجابة للفقرة أُحادي، وثُلاثي المَعْلَمَة، بعد التأكد من مطابقة الأفراد، والفقرات لنموذج نظرية الاستجابة للفقرة أُحادي، وثُلاثي المَعْلَمَة.

خ) تم معادلة قيم تقديرات قُدرات الأفراد في صورة الاختبار الثاني(B) إلى قيم تقديرات القُدرة للأفراد في صورة

الاختبار الأول (A) باستخدام المعادلة التالية المشار إليها في رايت وماسترز (Wright & Masters, 1982):

$$\theta_{eq} = \frac{\sigma_{eq}}{\sigma_A}\left(\theta_A - \overline{\theta_A}\right) + \overline{\theta_{eq}} \quad \ldots\ldots\ldots\ldots\ldots(1)$$

حيث: θ_{eq}: قُدرات الأفراد في نموذج الاختبار المعادل إلى علامات الأفراد في صورة الاختبار (A).

σ_A: الانحراف المعياري المعدّل من الخطأ لتقديرات القُدرة للأفراد في نموذج الاختبار A.

θ_A: تقديرات القُدرة في نموذج الاختبار A.

$\overline{\theta_{eq}}$: متوسط تقديرات القُدرة للأفراد في نموذج الاختبار المعادل.

$\overline{\theta_A}$: متوسط تقديرات القُدرة للأفراد في نموذج الاختبار A.

σ_{eq}: الانحراف المعياري المعدّل من الخطأ لتقديرات القُدرة للأفراد في نموذج الاختبار المعادَل.

ويتم حسابه من خلال المعادلة:

$$\sigma_{eq} = \sqrt{\sigma_p^2 - MSE_p} \quad \ldots\ldots\ldots\ldots\ldots(2)$$

حيث: MSE_p: متوسط مربعات الأخطاء المعيارية في قياس قُدرات الأفراد (Mean Square Measurement Error) ويحسب من العلاقة التالية:

$$MSE_p = \frac{\sum_{n=1}^{N} S_n^2}{N} \quad \ldots\ldots\ldots\ldots\ldots(3)$$

حيث: *Sn*: الخطأ المعياري في تقدير قُدرات الأفراد.

N: عدد الأفراد.

نتائج الدراسة ومناقشتها

النتائج المتعلقة بالسؤال الأول الذي ينصُّ على: "ما قيم القُدرة المُعادَلة لصورة الاختبار المُولَّد وفق نموذج نظرية الاستجابة للفقرة أُحادي المَعْلَمَة، وثلاثي المعلمة باختلاف طريقة تعويض القيم المفقودة فيه؟ للإجابة عن هذا السؤال تم وفق نظرية الاستجابة للفقرة تقدير قُدرات الأفراد في صورة الاختبار الأول (A)، وكذلك تم تقدير قُدرات الأفراد في صورة الاختبار الثاني (B) باختلاف طريقة معالجة القيم المفقودة فيه (EM, MI, RF)، وذلك بعد حذف الأفراد، والفقرات غير المطابقة لنموذجي نظرية الاستجابة للفقرة الأُحادي المَعْلَمَة. ويُبيّن الملحق (2) تقدير القُدرة بوحدة اللوجيت للأفراد، وقيم الأخطاء المعيارية في قياسها، وفقاً للنموذج أُحادي المَعْلَمَة. ويُبيّن ملحق (3) تقدير القُدرة بوحدة اللوجيت للأفراد، وقيم الأخطاء المعيارية في قياسها، وفقاً للنموذج ثُلاثي المَعْلَمَة.

حيث يلاحظ من الملحقين (2،3) اختلاف تقدير قُدرات الأفراد باختلاف طرق تعويض القيم المفقودة، كما نلاحظ أن مدى الدرجات الخام في طريقة التعويض المتعدد (MI) أقل من المدى في طريقتي (EM, RF)، وفي كلا النموذجين، حيث تراوحت بين 7 و 41، في حين كان مدى القدرة لهذه الطريقة أعلى منه لباقي الطرق المستخدمة وفي كلا النموذجين، وربما يعود ذلك لعدد مرات التعويض التي تتم في طريقة (MI) وأخذ المتوسط الحسابي للقيم الناتجة مما يخفف من أثر القيم المتطرفة، وبالتالي يقلل المدى، وهذا يؤكد أن طرق التعويض المختلفة تقوم على أسس رياضية مختلفة.

كما يلاحظ أن أعلى مدى للعلامات الخام يعود لطريقة التعويض (RF)، ولكلا النموذجين المستخدمين (الأُحادي المَعْلَمَة، والثُلاثي المَعْلَمَة) حيث تراوحت الدرجات الخام لهذه الطريقة بين4 و 44، في حين أن مدى القدرة لهذه الطريقة أقل منه لباقي الطرق إذ تراوح مدى القدرة بين 2.3501− و 2.7853. إذ إن طريقة تعويض القيم المفقودة بطريقة (RF) أعطت دقة أكبر لمعادلة صور الاختبار، فإن نتيجة دقة المعادلة – كما أظهرت الدراسة– تتناسب عكسيا مع مدى الدرجات الخام، وطرديا مع مدى القدرة.

ثم تمت معادلة قيم تقديرات القُدرة للأفراد في صورة الاختبار الثاني (B) باختلاف طريقة معالجة القيم المفقودة فيه (EM, MI, RF) إلى قيم تقديرات القُدرة للأفراد في الصورة الأولى (A1) وفق نموذج نظرية الاستجابة للفقرة أُحادي المَعْلَمَة باستخدام المعادلة المشار إليها في رايت وماسترز (Wright & Masters, 1982)، المعادلات **(1 ,2, 3)**. ويبين الجدول (3) تقديرات القُدرة بوحدة اللوجيت المُعادَلة لصورة الاختبار الثاني (B) باختلاف طريقة معالجة القيم المفقودة فيه (EM, MI, RF) باستخدام نموذج نظرية الاستجابة للفقرة الأُحادي المَعْلَمَة والمُقابِلة للدرجة الخام الخاصة بصورة الاختبار الأول (A1).

جدول (3): تقديرات القُدرات المُعادَلة لصورة الاختبار الثاني (B) باختلاف طريقة تعويض القيم المفقودة فيه (EM, MI, RF) وفقاً للنموذج أُحادي المَعْلَمَة والمُقابلة للدرجة الخام الخاصة بصورة الاختبار الأول (A1)

الدرجة الخام	القُدرة في صورة الاختبار الأول (A)	الصورة B باختلاف طريقة تعويض القيم المفقودة		
		EM	MI	RF
		القُدرة المعادلة	القُدرة المعادلة	القُدرة المعادلة
7	-2.6772	-2.5559	-2.4030	-2.3489
8	-2.5089	-2.3925	-2.2451	-2.1904
9	-2.3857	-2.2636	-2.1231	-2.0670
10	-2.2402	-2.1268	-1.9940	-1.9417
11	-2.0586	-1.9548	-1.8230	-1.7728
12	-1.8542	-1.7453	-1.6204	-1.5755
13	-1.6564	-1.5539	-1.4423	-1.4008
14	-1.5181	-1.4258	-1.3131	-1.2764
15	-1.4100	-1.3159	-1.2075	-1.1778
16	-1.2953	-1.1997	-1.0983	-1.0636
17	-1.1236	-1.0379	-0.9429	-0.9115
18	-0.9120	-0.8265	-0.7431	-0.7122
19	-0.7181	-0.6345	-0.5658	-0.5386
20	-0.5747	-0.4959	-0.4321	-0.4081
21	-0.4708	-0.3920	-0.3244	-0.3077
22	-0.3531	-0.2808	-0.2203	-0.1937
23	-0.1868	-0.1134	-0.0604	-0.0454
24	0.0169	0.0790	0.1332	0.1395
25	0.2163	0.2842	0.3140	0.3301
26	0.3721	0.4364	0.4625	0.4712
27	0.4884	0.5480	0.5715	0.5845
28	0.6185	0.6706	0.6966	0.6978
29	0.7896	0.8449	0.8552	0.8612
30	0.9956	1.0434	1.0473	1.0468
31	1.1996	1.2410	1.2349	1.2371
32	1.3574	1.3908	1.3815	1.3822
33	1.4837	1.5223	1.5066	1.4961
34	1.6262	1.6593	1.6488	1.6342
35	1.8225	1.8545	1.8256	1.8096
36	2.0446	2.0639	2.0306	2.0102
37	2.2381	2.2578	2.2102	2.1919
38	2.3848	2.4070	2.3520	2.3289
39	2.5398	2.5623	2.4971	2.4768
40	2.7343	2.7507	2.6786	2.6567
41	2.9720	2.9767	2.8978	2.8688

يُبيّن جدول (4) تقديرات القُدرة بوحدة اللوجيت المُعادَلة لصورة الاختبار الثاني (B) باختلاف طريقة معالجة القيم المفقودة فيه (EM, MI, RF) باستخدام نموذج نظرية الاستجابة للفقرة الثُلاثي المَعْلَمَة والمُقابِلة للدرجة الخام الخاصة بصورة الاختبار الأول (A3).

جدول (4): تقديرات القُدرات المُعادَلة لصورة الاختبار الثاني (B) باختلاف طريقة تعويض القيم المفقودة فيه (EM, MI, RF) وفقاً للنموذج ثُلاثي المَعْلَمَة والمُقابِلة للدرجة الخام الخاصة بصورة الاختبار الأول (A)

الدرجة الخام	القُدرة في صورة الاختبار الاول (A3)	الصورة B باختلاف طريقة تعويض القيم المفقودة		
		EM	MI	RF
		القُدرة المعادلة	القُدرة المعادلة	القُدرة المعادلة
7	-2.7359	-2.6836	-2.5688	-2.5557
8	-2.5564	-2.4958	-2.3900	-2.3800
9	-2.4484	-2.3904	-2.2896	-2.2813
10	-2.1741	-2.1226	-2.0346	-2.0307
11	-2.0503	-2.0017	-1.9195	-1.9176
12	-1.8742	-1.8298	-1.7558	-1.7568
13	-1.7446	-1.7033	-1.6354	-1.6384
14	-1.5881	-1.5505	-1.4899	-1.4954
15	-1.4444	-1.4102	-1.3563	-1.3641
16	-1.2719	-1.2418	-1.1959	-1.2065
17	-1.0849	-1.0592	-1.0221	-1.0356
18	-0.8859	-0.8649	-0.8371	-0.8538
19	-0.7300	-0.7127	-0.6922	-0.7114
20	-0.5877	-0.5738	-0.5599	-0.5814
21	-0.4574	-0.4466	-0.4388	-0.4624
22	-0.3321	-0.3242	-0.3223	-0.3479
23	-0.1508	-0.1472	-0.1538	-0.1822
24	0.0594	0.0580	0.0416	0.0098
25	0.2574	0.2513	0.2257	0.1907
26	0.3859	0.3768	0.3451	0.3081
27	0.4982	0.4864	0.4495	0.4107
28	0.6049	0.5906	0.5487	0.5082
29	0.8039	0.7849	0.7337	0.6900
30	1.0368	1.0122	0.9502	0.9028
31	1.2390	1.2096	1.1382	1.0875
32	1.3639	1.3316	1.2543	1.2016
33	1.4485	1.4142	1.3329	1.2789
34	1.5517	1.5149	1.4289	1.3732
35	1.7135	1.6729	1.5793	1.5210
36	1.9957	1.9484	1.8416	1.7788

الدرجة الخام	القُدرة في صورة الاختبار الاول (A3)	الصورة B باختلاف طريقة تعويض القيم المفقودة		
		EM	MI	RF
		القُدرة المعادلة	القُدرة المعادلة	القُدرة المعادلة
37	2.2144	2.1619	2.0449	1.9786
38	2.3237	2.2686	2.1465	2.0785
39	2.3852	2.3287	2.2037	2.1347
40	2.4752	2.4166	2.2873	2.2169
41	2.6573	2.5943	2.4566	2.3833

يُلاحظ من الجدولين (3، 4) وجود اختلاف ظاهري بين قيم القدرات المعادلة للأفراد في صورة الاختبار الثاني (B) باختلاف طريقة تعويض القيم المفقودة فيه، وهاتان النتيجتان تدلان على اختلاف تقدير قُدرات الأفراد باختلاف طرق تعويض القيم المفقودة، وهذا ما أيدته معظم الدراسات التي تناولت أثر طرق تعويض القيم المفقودة على تقدير قُدرات الأفراد كدراسة: الدرابسة (2012)، والزعبي (2013). مما يؤكد اختلاف القُدرات المُعادَلة باختلاف طريقة التعويض للقيم المفقودة. كما يلاحظ أن مدى العلامات الخام في طريقة التعويض المتعدد (MI) أقل من المدى في طريقتي (EM, RF)، وفي كلا النموذجين، في حين كان مدى القدرة لهذه الطريقة أعلى منه لباقي الطرق المستخدمة، وفي كلا النموذجين، وربما يعود ذلك لعدد مرات التعويض التي تتم في طريقة (MI)، وأخذ المتوسط الحسابي للقيم الناتجة مما يخفف من أثر القيم المتطرفة، وبالتالي يقلل المدى, وهذا يؤكد أن طرق التعويض المختلفة تقوم على أسس رياضية مختلفة.

النتائج المتعلقة بالسؤالين الثاني، والثالث وهما: "هل تختلف دقة معادلة الاختبار وفاعليتها في كل من نموذجي نظرية الاستجابة للفقرة الأُحادي المَعْلَمَة باختلاف طريقة تعويض القيم المفقودة ؟". "هل تختلف دقة معادلة الاختبار وفاعليتها في نموذج نظرية الاستجابة للفقرة الثلاثي المَعْلَمَة باختلاف طريقة تعويض القيم المفقودة ؟".

تم التحقق من دقة، وفاعلية معادلة الاختبار باختلاف طريقة تعويض القيم المفقودة فيه وفق نموذج نظرية الاستجابة للفقرة أُحادي المَعْلَمَة، وثلاثي المَعْلَمَة باستخدام محكي: الخطأ المعياري للمعادلة، والصدق التقاطعي.

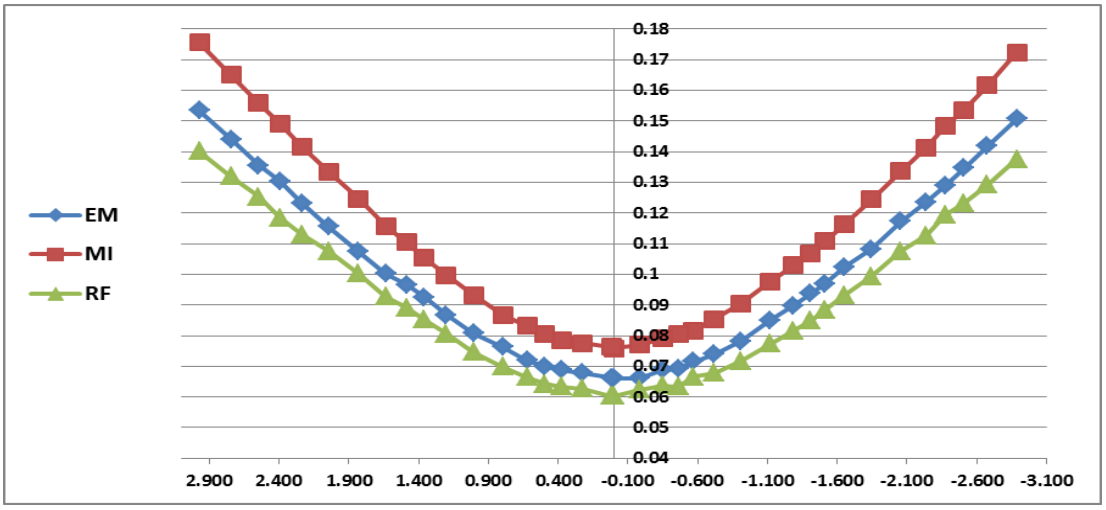
محك الخطأ المعياري للمعادلة (Standard Error of Equating): تم إيجاد قيم الخطأ المعياري للمعادلة عند قيم مختارة من الدرجات المعيارية لصورة الاختبار الثاني (B) باختلاف طريقة تعويض القيم المفقودة فيه (EM, MI, RF) في نموذج نظرية الاستجابة للفقرة أُحادي المَعْلَمَة، ويُبيّن جدول (5) الخطأ المعياري للمعادلة الخطية الأفقية لصورة الاختبار الثاني **(B)** باختلاف طريقة تعويض القيم المفقودة فيه (EM, MI, RF) عند علامات معيارية مختارة في نموذجي نظرية الاستجابة للفقرة أُحادي المَعْلَمَة، وثلاثي المعلمة.

جدول (5): الخطأ المعياري للمعادلة الخطية الأفقية لصورة الاختبار الثاني (B) باختلاف طريقة تعويض القيم المفقودة فيه (EM, MI, RF) عند علامات معيارية مختارة في نموذجي نظرية الاستجابة للفقرة أُحادي المَعْلَمَة وثلاثي المعلمة

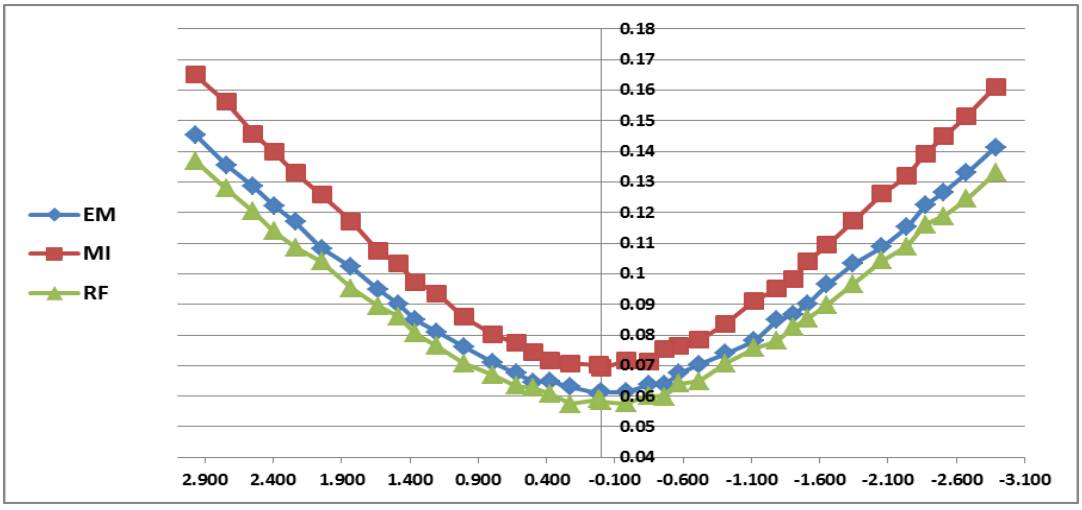
الدرجات المختارة Z	نموذج نظرية الاستجابة للفقرة أحادي المعلمة			نموذج نظرية الاستجابة للفقرة ثلاثي المعلمة		
	EM	MI	RF	EM	MI	RF
-2.890	0.1507	0.1724	0.1375	0.1414	0.1611	0.1330
-2.675	0.1418	0.1616	0.1292	0.1330	0.1516	0.1244
-2.507	0.1346	0.1537	0.1232	0.1267	0.1449	0.1188
-2.375	0.1288	0.1484	0.1194	0.1224	0.1393	0.1161
-2.237	0.1236	0.1414	0.1127	0.1153	0.1319	0.1090
-2.057	0.1172	0.1336	0.1076	0.1088	0.1261	0.1044
-1.846	0.1083	0.1244	0.0993	0.1033	0.1173	0.0967

الدرجات المختارة Z	نموذج نظرية الاستجابة للفقرة أحادي المعلمة			نموذج نظرية الاستجابة للفقرة ثلاثي المعلمة		
-1.655	0.1025	0.1165	0.0932	0.0964	0.1094	0.0898
-1.515	0.0970	0.1108	0.0885	0.0901	0.1041	0.0854
-1.407	0.0938	0.1069	0.0849	0.0868	0.0984	0.0826
-1.286	0.0899	0.1030	0.0816	0.0848	0.0951	0.0781
-1.119	0.0850	0.0976	0.0774	0.0780	0.0912	0.0757
-0.911	0.0783	0.0905	0.0717	0.0739	0.0837	0.0707
-0.717	0.0740	0.0854	0.0679	0.0703	0.0784	0.0650
-0.572	0.0715	0.0816	0.0666	0.0674	0.0763	0.0641
-0.462	0.0691	0.0804	0.0634	0.0638	0.0753	0.0598
-0.346	0.0688	0.0790	0.0638	0.0637	0.0712	0.0602
-0.184	0.0661	0.0771	0.0624	0.0614	0.0716	0.0576
0.000	0.0661	0.0759	0.0603	0.0613	0.0691	0.0583
0.021	0.0663	0.0763	0.0603	0.0609	0.0702	0.0589
0.222	0.0679	0.0776	0.0627	0.0631	0.0708	0.0574
0.377	0.0688	0.0786	0.0634	0.0649	0.0717	0.0609
0.496	0.0701	0.0805	0.0642	0.0645	0.0743	0.0627
0.623	0.0719	0.0832	0.0665	0.0674	0.0774	0.0634
0.794	0.0763	0.0867	0.0698	0.0711	0.0802	0.0670
1.004	0.0808	0.0933	0.0746	0.0762	0.0859	0.0707
1.204	0.0868	0.0998	0.0806	0.0810	0.0935	0.0763
1.358	0.0926	0.1054	0.0852	0.0848	0.0973	0.0805
1.487	0.0967	0.1105	0.0892	0.0901	0.1033	0.0860
1.635	0.1003	0.1158	0.0928	0.0949	0.1075	0.0895
1.829	0.1074	0.1247	0.1002	0.1023	0.1171	0.0953
2.046	0.1155	0.1335	0.1076	0.1083	0.1258	0.1039
2.237	0.1232	0.1417	0.1131	0.1170	0.1330	0.1086
2.392	0.1304	0.1492	0.1185	0.1222	0.1398	0.1139
2.547	0.1355	0.1558	0.1253	0.1286	0.1457	0.1204
2.742	0.1440	0.1650	0.1320	0.1356	0.1561	0.1278
2.971	0.1535	0.1757	0.1402	0.1455	0.1650	0.1369

يُلاحظ من جدول (5) أن دقة المعادلة الخطية تبلغ أكبر قيمة لها عند الدرجة المعياري (Z = 0.00) ، أي أن قيمة الخطأ المعياري يكون أقل ما يمكن، ويزداد هذا الخطأ بالابتعاد عن الوسط الحسابي من الجهتين ويعود ذلك الى أخطاء المعاينة التي تؤثر على تقديرات المتوسطات الحسابية والانحرافات المعيارية المحسوبة. وقد تم تمثيل أخطاء المعادلة الناتجة من معادلة صورة الاختبار الثاني (B) باختلاف طريقة تعويض القيم المفقودة فيه (EM, MI, RF) عند قيم مختارة من الدرجات المعيارية لنموذجي **نظرية الاستجابة للفقرة أُحادي المعلمة وثلاثي المعلمة**، ويبين الشكلين (1، 2) ذلك.



شكل (1): التمثيل البياني لأخطاء المعادلة عند قيم معيارية مختارة تبعاً لمتغير طريقة تعويض القيم المفقودة في نموذج نظرية الاستجابة للفقرة الأُحادي المَعْلَمَة



شكل (2): التمثيل البياني لأخطاء المعادلة عند قيم معيارية مختارة تبعاً لمتغير طريقة تعويض القيم المفقودة في نموذج نظرية الاستجابة للفقرة الثُلاثي المَعْلَمَة

يُلاحظ من الشكلين (1، 2) انخفاض قيمة الخطأ المعياري للمعادلة كلما اقتربنا من الصفر لقيم الدرجات المعيارية للقُدرات المعادلة، ولجميع طرق التعويض الثلاث (EM, MI, RF)، مع أفضلية لطريقة (RF)، وذلك في النموذجين الأُحادي المَعْلَمَة، والثلاثي المَعْلَمَة.

ولمعرفة الدلالة الإحصائية للفروقات الظاهرية في قيم الأخطاء المعيارية للمعادلة، باختلاف متغير طريقة تعويض القيم المفقودة (EM, MI, RF) في نموذج نظرية الاستجابة للفقرة الأُحادي المَعْلَمَة، تم حساب المتوسطات الحسابية، والانحرافات المعيارية للأخطاء المعيارية لمعادلة صورة الاختبار الثاني (B) وفقاً لمتغير طريقة معالجة القيم المفقودة فيه (EM, MI, RF)، وجدول (6) يُبيّن ذلك.

جدول (6): المتوسطات الحسابية والانحرافات المعيارية للأخطاء المعيارية لمعادلة صورة الاختبار الثاني (B) وفقاً لمُتغير طريقة معالجة القيم المفقودة فيه(EM, MI, RF) في نموذج نظرية الاستجابة للفقرة (الأُحادي المَعْلَمَة و الثلاثي المَعْلَمَة)

النموذج	طريقة المعالجة	المتوسط الحسابي	الانحراف المعياري
الأُحادي المَعْلَمَة	**EM**	.0988	.0277
	MI	.1133	.0316
	RF	.0907	.0252
الثلاثي المَعْلَمَة	**EM**	.0926	.0264
	MI	.1057	.0303
	RF	.0873	.0247

يُلاحظ من بيانات جدول (6) وجود فروق ظاهرية بين المتوسطات الحسابية للأخطاء المعيارية لمعادلة صورة الاختبار الثاني (B)، تبعاً لمُتغير طريقة معالجة القيم المفقودة فيه في نموذج نظرية الاستجابة للفقرة. ولمعرفة الدلالة الإحصائية لتلك الفروق الظاهرية وفقاً لمُتغير طريقة معالجة القيم المفقودة فيه (EM, MI, RF) في نموذج نظرية الاستجابة للفقرة الأُحادي المَعْلَمَة، تم استخدام تحليل التباين الأُحادي للقياسات المتكررة لكل من النموذجين: الأحادي والثلاثي المعلمة على حده، وكما هو موضح في الجدولين (7, 9).

جدول (7): نتائج تحليل التباين الأُحادي للقياسات المتكررة للمتوسطات الحسابية للأخطاء المعيارية لمعادلة صورة الاختبار الثاني (B) وحسب مُتغير طريقة معالجة القيم المفقودة فيه(EM, MI, RF) في نموذج نظرية الاستجابة للفقرة (الأُحادي المَعْلَمَة)

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة ف	الدلالة الإحصائية
طريقة المعالجة	.00972	2	.00486	464.731	.000
الخطأ (طريقة المعالجة)	.00075	72	.00001		
الخطأ	.08567	36	.00238		
الكلي	**.09614**	**110**			

يُلاحظ من جدول (7) أن قيمة الدلالة الإحصائية لمتغير (طريقة تعويض القيم المفقودة) بلغت (0.000) وهي أقل من مستوى الدلالة الإحصائية ($\alpha = 0.05$)؛ مما يدل على وجود فروق دالة إحصائياً بين المتوسطات الحسابية للأخطاء المعيارية للمعادلة تُعزى لمتغير (طريقة تعويض القيم المفقودة)، وللكشف عن مواقع تلك الفروق الدالة تم استخدام اختبار بونفيروني (Bonferroni) المُبيّن في جدول (8).

جدول (8): نتائج اختبار بونفيروني (Bonferroni) لمقارنة المتوسطات الحسابية لأخطاء المعادلة وفقاً لمتغير (طريقة تعويض القيم المفقودة)

طريقة تعويض القيم المفقودة	المتوسط الحسابي	الخطأ المعياري	الفرق بين المتوسطين الحسابيين	
			MI	RF
EM	.099	.005	*0.004**	*0.008**
MI	.113	.005		*0.012**
RF	.091	.004		

يُلاحظ من جدول (8) وجود فرق دالّ إحصائياً بين طريقة تعويض القيم المفقودة (EM) مقارنة بطريقة تعويض القيم المفقودة (MI)، ولصالح طريقة تعويض القيم المفقودة (EM). ووجود فرق دالّ إحصائياً بين طريقة تعويض القيم المفقودة (EM) مقارنة بطريقة تعويض القيم المفقودة (RF) ولصالح طريقة تعويض القيم المفقودة (RF). ووجود فرق دالّ إحصائياً

بين طريقة تعويض القيم المفقودة (MI) مقارنة بطريقة تعويض القيم المفقودة (RF)، ولصالح طريقة تعويض القيم المفقودة (RF). وهذا يعني أن طريقة تعويض القيم المفقودة باستخدام دالة الاستجابة هي أكثر دقة من طريقة التعويض باستخدام المتوسطات، واتفقت هذه النتيجة مع دراسة جيمسي وآخرون (Gemici et al., 2011)، ودراسة وايمان (Wayman, 2003)، ودراسة بني عواد (2010)، ودراسة أليسون (Allison, 2006)، واختلفت مع طريقة إندرز (Enders,2004)، ودراسة فينج (Finch, 2008)، كما تبين أن طريقة تعويض القيم المفقودة باستخدام المتوسطات هي أكثر دقة من طريقة التعويض باستخدام خوارزمية تعظيم التوقعات، واتفقت هذه النتيجة مع دراسة جيمسي وآخرون (Gemici et al., 2011)، ودراسة بني عواد (2010)، ودراسة أليسون (Allison, 2006)، ، واختلفت مع دراسة وايمان (Wayman, 2003)، ويمكن تفسير السبب في تفوق طريقة دالة الاستجابة (RF) لكون عملية التعويض للقيم المفقودة للمفحوص تكون من خلال نفس استجاباته، ولا تتأثر استجابات المفحوص باستجابات مفحوص آخر ولا علاقة بينهما فاذا تم تغيير أي من الفقرات لمفحوص معين، فهذا لا يؤثر على المفحوص الآخر؛ بمعنى آخر لا تعتمد طريقة التعويض على المتوسطات الحسابية للفقرات، ولا تفترض أي شيء حول معالم الفقرات، ولا تحاول أن تقدرها، كما هو الحال في طريقة خوارزمية تعظيم التوقعات (EM)، وطريقة حساب قيم تعويضية لتقدير البيانات المفقودة، والتي يتم خلالها استبدال كل قيمة مفقودة بمتوسط قيم من القيم المختارة عشوائياً، وبناءً على هذه النتيجة فانَّ الباحثين يوصيان باستخدام طريقة تعويض دالة الاستجابة أولاً، وطريقة المتوسطات ثانياً، وأخيراً طريقة التعويض باستخدام خوارزمية تعظيم التوقعات، وذلك عندما يرغب الباحث بمعادلة صور اختبار نسبة الفقد فيها بحدود 20% باستخدام المعادلة الأفقية.

كما يُلاحظ من بيانات جدول (5) وجود فروق ظاهرية بين المتوسطات الحسابية للأخطاء المعيارية لمعادلة صورة الاختبار الثاني (B)، وحسب مُتغير طريقة معالجة القيم المفقودة فيه. ولمعرفة الدلالة الإحصائية لتلك الفروق الظاهرية وفقاً لمُتغير طريقة معالجة القيم المفقودة فيه(EM, MI, RF) في نموذج نظرية الاستجابة للفقرة الثُلاثي المَعْلَمَة، فقد تم استخدام تحليل التباين الأحادي للقياسات المتكررة، وكما هو موضح في جدول (9).

جدول (9): نتائج تحليل التباين الأحادي للقياسات المتكررة للمتوسطات الحسابية للأخطاء المعيارية لمعادلة صورة الاختبار الثاني (B) وحسب مُتغير طريقة معالجة القيم المفقودة فيه(EM, MI, RF) في نموذج نظرية الاستجابة للفقرة (الثُلاثي المَعْلَمَة)

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة ف	الدلالة الإحصائية
طريقة المعالجة	.00663	2	.00332	385.287	.000
الخطأ (طريقة المعالجة)	.00062	72	.00001		
الخطأ	.07954	36			
الكلي	.08680	110			

يُلاحظ من جدول (9) أن قيمة الدلالة الإحصائية لمتغير (طريقة تعويض القيم المفقودة) بلغت (0.000) وهي أقل من مستوى الدلالة الإحصائية (α = 0.05)؛ مما يدل على وجود فروق دالة إحصائياً بين المتوسطات الحسابية لأخطاء المعادلة تُعزى لمتغير (طريقة تعويض القيم المفقودة)، ولمعرفة لصالح من تلك الفروق الدالة تم استخدام اختبار بونفيروني (Bonferroni) المُبيّن في جدول (10).

جدول (10): نتائج اختبار بونفيروني (Bonferroni) لمقارنة المتوسطات الحسابية لأخطاء المعادلة وفقاً لمتغير (طريقة تعويض القيم المفقودة)

طريقة تعويض القيم المفقودة	المتوسط الحسابي	الخطأ المعياري	الفرق بين المتوسطين الحسابيين	
			MI	RF
EM	.093	.004	0.013*	0.006*
MI	.106	.005		0.019*
RF	.087	.004		

يُلاحظ من جدول (10) وجود فرق دالّ إحصائياً بين طريقة تعويض القيم المفقودة (EM) مقارنة بطريقة تعويض القيم المفقودة (MI)، ولصالح طريقة تعويض القيم المفقودة (EM). ووجود فرق دالّ إحصائياً بين طريقة تعويض القيم المفقودة (EM) مقارنة بطريقة تعويض القيم المفقودة (RF)، ولصالح طريقة تعويض القيم المفقودة (RF). ووجود فرق دال إحصائيا بين طريقة تعويض القيم المفقودة (MI) مقارنة بطريقة تعويض القيم المفقودة (RF)، ولصالح طريقة تعويض القيم المفقودة (RF).

محك الصدق التقاطعي (Cross Validation Criterion):

للتحقق من دقة معادلة الاختبارات وفاعليتها باستخدام محك الصدق التقاطعي، والذي يُبيّن درجة استقرار الدرجات المكافئة، أو المعادلة في عينتين عشوائيتين مستقلتين عن بعضهما البعض، فقد تم اختيار عينة عشوائية من صورة الاختبار الأولى (A) وثلاث عينات عشوائية مستقلة من صورة الاختبار الثانية (B)، كل منهما تحوي (300) فرد من مجتمع الدراسة، ثم تمت معادلة قُدرات هذه العينات خطياً وفق نموذج نظرية الاستجابة للفقرة أُحادي المَعْلَمَة، ثم تم حساب معاملات الصدق التقاطعي ، ويُبيّن جدول (11) قيم معاملات الصدق التقاطعي.

جدول (11): معاملات الصدق التقاطعي للمعادلة الخطية لمعادلة للصورة الجديدة للاختبار (B) مع الصورة الأولى للاختبار (A) باختلاف متغير طريقة تعويض القيم المفقودة في نموذج نظرية الاستجابة للفقرة الأُحادي المَعْلَمَة والثُلاثي المَعْلَمَة

نموذج نظرية الاستجابة للفقرة	طريقة المعالجة	معامل الصدق التقاطعي
الأُحادي المَعْلَمَة	EM	0.00028
	MI	0.00039
	RF	0.00023
الثُلاثي المَعْلَمَة	EM	0.00024
	MI	0.00036
	RF	0.00011

يُلاحظ من جدول (11) أن أقل قيمة لمعامل الصدق التقاطعي كانت للمعادلة في الصورة (B) بطريقة (RF) وفق نموج نظرية الاستجابة للفقرة أُحادي المَعْلَمَة، وهذا يشير إلى أن دقة المعادلة بعد تعويض القيم المفقودة بطريقة (RF)، وفق نموج نظرية الاستجابة للفقرة أُحادي المَعْلَمَة، وثلاثي المَعْلَمَة كانت الأكثر فاعلية من بين الطرق الثلاث المستخدمة. يليها المعادلة بعد تعويض القيم المفقودة بطريقة (EM)، وأخيراً المعادلة بعد تعويض القيم المفقودة بطريقة (MI).

الخلاصة والتوصيات: تأتي أهمية هذه الدراسة في محاولة لمعرفة أثر طريقة تعويض القيم المفقودة على دقة المعادلة الأفقية للاختبار، وتحديد الطريقة الفضلى من بين الطرق المستخدمة وهي (EM, MI, RF)، حيث أظهرت النتائج أفضلية لطرق تعويض القيم المفقودة المستخدمة في النموذج الثلاثي المعلمة على النموذج الأحادي المعلمة، ولجميع الطرق المستخدمة، أما بالنسبة لطرق التعويض فقد بينت النتائج أفضلية لطريقة (RF) على طريقتي (EM, MI)، وخاصة في النموذج الثلاثي المعلمة، وبناءً على نتائج الدراسة فإن الباحثين يوصيان باستخدام النموذج الثلاثي المعلمة في تحليل البيانات في حالة البيانات المفقودة، وإجراء المزيد من الدراسات باستخدام طرق أخرى لتعويض القيم المفقودة، وباستخدام نماذج أخرى لنظرية الاستجابة للفقرة المعلمية كالنموذج الثنائي المعلمة واللامعلمية كنموذج موكن، وباختلاف نسبة البيانات المفقودة، وباختلاف نوع المعادلة، إذ ربما أن اختلاف نوع المعادلة يكون لها دور أيضاً، كما يوصي الباحثان بإجراء دراسة تتناول أثر طرق تعويض القيم المفقودة على افتراض نظرية الاستجابة للفقرة المتعلق بالاستقلال الموضعي للفقرات

باستخدام إما بيانات حقيقية، أو مولدة، حيث لاحظ الباحثان من خلال التحليل الإحصائي اختلاف نتائج الاستقلال الموضعي للفقرات باختلاف طريقة تعويض القيم المفقودة.

المصادر والمراجع

المراجع العربية

أيوب، ح .(1994). *المقارنة بين أربع طرق للمعادلة عندما يكون التصميم من مجموعات متكافئة وغير متكافئة*، رسالة دكتوراه غير منشورة، الجامعة الأردنية، عمان، الأردن.

بني عواد، ع. (2010). *مقارنة طرق التعامل مع البيانات المفقودة في تقدير معالم الفقرات وقُدرات الأفراد*. رسالة دكتوراه غير منشورة، جامعة اليرموك، الأردن.

الدرابسة، ر .(2012). *أثر طريقة تقدير القُدرة وطريقة التعامل مع القيم المفقودة على دقة تقدير معالم الفقرات والأفراد*. رسالة دكتوراه غير منشورة، جامعة اليرموك، اربد، الأردن.

الرحيل، ر. (2013). اثر وجود أداء تفاضلي في الفقرات المرساوية على دقة المعادلة العمودية لاختبار اوتيس لينون للقدرة العقلية. المجلة الدولية التربوية المتخصصة، 2 (8)، 754 – 771.

الزعبي، ع. (2013). *أثر نسبة البيانات المفقودة وطريقة التعويض عنها في دقة تقدير معالم الفقرات والأفراد*. رسالة دكتوراه غير منشورة، جامعة اليرموك، اربد، الأردن.

الشريفين، ن. **(2009)**. معادلة درجات نماذج مختلفة من اختبار الكفاءة اللغوية في اللغة الانجليزية لدى طلبة جامعة اليرموك. مجلة جامعة أم القرى للعلوم التربوية والنفسية، **1 (2)، 11 – 62**.

الشريفين، ن .(2003). *مدى تحقق معايير الفاعلية في معادلة اختبارين أحدهما ثنائي التدريج والآخر متعدد التدريج وفق النظرية الكلاسيكية والنظرية الحديثة في القياس*. رسالة دكتوراه غير منشورة، جامعة عمان العربية للدراسات العليا، عمان، الأردن.

علام، ص. (1987). دراسة موازنة ناقدة لصور السمات الكامنة، والصور الكلاسيكية في القياس النفسي والتربوي. جامعة الكويت، *المجلة العربية للعلوم الإنسانية، (27)، 44-18.*

علام، ص. (2005). *صور الاستجابة للفقرات الاختبارية أُحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي*. القاهرة: دار الفكر العربي.

الكوفحي، م .(2013). *أثر طريقة تصفية المموهات في فقرات الاختيار من متعدد على معادلة الاختبار*. رسالة دكتوراه غير منشورة، جامعة اليرموك، اربد، الأردن

المراجع الأجنبية

Acock, A. (2005). Working with Missing Values. *Journal of Marriage and Family, 67*(4), 1012- 1028.

Allison, P., D. (2006). *Imputation of Categorical Variables with PROC MI*. Paper presented at the Annual Meeting of the SAS users Group International, San Francisco, CA.

Angoff, W.H. & Cowell, W. R. (1986). An examination of the assumption that the equating of parallel forms is population-independent. *Journal of Educational Measurement, 23*, 327-345

Cokluk, O. & Kayri, M. (2011). The effect of Methods of Imputation for Missing Values on the Validity and Reliability of Scales. *Educational Science: Theory and Practice*, 11(1), 303-310.

Crocker, L. & Algina, J. (1986). Int*roduction to classical and modern test theory*. New York: Holt, Rinehart & Winston.

De-Ayala, R., J., Plake, B., S., & Impara, J., C. (2006). The Impact of Omitted Responses on the Accuracy of Ability Estimation in Item Response Theory. Journal of Educational Measurement, 38(3), 213 – 234.

Dempster, A. P. Laird, N. M. & Rubin, D. R. (1977). Maximum Likelihood Estimation from Incomplete Data Via the EM Algorithm. *Journal of the Royal Statistical Society, 39* (1), 1-38.

Enders, C. K. (2004). The Impact of Missing Data on Sample Reliability Estimates: Implication for Reliability Reporting Practices. *Educational and Psychological Measurement, 64*(3), 419 – 436.

Enders, C. K. (2010). *Applied Missing Data Analysis*. New York: Guilford Press.

Finch, H. (2008). Estimation of Item Response Theory Parameters in the Presence of Missing Data. *Journal of Educational Measurement, 45*(3), 225-245.

Gemici, S., Bednarz, A., & Lim, P. (2012). 'A primer for handling missing values in the analysis of education and training data', *International Journal of Training Research, 10*(3), 233-250.

Green, B. F. (Ed). (1981). I*ssues in testing: Coaching, disclosure, and ethnic bias*. San Francisco: Jssey-Basslnc. Publishers.

Hambelton, R. & Swaminthan, H. & Rogers, J. (1991). *Fundamentals of Item Response Theory*, Newbury Park California: Stage Publications.

Hambelton, R. & Swaminthan, H. (1985). *Item response theory principles and application*. Boston: Kluwer – Nijgoff Publishing.

Hattie, J. (1985). Methodology Review: Assessing Unidimenionality of Tests and Items. *Applied psychological Measurement, 9*(2), 139-164.

Hawthorn, G, & Elliott, P. (2004). Imputing Cross-Sectional Missing Data: Comparison of Common Techniques.

Patrick, E. M., (2007). *Missing Data a Gentle Introduction*. New York: Gulford Press.

Peterson, N. S., Kolen, M. J. & Hoover, H. D. (1989). *Scaling, Norming, and Equating*. In R. L. Linn (Ed.), Educational Measurement (3rd ed.). New York: Macmillan.

Peugh, J. L., & Enders, C.K. (2004). Missing Data in Educational Research: A Review of Reporting Practices and Suggestions for Improvement. *Review of Educational Research, 74*(4), 525 556.

Pigott, T. D. (2001). A Review of Methods for Missing Data. *Educational Research and Evaluation, 7* (4), 353-383.

Schafer, J. L. & Graham, J. W. (2002). Missing Data: Our View of the State of the Art. *Psychological Methods, 7* (2), 147-177.

Vander Linden, W. (2006). Equating scores from adaptive to linear tests, Applied Psychological Measurement, 30 (6), 493 – 508

Wayman, J., C. (2003). *Multiple Imputations for Missing Data: What is it and How Can I use it*. Paper presented at the 2003 Annual Meeting of the American Educational Research Association, Chicago, IL.

Wiley, A. (1999). *An Investigation into Two Models for Equating Examinations with Multiple Item Formats.* Dissertation Abstract International, October.

Wright, B. D. & Masters, G. N., (1982). *Rating Scale Analysis. Rasch Measurement*. MESA Press, kimbark Avenue, Chicago, P. 59

*Australian and New Zealand Journal of Psychiatry, 3*9(7), 583 – 590.

Hills, J., R., Subhiyah, R., G., & Hirsch, T., M. (1988). Equating minimum competency tests: comparisons of instructional method. *Journal of Educational Measurement*, 25,221-232.

Holland, P. & Rubin, D. (1982). *Test equating. New York*: Academic Press.

Kolen, M. J. & Bernan, R. L. (2004). Test Equating: methods and practices. New York: spring.

Kolen, M. J. (1981). Comparison of Traditional and Item Response Theory Methods for Equating Tests. *Journal of Educational Measurement, 18, 1*-11.

Lissitz, J., M., & Yang, Y., N. (1999). *Estimating IRT equating coefficients with polytomously and dichotomously scored items*. Paper presented at the annual meeting of National of Council on Measurement in Education, Montreal. (ERIC Document Reproduction Service Number ED431800).

Little, R. J. A., & Rubin, D. B. (1987). S*tatistical Analysis with Missing Data*. New York: Wiley.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data*. 2nd edition. Hoboken, New York: John Wiley.

Lord, F. M. (1980). Practical Applications of Item Characteristics Curve Theory. *Journal of Educational Measurement. 14*, 117-138.

The Effect of the Method of Imputation for the Missing-Responses on the Accuracy of

Equating Test Different Forms Using Paragraph Response Theory Models

Nedal K. Shraifin*

ABSTRACT

This study aimed to detect the Effect of the Method of Imputation for the Missing-Responses on the Accuracy of Equating Test Models Using Paragraph Response Theory. To achieve the aim of the study, (WINGEN) program was used to generate (1000) responses on two tests consisting of (50) dichotomous response paragraphs, and conform to the one and three logistic paragraph response theory (PRT), with paragraph discrimination between (0.1) and (2.0), paragraphs difficulty between (-3.50) and (3.50), and normal distribution of ability (mean zero, one standard deviation). Data with 20% missing scores were generated by using (SPSS) and (EXCEL); these data were treated by three imputation methods: Maximizing Expectations (ME), multiple imputations (MI) and response function imputation (RF). After verifying unidimensionality using factor analysis, and eliminated a number of paragraphs and individuals (6 items, 14 individuals) for 1PL, and (2 items, 9 individuals) for 3PL after paragraph-person fit, by using (BILOG-MG3) paragraphs parameter and person ability was estimated. Standard error of equating and cross validation index were used to detect the accuracy of test horizontal equating. The results of the study showed a different estimation of persons ability by using different methods for imputation missing data, that the linear equation accuracy approached its greatest at the standard score (zero), moreover, the results showed statistical significant difference between the mean errors of equating due to the (imputation methods of missing data) for the imputation method (EM) compared to (MI), and for the imputation method (RF) compared to imputation method (EM). And for the imputation method (RF) compared to imputation method (MI). Finally, the results of the cross validation confirmed the existence of the three-logistic parameter model after impute missing data by (RF) method.

Keywords: Missing responses, Expectation Maximization (EM), Multiple Imputations (MI), Response Function (RF), and Horizontal Equating Accuracy.

*Yarmouk University. Received on 24/6/2015 and Accepted for Publication on 26/8/2015.