المقارنة بين نتائج طرق المعادلة الكلاسيكية لدرجات نماذج اختبار القدرات العامة باستخدام تصميم الجماعات المتكافئة

د. راشد سيف المحرزي

جامعة السلطان قابوس

ملخص

تهدف الدراسة الحالية إلى إجراء معادلة لأربعة نماذج من اختبار القدرة اللفظية واختبار القدرة الكمية باستخدام ثلاث طرق كلاسيكية للمعادلة في تصميم الجماعات المتكافئة والمقارنة بينها. وتمثلت الطرق الكلاسيكية الثلاث للمعادَلة في المعادَلة الخطية والمعادَلة المئينية والمعادَلة المئينية المهدة بعديا. وتظهر الدراسة عدم تكافؤ النماذج الأربعة في الاختبارين والحاجة إلى إجراء المعادلة لهذه النماذج. وتظهر نتائج التحيز (الفروق) اختلاف الدرجات المعادِلة (الخام والمعيارية) لنماذج الاختبار الناتجة من طرق المعادلة الثلاث عن الدرجات الناتجة من المعادلة المساوية، وزيادة هذه الفروق بزيادة الاختلاف في مستوى صعوبة النموذجين المراد معادلتهما. كما تظهر الدرجات المعادِلة الناتجة من طرق المعادلة الثلاث أنها مختلفة. وتظهر نتائج الأخطاء المعيارية للمعادَلة أن المعادَلة المئينية المهدة بعديا تحقق أفضل النتائج عن بقية طرق المعادَلة، حيث إنها تحقق أقل مقدار للأخطاء المعيارية عبر جميع درجات الاختبار مع تحقيق علاقة ممهدة بين الدرجات الخام للنموذج الجديد والدرجات المعادِلة لها في النموذج المرجعي.

الكلمات المفتاحية: معادَلة درجات نماذج الاختبار، النظرية الكلاسيكية، تصميم الجماعات المتكافئة، اختبار القدرات العامة.

الاطار النظري والدراسات السابقة

تحتاج المؤسسات المعنية ببناء اختبارات متعددة النماذج إلى إجراء عملية معادَلة نماذج الاختبار Test form Equating لضمان تحقيق صدق هذه الاختبارات وتحقيق العدالة بين مقدمي مختلف نماذج الاختبار. وتعرف عملية معادَلة نماذج الاختبار بالعملية الإحصائية التي تستخدم لتعديل الدرجات في النماذج المختلفة للاختبار؛ بحيث يمكن استخدام درجات تلك النماذج بطريقة تبادلية (Kolen & Brennan، 2004). وحيث إن السبب الأساسي لبناء نماذج مختلفة للاختبار هو صعوبة تطبيق الاختبار على جميع الأفراد في توقيت واحد والخوف من تسرب فقرات الاختبار في حالة تقديم النموذج نفسه عددا من المرات. بينما يتطلب تقديم نماذج مختلفة من الاختبار ضمان تساوي هذه النماذج في مستوى صعوبتها وتكافئها؛ للتمكن من إصدار أحكام متشابهة (للطالب أو المؤسسات المستخدمة لنتائج الاختبار) باستخدام أي من نماذج الاختبار. ويعود السبب وراء ذلك أن تطبيق نماذج الاختبار على عينات مختلفة من الأفراد في حالة عدم تكافؤ هذه النماذج ، قد يؤدي في غالب الأحيان إلى اختلاف المتوسط الحسابي لهذه النماذج. وقد يكون مرد اختلاف المتوسط الحسابي للنماذج غير المتكافئة إلى اختلاف صعوبة نماذج الاختبار أو اختلاف مستوى قدرة عينات مقدمي نماذج الاختبار. ولا يمكن الجزم بالسبب الرئيسي وراء اختلاف المتوسط الحسابي لنماذج الاختبار. فهناك عدد من العوامل

المرتبطة بعملية تطبيق الاختبار ومستويات الطلبة وسلوك المفردات أثناء التطبيق وتفاعل الطلبة معها تجعل درجات الطلاب التي يُتحصل عليها من تطبيق نماذج الاختبار غير متكافئة. فتشابه نماذج الاختبار في المحتوى المقاس من خلال جدول المواصفات لا يضمن تساوي مستوى صعوبة نماذج الاختبار (Angoff، 1987). وبالتالي يمكن القول بأن عملية معادَلة نماذج الاختبار ضرورية لضمان تساوي مستوى صعوبة هذه النماذج واستخدامها بشكل تبادلي للحصول على تدريجات متكافئة لدرجات الطلاب (Chulu & Sireci، 2011).

وينتج من معادلة نماذج الاختبارات علاقة اقتران بين درجات نماذج الاختبار، بحيث تقترن كل درجة من درجات أحد نماذج الاختبار (يطلق عليه مسمى الاختبار الجديد) بما يعادلها من درجات نموذج آخر (يطلق عليه مسمى الاختبار المرجعي). وتقوم عملية معادَلة نماذج الاختبار على مبدأ تعديل الاختلافات في مستوى صعوبة نماذج الاختبار لكي تكون متساوية في الصعوبة ولتستخدم بطريقة تبادلية. وتتطلب هذه العملية تطبيق نماذج الاختبار بتصميم خاص (يسمى تصميم المعادَلة Equating design) يهدف إلى ضبط مستوى القدرة والسمة التي يقيسها الاختبار بين العينات التي يطبق عليها الاختبار؛ ليتسنى بعد ذلك التأكد من التوزيع التكراري والمتوسط الحسابي للنماذج المختلفة. فإذا تبقى اختلاف في المتوسط الحسابي (أو التوزيع التكراري) لنماذج الاختبار، فيرجع ذلك إلى اختلاف مستوى صعوبتها، ثم تجرى معادَلة إحصائية للحصول على الدرجات المعادِلة للنموذج الجديد مع النموذج المرجعي.

وتعرف طريقة معادلة درجات نماذج الاختبار بالمعادلة الأفقية (Horizontal Equating)، بينما يوجد نوع آخر لمعادلة درجات الاختبار يعرف بالمعادلة الرأسية (Vertical Equating)، والذي يعرف أيضا بالتدريج الرأسي (Vertical Scaling). وفي المعادلة الرأسية، تكون نماذج الاختبار متدرجة في مستوى صعوبتها بشكل مقصود، ويتم تقديمها من قبل مجموعات طلابية مختلفة في مستوى قدراتها، وغالبا ما يكون الطلاب في مستويات دراسية أو صفوف دراسية مختلفة. ويقوم هذا النوع من المعادلة على افتراض أن القدرة المقيسة على تدريج متصل واسع المدى وممتد عبر مستويات دراسية متعددة (صفوف دراسية مثلا)، كما هو الحال في القدرات والسمات التطورية. ويتطلب توفير نماذج للاختبار متباينة في مستوى صعوبتها بسبب عدم القدرة على قياس هذه السمة الممتدة عبر المستويات والصفوف الدراسية باختبار واحد. ويشترط في هذه النماذج أن تقيس سمة واحدة؛ وبحيث يقيس كل نموذج مستوى أو جزءًا واحدا من المدى المتصل للقدرة المقيسة. وتهدف المعادلة الرأسية إلى وضع تدريج جميع نماذج الاختبار ذات الصعوبة المختلفة على متصل وتدريج موحد للقدرة يسهل مقارنة درجات الطلاب عبر المستويات الدراسية المختلفة. وتفيد الاختبارات المعادلة رأسيا في قياس مستوى التقدم الذي يحققه الطالب في السمة المقيسة عبر انتقاله بين المستويات الدراسية. ومن بين الاختبارات المشهورة التي تستخدم المعادلة الرأسية اختبار القدرات المعرفية Cognitive Ability Test (CogAT) الذي أعده ثورندايك وهايجن Thorndike & Hagen وطوره لاحقا لوهمان وهايجن Lohman & Hagen (Lohman & Gambrell، 2012). ويهدف هذا الاختبار قياس القدرات المعرفية من الصف الدراسي التمهيدي وحتى الصف الدراسي الثاني عشر في أغلب الولايات المتحدة الأمريكية. ومن بين الاختبارات في الوطن العربي التي تستخدم المعادلة الرأسية مقياس الخليج للقدرات العقلية المتعددة – الذي أعده فريق من الباحثين بإشراف من مكتب التربية العربي لدول الخليج. ويوظف هذا الاختبار المعادلة الرأسية لقياس القدرات العقلية من الصف الدراسي الأول وحتى الصف الدراسي السادس (الزيات والمحرزي، 2011).

وتتطلب عملية إجراء المعادلة بين نموذجي الاختبار مجموعة من الشروط التي يجب توفرها وعدم الاخلال بها والتي تتمثل في (Kolen & Brennan، 2004;):

- أن يقيس نموذجا الاختبار القدرة نفسها.
- أن تكون معادلة درجات نموذجي الاختبار مستقلة عن عينة الأفراد المطبق عليهم نماذج الاختبار.
- أن يتحقق ما يسمى بشرط المساواة (Equity)، والمتمثل في تشابه خصائص التوزيع التكراري المشروط لدرجات النموذجين المراد معادلتهما.
- أن يتحقق ما يسمى بشرط التماثل (Symmetry)، والمتمثل في تشابه الدرجات المعادِلة الناتجة من تحويل الدرجات من النموذج الأول إلى النموذج الثاني مع الدرجات المعادلة الناتجة من تحويل الدرجات من النموذج الثاني إلى النموذج الأول.
- أن يتمتع نموذجا الاختبار بالقدر نفسه من ثبات الدرجات إذا طبقا على عينات عشوائية من أفراد المجتمع المستهدف نفسه.

قد تتم عملية معادَلة نماذج الاختبار بشكل منفصل وسابق عن التطبيق الفعلي للاختبار، أو قد تجرى ضمن التطبيق الفعلي للاختبار. وتتصف الطريقة الأولى بصعوبات عملية في تطبيقها، بسبب الحاجة إلى تقديم نماذج الاختبار بوقت سابق مما يهدد سرية مفردات الاختبار. فنجد أن أغلب ممارسات معادَلة نماذج الاختبار تجرى ضمن التطبيق الفعلي للاختبار، والذي يتطلب مراعاة اشتراطات جمع البيانات في عملية المعادَلة أثناء تصميم التطبيق الفعلي للاختبار. وهناك ثلاثة تصاميم رئيسة لجمع البيانات بغرض إجراء عملية المعادَلة والتي تقوم كلها على مبدأ استبعاد اختلاف مستوى القدرة المقيسة لدى عينات الأفراد التي تقدم نماذج مختلفة للاختبار. وهذه التصاميم الثلاثة هي (Holland & Rubin، 1982; Kolen، 1988):

1. **تصميم الجماعة الواحدة** Single Group Design: ويتطلب تطبيق نماذج الاختبار (نموذجين أو أكثر) على نفس المجموعة من الأفراد، بحيث يقدم كل فرد جميع نماذج الاختبار المراد معادَلة درجاتها. ويتصف هذا التصميم (مقارنة بالتصاميم الأخرى) بالقدرة العالية على ضبط مستوى قدرة عينات الأفراد المقدمة للنماذج؛ إلا أنه يعاني من مشكلة الصعوبة العملية والميدانية في تطبيقه، من حيث الوقت المستغرق والجهد المبذول من قبل الطالب وإمكانية تعلم الطالب لمفردات الاختبار خلال انتقاله عبر النماذج، وشعوره بالملل والإرهاق.

2. **تصميم الجماعات المتكافئة** - Equivalent Groups Design: ويتطلب هذا التصميم انتقاء جماعات طلابية متكافئة في السمة المقيسة في الاختبار، بحيث تطبق كل جماعة طلابية أحد نماذج الاختبار. ويشترط هذا التصميم بشكل أساسي عشوائية وتكافؤ الجماعات الطلابية من حيث العدد ومستوى القدرة المقيسة، ووقت التطبيق وظروفه. ويتحقق ذلك من خلال توزيع جميع نماذج الاختبار على جماعات طلابية متساوية العدد في نفس التوقيت. ويتم توزيع نماذج الاختبار بطريقة عشوائية في كل قاعة اختبار (Kolen & Brennan، 2004; Lyrén & Hambleton، 2011). وينصح كولن وبرنن (Kolem & Brennan، 2004) باستخدام الطريقة التصاعدية (Spiraling) في توزيع نماذج الاختبار لتحقيق أعلى درجة التكافؤ بين الجماعات الطلابية. وتتطلب الطريقة التصاعدية ترتيب أوراق نماذج الاختبار بشكل تسلسلي (فمثلا في حالة أربعة نماذج، فيكون التوزيع كالتالي: 4،3،2،1؛ 4،3،2،1؛ وهكذا)، ثم يتم توزيعها على الطلاب بهذا الترتيب في كل قاعة امتحانية.

3. **تصميم المفردات المشتركة والجماعات غير المتكافئة** - Common-Item Non-equivalent Groups Design، يشبه هذا التصميم تصميم الجماعات المتكافئة، من حيث تطبيق نماذج الاختبار على جماعات مختلفة من الأفراد ، ولكن بدون اشتراط تكافؤ جماعات الأفراد. ويستخدم هذا التصميم عندما تتطلب سياسة الاختبار تطبيق نماذج الاختبار في أوقات وأيام مختلفة. ولتحقيق المبدأ الأساسي لعملية المعادَلة والقائم على استبعاد اختلاف مستوى قدرة عينات الأفراد المطبقة لنماذج الاختبار، فيتم وضع مفردات مشتركة بين نماذج الاختبار.. ويشترط ظهور جميع المفردات المشتركة في نماذج الاختبار المراد معادلتها في نفس المواقع والترتيب في الورقة الامتحانية. كما يشترط أن تشابه هذه المفردات المشتركة المفردات الأصلية في الاختبار بحيث يمكن اعتبارها كاختبار مصغر للاختبار الأصلي، من حيث جدول المواصفات ، والخصائص السيكومترية. ويمنح هذا التصميم مرونة كبيرة للتطبيق العملي للاختبار.

وعند مقارنة تصاميم المعادلة الثلاثة، يعد تصميم الجماعات المتكافئة من أقل التصاميم تعقيدا من حيث عمليات بناء نماذج الاختبار، حيث لا يتطلب بناء مفردات مشتركة والتي يصعب كتابتها لتكون عينة ممثلة لما يقيسه الاختبار ككل. كما يحقق هذا التصميم الافتراضات الإحصائية اللازمة لعملية المعادلة بشكل كبير من خلال التوزيع العشوائي للطلبة على نماذج الاختبار. وما قد يؤخذ على هذا التصميم هو الحجم الكبير للمجموعات الطلابية اللازمة لتطبيق نماذج الاختبار، والتي قد لا تتوفر دائما في بعض الاختبارات. كما يتطلب هذا التصميم في بعض الأحيان تطبيق النموذج المرجعي أكثر من مرة في أيام مختلفة، مما يهدد سرية مفردات النموذج الأصلي، والذي بدوره قد يدخل أخطاء منتظمة في عملية المعادَلة ويجعل أداء الطلبة على النموذج المرجعي – في حالة تسربها – مختلفا عن الأداء في الظروف الطبيعية.

وتختلف الطرق الاحصائية المتبعة لمعادَلة نماذج الاختبار في تصميم الجماعات المتكافئة حسب نظرية الاختبار التي تتبناها الجهة المسؤولة عن إعداد الاختبار وتقديمه. فهناك نظريتان أساسيتان في بناء الاختبارات ومعادلتها، هما: نظرية الاستجابة للمفردة والنظرية الكلاسيكية للاختبار (Kolen & Brennan، 2004؛ الدوسري، 2001). وتقوم نظرية الاستجابة على المفردة على تفسير العلاقة بين السمة التي يقيسها الاختبار والاستجابة على كل مفردة من مفردات الاختبار من خلال دالة إحصائية متزايدة. وتتعدد النماذج الرياضية التي تتنبأ بهذه الدالة حسب عدد السمات المقاسة ونوعية الخصائص السيكومترية للمفردات. وبالتالي تتطلب طرق المعادَلة في نظرية الاستجابة للمفردة ملاءمة البيانات لأحد نماذج الدالة الإحصائية قبل تطبيق طرق المعادَلة فيها. وفي حالة ملاءمة البيانات للنموذج الإحصائي المتبع في نظرية الاستجابة للمفردة، فتقدم هذه النظرية عددا من الطرق الإحصائية المناسبة لمعادلة نماذج الاختبار. ومن بين هذه الطرق الإحصائية ما يعرف بمعادلة درجات القدرة (Ability Score Equating)، ومعادلة الدرجات الحقيقية (True Score Equating)، ومعادلة الدرجات الظاهرية (الدرجات الخام) (Observed Score Equating) (Baker & Al-Karni، 1991; Harris & Kolen، 1990; Lord، 1980).

بينما تتميز طرق المعادَلة في النظرية الكلاسيكية بعدم استخدام نماذج رياضية يلزم التأكد من ملاءمة بيانات الاختبار لافتراضاتها كما هو الحال في نماذج نظرية الاستجابة للمفردة. فتعتبر الطرق الاحصائية للمعادَلة في النظرية الكلاسيكية للاختبار أكثر قابلية للتطبيق. وأوضحت الدراسات تقارب نتائج المعادَلة وفقا للنظريتين إلى حد كبير (Wang، Lee، Brennan، & Kolen، 2008). فمثلا أظهرت دراسة جلواكي (Glowacki، 1991)، عند مقارنته

بين الطرق الإحصائية في نظرية الاستجابة للمفردة والنظرية الكلاسيكية لمعادَلة اختبارات اللغة والقراءة والرياضيات في اختبار ولاية ألباما لتخريج طلبة المرحلة الثانوية بالولايات المتحدة الأمريكية، تقارب نتائج جميع الطرق المستخدمة في الدراسة. وأوصت الدراسة باستخدام طرق النظرية الكلاسيكية نظرا لسهولتها واستخدامها في بناء هذا الاختبار وعدم اختلاف نتائجها عن طرق نظرية الاستجابة للمفردة.

وتنقسم طرق معادَلة نماذج الاختبار حسب النظرية الكلاسيكية للاختبار في تصميم الجماعات المتكافئة إلى قسمين رئيسين: (1) الطرق الخطية (تفترض العلاقة الخطية بين درجات النموذجين)، (2) الطرق المئينية (تفترض العلاقة غير الخطية بين درجات النموذجين) (Kolen & Brennan، 2004). وتتمثل الطرق الخطية في:

1. **المعادَلة بالمتوسط** Mean Equating: وتقوم على افتراض أن جميع الإحصاءات الوصفية لدرجات النموذجين المراد معادلتهما متساوية ما عدا متوسطهما الحسابي. وتقوم طريقة المعادَلة بالمتوسط على معرفة المعادَلة الحسابية التي تجعل المتوسط الحسابي للنموذجين متساويا واستخدام هذه المعادَلة الحسابية لتحويل درجات النموذج الجديد إلى تدريج النموذج المرجعي.
2. **المعادَلة الخطية** Linear Equating: وتقوم على افتراض أن النموذجين المراد معادلتهما مختلفان في المتوسط الحسابي والانحراف المعياري؛ بينما تتقارب بقية الإحصاءات الوصفية وشكل التوزيع التكراري. وتستخدم طريقة المعادَلة الخطية كلا من المتوسط الحسابي والانحراف المعياري للنموذجين لتحويل درجات النموذج الجديد خطيا إلى تدريج النموذج المرجعي (Tanguma، 2000).

وتتمثل الطرق غير الخطية في:

أ. **المعادَلة المئينية غير المهدة** Unsmoothed Equipercentile Equating: تقوم هذه الطريقة على افتراض أن التوزيعات التكرارية لدرجات الطلاب في النموذجين مختلفة في جميع خصائصها، وبالتالي يختلف مستوى صعوبة النموذجين عبر الأجزاء المختلفة لتدريج درجات الاختبار. وتتم معادَلة درجات النموذجين بهذه الطريقة من خلال معرفة الدرجة المقابلة في التوزيع التكراري للنموذج المرجعي والتي لها نفس الرتبة المئينية للدرجة في التوزيع التكراري للنموذج الجديد المطلوب معادلته.

ب. **المعادَلة المئينية المهدة** Smoothed Equipercentile Equating: تعطي المعادَلة المئينية غير المهدة علاقة متعرجة وغير ممهدة بين درجات النموذجين، وخاصة في العينات صغيرة الحجم. ويتوقع عند توافر عينات كبيرة كما في المجتمع الأكبر للطلاب أن تكون هذه العلاقة ممهدة وليست متعرجة. وللحصول على علاقة مئينية ممهدة في حالة عدم توفر المجتمع، يتم استخدام طرق إحصائية لتمهيد المعادَلة المئينية. ويتوفر نوعان من طرق تمهيد المعادَلة المئينية، هما:

1. المعادَلة المئينية المهدة قبليا Pres-moothed Equipercentile Equating: وتقوم على تمهيد التوزيعات التكرارية لدرجات عينة الأفراد في النموذجين أولا، ثم إجراء المعادَلة المئينية للدرجات باستخدام التوزيع التكراري المهد.

2. المعادَلة المئينية المهدة بعديا Posts-moothed Equipercentile Equating: وتقوم على إجراء المعادَلة المئينية لدرجات النموذجين أولا باستخدام التوزيعات التكرارية الأصلية، ثم يتم تمهيد العلاقة الناتجة من المعادَلة المئينية لاحقا.

وقد تتشابه أو تختلف نتائج عملية المعادلة باستخدام الطرق الإحصائية المختلفة (Chulu & Sireci، 2011). فمثلا أشارت دراسة هاريس

وكولن (Harris & Kolen، 1990) إلى عدم اختلاف نتائج طرق المعادَلة المستخدمة في الدراسة، والتي تمثلت في المعادَلة الخطية والمعادَلة المئينية والمعادَلة وفق النموذج الثلاثي في نظرية الاستجابة للمفردة باستخدام خمسة نماذج لاختبارات المفاهيم الحسابية (Arithmetic Concepts Tests). بينما أظهرت دراسة سليند ولن (Slinde & Linn، 1997) أن المعادَلة المئينية تفوقت على طريقتي نموذج راش والنموذج الثلاثي في نظرية الاستجابة للمفردة في إجراء المعادَلة العمودية (Vertical Equating) (معادَلة نماذج الاختبار بين الصفوف الدراسية). وأظهرت دراسة كولن ووتني (Kolen & Whitney، 1982) لمقارنة أربع طرق لمعادَلة اثنتي عشر نموذجا من اختبارات التنمية التربوية العامة في تصميم الجماعات المتكافئة، أن كلا من طريقة المعادَلة الخطية والمعادَلة المئينية ونموذج راش حققت مؤشرات أفضل من طريقة النموذج ثلاثي المعالم في نظرية الاستجابة للمفردة.

وقام الشريفين (2009) بمقارنة طريقتي المعادَلة الخطية والمعادَلة المئينية في النظرية الكلاسيكية للقياس في تصميم الجماعات المتكافئة لإجراء معادَلة لسبعة نماذج مختلفة لاختبار الكفاءة في اللغة الإنجليزية الذي عقدته **جامعة اليرموك للعام** الدراسي 2006/2007م. واستخدمت الدراسة عينة بحجم 1199 طالبا موزعين عشوائيا على النماذج السبعة بأحجام مختلفة. وأظهرت النتائج تقارب الدرجات المعادِلة بين طرق النظرية الكلاسيكية وطريقة المعادَلة باستخدام نموذج راش في نظرية الاستجابة للمفردة.

ويمكن تفسير الاختلاف في نتائج الطرق الإحصائية للمعادلة من خلال إدراك أن هناك عدداً من العوامل التي تؤثر على نتائج كل طريقة من طرق المعادَلة، مما يؤدي إلى الاختلاف فيما بينها، ومن بين هذه العوامل مقدار الاختلاف بين نماذج الاختبار في الإحصاءات الوصفية والتوزيعات التكرارية، وعدد أسئلة الاختبار، وحجم العينات المستخدمة. ومن بين الإشكاليات في مقارنة نتائج المعادَلة بين التصاميم المختلفة وطرق المعادَلة عدم وجود محكات متفق عليها يمكن استخدامها لتحديد التصميم وطريقة المعادَلة التي تعطي أفضل النتائج. ومن بين المحكات الأكثر شيوعا هي: 1) حساب تحيز (Bias) الدرجات المعادِلة المتحصل عليها بالطرق المختلفة، وحساب متوسطاتها غير الموزونة ومتوسطاتها الموزونة، 2) الخطأ المعياري للمعادَلة Standard Error of Equating، والذي يحسب بالانحراف المعياري للدرجات المعادِلة الناتجة من استخدام عينات مختلفة في إجراء المعادَلة باستخدام أي طريقة معادَلة (Kolen & Brennan، 2004). ويفضل استخدام طريقة المعادَلة التي يكون خطؤها المعياري أصغر ما يمكن.

وتختلف مؤسسات إعداد الاختبارات في تصاميم وطرق معادَلة نماذج اختباراتها (Kolen & Brennan، 2004). فتستخدم اختبارات الكليات الأمريكية (American College Test (ACT مثلا طريقة المعادَلة المئينية في تصميم الجماعات المتكافئة، بينما يستخدم اختبار التقييم الدراسي (Scholastic Assessment Test (SAT طريقة المعادَلة المئينية في تصميم المفردات المشتركة. وتستخدم اختبارات التنمية التربوية العامة (General Educational Development (GED التي يقدمها المجلس الأمريكي في التربية American Council on Education طريقة المعادَلة المئينية المهدة قبليا في تصميم الجماعات المتكافئة لمعادَلة نماذج الاختبارات في العام الواحد وعبر الأعوام بتقديم النموذج المرجعي المعد في عام 2002م مع النماذج الجديدة (Ezzelle & Setzer، 2009).

وقام الشريفين (2009) في دراسته بالتأكد فيما كانت عشوائية توزع الطلاب على النماذج السبعة اختبار اللغة الإنجليزية في الاختبارات التوجيهية

بالتعليم العام كفيلة بتحقيق تكافؤ المجموعات الطلابية في السمة المقيسة بالاختبار في تصميم الجماعات المتكافئة. حيث قام الشريفين بمقارنة مستوى الطلاب في النماذج السبعة في اختبار اللغة الإنجليزية، ووجد عدم وجود فروق دالة إحصائيا بين مستويات المجموعات السبع التي أخذت النماذج السبعة. بينما بينت نتائج هذه الدراسة عدم تكافؤ النماذج السبعة لاختبار الكفاءة في متوسطاتها الحسابية وانحرافاتها المعيارية، واستنتج عدم صلاحية استخدام الدرجة 25 المعتمدة كدرجة نجاح في جميع النماذج السبعة نظرا لاختلاف مستوى صعوبة هذه النماذج. وباستخدام طريقتي المعادَلة الخطية والمعادَلة المئينية لستة نماذج مع النموذج المرجعي، قدمت الدراسة الدرجات المعادلة للدرجات الخام للنماذج السبعة التي يصلح استخدامها في تحويل الدرجات بين النماذج السبعة. وخلصت الدراسة إلى اختلاف الدرجة الخام في النماذج الستة المقابلة للدرجة 25 في النموذج المرجعي المستخدمة لتحديد نجاح الطالب في اختبار الكفاءة من عدمه.

وأكد الشريفين (2009) على أهمية إجراء معادَلة للنماذج المختلفة للاختبار، حيث إن عدم تكافؤ درجات نماذج المختلفة قد يؤدي إلى اتخاذ قرارات خاطئة لمن ينجح أو يرسب في الاختبار، فالطالب الذي قد ينجح في نموذج معين قد لا يتمكن من ذلك في حالة تقديمه لنموذج آخر لوتم استخدام درجة النجاح نفسها في النموذجين من دون إجراء معادَلة.

مشكلة الدراسة وأهميتها

يقع على عاتق المركز الوطني للقياس والتقويم إعداد اختبار القدرات العامة لتحقيق العدالة والمساواة في فرص القبول للطلاب بمؤسسات التعليم العالي بالمملكة العربية السعودية. ونظرا لخطورة وأهمية القرارات في تحديد قبول الطالب بمؤسسات التعليم العالي من دونه، يحرص المركز الوطني للقياس والتقويم في التعليم العالي على تحقيق أعلى درجات التكافؤ بين نماذج اختبار القدرات العامة. وفي هذا الصدد، يتبع المركز الوطني سياسة واضحة وصارمة في كتابة مفردات متكافئة في محتواها وبناء نماذج متكافئة من حيث جدول المواصفات و تقارب دالة المعلومات لهذه النماذج من خلال توظيف الكفاءة النسبية للنماذج باستخدام نظرية الاستجابة للمفردة، ويتم ذلك من خلال انتقاء المفردات من مفردات تجريبية وضعت في بعض النماذج السابقة للاختبار، والاستفادة من الخصائص السيكومترية لهذه المفردات لتحقيق تكافؤ نماذج الاختبار المكونة قبل تطبيقها.

كما يتبع المركز الوطني سياسة توزيع الطلاب على أيام ونماذج الاختبار بطريقة تضمن عشوائية توزع الطلاب على نماذج الاختبار، مما يزيد من فرص تكافؤ المجموعات الطلابية التي تقدم النماذج المختلفة للاختبار. بينما وحرصا من القائمين على اختبار القدرات العامة على تطوير الاختبار وبنائه بأفضل الطرق السيكومترية الملائمة، يدور في هاجس المسؤولين عن هذا الاختبار تساؤلات عما إذا كانت الطرق المتبعة حاليا في بناء نماذج الاختبار المختلفة تحقق تكافؤ نماذج الاختبار. وفي حالة عدم كفاية الإجراءات المتبعة حاليا، فما هي الطرق الملائمة لإجراء معادلة نماذج الاختبار. فيدرك المسؤولون أن هناك عدداً من العوامل المرتبطة بتطبيق نماذج الاختبار على الطلاب في أيام وظروف مختلفة تجعل افتراض تكافؤ نماذج الاختبار غير مسلم به. ومن أمثلة هذه العوامل التباين في مستوى قدرات الطلبة الذين يقدمون نماذج الاختبار، والتباين في ظروف التطبيق، والتباين في استعداد الطلبة بين أيام التطبيق، والتباين في عدد الطلبة الذين قدموا الاختبار لأكثر من مرة سابقا، وغيرها. وقد تؤدي هذه العوامل إلى وجود اختلافات في الإحصاءات الوصفية (المتوسطات الحسابية والتباين) وشكل التوزيعات التكرارية في بيانات نماذج الاختبار.

أهداف الدراسة:

تهدف هذه الدراسة إلى التحقق من تكافؤ نماذج اختبار القدرات العامة من حيث تساوي متوسطاتها الحسابية وتباين الدرجات وتشابه توزيعاتها التكرارية. وتسعى الدراسة بعد ذلك إلى تطبيق عدد من طرق المعادلة الكلاسيكية في تصميم الجماعات المتكافئة على نماذج الاختبار ومقارنتها في كل من اختبار القدرة اللفظية واختبار القدرة الكمية في اختبار القدرات العامة. واقتصرت الدراسة الحالية على تطبيق طرق المعادلة الكلاسيكية لما تتميز بها هذه الطرق من سهولة في التطبيق وسهولة تحقق افتراضات النظرية الكلاسيكية وطرق المعادلة الكلاسيكية. كما دعمت الدراسات المذكورة سابقا استخدام طرق المعادلة الكلاسيكية وتحقيقها لنتائج معادلة تقترب من نتائج طرق المعادلة باستخدام نماذج نظرية الاستجابة للمفردة، وفي بعض الدراسات تتفوق طرق المعادلة الكلاسيكية على نظيراتها في نظرية الاستجابة للمفردة (Kolen & Whitney، 1982; Slinde & Linn، 1997). كما يتم حساب الدرجات الخام والدرجات المعيارية في اختبار القدرات العامة باستخدام النظرية الكلاسيكية، مما يتوافق مع توظيف الطرق الإحصائية في النظرية الكلاسيكية لإجراء معادلة نماذج الاختبار. وتتلخص أهداف الدراسة في المحاولة للإجابة عن الأسئلة التالية:

1. هل تعتبر نماذج اختباري القدرة اللفظية والقدرة الكمية في اختبار القدرات العامة التي يستخدمها المركز الوطني للقياس في التعليم العالي متكافئة؟
2. هل تعطي طرق المعادَلة الكلاسيكية في تصميم الجماعات المتكافئة درجات معادِلة (الخام والمعيارية) مختلفة في كل من نماذج اختباري القدرة اللفظية ونماذج اختبار القدرة الكمية؟
3. ما هي أفضل طرق المعادَلة الكلاسيكية في تصميم الجماعات المتكافئة التي تحقق أصغر مقادير للأخطاء المعيارية للمعادَلة في كل من نماذج اختبار القدرة اللفظية ونماذج اختبار القدرة الكمية؟

أهمية الدراسة:

تكمن أهمية هذه الدراسة في مقارنة عدد من الطرق الإحصائية لإجراء معادلة نماذج اختبار القدرات العامة، مما يساعد في اتخاذ القرار المناسب في تبني الطريقة الإحصائية الأكثر ملاءمة لإجراء معادلة النماذج. كما تكتسب هذه الدراسة أهميتها من القيمة التي يمثلها اختبار القدرات العامة من حيث عدد المتقدمين وخطورة القرارات المبنية على نتائجه. كما يتوقع أن تسهم هذه الدراسة في نشر ثقافة معادلة نماذج الاختبارات في البيئة العربية.

مصطلحات الدراسة:

معادَلة نماذج الاختبار (Test Forms Equating): تحويل درجات أحد نماذج الاختبار إلى ما يناظرها من درجات النموذج المرجعي للاختبار، بحيث تصبح القياسات المستمدة من درجات كل من النموذجين متكافئة. وفي هذه الدراسة تم حساب الدرجات الخام لنماذج الاختبار (مجموع الإجابات الصحيحة على فقرات نموذج الاختبار)، والدرجات المعيارية المقابلة لها (تحويل خطي للدرجات الخام بحيث يصبح متوسطها الحسابي 65 وانحرافها المعياري 10).

نماذج الاختبارات المتكافئة (Equivalent Test Forms): نماذج الاختبار التي يتم بناؤها بحيث تحتوي على مفردات مختلفة وفي ذات الوقت تقيس سمة واحدة وتتبع جدول مواصفات موحد، وفي حالة تطبيقها على مجموعات كبيرة جدا ومتكافئة في السمة المقيسة، يكون لها نفس المتوسطات الحسابية وتباين درجات الأفراد، وتتشابه توزيعاتها التكرارية.

طرق المعادَلة في النظرية الكلاسيكية (Classical Theory Equating Methods): طرق المعادلة

التي تستخدمها النظرية الكلاسيكية في القياس والتي تتميز ببساطة وسهولة تحقق افتراضاتها والتي تعتمد على الدرجات الخام للطلاب. وتتمثل هذه الطرق في المعادلة بالمتوسط والمعادلة الخطية والمعادلة المئينية والمعادلة المئينية المهدة (قبليا وبعديا). وفي هذه الدراسة تم توظيف كل من المعادلة الخطية والمعادلة المئينية والمعادلة المئينية المهدة بعديا.

اختبار القدرات العامة: يهدف اختبار القدرات العامة إلى قياس الاستعداد الدراسي للقبول بمؤسسات التعليم العالي بالمملكة العربية السعودية، ويسعى إلى تقديم معلومات صادقة عن قدرة الطالب على التعامل مع المهارات والعمليات التي تتطلبها البرامج الدراسية بمؤسسات التعليم العالي. ويقيس اختبار القدرات العامة قدرتين أساسيتين للنجاح في التعليم العالي وهي القدرة اللفظية والقدرة الكمية. ويهدف قياس القدرة اللفظية في الاختبار إلى تزويد الجامعات بمعلومات عن مستوى الطالب ومقدار تمكنه من فهم محتوى ومعنى ما يقرؤه وقدرته على استنباط الأفكار من النصوص والربط بينها والوصول إلى تعميمات صحيحة تستند إلى أدلة وشواهد. ويهدف قياس القدرة الكمية في الاختبار إلى تزويد الجامعات بمعلومات عن مستوى الطالب ومقدار تمكنه من فهم المفاهيم الرياضية والعمليات الحسابية المختلفة والجبر والهندسة والعلاقات الرياضية والقدرة على الوصول إلى تعميمات رياضية صحيحة تستند إلى أدلة وشواهد.

حدود الدراسة:

1. اقتصرت الدراسة على النماذج الأربعة لكل من اختبار القدرة اللفظية واختبار القدرة الكمية في اختبار القدرات العامة الذي يعده ويقدمه المركز الوطني للقياس والتقويم في التعليم العالي بالمملكة العربية السعودية.
2. اقتصرت الدراسة على عينة الدراسة التي طبق عليها النماذج الأربعة لاختبار القدرة اللفظية واختبار القدرة الكمية في العام الأكاديمي 1430/1431هـ.
3. اقتصرت الدراسة على مقارنة ثلاث طرق للمعادلة في النظرية الكلاسيكية للقياس، وهي: المعادلة الخطية، المعادلة المئينية، المعادلة المئينية المهدة بعديا.

المنهجية والإجراءات

العينة:

تكونت عينة الدراسة من طلبة المرحلة الثانوية (بالصفين الثاني الثانوي والثالث الثانوي) من مختلف مناطق المملكة العربية السعودية في الفترة الثانية بنهاية العام الأكاديمي 1430/1431هـ ممن طبق عليهم كل من نماذج الاختبار التالية: (1081)، (1082)، (1083)، (2034). وتم تطبيق النماذج الأربعة في ثلاثة أيام من نفس الأسبوع، على النحو الذي يبينه الجدول (1). وفي كل يوم من أيام التطبيق، اختيرت عشرة مراكز لتطبيق الاختبار موزعة على مختلف مناطق المملكة العربية السعودية للحصول على تمثيل ملائم لمجتمع طلاب المملكة العربية السعودية. وحيث يتطلب تصميم الجماعات المتكافئة تشابه ظروف تطبيق النموذجين، وُزع كل نموذجين يُراد معادلتهما على الطلبة بكل قاعة امتحانية بطريقة عشوائية، من خلال ترتيب أوراق النموذجين بالطريقة التصاعدية (Spiraling). ويساعد هذا النظام في توزيع أوراق النموذجين في تحقيق التشابه لأقصى درجة ممكنة لمستوى قدرات الطلبة وظروف التطبيق بين نموذجي الاختبار. وتحققت عشوائية عينة الدراسة من خلال اختيار الطلاب بأنفسهم لتوقيت الاختبار ومركز تقديم الاختبار والقاعات الامتحانية، من دون تدخل من قبل المركز الوطني للقياس في اختيار توقيت الاختبار للطالب، ما عدا تحديد الطاقة الاستيعابية لمركز تقديم الاختبار في أيام التطبيق.

جدول (1)

توزيع عينات الدراسة على نماذج الاختبار وأيام التطبيق

يوم التطبيق	النموذج الجديد		النموذج المرجعي	
	رقم النموذج	العدد	رقم النموذج	العدد
السبت	(١٠٨٢)	٢٠٣٩	(١٠٨١)	٢٣٤٩
الاثنين	(٢٠٣٤)	٢٥٧٠	(١٠٨١)	٢٥٧٠
الاثنين	(٢٠٣٤)	٢٥٦٤	(١٠٨٢)	٢٥٦٤
الخميس	(١٠٨٣)	٢١٢٠	(١٠٨٢)	٢١٢٠

أدوات الدراسة:

استخدمت الدراسة اختباري القدرة اللفظية والقدرة الكمية في اختبار القدرات العامة المعد من قبل المركز الوطني للقياس والتقويم بالمملكة العربية السعودية. وتتكون جميع نماذج اختبار القدرة اللفظية من 68 مفردة من نوع الاختيار من متعدد (أربعة بدائل) موزعة على أربعة مكونات بتوزيع معين ومتشابه عبر النماذج المختلفة. وتتمثل مكونات القدرة اللفظية في الاختبار في معاني الكلمات، ويشكل ما وزنه 20 ٪ تقريبا؛ وإكمال الجمل، ويشكل ما وزنه ٪23 تقريبا؛ والتناظر اللفظي، ويشكل ما وزنه 26 ٪ تقريبا؛ واستيعاب المقروء، ويشكل ما وزنه 31 ٪ تقريبا. وتتكون جميع نماذج اختبار القدرة الكمية من 52 مفردة من نوع الاختيار من متعدد (أربعة بدائل) تقيس كلا من العمليات الحسابية والجبر والهندسة والعلاقات الرياضية.

وتتراوح الدرجة الخام (عدد الإجابات الصحيحة للطالب في أسئلة كل نموذج) لاختبار القدرة اللفظية بين الدرجة صفر والدرجة 68، بينما تتراوح الدرجة الخام لاختبار القدرة الكمية بين الدرجة صفر والدرجة 52. تحول الدرجة الخام في كلا الاختبارين تحويلا خطيا إلى درجة معيارية بمتوسط حسابي وانحراف معياري محددين يساعد في تفسير الدرجات معياريا. وتحسب الدرجة المعيارية كدرجة زائية (باستخدام المتوسط الحسابي والانحراف المعياري لمجموعة الطلبة الذين طبقوا كل نموذج) وتحويلها خطيا لتأخذ متوسطا حسابيا قدره 65 وانحرافا معياريا قدره 10.

استخدمت أربعة نماذج من الاختبارين، وهي: (1081)، (1082)، (1083)، (2034). وتتشابه هذه النماذج في عدد المفردات بكل منها، وفي جدول مواصفاتها وخصائصها السيكومترية، بينما تختلف هذه النماذج من حيث أسئلتها، حيث لا يوجد بينها أي مفردات مشتركة. وتم التأكد من تشابه معامل ثبات هذه النماذج من خلال حساب معامل ثبات ألفا لكرونباخ لنتائج جميع نماذج اختباري القدرة اللفظية والقدرة الكمية، حيث تراوحت بين 0,868 و 0,896، مما يشير إلى ثبات عالٍ وتشابهه بين النماذج الأربعة في كلا الاختبارين.

المنهجية والاجراءات:

طبق تصميم الجماعات المتكافئة لمعادلة درجات النماذج الأربعة لاختباري القدرات العامة، بحيث طبق نموذجان من كلا الاختبارين في كل تطبيق من التطبيقات الأربعة كما هو موضح في الجدول (1)، بحيث يمثل أحد النموذجين النموذج الجديد الذي يراد معادَلة درجاته ولنرمز أنه يمثل المتغير س، ويمثل النموذج الثاني النموذج المرجعي الذي يشكل التدريج المرجعي التي تُعادَل درجات النموذج الجديد إليه، ولنقل أنه يمثل المتغير ص.

وللحصول على الدرجات المعادِلة، طبقت ثلاث طرق إحصائية لمعادَلة النموذج الجديد مع النموذج المرجعي عن طريق تحويل درجات

النموذج الجديد (س) إلى تدريج النموذج الجديد (ص) في التطبيقات الأربعة باستخدام برنامج المعادَلة RAGE-RGEQUATE من إعداد كولن (Kolen، 2004). وتمثلت الطريقة الإحصائية الأولى في المعادَلة الخطية (Linear Equating)، والتي يتم فيها تحويل الدرجة س إلى الدرجة ص بتحويل خطي باستخدام المتوسط الحسابي للنموذجين (مس و مص) والانحراف المعياري لهما (عس و عص). وبالتالي فإن الدرجة المعادِلة لأي درجة (س) في النموذج الجديد في تدريج النموذج المرجعي هي صم، وتحسب كالتالي:

$$ص_م = م_{ص} + (ع_{ص} / ع_{س})(س - م_{س})$$

وينتج من المعادلة الخطية تساوي المتوسط الحسابي والانحراف المعياري للدرجات المعادِلة لدرجات النموذج الجديد (س) مع المتوسط الحسابي والانحراف المعياري لدرجات النموذج المرجعي (ص)، على الترتيب. بينما تحتفظ الدرجات المعادِلة لدرجات النموذج الجديد بشكل توزيعها التكراري قبل المعادَلة.

وكانت الطريقة الإحصائية الثانية هي المعادَلة المئينية (Equipercentile Equating). وتستخدم هذه الطريقة الرتب المئينية للدرجات في النموذجين. والرتبة المئينية هي التكرار النسبي المتجمع الصاعد لعدد الطلاب الذين حصلوا على درجات أكبر من أو تساوي كل درجة من درجات الاختبار، وتتراوح بين 1 وحتى 99. وتُعرف المعادلة المئينية الدرجة المعادِلة لأي درجة (س) في النموذج الجديد في تدريج النموذج المرجعي(صم) بأنها الدرجة (ص) في النموذج المرجعي التي لها نفس الرتبة المئينية للدرجة (س) في النموذج الجديد المراد معادلته. وينتج من المعادلة المئينية تشابه الإحصاءات الوصفية (المتوسط الحسابي والانحراف المعياري) وشكل التوزيع التكراري للدرجات المعادلة لدرجات النموذج الجديد (س) مع تلك التي يقابلها في درجات النموذج المرجعي.

وتمثلت الطريقة الإحصائية الثالثة بالمعادَلة المئينية المهدة بعديا (Postsmoothed Equipercentile Equating). وتعالج هذه الطريقة مشكلة التعرج في تحويل الدرجات بين النموذجين الناتج من المعادلة المئينية، حيث يتم الحصول على الدرجات المعادِلة لدرجات النموذج الجديد بنفس طريقة المعادلة المئينية، ثم يتم توظيف أحد أساليب التمهيد الإحصائي لتمهيد التحويل الناتج. ومن بين أساليب التمهيد الإحصائي البعدي الأكثر استخداما هو أسلوب "التمهيد بالمكعبات Cubic Spline"، والذي يتوفر فيه عدة درجات للتمهيد. وتؤثر درجة التمهيد في مستوى التمهيد ومقدار قرب أو بعد المعادلة المئينية المهدة عن المعادلة المئينية الأصلية، فكلما زادت درجة التمهيد المستخدمة، كلما كان التحويل أكثر تمهيداً ولكنه يبتعد كثيرا عن التحويل الأصلي بالمعادلة المئينية. ولا توجد طريقة إحصائية متفق عليها لتحديد درجة التمهيد المناسب، ولكن يمكن تحديد درجة التمهيد المناسب عن طريق الحكم الشخصي على مستوى التمهيد المقبول من دون الابتعاد كثيرا عن التحويل الأصلي. وبعد تجريب عدة درجات لتمهيد المعادَلة المئينية، أظهرت الدرجة 0,1 تمهيدا ملائما للمعادَلة المئينية في جميع النماذج مع عدم ابتعاد المعادَلة المئينية المهدة بعديا بالقدر الكبير عن المعادَلة المئينية الأصلية.

بعد ذلك حسبت الدرجات الخام المعادِلة لدرجات النموذج الجديد حسب النموذج المرجعي بالطرق الإحصائية الثلاث حسب ما تم وصفه، ثم حسبت الدرجة المعيارية لهذه الدرجات الخام المعادِلة للنموذج الجديد من خلال استخدام الدرجات المعيارية للنموذج المرجعي المتبعة في المركز الوطني للقياس، وهي درجات معيارية خطية بمتوسط حسابي 65 وانحراف معياري 10، بحيث تتراوح بين صفر - 100. وفي حالة تجاوز الدرجة المعيارية عن 100، يتم تثبيتها عند الدرجة 100 كسقف علوي للدرجات المعيارية. كما حسبت الدرجة المعيارية لكل

نموذج جديد يراد معادلته باستخدام بياناته الأصلية (بمتوسط حسابي 65 وانحراف معياري 10)، والذي يمثل تحويل درجات النموذج إلى درجات معيارية من دون إجراء معادَلة عليه (أطلق عليه اسم الدرجات المعيارية بدون معادلة) من أجل المقارنة.

المعالجات الاحصائية:

وللإجابة عن أسئلة الدراسة، حسبت المتوسطات الحسابية والانحرافات المعيارية والتوزيعات التكرارية للدرجات الخام في جميع نماذج الاختبارين (القدرة اللفظية والقدرة الكمية) في التطبيقات الأربعة. كما تم استخدام اختبارت لمجموعتين مستقلتين لدراسة الدلالة الإحصائية للفروق في المتوسطات الحسابية لنموذجي الاختبار المراد معادلتهما في التطبيقات الأربعة، وتم استخدام اختبار ليفين (Leivn's F test) لدراسة الدلالة الإحصائية للفروق بين تباين درجات الأفراد في النموذجين للحكم على ما إذا كانت نماذج الاختبارات متكافئة. كما قورنت قيم الالتواء والتفلطح بين النموذجين في كل من اختبار القدرة اللفظية والقدرة الكمية.

ولمقارنة الدرجات المعادِلة للنماذج الأربعة الجديدة بين الطرق الإحصائية للمعادلة، تم حساب تحيز الدرجات المعادِلة وابتعادها عن نتائج المعادلة المساوية Identity Equating. وتفترض المعادلة المساوية تكافؤ النموذجين بشكل تام وتساوي الدرجات الخام للنموذجين. ويمكن استخدام الدرجات المعيارية للنموذج المرجعي بشكل مباشر لتفسير الدرجات في النموذج الجديد. ويعتبر كولن وبرنن (Kolen & Brennan، 2004) أن المعادَلة المساوية تقدم محكا مشتركا بين جميع طرق المعادَلة ومستقلا عن أي منها، بحيث لا تتحيز لأي طريقة من طرق المعادَلة. ويساعد حساب تحيز الدرجات المعادِلة بهذه الطريقة معرفة مقدار اختلاف صعوبة النموذجين عبر درجات الاختبار ومقدار اختلاف الدرجات المعادِلة عن الدرجات الأصلية للنموذج الجديد. وتم حساب التحيز لكل من الدرجات الخام المعادِلة والدرجات المعيارية المعادِلة من خلال حساب الفرق بين الدرجة المعادِلة (الخام والمعيارية) الناتجة من طريقة المعادَلة والدرجة (الخام والمعيارية) الناتجة من المعادَلة المساوية. أي استخدمت العملية الحسابية التالية لحساب التحيز لكل درجة من درجات التدريج:

التحيز لكل درجة = الدرجات المعادلة لكل درجة في النموذج الجديد – الدرجة الناتجة من المعادَلة المساوية.

وطبقت المعادلة السابقة على كل من الدرجات الخام المعادِلة والدرجات المعيارية المعادِلة الدرجات المعيارية بدون معادَلة لكل نموذج من نماذج الاختبار.

ولتلخيص تحيزات الدرجات المعادِلة (الخام والمعيارية) عبر درجات الاختبار، حُسب المتوسط غير الموزون للفروق (Unweighted Mean of Signed Difference) كالتالي:

المتوسط غير الموزون للفروق = مجـ (التحيز لكل درجة) / (عدد الأسلة + 1)،

حيث تم حساب مجموع التحيزات عبر جميع درجات تدريج الاختبار. وحُسب أيضا المتوسط الموزون للفروق (حسب تكرار الدرجة في النموذج الجديد) (Weighted Mean of Signed Dif-ference) كالتالي:

المتوسط الموزون للفروق = مجـ (التحيز لكل درجة X تكرار الدرجة) / (حجم العينة).

كما حُسب الخطأ المعياري للمعادَلة (Stan-dard Error of Equating) لكل درجة خام معادِلة ولكل درجة معيارية معادِلة كمؤشر للخطأ العشوائي في المعادَلة الناتج من اختلاف العينات، وحُسب الخطأ المعياري للمعادَلة بطريقة Boot-strap (Efron، 1982; Efron & Tibshirani، 1993). ويحسب الخطأ المعياري للمعادلة بطريقة Bootstrap من خلال حساب الانحراف المعياري

للدرجـات المعادلـة الناتجة من 500 إعـادة للمعادَلة بـين النموذجـين. وفي كل إعـادة للمعادَلـة، سُـحبت عينة عشـوائية مع الإرجاع بنفس حجم عينة الطلاب مـن درجـات كل مـن النموذجين الجديـد والمرجعي. وتم تكـرار هـذه العمليـة لـكل طريقـة إحصـائية في النمـاذج الأربعة لكلا الاختبارين. وتم توظيف برنامج (Equating error" (Hanson، 2004") في حسـاب الخطأ المعيـاري بهذه الطريقـة. والعملية الحسابية التالية تصف طريقة حساب الخطأ المعياري لكل درجة:

الخطـأ المعيـاري للمعادَلـة لـكل درجة= الجـذر التربيعـي (مجـ (الدرجـة المعادِلـة لـكل درجـة – المتوسط) / 500) ،

حيث حسب المتوسط في المعادلة كمتوسط حسابي للدرجـات المعادِلة في 500 إعادة للمعادلة. ولتلخيص الأخطـاء المعيارية للمعادلة، حُسـب المتوسـط الموزون لهـذه الأخطـاء المعيارية للمعادَلة عـبر جميع درجات الاختبار للحكم الكلي على الدرجات المعادِلة المتحصل عليهـا مـن كل طريقة إحصـائية باسـتخدام العملية الحسابية التالية:

المتوسـط المـوزون للخطـأ= مجـ (مربـع الخطـأ المعيـاري للمعادلـة لـكل درجـة X تكـرار الدرجة) / (حجم العينة) .

نتائج الدراسة

الإحصاءات الوصفية:

يعرض الجدول (2) والجدول (3) الإحصـاءات الوصـفية للنمـاذج الأربعة لاختباري القـدرة اللفظية والقـدرة الكمية، على التوالي. وتظهر النتائج اختلافا كبيرا نسـبيا بـين المتوسـطات الحسـابية للنموذجين الجديد والمرجعـي المراد معادلتهمـا في أيام التطبيق الثلاثـة. وعنـد اختبار دلالـة الفروق في المتوسـطات الحسـابية بين النموذجين المراد معادلتهما في كل يوم مـن أيـام التطبيق لكل مـن القدرة اللفظيـة والقدرة الكمية باسـتخدام اختبـار ت لمجموعتين مسـتقلتين، وُجـد أنها جميعا دالة إحصـائيا عند مسـتوى الدلالة 0,05 (تراوح حجم الأثـر بين 0,085 و0,351 في نمـاذج القـدرة اللفظية، وبـين 0,367 و0,773 في نمـاذج القدرة الكميـة). ففي نماذج اختبـار القدرة اللفظية، يزيد المتوسـط الحسابي للنموذج (1082) عن المتوسـط الحسابي للنموذج (1081)، مما يشير إلى أن النمـوذج الجديـد (1082) أقل صـعوبة عن النمـوذج المرجعـي 1081. وكذلك النمـوذج الجديد (2034) أقل صعوبة من النموذج المرجعي (1081)، بينما كان النموذج الجديد (1083) أكثر صعوبة من النمـوذج المرجعـي (1082)، وكذلـك كان النموذج الجديد (2034) أكثر صـعوبة من النموذج المرجعي (1082). وعنـد تفحص تكافؤ درجـات النموذجين المـراد معادلتهمـا مـن حيث تبايـن درجـات الأفراد باسـتخدام اختبار ليفين (Levine's F test)، وُجد عـدم تسـاوي تباين درجـات كل من القـدرة اللفظية والقدرة الكميـة بين النموذجين (2034) و(1082) و النموذج (1081).

جدول (2). الإحصاءات الوصفية للنماذج الأربعة لاختبار القدرة اللفظية

	النموذج الجديد					النموذج المرجعي				
يوم التطبيق	النموذج	م	ع	الالتواء	التفلطح	النموذج	م	ع	الالتواء	التفلطح
السبت	١٠٨٢	٣٥,٩٤ ×	١٠,٥٤	٠,١٠	٢,٢٩	١٠٨١	٣٢,٢٢	١٠,٦٤	٠,٠٢٨	٢,٤٢
الاثنين	٢٠٣٤	٣٦,٠٧ ×	١٠,١٥ ×	– ٠,٠٥	٢,١٣	١٠٨١	٣٢,٩٢	١٠,٥٩	٠,١٤	٢,٢١
الاثنين	٢٠٣٤	٣٦,٠٧ ×	١٠,١٥ ×	– ٠,٠٥	٢,١٣	١٠٨٢	٣٦,٩٦	١٠,٧٢	– ٠,٠٥	٢,٢١
الخميس	١٠٨٣	٣٦,٢٤ ×	١٠,٦١	– ٠,٠٤	٢,٣٣	١٠٨٢	٣٨,١٩	١٠,٩٩	– ٠,١٧	٢,٢٢

م: المتوسط الحسابي، ع: الانحراف المعياري × الفروق بين النموذجين (الجديد والمرجعي) دالة عند مستوى a = 0,05

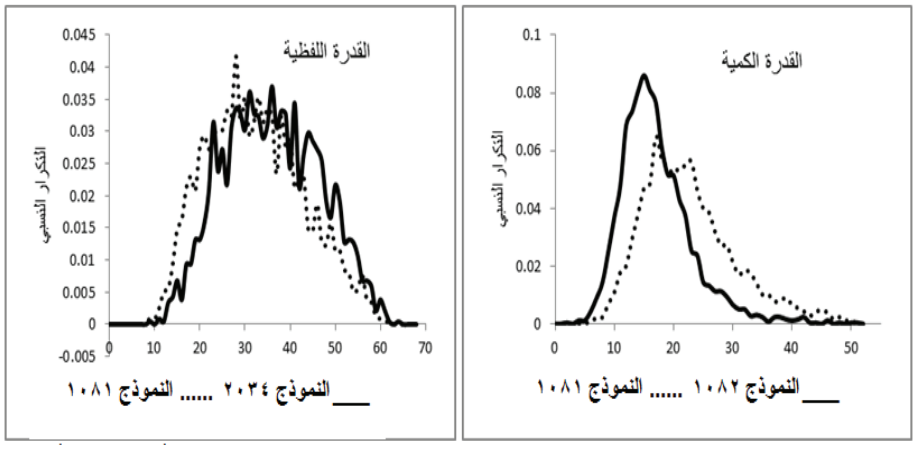
جدول (3). الإحصاءات الوصفية للنماذج الأربعة لاختبار القدرة الكمية

النموذج الجديد					النموذج المرجعي				
يوم التطبيق	النموذج	م	ع	الالتواء	التفلطح	النموذج	م	ع	الالتواء	التفلطح
السبت	١٠٨٢	×١٩,٤٩	×٧,٠٨	٠,٩٩	٤,٢٤	١٠٨١	٢٢,٢٤	٧,٨٢	٠,٩٦	٤,٠٢
الاثنين	٢٠٣٤	×١٧,١٣	×٦,٠٧	١,٢٩	٦,٠٣	١٠٨١	٢٢,٥٤	٧,٨٢	٠,٨٣	٣,٥٨
الاثنين	٢٠٣٤	×١٧,١٣	×٦,٠٧	١,٢٩	٦,٠٣	١٠٨٢	٢٠,٣٣	٧,٢٦	٠,٩٩	٤٠٣٣
الخميس	١٠٨٣	×١٧,١٩	×٥,٩٠	١,١١	٥,٣٨	١٠٨٢	٢٠,٩٦	٧,٤٥	٠,٧٥	٣,٢٨

م: المتوسط الحسابي، ع: الانحراف المعياري × الفروق بين النموذجين (الجديد والمرجعي) دالة عند مستوى a = 0,05

كما يختلف مستوى الالتواء والتفلطح للتوزيعات التكرارية لنماذج اختبار القدرة الكمية عن بعضها البعض بدرجة أكبر مقارنة بنماذج اختبار القدرة اللفظية. ويظهر الشكل (1) مثالين على عدم تطابق شكل التوزيعات التكرارية لدرجات الطلاب في نماذج الاختبارين. وتشير هذه النتائج في مجملها إلى أن النموذجين المراد معادلتهما في جميع أيام التطبيق غير متكافئين من حيث متوسطاتهما الحسابية وتباينهما وبقية الإحصاءات الوصفية لتوزيعاتهما التكرارية.

شكل (1). التوزيع اتكراري للنموذجين (1082) و (1081) يوم السبت لاختباري القدرة اللفظية والقدرة الكمية .



الفروق بين نتائج طرق المعادَلة (مقدار التحيز):

معادَلة النموذجين (1082) و (1081): يظهر الجدول (4) والجدول (5) الإحصاءات الوصفية لكل من الدرجات المعيارية بدون معادلة للنموذج (1082)، و الدرجات المعادِلة (الخام والمعيارية) للقدرة اللفظية والقدرة الكمية الناتجة من طرق المعادَلة الثلاث للنموذج (1082) مع النموذج (1081). و تظهر النتائج أن جميع الطرق الثلاث والدرجات المعيارية للنموذج الجديد بدون معادلة وحدت المتوسطات الحسابية والانحرافات المعيارية للدرجات المعادِلة للنموذج (1082) لتتطابق مع النموذج المرجعي (1081). بينما ظلت الخصائص الأخرى (الالتواء والتفلطح) للتوزيع التكراري للدرجات المعادِلة كما هي عليه في الطريقة الخطية و الدرجات المعيارية للنموذج الجديد بدون معادلة ، وتطابقت هذه الخصائص مع درجات النموذج المرجعي في الطريقة المئينية والطريقة المئينية المهدة بعديا.

جدول (4). الاحصاءات الوصفية لدرجات القدرة اللفظية في النموذج المرجعي (1081) والدرجات المعادِلة للنموذج الجديد (1082).

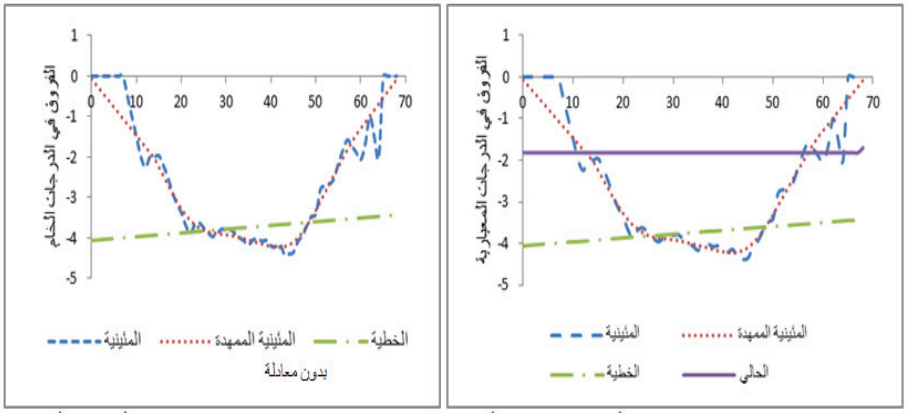
الدرجة	الطريقة (النموذج)	المتوسط الحسابي	الانحراف المعياري	الالتواء	التفلطح
الدرجة الخام	النموذج المرجعي	٣٢,٢١٦	١٠,٦٣٣	٠,٢٨٠	٢,٤١٦
	النموذج الجديد قبل المعادلة	٣٥,٩٤٠	١٠,٥٣٥	٠,١٠٠	٢,٢٩٣
الدرجات الخام المعادِلة	المعادَلة المئينية	٣٢,٢١٥	١٠,٦٤٢	٠,٢٨٥	٢,٤٢١
	المعادَلة المئينية المهدة بعديا	٣٢,٢٢٠	١٠,٦٢٧	٠,٢٨٠	٢,٤١٦
	المعادَلة الخطية	٣٢,٢١٦	١٠,٦٣٣	٠,١٠٠	٢,٢٩٣
الدرجة المعيارية	النموذج المرجعي	٦٤,٣٣٦	١٠,٦٣٣	٠,٢٨٠	٢,٤١٦
	النموذج الجديد بدون معادلة	٦٥,٥٣٤	١٠,٥٣٧	٠,١٠٠	٢,٢٩٣
الدرجات المعيارية المعادِلة	المعادَلة المئينية	٦٤,٣٣٥	١٠,٦٢٧	٠,٢٨٠	٢,٤١٦
	المعادَلة المئينية المهدة بعديا	٦٤,٣٤٠	١٠,٦٤٢	٠,٢٨٥	٢,٤١٦
	المعادَلة الخطية	٦٤,٣٣٦	١٠,٦٣٣	٠,١٠٠	٢,٢٩٣

جدول (5). الاحصاءات الوصفية لدرجات القدرة الكمية في النموذج المرجعي (1081) والدرجات المعادِلة للنموذج الجديد (1082).

الدرجة	الطريقة (النموذج)	المتوسط الحسابي	الانحراف المعياري	الالتواء	التفلطح
الدرجة الخام	النموذج المرجعي	٢٢,٢٤٧	٧,٨١٧	٠,٩٥٩	٤,٠١٦
	النموذج الجديد قبل المعادلة	١٩,٤٩٠	٧,٠٨٠	٠,٩٨٧	٤,٢٤٢
الدرجات الخام المعادِلة	المعادَلة المئينية	٢٢,٢٤٧	٧,٨١١	٠,٨٢٨	٤,٠٢٣
	المعادَلة المئينية المهدة بعديا	٢٢,٢٥٥	٧,٨٢٠	٠,٨٣٩	٣,٩٩٣
	المعادَلة الخطية	٢٢,٢٤٧	٧,٨١٧	٠,٩٨٧	٤,٢٤٢
الدرجة المعيارية	النموذج المرجعي	٦٤,٤٩٦	٩,٧٦٨	٠,٨٢٥	٤,٠٠١
	النموذج الجديد بدون معادلة	٦٤,٨٩٤	٨,٨٥٢	٠,٩٨٧	٤,٢٤٢
الدرجات المعيارية المعادِلة	المعادَلة المئينية	٦٤,٤٩٥	٩,٧٦٠	٠,٨٢٨	٤,٠٠٦
	المعادَلة المئينية المهدة بعديا	٦٤,٥٠٥	٩,٧٧٠	٠,٨٣٨	٣,٩٧٦
	المعادَلة الخطية	٦٤,٤٨٤	٩,٧٢٣	٠,٩٣٩	٤,٠٢٤

يظهر الجدول (6) المتوسطات غير الموزونة والموزونة للفروق بين الدرجات المعادِلة (الخام والمعيارية) للنموذج (1082) في النموذج المرجعي (1081) والدرجات المعادِلة (الخام والمعيارية) الناتجة من المعادَلة المساوية في اختباري القدرة اللفظية والقدرة الكمية. وتشير المتوسطات غير الموزونة والموزونة للفروق بين الدرجات المعادلة في اختبار القدرة اللفظية إلى أنها سالبة في جميع طرق المعادَلة، وتدل على أن مجمل الدرجات المعادِلة (الخام والمعيارية) للنموذج (1082) تقل عن مجمل الدرجات بالمعادلة المساوية، فمثلا تقل المتوسطات غير الموزونة لمجمل الدرجات الخام المعادِلة للنموذج (1082) عن مجمل درجاتها بالمعادلة المساوية بمقدار 2,518 في طريقة المعادَلة المئينية، وبمقدار 2,556 في طريقة المعادَلة المئينية المهدة بعديا وبمقدار 3,742 في طريقة المعادَلة الخطية. وترتبط

هـذه الفـروق السـالبة مع سـهولة مفـردات النموذج الجديـد مقارنـة مـع مفـردات النمـوذج المرجعـي. وتتقارب مقادير التحيز في الدرجات المعيارية المعادِلة مـع مقادير التحيز في الدرجات الخام المعادِلة للقدرة اللفظيـة في جميـع طـرق المعادَلـة. ويعود السـبب في ذلـك إلى أن الدرجات المعيارية هـي تحويلات خطية والانحرافـات المعياريـة للدرجات الخـام (ع = 10) مسـاوية مع الانحرافات المعيارية للدرجـات المعيارية (ع =10).



شكل (2). الفروق بين الدرجات المعادِلة (الخام والمعيارية) للنموذج (1082) مع (1081) ودرجاتها بالمعادلة المساوية في اختبار القدرة اللفظية

ويظهـر الشـكل (2) الفـروق التفصـيلية لجميع الدرجـات المعادِلـة (الخـام والمعياريـة) في اختبـار القـدرة اللفظيـة للنمـوذج (1082) مـع النمـوذج (1081). وتظهر هذه الفروق أنها تقل عن الدرجات الخام بالمعادلة المساوية. وتزيد هذه الفروق للدرجات المعادِلـة في المعادَلـة المئينيـة والمئينية المهـدة بعديا عنـد الدرجات الوسـطية التـي يتمركز حولهـا أغلب الطلبـة، بينما تقل هذ الفروق عند الدرجات الطرفية في تدريـج الدرجات. وتظهر الفروق للدرجات المعادِلة في المعادَلة الخطية الانتظام الخطي بحيث أن الفروق تقـل بالانتقال من الدرجات المنخفضـة إلى الدرجات المرتفعة مـع بقاء الاتجاه السـالب للفـروق. وتتقارب الدرجات المعادِلة في وسـط تدريج الدرجات (بين 22 و50) المتحصـل عليها مـن المعادَلـة الخطية مع تلك المتحصـل عليها من المعادَلة المئينيـة والمئينية المهدة بعديـا، بينمـا تختلف كثـيرا عند الانتقـال إلى طرفي تدريج الدرجات. فابتعـدت الدرجات المعادِلة (الخام والمعيارية) حسب المعادَلة الخطية كثيرا عن الدرجات بالمعادلة المسـاوية، حيث أنها تخطـت تدريج الاختبار مـن جهـة الدرجـات المنخفضـة (الدرجـات الخـام المعادِلة أقل من صفر).

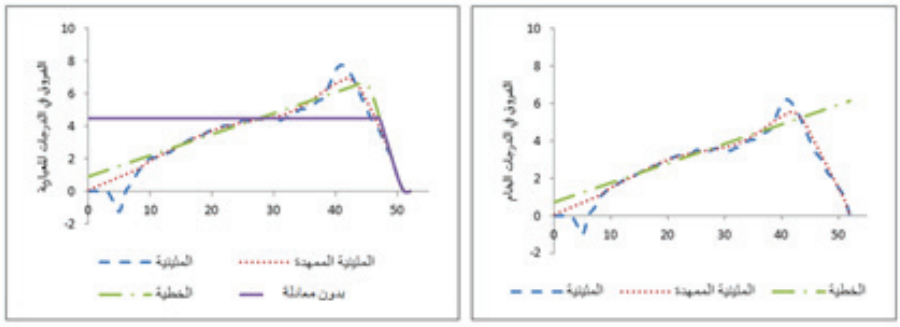
ويقـل المتوسـط غير المـوزون والمـوزون للدرجات المعيارية بدون معادلة لدرجات النموذج (1082) عن مثيلاتها في جميع الدرجات المعادِلة الناتجة من طرق المعادَلـة الثلاث. وتشـير هـذه النتيجـة إلى اختلاف نتائج إجـراء المعادَلة لنماذج اختبـار القدرة اللفظية مـن عدم إجرائها ، مما يشـير إلى تأكيد الحاجة إلى إجـراء معادلة لنماذج الاختبـار. كما تدل الفروق بين الدرجـات المعيارية بدون معادلـة والدرجات بالمعادلة المساوية على أنها تقل بشكل منتظم عبر جميع درجات الاختبار بمقدار 1,818. وهذه النتيجة متوقعة نظرا لاختلاف المتوسطات الحسابية للنموذجين (1082) و(1081) في اختبار القدرة اللفظية.

جدول (6)
المتوسطات غير الموزونة والموزونة لفروق الدرجات المعادِلة للنموذج (1082) و(1081) في اختبار القدرات العامة

الدرجة	الطريقة	القدرة اللفظية		القدرة الكمية	
		المتوسط غير الموزون للفروق	المتوسط الموزون للفروق	المتوسط غير الموزون للفروق	المتوسط الموزون للفروق
الدرجة الخام المعادِلة	المعادَلة المئينية	-٢,٥١٨	-٣,٧٢٦	٢,٧٠٤	٢,٧٥٦
	المعادَلة المئينية المهدة بعديا	-٢,٥٥٦	-٣,٧٢٦	٢,٨٣١	٢,٧٥٦
	المعادَلة الخطية	-٣,٧٤٢	-٣,٧٢٦	٣,٤٣٦	٢,٧٥٦
الدرجة المعيارية المعادِلة	المعادَلة المئينية	-٢,٥١٨	-٣,٧٢٦	٣,٣٢٥	٣,٤٤٥
	المعادَلة المئينية المهدة بعديا	-٢,٥٥٧	-٣,٧٢٦	٣,٤٨٢	٣,٤٤٥
	المعادَلة الخطية	-٣,٧٤٠	-٣,٧٢٦	٣,٦٤٠	٣,٤٤٥
الدرجة المعيارية بدون معادلة		-١,٨١٨	-١,٧٩٤	٤,١٨١	٤,٥٠٥

وبما أن النموذج الجديد (1082) في القدرة الكمية أكثر صعوبة من النموذج المرجعي (1081)، يظهر الجدول (6) أن المتوسطات غير الموزونة والموزونة للفروق في الدرجات المعادِلة (الخام والمعيارية) للنموذج (1082) موجبة في جميع طرق المعادَلة، وتشير إلى أن مجمل الدرجات المعادِلة للقدرة الكمية أكبر من الدرجات بالمعادلة المساوية. كما زادت المتوسطات غير الموزونة والموزونة للفروق في الدرجات المعيارية المعادِلة للنموذج عن مثيلاتها في الدرجات الخام المعادِلة حسب جميع طرق المعادَلة كنتيجة لزيادة الانحراف المعياري للقدرة الكمية في النموذج المرجعي (1081) عن الانحراف المعياري في النموذج الجديد (1082).

ويظهر الشكل (3) الفروق التفصيلية لجميع الدرجات المعادِلة (الخام والمعيارية) في اختبار القدرة الكمية للنموذج (1082) مع النموذج (1081). وأظهرت الفروق التفصيلية للدرجات المعيارية المعادِلة عبر درجات الاختبار أنها تتبع نمطا مشابها لنمط الفروق للدرجات الخام. بشكل كبير، تقاربت مقادير الفروق التفصيلية للدرجات المعادِلة الناتجة من طرق المعادَلة الثلاث عبر درجات الاختبار. ويلاحظ أن الفروق في طرق المعادَلة الثلاث زادت بالانتقال من الدرجات المنخفضة إلى الدرجات المرتفعة حتى وصلت عند حدود الدرجة 43، حينها بدأت الفروق للمعادَلة المئينية والمئينية المهدة بالتناقص التدريجي . بينما استمرت الفروق الناتجة من المعادَلة الخطية بالزيادة، وتخطت الدرجة الخام المعادِلة للدرجة القصوى (52) لتدريج درجات الاختبار، وكذلك الدرجات المعيارية المعادِلة، والتي تم تثبيتها عند الدرجة 100 في تلك الدرجات (والذي يظهر على شكل انكسار في الخط المستقيم في الدرجة المعيارية الناتجة من المعادلة الخطية والناتجة من دون معادلة). كما تدل الفروق بين الدرجات المعيارية بدون معادلة والدرجات بالمعادلة المساوية على أنها تقل بشكل منتظم عبر جميع درجات الاختبار بمقدار 4,181. وهذه النتيجة متوقعة نظرا لاختلاف المتوسطات الحسابية للنموذجين (1082) و(1081) في اختبار القدرة الكمية.



شكل (3). الدرجات المعادِلة (الخام والمعيارية) للنموذج (1082) مع النموذج (1081) في اختبار القدرة الكمية

معادَلة النموذجين (2034) و (1081): يظهر الجدول (7) المتوسطات غير الموزونة والموزونة للفروق بين الدرجات المعادِلة (الخام والمعيارية) للنموذج (2034) مع النموذج المرجعي (1081) والدرجات الناتجة من المعادَلة المساوية للقدرتين اللفظية والكمية. ففي اختبار القدرة اللفظية، تظهر نتائج معادَلة هذين النموذجين تقاربا مع نتائج معادَلة النموذجين (1082) و (1081). فتشير المتوسطات غير الموزونة والموزونة للفروق في جميع طرق المعادَلة إلى أنها سالبة، فمثلا تقل المتوسطات غير الموزونة لمجمل الدرجات الخام المعادِلة للنموذج 2034 عن مجمل درجاتها المعادِلة الخام بالمعادلة المساوية بمقدار 1,274 في طريقة المعادَلة المئينية، وبمقدار 2,122 في طريقة المعادَلة المئينية المهدة بعديا و وبمقدار 3,241 في طريقة المعادَلة الخطية.

جدول (7)

المتوسطات غير الموزونة والموزونة لفروق الدرجات المعادِلة للنموذجين (2034) و (1081) في اختبار القدرات العامة

الدرجة	الطريقة	القدرة اللفظية		القدرة الكمية	
		المتوسط غير الموزون للفروق	المتوسط الموزون للفروق	المتوسط غير الموزون للفروق	المتوسط الموزون للفروق
الدرجة الخام المعادِلة	المعادَلة المئينية	-١,٢٧٤	-٣,١٥٠	٥,١٩٦	٥,٤٠٦
	المعادَلة المئينية المهدة بعديا	-٢,١٢٢	-٣,١٦٤	٥,١٩٦	٥,٤٠٦
	المعادَلة الخطية	-٣,٢٤١	-٣,١٥١	٧,٩٧٦	٥,٤٠٦
الدرجة المعيارية المعادِلة	المعادَلة المئينية	-١,٢٧٥	-٣,١٧٤	٦,٥٠١	٦,٧٨٤
	المعادَلة المئينية المهدة بعديا	-٢,١٢١	-٣,١٧٤	٦,٤١٣	٦,٧٣١
	المعادَلة الخطية	-٣,٢٤٠	-٣,١٧٤	٧,٢١٠	٦,٦٧٨
الدرجة المعيارية بدون معادلة		-٢,١٦٦	-١,٩٣٢	٦,٢٠٩	٧,٦٣٢

وتظهر الفروق التفصيلية عبر جميع الدرجات الخام المعادِلة للقدرة اللفظية للنموذج (2034) أن اتجاه ومقدار الفروق للدرجات الخام المعادِلة بطرق المعادَلة الثلاث يختلف باختلاف موقع الدرجات الخام (لم يتم وضع شكل الفروق بسبب حجم الدراسة). فيظهر أن الفروق للدرجات المعادِلة الناتجة من المعادَلة المئينية المهدة بعديا كانت سالبة عبر جميع الدرجات الخام، وكانت كبيرة عند الدرجات المتوسطة، بينما قلت عند الدرجات الطرفية حتى اقتربت من الصفر عند نهايتي تدريج الدرجات. ويتكرر هذا النمط مع المعادَلة المئينية ما عدا أن الفروق كانت موجبة عند الدرجات الطرفية، وخاصة المنخفضة منها (أي أن الدرجات المعادِلة بالمعادَلة المئينية أكبر من الدرجات الخام بالمعادلة المساوية). بينما تظهر الفروق الناتجة من المعادَلة الخطية انتظاما خطيا مع بقاء الاتجاه السالب للفروق، بحيث أن الفروق كانت أكبر ما يمكن عند الدرجات المنخفضة، وقلت بشكل منتظم عند الانتقال إلى الدرجات الأعلى.

وتظهر الفروق التفصيلية للدرجات المعيارية المعادِلة للقدرة اللفظية في النموذجين (2034) و(1081) نمطا متطابقا مع تلك المشاهدة في الدرجات الخام المعادِلة. ويعود السبب في ذلك إلى أن الدرجات المعيارية هي تحويلات خطية مع تساوي الانحرافات المعيارية للدرجات الخام مع الدرجات المعيارية. وتظهر المعادَلة الخطية تقاربا للفروق بين الدرجات المعيارية المعادِلة مع تلك الفروق المتحصل عليها من المعادَلة المئينية، مع عدم تخطي الدرجات المعيارية المعادِلة لها عن حدود التدريج من الجهتين. بينما تظهر الدرجات المعيارية بدون معادَلة نمطا مختلفا، حيث كانت الفروق لها مرتفعة في الدرجات المنخفضة للاختبار وقلت هذه الفروق بالانتقال الى الدرجات المرتفعة؛ نظرا لاختلاف الانحراف المعياري للنموذجين (2034) و(1081) واختلاف متوسطاتها الحسابية.

وبما أن النموذج الجديد (2034) في اختبار القدرة الكمية أكثر صعوبة من النموذج المرجعي (1081)، يظهر الجدول (7) أن المتوسطات غير الموزونة والموزونة للفروق في الدرجات المعادِلة (الخام والمعيارية) للنموذج (1082) كانت موجبة في جميع طرق المعادَلة، وتشير إلى أن مجمل الدرجات المعادِلة للقدرة الكمية أكبر من الدرجات بالمعادلة المساوية. كما زادت المتوسطات غير الموزونة والموزونة للفروق في الدرجات المعيارية المعادِلة للنموذج عن مثيلاتها في الدرجات الخام المعادِلة حسب جميع طرق المعادَلة كنتيجة لزيادة الانحراف المعياري للقدرة الكمية في النموذج المرجعي (1081) عن الانحراف المعياري في النموذج الجديد (2034). بينما تشير الفروق التفصيلية للدرجات المعيارية المعادِلة عبر درجات الاختبار أنها تتبع نمطا مشابها لنمط الفروق للدرجات الخام (لم يتم وضع شكل الفروق بسبب حجم الدراسة). كما تقاربت مقادير الفروق التفصيلية للدرجات المعادِلة الناتجة من طرق المعادَلة الثلاث عبر درجات الاختبار بشكل كبير. ويلاحظ أن الفروق في طرق المعادَلة الثلاث زادت بالانتقال من الدرجات المنخفضة إلى الدرجات المرتفعة حتى وصلت عند حدود الدرجة 33، حينها بدأت الفروق للمعادَلة المئينية والمئينية المهدة بالتناقص تدريجيا. بينما استمرت الفروق الناتجة من المعادَلة الخطية بالزيادة وتخطت الدرجات الخام المعادِلة للدرجة القصوى لتدريج الدرجات (52)، وكذلك الحال مع الدرجات المعيارية المعادِلة، والتي تم تثبيتها عند الدرجة 100 في تلك الدرجات.

معادَلة النموذجين (2034) و (1082): يظهر الجدول (8) المتوسطات غير الموزونة والموزونة للفروق بين الدرجات المعادِلة (الخام والمعيارية) في اختباري القدرة اللفظية والقدرة الكمية للنموذج (2034) مع النموذج المرجعي (1082)، والدرجات الناتجة من المعادلة المساوية (الخام والمعيارية) للنموذج (2034). وتشير المتوسطات غير الموزونة والموزونة

للفروق في اختبار القدرة اللفظية بجميع طرق المعادَلة أنهـا موجبـة، والتي تـدل علـى أن مجمـل الدرجات الخـام المعادِلة للنمـوذج (1082) بمجملها تزيد عن مجمل الدرجات الناتجة من المعادلة المسـاوية، وذلك نتيجة لزيادة صـعوبة النموذج الجديد (2034) عن النموذج المرجعي (1082). فمثلا زادت المتوسـطات غير الموزونة لمجمل الدرجـات الخام المعادِلة للنموذج (2034) عن مجمل درجاتها الخام بالمعادلة المساوية بمقـدار 1,866 في طريقة المعادَلـة المئينية، وبمقدار 0,708 في طريقـة المعادَلـة المئينيـة المهـدة بعديا وبمقـدار 0,778 في طريقـة المعادَلـة الخطية. وعند وزن هذه الفروق بتكـرارات الدرجات الخام، نجد أن المتوسـطات الموزونة لهذه الفروق تنقص بشـكل كبير عن المتوسـطات غير الموزونـة ، ويظهر هذا النمط في طـرق المعادَلة الثلاث، مع اسـتمرار اتجـاه الفرق في ارتفـاع الدرجات الخام المعادِلة عـن الدرجات الخام الناتجة من المعادلة المساوية المقابلة لها.

جدول (8)

المتوسطات غير الموزونة والموزونة لفروق الدرجات المعادِلة للنموذجين (2034) و (1082) في اختبار القدرات العامة

الدرجة	الطريقة	القدرة اللفظية		القدرة الكمية	
		المتوسط غير الموزون للفروق	المتوسط الموزون للفروق	المتوسط غير الموزون للفروق	المتوسط الموزون للفروق
الدرجة الخام المعادِلة	المعادَلة المئينية	١,٨٦٦	٠,٨٩٧	٣,٤٢٧	٣,١٨٠
	المعادَلة المئينية المهدة بعديا	٠,٧٠٨	٠,٨٢٨	٣,٣٢٩	٣,١٨٠
	المعادَلة الخطية	٠,٧٧٨	٠,٨٢٨	٤,٩٥٠	٣,١٨٠
الدرجة المعيارية المعادِلة	المعادَلة المئينية	١,٨٦٧	٠,٨٩٧	٤,٠٠١	٣,٩٧٥
	المعادَلة المئينية المهدة بعديا	٠,٧٠٩	٠,٨٩٧	٣,٩٠٠	٣,٩٧٥
	المعادَلة الخطية	٠,٧٦٤	٠,٨٩٧	٤,١٣٥	٣,٩٧٥
الدرجة المعيارية بدون معادلة		-٠,٣٥٧	- ٠,٠٦٩	٢,٠٢٨	٣,١٢٧

وتُظهر الفروق التفصيلية لجميع الدرجات الخام المعادِلـة للنموذج (2034) في اختبار القدرة اللفظية أن الفـروق للدرجـات الخـام المعادِلـة كانـت موجبة عـبر جميع درجـات الاختبار حسـب المعادَلـة المئينية والمئيينيـة المهـدة بعديـا، وازدادت بالانتقـال الى الدرجات المرتفعة (لم يتم وضـع شـكل الفروق بسبب حجـم الدراسـة). وأظهـرت المعادَلة المئينيـة فروقا شاذة عند الدرجات المنخفضة، حيث كانت الدرجات المعادِلـة أكبر بكثـير من الدرجـات الخام الأصـلية، وقـد يعود ذلـك إلى قلة عدد الطلاب الذين حصـلوا علـى الدرجات المنخفضـة. وعالجـت المعادَلة المئينية المهدة بعديـا هذه الفروق الشـاذة للمعادَلة المئينية. وتقاربـت الفروق في المعادَلة الخطيـة مع تلك الفروق في المعادَلة المئينية عبر معظم درجات الاختبار ما عدا الدرجات الطرفية، وخاصة عند الدرجات المنخفضة، حيث كانت الفروق سـالبة في المعادَلة الخطية. وتظهر الدرجـات المعياريـة المعادِلـة للنمـوذج (2034) مع النمـوذج (1082) فروقـا لهـا نمـط مشـابه للنمط المشـاهد في الدرجـات الخـام المعادِلة نظـرا لتقارب قيـم الانحـراف المعياري للدرجات الخـام والدرجات المعياريـة للقـدرة اللفظيـة في النمـوذج (2034). بينما تختلف الفـروق للدرجات المعيارية بدون معادلة عن تلك الفروق المشـاهدة من طـرق المعادلة الثلاث. فكانت الفروق للدرجات المعيارية بدون معادلة سـالبة في الدرجات التي تقل عن المتوسـط الحسابي للنموذج (2034)، بينمـا كانت موجبة في الدرجات التي تزيد

عن المتوسط الحسابي. كما تخطت الدرجات المعيارية بدون معادلة الدرجة المعيارية القصوى (100) والتي تم تثبيتها.

وفي اختبار القدرة الكمية، أظهرت الفروق غير الموزونة والموزونة للدرجات المعادِلة (الخام والمعيارية) للنموذج (2034) مع النموذج (1082) أنها موجبة في جميع طرق المعادَلة، نتيجة لزيادة صعوبة النموذج الجديد عن النموذج المرجعي. وتزيد الفروق في الدرجات المعيارية المعادِلة عن تلك الفروق في الدرجات الخام المعادِلة في جميع طرق المعادَلة. وتُظهر الفروق التفصيلية في المعادَلة المئينية والمئينية المهدة بعديا نمطا متشابها عبر درجات الاختبار، حيث تزيد الفروق عند الدرجات الوسطية؛ بينما تقل عند الدرجات الطرفية في الاتجاهين (لم يتم وضع شكل الفروق بسبب حجم الدراسة). بينما تظهر المعادَلة الخطية زيادة منتظمة في الفروق بالانتقال من الدرجات المنخفضة للاختبار إلى الدرجات المرتفعة. وعند الدرجات المرتفعة، تُظهر الفروق للمعادَلة الخطية أن الدرجات الخام المعادِلة والمعيارية تخطت الحدود القصوى لتدرج الدرجات عند الدرجات أكبر من 39، وتم تثبيت الدرجات المعيارية المعادِلة عند الدرجة 100. وتظهر الدرجات المعيارية بدون معادلة نمطا معاكسا لطرق المعادَلة الثلاث، حيث كانت الفروق مرتفعة في الدرجات المنخفضة للاختبار ثم قلت هذه الفروق بالانتقال إلى الدرجات المرتفعة، كما تخطت الدرجات المتحصل عليها الحدود القصوى لتدرج الدرجات المعيارية.

معادَلة النموذجين (1083) و (1082): يظهر الجدول (9) المتوسطات غير الموزونة والموزونة للفروق بين الدرجات المعادِلة (الخام والمعيارية) للنموذج (1083) مع النموذج المرجعي (1082) في اختباري القدرة اللفظية والقدرة الكمية والدرجات الناتجة من المعادلة المساوية (الخام والمعيارية) للنموذج (1083). وتُشير المتوسطات غير الموزونة والموزونة للفروق في جميع طرق المعادَلة إلى أنها جاءت موجبة، ويدل ذلك على أن الدرجات الخام المعادِلة للنموذج (1083) بمجملها تزيد عن مجمل الدرجات الناتجة من المعادلة المساوية، فمثلا تزيد المتوسطات غير الموزونة لمجمل الدرجات الخام المعادِلة للنموذج (1083) عن مجمل درجاتها الخام بالمعادلة المساوية بمقدار 1,816 في طريقة المعادَلة المئينية، وبمقدار 1,202 في طريقة المعادَلة المئينية المهدة بعديا وبمقدار 1,866 في طريقة المعادَلة الخطية. وعند وزن هذه الفروق بتكرارات الدرجات الخام، نجد أن المتوسطات الموزونة للفروق تنقص عن المتوسطات غير الموزونة للفروق، ويتكرر هذا النمط في طرق المعادَلة الثلاث، مع استمرار اتجاه الفرق في زيادة الدرجات الخام المعادِلة عن الدرجات الخام الناتجة من المعادلة المساوية المقابلة لها (0,028).

وعند تفحص الفروق التفصيلية لجميع الدرجات الخام المعادِلة للنموذج (1083)، يلاحظ أن جميعها جاءت موجبة، حيث تزيد الدرجات الخام المعادِلة عن الدرجات الخام الناتجة من المعادلة المساوية (لم يتم وضع شكل الفروق بسبب حجم الدراسة). وتزيد هذه الفروق للدرجات المعادِلة الناتجة من المعادَلة المئينية والمئينية المهدة بعديا عند الدرجات الوسطية التي يتمركز حولها أغلب الطلبة، بينما تقل هذ الفروق عند الدرجات الطرفية في تدريج الدرجات. كما تزيد الدرجات الخام المعادِلة الناتجة من المعادَلة المئينية عن الدرجات الناتجة من المعادلة المساوية بمقادير كبيرة عند الدرجات المنخفضة. بينما تُظهر المعادَلة الخطية انتظاما خطيا للفروق مع بقاء الاتجاه الموجب لهذه الفروق (أي أن الدرجات المعادِلة أكبر من الدرجات الخام الناتجة من المعادلة المساوية). وكانت الفروق أصغر ما يمكن عند الدرجات المنخفضة وازدادت بشكل منتظم عند الانتقال إلى الدرجات الأعلى. وتتخطى المعادَلة الخطية أيضا الحد الأقصى لدرجات الاختبار (الدرجة 68). وتُظهر الفروق للدرجات المعيارية المعادِلة للنموذج (1083) مع (1082) نمطا متشابها تماما لتلك

المشـاهدة في الدرجات الخام المعادِلة. وتُظهر النتائج تشـابه نمط الفـروق للدرجـات المعياريـة المعادِلة مع تلك المشـاهدة مـع الدرجات الخـام المعادِلة في طرق المعادَلـة الثـلاث. وتُظهر الفـروق للدرجـات المعادِلة (الخام والمعيارية) لاختبار القدرة الكمية للنموذجين (1083) مـع (1082) نفـس النمـط المشـاهَد مـع النموذجـين (2034) مـع (1082) لجميـع طـرق المعادَلة والدرجات المعيارية بدون معادلة.

جدول (9)

المتوسطات غير الموزونة والموزونة لفروق الدرجات المعادِلة للنموذجين (1083) و (1082) في اختبار القدرات العامة

الدرجة	الطريقة	القدرة اللفظية		القدرة الكمية	
		المتوسط غير الموزون للفروق	المتوسط الموزون للفروق	المتوسط غير الموزون للفروق	المتوسط الموزون للفروق
الدرجة الخام المعادِلة	المعادَلة المئينية	١,٨١٦	١,٩٣٢	٣,٦٣١	٣,٧٦٣
	المعادَلة المئينية المهدة بعديا	١,٢٠٢	١,٩٣٢	٣,٦٩٨	٣,٧٦٣
	المعادَلة الخطية	١,٨٦٦	١,٩٣٢	٦,٠٨٨	٣,٧٦٣
الدرجة المعيارية المعادِلة	المعادَلة المئينية	١,٨١٥	١,٩٣٢	٤,٤١٥	٤,٧١٧
	المعادَلة المئينية المهدة بعديا	١,٢٠٢	١,٩٣٢	٤,٤٠٣	٤,٧١٧
	المعادَلة الخطية	١,٨٤١	١,٩٣٢	٤,٧٦٣	٤,٧١٧
الدرجة المعيارية بدون معادلة		-٠,١١٠	-٠,١٣٨	٢,٠٢٨	٣,١٢٧

تأثير الفروق بين نماذج الاختبار على تحيز المعادَلة: عنـد مقارنـة متوسـطات الفـروق بـين الدرجـات المعادِلـة (الخـام والمعياريـة) والدرجـات الناتجة من المعادِلة المساوية في نتائج معادلة النماذج الأربعـة حسـب الجـداول (6)، (7)، (8)، (9)، وربطها بمقدار اختلاف مسـتوى صـعوبة النموذجين المسـتخدمين في كل معادَلـة كمـا في الجدولين (2) و (3)، يظهر أن متوسـطات هذه الفروق في جميع طرق المعادَلة، وخاصة المهدة منها، تزيد بزيادة الاختلاف في مسـتوى صـعوبة النموذجـين. فـكان ترتيب مطلق المتوسـطات غير الموزونـة للفروق في القـدرة اللفظية تشـير إلى الترتيـب التنـازلي التـالي: النموذجـان (1082) و(1081)، ثـم النموذجـان (2034) و(1081)، ثـم النموذجان (1083) و(1082)، ثم النموذجـان (2034) و(1082)، وهـذا يتوافـق مع الترتيب التنازلي لمطلق الفروق في مسـتوى صعوبة كل نموذجـين (انظر الجدول 2). وكان الترتيب التنازلي لمطلق متوسـطات الفروق في القـدرة الكمية كالتالي: النموذجـان (2034) و(1081)، ثـم النموذجـان (1083) و(1082)، ثـم النموذجـان (2034) و(1082)، ثـم النموذجـان (1082) و(1081)، والذي يتوافق أيضا مع الترتيب التنازلي لمطلق الفروق في مستوى صعوبة كل نموذجين (انظر الجدول 3).

الأخطاء المعيارية لطرق المعادَلة:

يعرضـ الجـدول (10) والجـدول (11) متوسـطات الأخطـاء المعياريـة للدرجـات المعادِلـة (الخام والمعيارية) لنماذج اختبـاري القدرة اللفظية والقـدرة الكميـة علـى التـوالي. وتظهـر النتائـج أن متوسـط الأخطاء المعيارية للدرجـات المعادِلة (الخام والمعياريـة) لجميـع طـرق المعادَلة كانـت متقاربة في نتائج معادلة النماذج الأربعة لكل من اختباري القدرة اللفظيـة والقدرة الكميـة((1082) مـع (1081)، (2034) مـع (1081)، (2034) مـع (1082)، (1083) مـع (1082)). وتقل متوسـطات الأخطاء المعيارية في جميع طرق المعادَلة للقدرة الكمية عن تلك

المشاهدة مع القدرة اللفظية. كما أن متوسط الأخطاء المعيارية للدرجات المعادِلة (الخام والمعيارية) الناتجة من المعادلة المئينية بنوعيها أقل من متوسط الأخطاء المعيارية للمعادَلة بالطريقة الخطية في معادلة النماذج الأربعة .

جدول (10)

المتوسط الموزون للأخطاء المعيارية للدرجات المعادِلة (الخام والمعيارية) لنماذج اختبار القدرة اللفظية.

الدرجة	الطريقة	نماذج الاختبار			
		(١٠٨٢) مع (١٠٨١)	(٢٠٣٤) مع (١٠٨١)	(٢٠٣٤) مع (١٠٨٢)	(١٠٨٣) مع (١٠٨٢)
الدرجة الخام المعادِلة	المعادَلة المئينية	٠,٤٥٨	٠,٤٢٦	٠,٤٤٦	٠,٤٩٠
	المعادَلة المئينية المهدة بعديا	٠,٤٥٨	٠,٤٢٦	٠,٤٤٦	٠,٤٩٠
	المعادَلة الخطية	٠,٥٤٧	٠,٥٢١	٠,٥٣٥	٠,٥٧١
الدرجة المعيارية المعادِلة	المعادَلة المئينية	٠,٣٥٩	٠,٤٠٢	٠,٤٢٢	٠,٤٦١
	المعادَلة المئينية المهدة بعديا	٠,٣٥٩	٠,٤٠٢	٠,٤٢٢	٠,٤٦١
	المعادَلة الخطية	٠,٤٩٥	٠,٥٠٥	٠,٥١٤	٠,٥٤٤

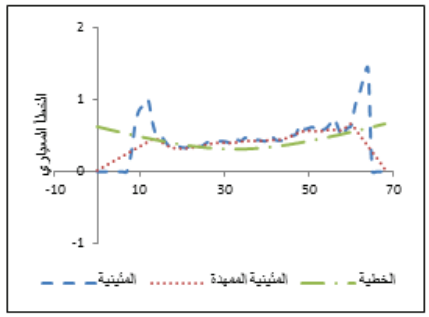
جدول (11)

المتوسط الموزون للأخطاء المعيارية للدرجات المعادِلة (الخام والمعيارية) لنماذج اختبار القدرة الكمية.

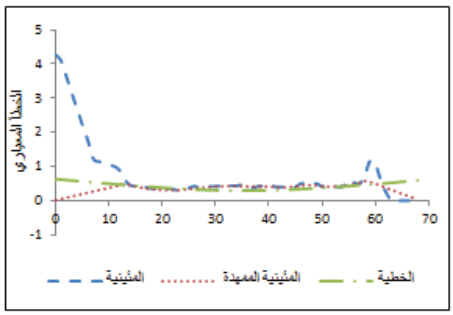
الدرجة	الطريقة	نماذج الاختبار			
		(١٠٨٢) مع (١٠٨١)	(٢٠٣٤) مع (١٠٨١)	(٢٠٣٤) مع (١٠٨٢)	(١٠٨٣) مع (١٠٨٢)
الدرجة الخام المعادِلة	المعادَلة المئينية	٠,٣٤٦	٠,٣٢٤	٠,٣١٥	٠,٣٤١
	المعادَلة المئينية المهدة بعديا	٠,٤٣٢	٠,٤٠٥	٠,٣٨٩	٠,٤٢٤
	المعادَلة الخطية	٠,٥٢٤	٠,٤٨٧	٠,٤٨٩	٠,٥٠٥
الدرجة المعيارية المعادِلة	المعادَلة المئينية	٠,٣٢٤	٠,٣٠٧	٠,٣٠٠	٠,٣٢٢
	المعادَلة المئينية المهدة بعديا	٠,٤٠٩	٠,٣٨٤	٠,٣٧٢	٠,٤٠٠
	المعادَلة الخطية	٠,٥٠٤	٠,٤٧٤	٠,٤٧١	٠,٤٨٧

وعند استعراض الأخطاء المعيارية للدرجات الخام المعادِلة لكل نموذجين تم معادلتهما في اختبار القدرة اللفظية عبر درجات الاختبار (الشكل (4أ) يظهر الأخطاء المعيارية لمعادلة النموذجين (1082) و (1081)، ولم توضع البقية بسبب حجم الدراسة)، تُشير النتائج إلى تقارب مقادير الأخطاء المعيارية للمعادَلة بين طرق المعادَلة الثلاث. ويلاحظ أن الأخطاء المعيارية للطريقة المئينية مرتفعة عند الدرجات الطرفية (المرتفعة والمنخفضة) بسب العدد القليل نسبيا للطلاب عند الدرجات الطرفية، بينما ضبطت عملية التمهيد التي أجريت على الدرجات المعادِلة بالطريقة المئينية الأخطاء المعيارية للمعادَلة، حيث كانت صغيرة في الطريقة المئينية المهدة بعديا عند الدرجات الطرفية. كما تزيد الأخطاء المعيارية للمعادَلة الخطية في الدرجات الطرفية (المرتفعة والمنخفضة) للقدرة اللفظية والقدرة الكمية. وتتكرر نفس النتائج في الأخطاء المعيارية في النماذج الأربعة لاختبار القدرة الكمية (الشكل (4ب) يظهر الأخطاء المعيارية لمعادلة النموذجين (1082) و (1081)، ولم توضع البقية بسبب حجم الدراسة).

(1) النموذجان (1082) و (1081)



(ب) النموذجان (2034) و (1081)



شكل (4). الأخطاء المعيارية للدرجات الخام المعادِلة لنماذج اختبار القدرة اللفظية

تأثير الفروق بين نماذج الاختبار على الأخطاء المعيارية لطرق المعادَلة:

عند مقارنة متوسطات الأخطاء المعيارية للدرجات المعادِلة (الخام والمعيارية) للقدرتين اللفظية والكمية في نتائج النماذج الأربعة في الجدولين (10) و (11)، وربطها بمقدار اختلاف مستوى صعوبة النموذجين المستخدمين في كل معادَلة كما في الجدولين (2) و (3)، تشير النتائج إلى تقارب قيم الأخطاء المعيارية في النماذج الأربعة برغم اختلاف مطلق الفرق بين صعوبة كل نموذجين. ويشير ذلك إلى عدم تأثر الأخطاء المعيارية بمقدار الاختلاف في صعوبة النموذجين المراد معادلتهما.

المناقشة والتوصيات

هدفت الدراسة الحالية إلى مقارنة نتائج معادَلة نماذج اختبار القدرة اللفظية واختبار القدرة الكمية باستخدام ثلاث طرق إحصائية للمعادَلة في تصميم الجماعات المتكافئة. طبقت الدراسة من خلال إجراء عملية معادلة لدرجات أربعة نماذج للاختبارين. وأظهرت الدراسة عدم تكافؤ النماذج الأربعة في كل من الاختبارين، حيث تختلف فيما بينها في متوسطاتها الحسابية، وانحرافها المعياري وشكل توزيعها التكراري. وكان عدم التكافؤ بين نماذج الاختبارات في اختبار القدرة الكمية أكثر وضوحا من اختبار القدرة اللفظية. وتتفق هذه النتيجة مع ما ذهب إليه العديد من الباحثين إلى أن بناء نماذج متكافئة في جدول مواصفاتها شرط أساسي لتكافؤ نماذج الاختبار؛ ولكنه لا يؤدي بالضرورة إلى تكافؤ هذه النماذج وتشابه إحصاءاتها الوصفية (Angoff، 1987; Kolen & Brennan، 2004).

وأشارت النتائج إلى أن المتوسطات غير الموزونة والموزونة للفروق في جميع طرق المعادَلة مرتفعة نسبيا مقارنة بالدرجات الناتجة من المعادلة المساوية لكل نموذج، مما يشير إلى وجود اختلاف بين تدريج الاختبار في حالة إجراء المعادَلة وبين عدم إجراء المعادَلة. وتشير نتائج معادَلة الدرجات الخام والدرجات المعيارية بالطرق الإحصائية الثلاث إلى أن جميع طرق المعادَلة للنماذج الأربعة في اختبار القدرة اللفظية تعمل فرقا في تدريج الدرجات سواء تدريج الدرجات الخام أو تدريج الدرجات المعيارية، وأن استخدام الدرجات المعادِلة للنموذج الجديد تعطي دلالات مختلفة عن مستوى الطلبة مقارنة بالدرجات الناتجة من المعادلة المساوية. وكان اتجاه الفروق سالبا عندما كانت صعوبة النموذج الجديد أقل من صعوبة النموذج المرجعي، وتشير إلى أن درجات النموذج الجديد تحولت إلى درجات معادِلة أصغر منها؛ بينما كان اتجاه الفروق موجبا عندما كانت صعوبة النموذج الجديد أكبر من صعوبة النموذج المرجعي، وتشير

إلى أن درجات النموذج الجديد تحولت إلى درجات معادِلة أكبر منها.

كما أشارت النتائج إلى ارتباط مقدار الفروق في الدرجات المعادِلة بين النموذجين في جميع طرق المعادَلة مع مقدار الاختلاف في مستوى صعوبة النموذجين. فكلما اختلفت صعوبة النموذجين كلما كانت المتوسطات غير الموزونة والموزونة للفروق أكبر. وتتسق هذه النتيجة مع المبدأ التي تقوم عليه معادَلة نماذج الاختبار، من حيث أنها تضبط الاختلاف بين مستوى صعوبة نماذج الاختبار، فكلما اختلف النموذجان في مستوى صعوبتهما، كلما احتاج النموذج الجديد إلى تعديل تدريج درجاته ليتكافأ مع النموذج المرجعي، وبالتالي اختلفت الدرجات المعادِلة لها عن الدرجات الناتجة من المعادلة المساوية بمقدار أكبر. وظهرت هذه العلاقة في جميع طرق المعادَلة وخاصة المهدة منها (المعادَلة الخطية والمئينية المهدة بعديا).

تتقارب الدرجات المعادِلة (الخام والمعيارية) الناتجة من المعادَلة الخطية مع تلك الناتجة من المعادَلة المئينية عند الدرجات المحيطة بالمتوسط الحسابي لدرجات الطلاب، بينما تختلف عند الدرجات الطرفية (الدرجات المرتفعة والمنخفضة). وتتفق هذه النتيجة مع ما ذهبت إليه العديد من الدراسات (Kolen & Brennan، 2004؛ Wang، et. al.، 2008) إلى أن المعادلات المئينية يصلح استخدامها أكثر عند اتخاذ قرارات حول أداء الطلبة باستخدام جميع درجات الاختبار؛ بينما يصلح استخدام الطريقة الخطية عند توظيف درجة المتوسط الحسابي وما حولها فقط في اتخاذ قرارات حول أداء الطلبة.

كما أظهرت الطريقة المئينية (غير المهدة) تباينا غير خطي ومتذبذب في الفروق عبر درجات الاختبار، وأظهرت فروقا كبيرة في الدرجات المنخفضة للاختبار، حيث لا يوجد عدد كبير من الطلبة. بينما أظهرت الطريقة المئينية المهدة بعديا تباينا غير متذبذب وممهدا للمعادَلة، مما يجعلها أكثر تفضيلا عن الطريقة المئينية الاعتيادية، من حيث أنها تعالج التأثير الناتج من عدم وجود عينة كافية في جزء من تدريج الدرجات.

وأظهرت النتائج اختلاف طرق المعادَلة في متوسط الأخطاء المعيارية للمعادَلة، وحصلت المعادَلة المئينية على متوسطات أخطاء معيارية أصغر نسبيا من المعادَلة الخطية. كما أظهرت المعادَلة المئينية المهدة بعديا أخطاء معيارية أصغر نسبيا من المعادَلة المئينية والمعادَلة الخطية عند الدرجات الطرفية (المرتفعة والمنخفضة)، مما يتماشى مع أفضلية الطريقة المئينية المهدة بعديا في أنها تعطي نتائج أكثر دقة وأقل أخطاء معيارية عند جميع درجات الاختبار وليس حول المتوسط الحسابي فقط. وأظهرت النتائج عدم تأثر الأخطاء المعيارية لطرق المعادَلة بمقدار الفروق في صعوبة النموذجين المراد معادلتهما.

وأظهرت النتائج عدم كفاية الدرجات المعيارية المستخدمة حاليا (الدرجات المعيارية بدون معادَلة) لتدريج نماذج اختبار القدرات العامة في ضمان تكافؤ درجات نماذج الاختبار، حيث لم تختلف كثيرا عن الدرجات المعيارية المعادِلة الناتجة من المعادَلة المساوية، حيث كان نمط تلك الدرجات المعيارية منتظما في أغلب المعادلات لنماذج الاختبار، وتختلف عن الدرجات المعيارية بالمعادلة المساوية بمقدار ثابت.

ويمكن تلخيص نتائج المعادَلة للنماذج الأربعة لاختبار القدرة اللفظية واختبار القدرة الكمية أن عملية المعادَلة أدت إلى نتائج مختلفة عن الدرجات الخام والمعيارية الناتجة من المعادلة المساوية للنماذج. كما يتضح أن المعادَلة بالطريقة المئينية والمئينية المهدة بعديا حققت مؤشرات إحصائية أفضل عن بقية طرق المعادَلة، فالأخطاء المعيارية للمعادَلة للطريقة المئينية المهدة بعديا تساوي الأخطاء المعيارية للمعادَلة للطريقة المئينية، وكلاهما

أصـغر من تلك المشاهدة مع الطريقة الخطية. ويمكن أن يرجـع ذلـك إلى أن المعادَلة المئينيـة تراعي جميع خصـائص التوزيـع التكـراري للنمـاذج أثنـاء عمليـة المعادَلة، بالإضـافة إلى أن المعادَلة المهدة بعديا ليس بها قفزات كما هو الحال في الطريقة المئينية.

وبالنظر في اسـتخدامات اختبار القدرات العامة كاختبار قبول بجميع مؤسسات التعليم العالي بالمملكة العربيـة السـعودية، واسـتخدامه لانتقاء المؤسسـات للطـلاب ذوي الدرجـات المرتفعـة في الاختبـار ممن رغبـوا في القبـول بهـذه المؤسسـات، فيمكـن اعتبار الدرجـات المرتفعة في اختباري القـدرات العامة على أنها أكثر حرجا في القرارات المبنية على الاختبار لكل من الطالب ومؤسسات التعاليم العالي. وبالتالي تظهر الحاجة ماسـة أكـثر إلى اختيار طريقـة المعادلة التي تعطي درجـات معادلة أكثر دقة وأقل أخطاء معيارية. ومن خـلال نتائج الدراسـة، تتوفر هذه الخصـائص المطلوبـة في المعادلة المئينية والمئينيـة المهدة لإجراء معادلة نمـاذج الاختبارات، حيث أنهـا تعطي درجات معادِلـة داخـل إطار تدريج الاختبـار، ولا تتطلب عمل تثبيـت للدرجـة (Truncation)، مع حصـولها على أخطـاء معيارية صـغيرة للمعادلة عند طـرفي تدريج الاختبار. ولذا توصـي الدراسـة باسـتخدام الطريقة المئينيـة أو الطريقـة المئينيـة المهـدة بعديـا لمعادلة نماذج اختبارات القدرات العامة باسـتخدام تصميم الجماعات المتكافئة.

وتقـترح الدراسـة إجـراء المزيد من الدراسـات على نمـاذج أخـرى لاختبار القـدرات العامـة للتأكد من فعالية الطريقـة المئينية والمهدة بعديا في معادَلة نماذج الاختبارات. كما يمكن إجراء دراسات تجريبية للتحقق من صـدق الدرجات المعادِلة المتحصـل عليها مـن خلال القدرة التنبؤية لهـا مع المعدل الجامعي في بعض مؤسسـات التعليم العالي. كما تقترح الدراسـة إجراء دراسات مستقبلية لمقارنة معادَلة نماذج اختبار القـدرات العامـة باسـتخدام الطرق الإحصـائية في النظرية الكلاسيكية مع نظرية الاستجابة للمفردة.

المراجع

- الدوسـري، راشـد حمـاد (2001). معادَلـة الاختبـارات: مفهومهـا، وطرقها، ومشـكلات تطبيقهـا. مجلـة العلـوم التربوية والنفسـية، جامعة البحرين، 2، 105-141.
- الزيـات؛ فتحي؛ المحـرزي، راشـد (2011). مقياس الخليـج للقـدرات العقليـة المتعـدد (جماس): التقريـر الفنـي. مكتـب التربيـة العربي لدول الخليج، الرياض.
- الشـريفين، نضـال كمـال (2009). معادَلـة درجـات نمـاذج مختلفة مـن اختبـار الكفاءة اللغويـة في اللغـة الانجليزيـة لـدى طلبـة جامعـة اليرمـوك. مجلـة جامعـة أم القـرى للعلـوم التربويـة والنفسـية، المملكـة العربيـة السعودية،1، 11-62.

- Angoff, W. H. (1987). Technical and practical issues in equating: A discussion of four papers. Applied Psychological Measurement, 11, 291-300.
- Baker, F. B. & Al-Karni, A. (1991). A comparison of two procedures for computing IRT equating coefficients. Journal of Educational Measurement, 28, 147-162.
- Chulu, B. W. & Sireci, S. G. (2011). Importance of equating high stakes educational measurements. International Journal of Testing, 11, 38-52.
- Glowacki, M. L. (1991). An analysis of test equating models for the Alabama high school graduation examination. Paper presented at the annual meeting of Mid-South Educational Research Association, Lexington. (ED 340720).
- Hanson, B. (2004). Manual of Equating error: A program for computing equating error using the bootstrap (GUI Version 2.0). Iowa City, IA: Center for Advanced Studies in Measurement and Assessment, The University

response theory to practical testing problems. New Jersey: Lawrence Erlbaum.

- Lyrén, P. & Hambleton, R. K. (2011). Consequences of violated equating assumptions under the equivalent groups design. International Journal of Testing, 11, 308-323.
- Slinde, J. A. & Linn, R. L. (1997). An exploration of adequacy of the rasch model for the problem of vertical equating. Journal of Educational Measurement, 15, 23-35.
- Tanguma, J. (2000). Equating test scores using the linear method: A primer. Paper presented at the Annual Meeting of Southwest Educational Research Association, Dallas, Texas.
- Wang, T., Lee, W., Brennan, R. L., & Kolen, M. J. (2008). A comparison of the frequency estimation and chained equipercentile methods under the common-item non-equivalent groups design. Applied Psychological Measurement, 32, 632-651.

of Iowa. (Available on http://www. Education.uiowa.edu/casma).

- Harris, D. J. & Kolen, M. J. (1990). Effect of examinee group on equating relationships. Applied Psychological Measurement, 10, 35-43.
- Holland, P. W. & Rubin, D. B. (1982). Test equating. New York: Academic.
- Efron, B. (1982).The Jackknife, the bootstrap and other resampling plans. Philadelphia, PA: Society for Industrial and Applied Mathematics.
- Efron, B. & Tibshirani, R. J. (1993). An introduction to the bootstrap (Monographs on Statistics and Applied Probability 57). New York: Chapman & Hall.
- Ezzelle, C. & Setzer, J. C. (2009). Technical manual: 2002 Series GED Tests. Washington, DC: GED Testing Service., American Council on Education.
- Kolen, M. J. (1988). An NCME instructional module on traditional equating methodology. Educational Measurement: Issues and Practice, 7, 29-36.
- Kolen, M. J. (2004). Rage-RGEQUATE manual (Console version). Iowa City, IA: Center for Advanced Studies in Measurement and Assessment, The University of Iowa. (Available on http://www. Education.uiowa.edu/casma).
- Kolen, M. J. & Brennan, R. L. (2004). Test equating, scaling, and linking: Methods and practices (2nd ed.). New York: Springer-Verlag.
- Kolen, M. J. & Whitney, D. R. (1982). Comparison of four procedures for equating the Tests of General Educational Development. Journal of Educational Measurement, 19, 279-293.
- Lohman, D. F., & Gambrell, J. (2012). Use of nonverbal measures in gifted identification. Journal of Psychoeducational Assessment, 30, 25-44.
- Lord, F. M. (1980). Application of item