أثر طريقة تصفية المموهات في فقرات الاختيار من متعدد على معادلة الاختبار

*موسى محمد محمود كوفحي، ريما فايز علي زواهرة**

ملخص

هدفت هذه الدراسة إلى استقصاء أثر طريقة تصفية المموهات في فقرات الاختيار من متعدد على معادلة الاختبار. ولتحقيق هدف الدراسة، تم إعداد اختبار في الرياضيات من (40) فقرة من نوع الاختيار من متعدد بستة بدائل. وبعد إعداده بصورته الأولية تم تطبيقه على عينة استطلاعية من (130) طالب، لحساب معاملات الصعوبة والتمييز للفقرات، وقد تكون الاختبار بصورته النهائية من (34) فقرة من ستة بدائل لكل فقرة، تم تكوين أربعة نماذج لكل منها أربعة بدائل للفقرة بعد حذف مموهين من بدائل فقرات الصورة النهائية للاختبار لتكون على النحو الآتي: نموذج بطريقة التصفية العشوائية للمموهات، نموذج بطريقة التصفية التجريبية، نموذج بطريقة التصفية بالتحكيم، ونموذج بطريقة التصفية المنطقية، وطبقت النماذج الأربعة على عينة مؤلفة من (880) طالباً وطالبة من طلاب الصف السابع الأساسي خلال الفترة الواقعة ما بين 2013/4/10م ولغاية 2013/4/16م، واستخدمت طريقة المعادلة وفق النموذج أحادي المعلم (راش) من الطرق القائمة على النظرية الحديثة في القياس. وقامت إجراءات المعادلة على تصميم المجموعات العشوائية لجمع البيانات، وبعد تحليل البيانات أظهرت النتائج التأكيد على ضرورة معادلة الدرجات على نماذج تصفية المموهات، وأظهرت كذلك تأثر دقة وفاعلية المعادلة باختلاف طريقة تصفية المموهات، وأظهرت النتائج أن فاعلية المعادلة وفق نموذج راش تقل عندما لا تكون البيانات متطابقة مع النموذج، وتزداد دقة وفاعلية المعادلة وفق نموذج راش في حال مطابقة البيانات لافتراضات النموذج.

الكلمات الدالة: تصفية المموهات، معادلة الاختبار، فقرات الاختيار من متعدد.

المقدمـة

تعد اختبارات الاختيار من متعدد من أكثر الاختبارات انتشاراً في التربية، والمجالات التنظيمية الأخرى، وهذا النوع من الاختبارات قد تفوقت على كل أشكال الفقرات الموضوعية الأخرى، وذلك بسبب كفاءتها وتعدد استعمالاتها، إذ يمكن بواسطتها قياس أهداف بسيطة وأخرى مركبة في مختلف المواضيع الدراسية (Aiken, 2003).

ومهما يكن عدد البدائل للاختبار يجب أن يراعى في صياغتها مجموعة من المعايير وذلك للحصول على مموهات جذابة للطلبة (ثورندايك، وهيجن، 1986).

وتعد عملية تحليل فقرات الاختبار من الأمور ذات الأهمية في معرفة أداء المفحوص على الاختبار وتعتمد عملية تحليل الفقرات على الغرض من الاختبار أو طريقة تفسير نتائجه. ويوجد بعض المؤشرات الإحصائية التي تستخدم في تحليل الفقرات مثل صعوبة الفقرة وتمييزها وجاذبية المموهات، ويشير مفهوم صعوبة الفقرة وفق النظرية الكلاسيكية في القياس إلى نسبة الطلبة الذين أجابو على الفقرة إجابة صحيحة، ويعتبر أفضل قيمة لمعامل الصعوبة بغياب التخمين (0.50) وبما أنه من الصعب أن تكون جميع الفقرات بهذا المستوى من الصعوبة، فإن أي فقرة ضمن توزيع لمعاملات الصعوبة يتراوح بين (0.20–0.80) بمتوسط (0.50) يمكن أن تكون مقبولة.

ويشير التمييز إلى قدرة الفقرة التمييزية بين من يمتلكون المعرفة وبين من لا يمتلكونها في اختيار البديل الصحيح، ويعتبر أنسب القيم لمعامل التمييز هي القيم التي تتراوح بين (0.20) والواحد صحيح إذ إنها تعتبر فقرات مقبولة ويمكن الاحتفاظ بها في الاختبار، وفيما يتعلق بجاذبية المموه فيتم حسابه من خلال إيجاد نسبة الطلبة الذين اختاروا المموه من الفئة العليا مطروحا منها نسبة الطلبة الذين اختاروا المموه من الفئة الدنيا، وفي حال كان عدد الطلبة ممن أجابوا على المموه من الفئة العليا أكثر من الطلبة من الفئة الدنيا فإن مثل هذا المموه يفضل حذفه، وبالتالي فان أي تغيير يطرأ على هذه المؤشرات يؤثر على عملية تحليل فقرات الاختبار مما يؤثر على أداء المفحوص على الاختبار(عودة، 2010). ويبقى السؤال مطروحاً حول تأثير هذا التغيير في الشكل والمضمون على خصائص الاختبار وخصائص درجات المفحوصين قائماً.

ويرى عودة (2010، المرجع السابق) بأن المموهات يجب أن تكون جذابة للطلبة، أي أنه من الأفضل أن لا يقل عدد

* وحدة القياس والتقويم، عمادة الجودة والاعتماد الأكاديمي، الجامعة الإسلامية بالمدينة المنورة، السعودية؛ وقسم الامتحانات التنافسية، ديوان الخدمة المدني، عمان، الأردن. تاريخ استلام البحث 2014/10/14، وتاريخ قبوله 2015/3/15.

الطلبة الذين اختاروا المموه عن (0.05) من الطلبة ككل، وبما أن المموه يعتبر إجابة خاطئة فمن البديهي أن يكون عدد الطلبة الذين يختارون أي منها في الفئة العليا أقل منه من الفئة الدنيا.

ويمكن القول أنه إذا كان عدد الطلبة الذين اختاروا المموه قليلاً يجب حذفه، وبالتالي، فإنه من الممكن اعتبار الطريقة التي يتم بها تصفية (حذف) المموه متغيرا يؤثر على خصائص الفقرة من صعوبة وتمييز وهذا التأثير أكدت عليه بعض الدراسات مثل دراسة ظاظا (2000)، ودراسة أبو جراد (1997)، وقد أشارت نتائجهم بوجود أثر لطريقة تصفية المموهات على معاملات الصعوبة والتمييز لفقرات اختبار الاختيار من متعدد.

وبسبب الاختلاف الذي قد يظهر في خصائص الفقرات والاختبار لنماذج أو صور اختبارات الاختيار من متعدد التي تقيس نفس النطاق من المحتوى، الناتجة عن طريقة تصفية المموهات، مثل هذه الاختبارات قد لا تكون عادلة في قياسها وخاصة أنها تعد نماذج أو صور مختلفة للاختبار نفسه، وقد غاب عن عين الباحثين بإجراء أي معادلة للاختبارات التي يتم تصفية مموهاتها بإحدى طرق التصفية المختلفة، وهذا ما جاءت به هذه الدراسة في معادلة النماذج المختلفة للاختبار الناتجة عن الطريقة التي يتم بها تصفية المموهات ودراسة أثر طريقة تصفية المموهات على معادلة هذه النماذج من الاختبار.

مشكلة الدراسة

تعتبر كتابة فقرات الاختبار من نوع الاختيار من متعدد ببدائل جذابة ومتجانسة مهمة ليست سهلة على كثير من المدرسين، وهذه المهمة تصبح أسهل كلما قل عدد بدائل الفقرات التي يعدونها، كما قد يختلف مصممي الاختبارات من نوع الاختيار من متعدد في عدد البدائل التي يضعونها لفقرات هذه الاختبارات، وقد يختلفون في الطريقة المستخدمة في تصفية المموهات، ونظراً لإمكانية اختلاف الخصائص السيكومترية للفقرات والاختبار بحسب اختلاف الطريقة المستخدمة في تصفية المموهات، الأمر الذي يؤدي إلى اختلاف درجات المفحوصين على صور الاختبار المكونة من طرق التصفية.

يشير كروكر وألجينا (Crocker & Algina, 1986) إلى ضرورة معادلة الدرجات المختلفة للنماذج المختلفة للاختبار نفسه للتخلص من الأخطاء التي يقع فيها المحللين لنتائج الدرجات عند تفسيرها، كما يشير سون (Suen, 1990) أنه عندما يكون لدينا أكثر من صورة واحدة من الاختبار نفسه تكون معادلة صور الاختبار قضية مهمة للحصول على التفسير المناسب والصحيح لدرجات الاختبار. ولما كانت الطريقة المستخدمة في تصفية المموهات تؤثر على خصائص الفقرات (الصعوبة والتمييز) في اختبار الاختيار من متعدد وبالتالي التأثير على درجات الأفراد في الاختبار كما أكدت عليه بعض الدراسات، كان من الضروري تعديل الفروق الناجمة عن اختلاف خصائص الفقرات ودرجات الأفراد في نماذج اختبار الاختيار من متعدد التي يتم تكوينها بالطرق المختلفة لتصفية المموهات، ومعرفة أثر طرق تصفية المموهات على فاعلية ودقة معادلة الدرجات للنماذج المختلفة في طريقة تصفية مموهاتها، وعلى هذا الأساس جاءت مشكلة الدراسة تبحث في أثر طريقة تصفية المموهات على معادلة نماذج اختبار الاختيار من متعدد التي تختلف في طريقة تصفية مموهاتها.

أسئلة الدراسة

تحديداً فإن هذه الدراسة سعت للإجابة عن الأسئلة التالية:

أ. **ما أثر طريقة تصفية المموهات في فقرات الاختيار من متعدد على دقة وفاعلية معادلة الاختبار؟** ويتفرع من هذا السؤال الأسئلة الفرعية الآتية:

- ما المؤشرات الاحصائية المتحققة لنماذج الاختبار تبعاً لطريقة تصفية المموهات وفق النظرية الكلاسيكية في القياس؟
- هل توجد فروق دالة إحصائياً عند مستوى الدلالة (α=0.05) للمتوسطات الحسابية لدرجات المفحوصين والمتوسطات الحسابية لمعاملات الصعوبة ومعاملات التمييز تعزي لطريقة تصفية المموهات وفق النظرية الكلاسيكية في القياس؟
- ما مؤشرات مطابقة الفقرات والافراد لنماذج الاختبار المكونة بطرق تصفية المموهات وفق نموذج راش؟
- هل تختلف دقة وفاعلية المعادلة لنماذج الاختبار (من خلال قيم محك الخطأ المعياري للمعادلة ومحك الصدق التقاطعي) باختلاف طريقة تصفية المموهات؟

أهمية الدراسة

تظهر أهمية هذه الدراسة من خلال معرفة أثر طريقة تصفية المموهات على معادلة درجات الاختبار، وفقاً لنظرية الاستجابة للفقرة، كما تظهر أهمية الدراسة في تقديمها الدليل العلمي على تأثير طريقة تصفية المموهات في فقرات الاختيار من متعدد على معادلة درجات الاختبار، وعدم الاكتفاء

بالنصائح المبنية على المظهر أو التوصيات التي تخلو من المراجع الداعمة والأبحاث، خصوصاً في ظل تزايد الحديث عن النماذج المختلفة للاختبارات في الاختبارات الدولية والاختبارات واسعة النطاق وكذلك في الاختبارات الصفية والاختبارات التي تُعد في المؤسسات التعليمية المختلفة، وقد تكون إحدى طرق تكوين هذه النماذج قد تمت بطرق مختلفة لتصفية المموهات، ودائما الادعاء وارد بين المفحوصين المطبق عليهم نماذج الاختبار المختلفة، بأنه وقع ظلم علينا لأن النموذج المطبق علينا أصعب من النماذج الأخرى، وبالتالي فقد يُسهم البحث الحالي في تقديمه تأكيد علمي على ضرورة معادلة الدرجات على نماذج الاختبار المختلفة بطريقة تصفية مموهاتها، حتى تتحقق العدالة بين المفحوصين. كما يمكن أن يسهم البحث الحالي في تقديم أداة اختبار جيدة لمعلمي الرياضيات.

تعريف المصطلحات:

1. **اختبار الاختيار من متعدد:** هو اختبار مؤلف من فقرات كل فقرة مكونة من متن وله عدة بدائل، وتكون الاستجابة لكل منها ثنائية تأخذ قيمتين: واحد للإجابة الصحيحة، وصفر للإجابة الخاطئة.
2. **التصفية العشوائية للمموهات** (Random Filtering) هو حذف المموه عشوائيا مع إرجاعه عند إعادة الحذف مرة ثانية.
3. **التصفية بالتحكيم للمموهات** (Judicial Filtering) هو حذف المموه غير الفعال الذي له أعلى نسبة إجماع بين المحكمين.
4. **التصفية التجريبية للمموهات** (Experimental Filtering) هو حذف المموه الأقل تمييزاً بين المفحوصين والذي يمثل نسبة الطلبة الذين أجابوا على الفقرة من الفئة العليا مطروحاً منه نسبة الطلبة الذين أجابوا على الفقرة من الفئة الدنيا، أو يمثل معامل ارتباط بايسيريل (biserial).
5. **التصفية المنطقية للمموهات** (logical Filtering) وهو حذف المموه الذي اختاره أقل عدد من الطلبة.
6. **معادلة الاختبار** (Test Equating) تحويل نظام وحدات القياس الخاص بإحدى الصورتين إلى نظام وحدات القياس الخاص بالصورة الأخرى، بحيث تصبح القياسات المستمدة من درجات كل من الصورتين متكافئة بعد إجراء هذا التحويل.
7. **المعادلة الأفقية** (Horizontal Equating) يتم من خلالها معادلة درجات صور مختلفة للاختبار متماثلة في المحتوى ومتقاربة في مستوى صعوبة فقراتها وتقيس السمة نفسها في المجتمع.
8. **المعادلة المئينية** (Equipercentile Equating) تعتبر درجات المفحوصين متعادلة إذا تساوت رتبها المئينية بالنسبة لمجموعة معينة من الأفراد.
9. **محك الخطأ المعياري للمعادلة** (Standard Error of Equating) ويمثل الانحراف المعياري للدرجات المعادلة من إحدى صورتي الاختبار للأخرى من تحويل مبني على عينات اختيرت بشكل مستقل.
10. **محك الصدق التقاطعي** (Cross-Validation) هو متوسط مربعات الانحرافات للدرجات المتعادلة عندما يطبق على عينات عشوائية مستقلة، ويكشف عن مدى استقرار الدرجات المتعادلة في عينتين عشوائيتين مستقلتين عن بعضهما البعض.

محددات الدراسة

عند تعميم نتائج هذه الدراسة يجب الأخذ بعين الاعتبار المحددات التالية:

أ. اقتصرت الدراسة على طريقة واحدة من طرق المعادلة القائمة على النظرية الحديثة في القياس وهي طريقة المعادلة باستخدام النموذج أحادي المعلم (راش)، بسبب وجود طرق أخرى للمعادلة من كل نظرية من نظريات القياس غير الطرق المستخدمة في الدراسة.

ب. اقتصرت الدراسة على الاستعانة ببيانات (50) فرد من أفراد العينة الاستطلاعية، لتحليلها في برنامج (Iteman) لاستخراج خصائص المموهات، والتي استخدمت في بناء نموذج التصفية التجريبية للمموهات، ونموذج التصفية المنطقية للمموهات، لأسباب فنية واقتصادية، والافتراض هنا أن البيانات المختارة تمثل شريحة ممثلة للعينة الاستطلاعية لإيجاد خصائص المموهات.

ج. اقتصرت الدراسة على معياريين من معايير الفاعلية وهما معيار الخطأ المعياري للمعادلة ومعيار الصدق التقاطعي.

د. اقتصرت عينة الدراسة على طلبة الصف السابع الأساسي في مديرية التربية والتعليم للواء قصبة إربد.

الدراسات السابقة

أجرى الحياصات (2011) دراسة هدفت إلى الكشف عن مدى تحقق معايير الفاعلية في معادلة نموذجي اختبار مع بقاء الفقرات ذات الأداء التفاضلي للجنس وحذفها. ولتحقيق هدف الدراسة، استخدمت البيانات التي تم الحصول عليها من خلال

استجابات طلبة الصف الرابع الأساسي على نموذجي الاختبار الوطني في الرياضيات لعامي (2006 و2009). وقد استخدمت طريقتان في المعادلة، الأولى تتبع النظرية الكلاسيكية في القياس، وهي: طريقة تكر الخطية، والثانية تتبع النظرية الحديثة في القياس، وهي: طريقة معادلة الدرجات الحقيقية. وللحكم على دقة المعادلة تم استخدام ستة معايير. وأشارت الدراسة في جزء من نتائجها أن حذف الفقرات ذات الأداء التفاضلي للجنس يؤدي الى زيادة دقة المعادلة وفاعليتها وفقاً لطريقة تكر الخطية وطريقة معادلة الدرجات الحقيقية. كما أظهرت النتائج أن المعادلة باستخدام طريقة تكر الخطية أكثر دقة وفاعلية من طريقة معادلة الدرجات الحقيقية قبل حذف الفقرات ذات الأداء التفاضلي للجنس وبعده.

وقامت الحواري (2008) بدراسة الخصائص السيكومترية لاختبار الرخصة الدولية لقيادة الحاسوب ومعادلة درجات الصور المختلفة من الاختبار، وقد استخدمت طريقتي المعادلة الخطية الأفقية والمعادلة المئينية الأفقية كطريقتين وفق النظرية الكلاسيكية واستخدام نموذج راش كطريقة من طرق النظرية الحديثة في القياس، وقد تم الاعتماد على معياري الخطأ المعياري ومعامل الصدق التقاطعي للحكم على فاعلية المعادلة. وأظهرت النتائج أن المعادلة وفق نموذج راش هي الأفضل والأكثر فاعلية في معادلة درجات المفحوصين من خلال وضعهم على ميزان واحد مشترك، ثم يليها المعادلة الخطية وأخيراً المعادلة المئينية.

وأما المدانات فقد أجرى عام (2008) دراسة هدفت إلى تقصي أثر طريقة المعادلة باستخدام اختبار مشترك وعدد فقراته وحجم العينة على القيم المعادلة والخطأ في المعادلة بين صورتي اختبار في الفيزياء. ولتحقيق هذا الهدف، تم بناء صورتين متكافئتن لاختبار الفيزياء عدد فقرات كل منها (40) فقرة، وتم استخدام أربع طرق من المعادلة تتبع النظرية الكلاسيكية وهي: طريقة تكر الخطية، طريقة ليفن الخطية، وطريقة براون، وطريقة هولند الخطية، وطريقتان ذات علاقة بالنظرية الحديثة وهما: طريقة معادلة الدرجات الحقيقية، وطريقة معادلة الدرجات المشاهدة. كما تم استخدام ثلاث أحجام مختلفة من العينات وثلاثة أعداد مختلفة من فقرات الجذع المشترك وقد أظهرت النتائج فيما يخص بدقة المعادلة وجود فروق لصالح طريقة معادلة الدرجات الحقيقية، وطريقة معادلة الدرجات المشاهدة، ولصالح العدد الأكبر من فقرات الاختبار المشترك بينما لم تظهر النتائج فروقا تُعزى إلى حجم العينة.

وقام عيلبوني (2007) بدراسة هدفت لمعرفة أثر عدد البدائل وتمييز المموهات في فقرات الاختيار من متعدد على التوافق مع النموذج ثلاثي المعلم. ولتحقيق أهداف الدراسة تم بناء اختبار تحصيلي في مادة الرياضيات للصف التاسع الأساسي، وتم تكوين ثلاثة أشكال اختبارية لكل منها (50) فقرة، بحيث يحتوي الشكل الأول على خمسة بدائل للفقرة، والشكل الثاني على ثلاثة بدائل وذلك بحذف مموهين بشكل عشوائي من الشكل الأول، والشكل الثالث له ثلاث بدائل وذلك بحذف مموهين بالطريقة التجريبية (حذف المموهين الأقل تمييزاً) من الشكل الأول. طبق الاختبار على عينة مكونة من (1656) طالباً وطالبة من طلاب الصف التاسع الأساسي التابعين لمديرية التربية والتعليم لمنطقة إربد الثانية. وبعد تحليل البيانات أظهرت الدراسة في جزء من نتائجها بأنه لا يوجد أثر لطريقة حذف المموهات في فقرات الاختيار من متعدد للأشكال ذات الثلاثة بدائل على خصائص الاختبار وفقراته وفق النموذج ثلاثي المعلم.

وأجرى هيه (heh, 2007) دراسة بحثت في دقة المعادلة باستخدام عينات صغيرة وتصميم المجموعات العشوائية، واستخدم ثلاث معايير للحكم على دقة المعادلة وهي: الجذر التربيعي لمتوسط مربعات الخطأ (RMSE)، والخطأ المعياري (SE)، والتحيز (Bias). وأشارت النتائج الخاصة بأطوال الاختبارات وتصاميم المعادلة المستخدمة في الدراسة أن دقة المعادلة بالنسبة للعينات الصغيرة تعتمد على الفروقات في الصعوبة بين صور الاختبار، وكذلك تعتمد على مدى العلامات التي تم دراسة دقة المعادلة عليها وعلى حجم العينة وطريقة المعادلة المستخدمة، ففي الاختبارات ذات الفروقات المتوسطة إلى العالية في الصعوبة فقد أظهرت النتائج عدم الجدوى في التركيز على خطأ العينة لتحديد إمكانية إجراء المعادلة عند العينات الصغيرة، فالعينات الصغيرة تعطي أخطاء عينة جوهرية مما يؤدي إلى فروقات صعوبة كبيرة نسبياً بين صور الاختبارات لقلة البيانات. لذلك وعند وجود اعتبارات عملية تصعب استخدام عينات كبيرة فإنه يمكن الحصول على فاعلية معادلة عند أحجام عينات صغيرة تصل إلى (25) فرداً، وفي أوضاع تختلف فيها الصعوبة بين صور الاختبار بمقدار لا يقل عن (0.3) انحرافات معيارية.

وأجرى الشريفين (2003) دراسة حول مدى تحقق معايير الفاعلية في معادلة اختبارين أحدهما ثنائي التدريج والآخر متعدد التدريج وفق نماذج النظرية الكلاسيكية والحديثة في القياس، حيث استخدم الباحث الطريقة الخطية والطريقة المئينية للمعادلة كإحدى طرق النظرية الكلاسيكية مقابل نموذج راش الأحادي كطريقة من طرق النظرية الحديثة في القياس، وقام

الباحث ببناء اختبارين تحصيلين في الفيزياء أحدهما ثنائي التدريج والآخر متعدد التدريج، وأظهرت النتائج أن نموذج راش كان لديه تفوق على طريقتي النظرية الكلاسيكية، باستخدام معياري الخطأ المعياري ومعامل الصدق التقاطعي كمحكات للحكم على فاعلية المعادلة.

وفي دراسة أجراها كيم وهانسون (kim & hanson, 2002) حول معادلة الاختبارات لنماذج مختلفة من اختبارات في الرياضيات من نوع الاختيار من متعدد، حيث تكونت هذه الاختبارات من (65) فقرة أعدت بصورة نموذجين مختلفين، بحيث تم تطبيق النموذج الأول في عام (1997) على عينة مكونة من (2163) مفحوص، وأما النموذج الثاني تم تطبيقه في عام (1998) على عينة مكونة من (2909) مفحوص، ولأغراض جمع البيانات استخدم الباحثان تصميم المجموعات العشوائية المتكافئة، ولإجراءات المعادلة استخدما طريقة منحنى خصائص الفقرة، وتم تقدير المعلمات المطبقة على نماذج الاختبارات المختلفة باستخدام برنامج (MULTILOG) وهو برنامج محوسب يتم من خلاله تقدير قدرة المفحوص باستخدام منحنى خصائص الفقرة بالاعتماد على الدرجات التي يحصل عليها الأفراد على كل فقرة من فقرات الاختبار، وقد أشارت النتائج إلى فاعلية طريقة منحنى خصائص الفقرة في معادلة نماذج الاختبارات المختلفة.

كما أجرى ظاظا (2000) دراسة بعنوان المقارنة بين أثر أربع طرق لحذف المموهات في فقرات اختبار الاختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته، ولتحقيق أهداف الدراسة تم استخدام اختبار الاستدلال الحسابي والمؤلف من (40) فقرة من نوع الاختيار من متعدد. وتم تكوين ثمانية نماذج من الاختبار الأصلي لتكون على النحو التالي: نموذجين بطريقة الحذف التجريبي، نموذجين بطريقة الحذف العشوائي للمموهات، ونموذجين بطريقة الحذف التتابعي للمموهات، وآخر نموذجين بطريقة الحذف بالتحكيم، وطبقت النماذج الثمانية على عينة مؤلفة من (436) من طلبة الصف العاشر، وبعد تحليل البيانات أشارت نتائج الدراسة إلى عدم وجود فروق ذات دلالة إحصائية في معاملات الصعوبة عند مستوى ($\alpha = 0.05$) تُعزى لطريقة الحذف. كما أشارت النتائج إلى وجود فروق ذات دلالة إحصائية في معاملات التمييز بين النموذج المكون وفق الطريقة التجريبية والنموذج المكون وفق طريقة الحذف بالتحكيم والذي لفقرات كل منها ثلاثة بدائل تُعزى لطريقة الحذف. وأما النتائج المتعلقة بمعاملات ثبات الاتساق الداخلي فقد تبين عدم وجود فروق ذات دلالة إحصائية بين النماذج المكونة وفق طرق الحذف المختلفة والتي لها نفس عدد البدائل. وأظهرت النتائج عدم وجود فروق ذات دلالة إحصائية بين جميع معاملات الصدق بين النماذج الأربعة المكونة بطرق الحذف المذكورة.

وقام أبو جراد (1997) بدراسة هدفت لمعرفة أثر الحذف التجريبي والعشوائي في فقرات اختبار التحصيل من نوع الاختيار من متعدد على الخصائص السيكومترية. ولتحقيق أهداف الدراسة تم بناء اختبار تحصيلي لطلبة الصف العاشر، وتكون الاختبار من (45) فقرة اختيار من متعدد لكل فقرة ستة بدائل، وتم تكوين ستة نماذج بطريقة الحذف العشوائي والتجريبي للمموهات بواقع ثلاثة نماذج لكل طريقة حذف، وتم تطبيق النماذج الستة على عينة مكونة من (467) طالباً وطالبة من طلبة الصف العاشر حيث أشارت النتائج إلى أن هناك فروقاً ذات دلالة إحصائية عند ($\alpha = 0.05$) في معاملات الصعوبة بين النموذجين اللذين لفقرات كل منهما أربعة بدائل لصالح النموذج المكون بطريقة الحذف العشوائي، وأظهرت النتائج الخاصة بمعاملات التمييز عدم وجود فروق دالة إحصائياً بين كل نموذجين لهما نفس عدد البدائل تُعزى لطريقة الحذف. أما فيما يتعلق بأثر طريقة الحذف التجريبي والعشوائي للمموهات في فقرات الاختبار على خصائصه السيكومترية فقد أشارت النتائج المتعلقة بالثبات إلى وجود فروق ذات دلالة إحصائية بين النموذجين اللذين لفقرات كل منهما خمسة بدائل لصالح النموذج المكون بطريقة الحذف العشوائي، ولا توجد فروق ذات دلالة إحصائية بين جميع معاملات الصدق تُعزى لطريقة الحذف، وأوصت الدراسة باستخدام الطريقة العشوائية في حذف المموهات.

وقام كل من سيزيك وأودي (Cizek & O'day, 1994) بإجراء دراسة هدفت إلى معرفة دقة التحكيم في تحديد البدائل غير الفعالة (Nonfunctioning) في فقرات اختبار الاختيار من متعدد ذات الخمسة بدائل مقارنة بالبيانات المتوفرة من تحليل الفقرات (نسبة المفحوصين الذين اختاروا البديل). ولتحقيق ذلك الهدف تم بناء مجمع للفقرات (Item Pool)من (800) فقرة في الطب الاختصاصي، وتم اختيار (32) فقرة ذات خمسة بدائل بناءً على بيانات توفرت من المحكمين. وتم عرض الاختبار على لجنة محكمين مكونة من (12) خبيراً في مجلس الامتحانات الطبي، وذلك لتحديد البديل غير الفعال في بدائل كل فقرة، وتم الاتفاق على أن البديل غير الفعال هو "البديل الذي من المحتمل ألا يختاره المفحوصين، وبعد تطبيق هذا الاختبار على (700) مفحوص، أظهرت النتائج وجود اتفاق بين طريقة التحكيم وطريقة التجريب لتحديد البديل غير الفعال في حوالي (75%) من فقرات الاختبار.

التعقيب على الدراسات السابقة

يتبين أن جزء من الدراسات السابقة قد بحث في دراسة أثر طرق تصفية (حذف) المموهات على الخصائص السيكومترية للاختبار مثل دراسة أبو جراد (1997) ودراسة ظاظا (2000) والتي أظهرت بوجود أثر لطريقة تصفية (حذف) المموهات على الخصائص السيكومترية للاختبار وفقراته، ودراسات لم تظهر أثر لطريقة تصفية المموهات على خصائص الاختبار وفقراته مثل دراسة سيزيك وأودي (Cizek & O'day, 1994) ودراسة عيلبوني (2007) وأن الجزء الآخر من الدراسات قد تناول موضوع معادلة الاختبارات في النظريتين الكلاسيكية والحديثة في القياس، وكانت تدور في معظمها حول المقارنة بين الطرق المختلفة المستخدمة في المعادلة وإجراءاتها من حيث طبيعة الاختبارات والعينات المستخدمة، وكانت نتائج هذه الدراسات متباينة إلا أن معظمها أكد على تفوق طرق النظرية الحديثة في فاعلية معادلة الاختبارات مثل دراسة كيم وهانسون (kim & hanson, 2002) ودراسة الشريفيين (2003) ودراسة الحواري (2008) ودراسة المدانات (2008). وأما دراسة هيه (heh, 2007) والتي أظهرت أن العينات الصغيرة تعطي أخطاء عينة جوهرية مما يؤدي إلى فروق صعوبة كبيرة نسبياً بين صور الاختبارات لقلة البيانات، وجزء قليل من الدراسات التي تناولت موضوع معادلة الاختبار تناولت تأثير بعض المواضيع على معادلة الاختبار منها من أخذ بعين الاعتبار موضوع الأداء التفاضلي للجنس وأثره على معادلة الاختبارات مثل دراسة الحياصات (2011) والتي أظهرت أن دقة وفاعلية المعادلة تزداد في حال حذف الفقرات ذات الأداء التفاضلي، ومنها من تناول موضوع عدد فقرات الاختبار المشترك في تصاميم ذات الاختبار المشترك وحجم العينة وأثرها على معادلة الاختبار مثل دراسة المدانات (2008).

ومع هذا العرض لأهم الدراسات السابقة ذات العلاقة بموضوع الدراسة لم تظهر أي دراسة بحثت بموضوع تصفية المموهات وتأثيره على معادلة الاختبارات، إذ يلاحظ أن معظم الباحثين الذين اهتموا بدراسة متغير تصفية المموهات قاموا بدراسة أثره على خصائص الاختبار وفقراته، وقد غاب عن أنظارهم أن الصور المختلفة للاختبار المكونة من طرق التصفية قد لا تكون عادلة في قياسها بسبب الاختلاف في خصائص الاختبار وفقراته، وبعض هذه الدراسات بحث في معادلة صور الاختبارات ودراسة بعض الموضوعات التي تؤثر على دقة وفاعلية المعادلة ولم يتم فحص الأثر الذي قد ينتج من طريقة تصفية المموهات على معادلة الاختبار، لذلك فقد تميزت هذه الدراسة عن الدراسات السابقة في أنها بحثت في أثر طريقة تصفية المموهات على معادلة الاختبار، بدراستها لمدى تحقق معايير الفاعلية لطرق المعادلة المستخدمة في الدراسة. وما مدى تأثير النماذج المختلفة للاختبار المكونة من طرق تصفية المموهات على درجات الأفراد وخصائص الاختبار وفقراته، والتحقق فيما لو كان هناك ضرورة لإجراء المعادلة للنماذج المختلفة لنماذج الاختبار المكونة من طرق تصفية المموهات.

الطريقة والإجراءات:

أ. **مجتمع الدراسة** (Accessible population): يتكون مجتمع الدراسة من طلاب وطالبات الصف السابع الأساسي الملتحقين في الدراسة في المدارس التابعة لمديرية تربية لواء قصبة إربد في الفصل الدراسي الثاني من العام الدراسي (2013/2012). ويبلغ عددهم (7937) طالباً وطالبة، منهم (3873) من الذكور و(4064) من الإناث، موزعين على (110) شعب للذكور و(115) شعبة للإناث، وذلك من خلال الرجوع إلى قسم التخطيط في المديرية، ويوضح الجدول (1) عدد الذكور والإناث وعدد الشعب لأفراد مجتمع الدراسة.

الجدول (1) توزيع أفراد مجتمع الدراسة حسب الجنس وعدد الشعب

الجنس	عدد الشعب	عدد الطلبة
ذكور	110	3873
إناث	115	4064
المجموع	225	7937

ب. **عينة الدراسة:** بعد تحديد المدارس التابعة لمديرية التربية والتعليم للواء قصبة إربد وحصرها للعام الدراسي (2013/2012)، ومن خلال الرجوع إلى قسم التخطيط في المديرية، تم تحديد المدارس التي يوجد فيها شعب للصف السابع الأساسي، وكان عدد هذه المدارس (170) مدرسة منها (71) مدرسة للذكور و(99) مدرسة للإناث، وبعد ذلك تم اختيار عينة الدراسة بالطريقة العشوائية العنقودية المتعددة المراحل (Multistage Cluster Random)، فقد تم اختيار (8) مدارس بشكل عشوائي تحتوي على شعب للصف السابع، وبعدها تم اختيار بعض الشعب عشوائياً من كل مدرسة وقع عليها الاختيار وبلغ عددها (25) شعبة، وذلك كما هو موضح في الجدول (2).

الجدول (2) توزيع أفراد عينة الدراسة حسب الجنس وعدد الشعب واسم المدرسة

الجنس	اسم المدرسة	عدد الشعب	عدد الطلبة
ذكور	مدرسة معاذ بن جبل الأساسية للبنين	4	168
	مدرسة عمر بن عبد العزيز الأساسية للبنين	3	142
	مدرسة عمر فائق الشلبي الأساسية للبنين	4	94
	مدرسة ابن زيدون الأساسية للبنين	2	49
إناث	مدرسة سكينة بنت الحسين الأساسية للبنات	3	115
	مدرسة ذات النطاقين الأساسية للبنات	4	157
	مدرسة خولة بنت الأزور الأساسية للبنات	3	94
مختلطة	مدرسة الملك عبدالله الثاني للتميز	2	61 (33) ذكور (28) إناث
المجموع		**25**	**880**

ج. **أداة الدراسة:** تكونت أداة الدراسة من اختبار لقياس المعرفة الرياضية لطلبة الصف السابع الأساسي، بمحتوى وحدة الجبر من المنهاج المقرر من قبل وزارة التربية والتعليم، وقد تكونت الصورة الأولية للاختبار من (40) فقرة من نوع الاختيار من متعدد، وقبل اختيار مموهات اختبار الاختيار من متعدد تم بناء الاختبار على شكل فقرات تكميل دون التغيير في أي متن لأي فقرة، وكان الهدف بتكوين الاختبار على شكل إكمال الفراغ بشكله الأولي الحصول على إجابات الطلبة (والذي تم عرضه على (70) طالباً وطالبة) وحصرها ليتم اختيار مموهات الاختبار من إجابات الطلبة. وبعد اعداد الاختبار على شكل فقرات اختيار من متعدد تم تطبيقه على عينة استطلاعية من (130) طالباً خارج عينة الدراسة ومن مجتمع الدراسة وبعد تحليل نتائج العينة الاستطلاعية وحذف الفقرات الغير ملاءمة بناءً على معايير الملاءمة للفقرات أصبح عدد فقرات الاختبار بصورته النهائية (34) فقرة.

د. **صدق الاختبار:** للتأكد من صدق الاختبار تم عرضه على مجموعة من الخبراء والمحكمين من ذوي التخصص والخبرة والكفاءة العلمية لإبداء رأيهم فيه، سواء من حيث مناسبة فقراته لأهداف البحث، أو مدى تغطيتها للجوانب والمجالات المقصودة في الدراسة، وبناءً على ملاحظات المحكمين فقد أجريت بعض التعديلات التي تم الأخذ بها ومراعاتها، الا أنه لم يجمع المحكمين على حذف أي فقرة من فقرات الصورة النهائية للاختبار وبذلك فقد بقي الاختبار بصورته النهائية مكون من (34) فقرة.

هـ. **ثبات الاختبار:** تم التحقق من دلالات الثبات للاختبار عن طريق حساب معامل ثبات الاختبار باستخدام معادلة كرونباخ ألفا (Cronbach Alpha) وكانت قيمته (0.876) مما يعني أن أداة الدراسة تتمتع بدرجة ثبات عالية، إذ أن جميع معاملات الثبات كانت أعلى من (0.60) وهي الحد الأدنى المقبول للثبات.

وبعد إعداد الاختبار بصورته النهائية من نوع الاختيار من متعدد من ستة بدائل، تم بناء أربع نماذج للاختبار كل منها يتكون من فقرات ذات أربع بدائل للإجابة بعد حذف مموهين من مموهات فقرات الاختبار الأصلي حسب كل طريقة من طرق تصفية المموهات المذكورة في الدراسة وبناءً على التعريف الإجرائي لكل طريقة من طرق التصفية المذكور سابقاً، وبذلك أصبح هنالك أربع نماذج تم بناءها لتحقيق غرض الدراسة وهي: نموذج التصفية العشوائية للمموهات، ونموذج التصفية التجريبية، ونموذج التصفية المنطقية، ونموذج التصفية بالتحكيم.

جمع البيانات: بعد إعداد الباحث نماذج الاختبار الأربعة المكونة بطرق تصفية المموهات بصورتها النهائية والمكونة من (34) فقرة، تم استخدام تصميم المجموعات العشوائية (Random (Groups Design)، أثناء تطبيق الاختبار على عينة الدراسة المكونة من (880) طالباً وطالبة في الفترة الواقعة ما بين 2013/4/10 ولغاية 2013/4/16، إذ تم توزيع نماذج الاختبار على أفراد عينة الدراسة بحيث يوزع كل نموذج على (220) طالباً وطالبة بشكل عشوائي لكلا الجنسين لضمان التجانس بين النماذج الأربعة، وبعد الانتهاء من التطبيق والحصول على النتائج تم ادخال البيانات إلى ملف نصي ولكل نموذج على حدا وترحيله على نسخة من برنامج الـ (SPSS).

المعالجات الإحصائية: للإجابة عن أسئلة الدراسة تم استخدام عدد من المعالجات الإحصائية وفق مجموعة من الإجراءات تم إتباعها وهي:

أولا: المؤشرات الإحصائية لنماذج الاختبار المختلفة في طريقة تصفية مموهاتها حسب النظرية الكلاسيكية في القياس

حيث تم التعرف على المواصفات الإحصائية المتحققة لنماذج

الاختبار المختلفة وفق النظرية الكلاسيكية في القياس من خلال دراسة مايلي:

1. **دراسة الخصائص السيكومترية:** تم تقدير معاملات الصعوبة والتمييز لفقرات النماذج المختلفة للاختبار حسب النظرية الكلاسيكية، بتحليل استجابات أفراد النماذج باستخدام البرمجية الإحصائية (SPSS)، فقد تم تقدير معامل الصعوبة لكل فقرة، وتم تقدير معامل التمييز للفقرات (معامل الارتباط الثنائي بين الفقرة والاختبار ككل) لكل نموذج تمت مطابقة فقرات النماذج المختلفة للاختبار وفق النظرية الكلاسيكية في القياس، كما تم إيجاد المتوسطات الحسابية لدرجات الأفراد والمتوسطات الحسابية لكل من معاملي الصعوبة والتمييز لكل نموذج بعد المطابقة الكلاسيكية.
2. **مؤشرات الثبات:** تم حساب معامل الثبات للنماذج المختلفة بواسطة معادلة كودر ريتشاردسون(20) والتي تعتبر مؤشراً على ثبات الاتساق الداخلي للاختبار، وكذلك تم حساب الخطأ المعياري للقياس لجميع النماذج، والذي يمكن استخدامه كدليل آخر على ثبات النماذج المختلفة للاختبار.
3. **تحليل التباين الأحادي:** تم تحليل البيانات باستخدام تحليل التباين الأحادي للكشف عن مدى الفروق في المتوسطات الحسابية لمعاملات الصعوبة والتمييز والمتوسطات الحسابية لدرجات المفحوصين للمجموعات المختلفة وعلى النماذج المختلفة، وذلك لضمان إجراء المعادلة.

ثانياً: معادلة درجات الاختبار وفقاً لنموذج راش

المعادلة وفقاً لنموذج أحادي المعلم (نموذج راش): تمت عملية المعادلة وفق نموذج راش من خلال مجموعة من الإجراءات وهي:

أ. تم فحص مطابقة الأفراد والفقرات للنموذج أحادي المعلم باستخدام البرنامج الإحصائي (Bilog-MG3) لتحليل البيانات لنماذج الاختبار الأربعة، وذلك عن طريق حساب تقديرات الأرجحية العظمى (ML) Maximum Likelihood Estimation، حيث تضمنت عملية التحليل إيجاد تقديرات قدرات الأفراد وحساب الخطأ المعياري لتقديرات قدرات الأفراد لكل نموذج على حدا، وإيجاد تقديرات معلمة الصعوبة والخطأ المعياري لتقديرات معلمة الصعوبة، بالاعتماد على النظرية الحديثة في القياس حسب النموذج أحادي المعلم.

ب. بعد استبعاد الأفراد والفقرات غير المطابقة للنموذج أحادي المعلم، تم التحقق من افتراض أحادية البعد من خلال فحص أحادية البعد للبيانات الناتجة من تطبيق النماذج المختلفة للاختبار باستخدام برنامج (SPSS) حيث أُجري التحليل بطريقة المكونات الأساسية (Principal Component Analysis) وقد استخدمت ثلاث مؤشرات للحكم على أحادية البعد وهي: ملاحظة نسبة ما يفسره العامل الأول من التباين الكلي (أكبر من (20)). ملاحظة نسبة قيمة الجذر الكامن (Eigen Value) للعامل الأول إلى الجذر الكامن للعامل الثاني (أكبر من (2)). ملاحظة التباين المفسر التراكمي بعد استخلاص مجموع مربعات التشبعات (أكبر من(54)).

ج. تمت معادلة تقديرات قدرات الأفراد للنماذج المختلفة باستخدام الرتب المئينية، وتعتبر القدرات متعادلة إذا كانت لها الرتبة المئينية نفسها، كما تم إيجاد الكفاءة النسبية للاختبارات عند مستويات القدرة المتعادلة من خلال حساب دالة معلومات الاختبار عند كل مستوى قدرة لكل من النماذج الأربعة، وقسمة دالة المعلومات لكل من نماذج التصفية التجريبية والتحكيم والمنطقية على دالة معلومات نموذج التصفية العشوائية، وللتحقق من معايير الفاعلية لدقة المعادلة المئينية للقدرات لنماذج المختلفة وفاعليتها، تم حساب محك الخطأ المعياري للمعادلة المئينية للقدرات لكل نموذج عند كل قدرة تمت معادلتها، ومعامل الصدق التقاطعي كمعيار إضافي للحكم على دقة وفاعلية معادلة الرتب المئينية للقدرات.

النتائج ومناقشتها

أولاً: النتائج المتعلقة بالسؤال الفرعي الأول (ما المؤشرات الإحصائية المتحققة لنماذج الاختبار تبعاً لطريقة تصفية المموهات وفق النظرية الكلاسيكية في القياس؟) لمعرفة الإحصائيات المتحققة لنماذج الاختبار حسب النظرية الكلاسيكية في القياس، تم تحليل البيانات والبالغ عددها(220) استجابة لكل نموذج، ثم حساب معاملات الصعوبة والتمييز لفقرات الاختبار لكل نموذج، بعد ذلك تمت مطابقة الفقرات وفق النظرية الكلاسيكية في القياس وحذف الفقرات غير المطابقة. كما تم إيجاد المتوسطات الحسابية والانحراف المعياري لدرجات المفحوصين ولمعاملات معلمتي الصعوبة والتمييز، وفيما يلي عرضاً للنتائج التي تم التوصل إليها في الدراسة.

ويبين الجدول (3) النتائج المتعلقة بمعاملات الصعوبة والتمييز والمتوسطات الحسابية لكل من معلمتي الصعوبة والتمييز وفق النظرية الكلاسيكية في القياس لكل نموذج من نماذج الاختبار المكونة بطرق تصفية المموهات:

الجدول(3) معاملات الصعوبة ومعاملات التمييز للنماذج المكونة بطرق تصفية المموهات

رقم الفقرة	نموذج تصفية المموهات								
	العشوائي		التجريبي		التحكيم		المنطقي		
	الصعوبة	التمييز	الصعوبة	التمييز	الصعوبة	التمييز	الصعوبة	التمييز	
1	0.60	0.34	0.70	0.46	0.75	0.39	0.71	0.52	
2	0.72	0.35	0.76	0.43	**0.85**	0.43	**0.84**	0.38	
3	0.35	**0.14**	0.64	0.48	0.39	0.39	0.47	0.42	
4	0.34	0.27	0.50	0.39	0.44	0.37	0.38	0.43	
5	0.61	0.43	0.67	0.38	0.70	0.50	0.65	0.33	
6	0.60	0.38	0.74	0.43	0.71	0.35	0.68	0.40	
7	0.61	0.40	0.75	0.23	0.72	0.32	0.75	0.28	
8	0.63	0.44	0.76	0.39	0.79	0.45	0.64	0.35	
9	0.61	0.40	0.80	0.51	0.60	0.30	0.76	0.31	
10	0.36	**0.14**	0.50	0.38	0.55	0.28	0.63	0.36	
11	0.61	0.50	0.69	0.53	0.67	0.47	0.68	0.44	
12	0.76	0.37	**0.86**	0.45	**0.87**	0.38	**0.86**	0.43	
13	0.66	0.51	0.80	0.55	0.75	0.45	0.76	0.53	
14	0.41	**0.16**	0.55	0.48	0.49	0.42	0.49	0.34	
15	0.31	**-0.01**	0.62	0.42	0.41	0.27	0.47	0.45	
16	0.79	0.43	**0.86**	0.49	**0.86**	0.40	**0.82**	0.41	
17	0.41	0.28	0.54	0.47	0.61	0.29	0.56	0.60	
18	0.46	0.41	0.62	0.46	0.49	0.29	0.61	0.48	
19	0.22	**0.19**	0.60	0.54	0.52	0.30	0.25	0.32	
20	0.53	**0.19**	0.56	0.33	0.51	0.35	0.58	0.48	
21	0.71	0.43	0.79	0.52	**0.81**	0.36	**0.85**	0.40	
22	0.43	0.35	0.67	0.45	0.62	0.24	0.72	0.37	
23	0.65	0.49	0.78	0.37	0.66	0.32	0.67	0.44	
24	0.73	0.52	**0.86**	0.50	0.68	0.37	0.74	0.52	
25	0.53	0.48	0.66	0.57	0.64	0.46	0.58	0.66	
26	0.35	0.43	0.57	0.52	0.52	0.48	0.54	0.61	
27	0.43	0.48	0.37	0.40	0.48	0.37	0.46	0.52	
28	0.58	0.43	0.71	0.54	**0.82**	0.42	**0.05**	**-0.19**	
29	0.67	0.48	0.75	0.52	0.61	0.40	0.70	0.45	
30	0.21	**-0.25**	0.20	**-0.29**	**0.11**	**-0.03**	**0.15**	**-0.29**	
31	0.48	0.28	0.67	0.51	0.71	0.32	0.56	0.36	
32	0.71	0.39	**0.85**	0.56	**0.85**	0.30	0.58	**0.19**	
33	**0.83**	**0.08**	**0.85**	0.27	**0.90**	**0.18**	**0.86**	0.27	
34	0.46	0.37	0.61	0.54	0.62	0.29	0.47	0.43	
المتوسط	0.54	0.33	0.67	0.43	0.64	0.35	0.60	0.38	

من خلال هذه المعاملات الواردة في الجدول (3)، وبناءً على المعايير المشار إليها في المقدمة فقد تم في النموذج العشوائي حذف الفقرات ذات الأرقام (3، 10، 14، 15، 19، 20، 30، 33) بسبب انخفاض القدرة التمييزية لهذه الفقرات، وأما النموذج التجريبي فقد تم حذف الفقرات (12، 16، 24، 32، 33) بسبب انخفاض معامل الصعوبة لهذه الفقرات وحذف الفقرة رقم (30) بسبب انخفاض القدرة التميزية لهذه الفقرة. وبالنسبة لنموذج التصفية بالتحكيم فقد تم حذف الفقرات ذات الأرقام (2، 12، 16، 21، 28، 32) بسبب انخفاض معامل الصعوبة لهذه الفقرات وحذف الفقرتين (30، 33) بسبب انخفاض القدرة التمييزية، وأما النموذج المنطقي فقد حذفت منه الفقرات ذات الأرقام (2، 12، 16، 21، 33) بسبب انخفاض معامل الصعوبة لهذه الفقرات وتم حذف الفقرات ذات الأرقام (28، 30، 32) بسبب انخفاض القدرة التمييزية لها، والجدول (4) يبين معاملات الصعوبة والتمييز والمتوسطات الحسابية لمعاملات الصعوبة والتمييز بعد حذف الفقرات غير المطابقة وفق النظرية الكلاسيكية ولكل نموذج من نماذج تصفية المموهات:

الجدول(4) معاملات الصعوبة والتمييز لنماذج تصفية المموهات بعد حذف الفقرات

رقم الفقرة	نموذج تصفية المموهات							
	العشوائي		التجريبي		التحكيم		المنطقي	
	الصعوبة	التمييز	الصعوبة	التمييز	الصعوبة	التمييز	الصعوبة	التمييز
1	0.60	0.36	0.70	0.45	0.75	0.37	0.71	0.51
2	0.72	0.40	0.76	0.44				
3			0.64	0.49	0.39	0.42	0.47	0.41
4	0.34	0.25	0.50	0.40	0.44	0.38	0.38	0.44
5	0.61	0.44	0.67	0.37	0.70	0.47	0.65	0.32
6	0.60	0.40	0.74	0.44	0.71	0.36	0.68	0.40
7	0.64	0.38	0.75	0.21	0.72	0.31	0.75	0.26
8	0.63	0.42	0.76	0.39	0.79	0.45	0.64	0.38
9	0.64	0.43	0.80	0.47	0.60	0.27	0.76	0.30
10			0.50	0.39	0.55	0.31	0.63	0.37
11	0.61	0.51	0.69	0.54	0.67	0.46	0.68	0.44
12	0.76	0.35						
13	0.66	0.50	0.80	0.55	0.75	0.42	0.76	0.50
14			0.55	0.49	0.49	0.43	0.49	0.36
15			0.62	0.44	0.41	0.29	0.47	0.48
16	0.79	0.44						
17	0.41	0.27	0.54	0.50	0.61	0.29	0.56	0.59
18	0.46	0.37	0.62	0.46	0.49	0.29	0.61	0.47
19			0.60	0.55	0.52	0.29	0.25	0.31
20			0.56	0.33	0.51	0.35	0.58	0.47
21	0.71	0.46	0.79	0.49				
22	0.43	0.34	0.67	0.46	0.62	0.25	0.72	0.36
23	0.65	0.52	0.78	0.35	0.66	0.31	0.67	0.43
24	0.73	0.56			0.68	0.36	0.74	0.50
25	0.53	0.49	0.66	0.54	0.64	0.46	0.58	0.65
26	0.35	0.44	0.57	0.53	0.52	0.47	0.54	0.62
27	0.43	0.46	0.37	0.42	0.48	0.36	0.46	0.54
28	0.58	0.45	0.71	0.53				
29	0.66	0.53	0.75	0.50	0.61	0.36	0.70	0.46
31	0.48	0.31	0.67	0.50	0.71	0.30	0.56	0.33
32	0.71	0.40						
34	0.46	0.38	0.61	0.55	0.62	0.29	0.47	0.47
المتوسط	0.58	0.42	0.66	0.46	0.60	0.36	0.60	0.44

ومن خلال الاطلاع على المعلومات الواردة في الجدول (4)، يتضح أن قيم معاملات الصعوبة والتمييز للفقرات لكل نموذج بعد حذف الفقرات غير المطابقة أصبحت فقرات مناسبة لأغراض الدراسة الحالية. ويبين الجدول (5) عدد الفقرات قبل وبعد حذف الفقرات غير المطابقة وفق النظرية الكلاسيكية في القياس ولكل نموذج من نماذج تصفية المموهات:

الجدول (5) عدد الفقرات قبل وبعد حذف الفقرات لنماذج تصفية المموهات

نموذج تصفية المموهات	قبل الحذف	بعد الحذف
	عدد الفقرات	عدد الفقرات
العشوائي	34	26
التجريبي	34	28
التحكيم	34	26
المنطقي	34	26

ويلاحظ من خلال الجدول (5) عدد الفقرات للصور النهائية لنماذج تصفية المموهات بعد حذف الفقرات غير المطابقة، ويحتوي الجدول (6) عرضاً تفصيلياً لملخص الإحصائيات الوصفية لنماذج الاختبار المكونة بطرق تصفية المموهات:

الجدول (6) ملخص الإحصائيات الوصفية للدرجات على نماذج تصفية المموهات

الدرجة الملاحظة	نموذج تصفية المموهات			
	العشوائي	التجريبي	التحكيم	المنطقي
المتوسط الحسابي	15.177	18.327	15.600	15.491
الانحراف المعياري	5.97	6.63	5.38	6.17
الالتواء	-0.25	-0.36	-0.05	0.10

- عدد الفقرات لكل من نماذج التصفية العشوائي، التحكيم، المنطقي (26) فقرة.
- عدد الفقرات لنموذج التصفية التجريبية (28) فقرة.

وفيما يتعلق بمعامل الثبات فقد تم استخراج معامل ثبات الاتساق الداخلي للنماذج المختلفة المكونة بطرق تصفية المموهات، وذلك بحساب معامل الاتساق الداخلي بواسطة معادلة كودر ريتشاردسون (20)، وكذلك تم حساب الخطأ المعياري للقياس ولجميع نماذج تصفية المموهات بعد حذف الفقرات غير المطابقة، والجدول (7) يبين معاملات الثبات والخطأ المعياري للقياس لنماذج تصفية المموهات:

الجدول (7) معاملات الثبات والخطأ المعياري للقياس والانحراف المعياري لنماذج تصفية المموهات

نموذج تصفية المموهات	KR-20	الانحراف المعياري	الخطأ المعياري للقياس لنماذج تصفية المموهات	عدد الفقرات
العشوائي	0.87	5.97	2.1782	26
التجريبي	0.89	6.63	2.1564	28
التحكيم	0.83	5.38	2.2426	26
المنطقي	0.88	6.17	2.1422	26

ويتضح من خلال الجدول (7) أن الاختبارات المكونة بطرق تصفية المموهات تتمتع بدرجة مقبولة من الثبات، ويتضح كذلك أن جميع القيم للخطأ المعياري للقياس كانت متدنية للنماذج المختلفة، مما يقدم مؤشراً على أن هناك اتساقاً في قياس الفقرات لقدرة معرفية واحدة على مستوى كل نموذج (Crocker & Algina, 1986)، مما يحقق شروط المعادلة لها. وللكشف عن الفروق في معاملات الثبات بين النماذج المختلفة في طريقة تصفية المموهات تم استخدام الإحصائي (M) المقترح من قبل هاكستين وولين (Hakstian & Whalen, 1976) والذي يتبع توزيع كاي تربيع بدرجات حرية تساوي (عدد معاملات الثبات – 1) وتحسب قيمة (M) من خلال المعادلة:

$$M = \frac{J-1}{1+3J}\left[\sum_{k=1}^{K} B_k - \frac{\left[\sum_{k=1}^{K}(1-\alpha_k)^{-2/3}\right]^2}{\sum_{k=1}^{K}(1-\alpha_k)^{-3/3}}\right] \quad \text{(18)}....$$

$$B_k = \frac{(9n_k-11)^2}{(n_k-1)} \quad \text{(19)}....$$

حيث: J: عدد الفقرات.

n: عدد الأفراد.

α_k: معامل ثبات النموذج K.

وبعد حساب قيمة (M) فإن الجدول (8) يبين نتائج الإحصائي (M) لفحص الفروق بين معاملات الثبات لنماذج تصفية المموهات (حيث تمثل القيمة بين [ ] القيمة الحرجة):

يتضح من خلال الجدول (8) وجود فروق ذات دلالة إحصائية بين معاملات الثبات لنماذج تصفية المموهات تُعزى لطريقة تصفية المموهات، حيث كانت قيمة (M) المحسوبة (13.47) وهي أكبر من القيمة الحرجة لتوزيع كاي تربيع والتي بلغت (7.81) عند مستوى الدلالة (α=0.05). وللكشف عن هذه الفروق تم إجراء مقارنات ثنائية بين معاملات ثبات كل نموذجين من نماذج تصفية المموهات على حدا، وذلك بحساب قيمة (M)

لكل نموذجين من نماذج تصفية المموهات ومقارنتها بالقيمة الحرجة، وقد أظهرت هذه المقارنات فروق ذات دلالة إحصائية بين معاملي الثبات لنموذج التصفية التجريبية والتصفية بالتحكيم يُعزى لطريقة تصفية المموهات ولصالح نموذج التصفية التجريبية، وفروق ذات دلالة إحصائية بين نموذج التصفية المنطقية ونموذج التصفية بالتحكيم ولصالح نموذج التصفية المنطقية عند مستوى الدلالة (α=0.05).وقد اتفقت هذه النتيجة مع ما توصلت إليه دراسة أبو جراد (1997) والتي أظهرت وجود فروق ذات دلالة إحصائية بين معاملات الثبات تُعزى لطريقة تصفية المموهات. كما وتتعارض مع دراسة سيزيك وأودي (Cizek & O'day, 1994) ودراسة ظاظا (2000) ودراسة عيلبوني (2007) والتي أظهرت بعدم وجود فروق ذات دلالة إحصائية بين معاملات الثبات تُعزى لطريقة تصفية المموهات.

ثانياُ: النتائج المتعلقة بالسؤال الفرعي الثاني (هل توجد فروق دالة إحصائياً عند مستوى الدلالة (α=0.05) للمتوسطات الحسابية لدرجات المفحوصين والمتوسطات الحسابية لمعاملات الصعوبة ومعاملات التمييز تعزي لطريقة تصفية المموهات وفق النظرية الكلاسيكية في القياس؟)

ولإجراءات الكشف عن مدى الفروق في المتوسطات الحسابية لدرجات المفحوصين على النماذج المختلفة تم توحيد المسطرة للدرجات الملاحظة للمفحوصين، وذلك بتحويل الدرجات الملاحظة للمفحوصين إلى نسب مئوية، والسبب في هذا التحويل اختلاف عدد الفقرات لكل نموذج، وبعدها تم حساب المتوسطات الحسابية والانحرافات المعيارية لدرجات المفحوصين المئوية (النسب المئوية)، والجدول (9) يبين المتوسطات الحسابية والانحرافات المعيارية للدرجات المئوية للمفحوصين ولنماذج المختلفة:

وبعد تحويل درجات المفحوصين إلى درجات مئوية وحساب المتوسطات الحسابية والانحرافات المعيارية لهذه الدرجات، يتبين وجود فروق ظاهرية بين متوسطات الدرجات المئوية، وللكشف عما اذا كانت هناك فروق دالة إحصائياً أجري تحليل التباين الأحادي لدرجات المفحوصين على نماذج الاختبار المكونة بطرق تصفية المموهات، وذلك لمعرفة تأثير الطريقة التي تم بها تصفية المموهات على درجات المفحوصين، وكذلك لضمان إجراء خطوات المعادلة.

والجدول (10) يبين عرضاً لنتائج تحليل التباين الأحادي للكشف عن الفروق بين المتوسطات الحسابية لدرجات المفحوصين على نماذج تصفية المموهات.

الجدول(8) نتائج الإحصائي (M) لفحص الاختلاف في معاملات الثبات لنماذج تصفية المموهات

نموذج تصفية المموهات	KR-20	حجم العينة	عدد الفقرات	قيمة الإحصائي M
العشوائي	0.87	220	26	**13.47**
التجريبي	0.89	220	28	[7.81]
التحكيم	0.83	220	26	
المنطقي	0.88	200	26	

الجدول (9) المتوسطات الحسابية والانحرافات المعيارية للدرجات المئوية للمفحوصين

علامة الطلبة المئوية على نموذج تصفية المموهات	العدد	المتوسط الحسابي	الانحراف المعياري
العشوائي	220	%58.374	22.97
التجريبي	220	%65.455	23.69
التحكيم	220	%60.000	20.69
المنطقي	220	%59.580	23.73

الجدول (10) تحليل التباين الأحادي للمتوسطات الحسابية لدرجات المفحوصين على نماذج تصفية المموهات

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط مجموع المربعات	قيمة ف المحسوبة	الدلالة الإحصائية
بين نماذج تصفية المموهات	6526.543	3	2175.514	**4.184**	0.006
داخل الأفراد	455489.819	866.333*	519.966		
الكلي	462016.362	879			
	* وفقاً لطريقة Brown-Forsythe				

ويتضح من خلال الجدول (10) وجود فروق ذات دلالة إحصائية بين المتوسطات الحسابية لدرجات المفحوصين لنماذج الاختبار المختلفة المكونة بطرق تصفية المموهات عند مستوى الدلالة ($\alpha = 0.05$) تُعزى لطريقة تصفية المموهات، وهذا يُظهر أن الطريقة التي تم بها تصفية المموهات قد أثرت بشكل جوهري على درجات المفحوصين، وهذا يُؤكد على ضرورة معادلة درجات المفحوصين على نماذج الاختبار التي تم تصفية مموهاتها بطرق مختلفة.

ويحتوي الجدول (11) على المتوسطات الحسابية والانحرافات المعيارية لمعاملات الصعوبة لنماذج تصفية المموهات.

وقد تم تحليل البيانات في الجدول (11) باستخدام تحليل التباين الأحادي للكشف عن مدى الفروق في المتوسطات الحسابية لمعامل الصعوبة للنماذج المختلفة. والجدول (12) يبين عرضاً لنتائج تحليل التباين الأحادي للكشف عن الفروق بين المتوسطات الحسابية لمعامل الصعوبة بين النماذج المختلفة.

يلاحظ من خلال الجدول (12) أنه لا يوجد فروق ذات دلالة إحصائية بين متوسطات معاملات الصعوبة للنماذج المختلفة عند مستوى الدلالة ($\alpha = 0.05$) تُعزى لطريقة تصفية المموهات، ومرد ذلك أن عدد الفقرات الداخلة في تحليل متوسطات معاملات الصعوبة للفقرات كانت قليلة فلم تُظهر اختلاف واضح على مستوى صعوبة الفقرات بين النماذج، إذا ما قورنت بعدد عينة الأفراد الداخلة في تحليل متوسطات الدرجات الملاحظة للأفراد والتي أظهرت اختلافاً واضحاً بين متوسطات الدرجات الملاحظة للأفراد. وقد اتفقت هذه النتيجة مع ما توصل إليه كل من سيزيك وأودي (Cizek & O'day, 1994) ودراسة ظاظا (2000) ودراسة عيلبوني (2007) والتي أشار كل منهما إلى عدم وجود فروق دالة إحصائياً بين متوسطات معاملات الصعوبة تُعزى لطريقة تصفية المموهات. وتعارضت هذه النتيجة مع نتيجة أبو جراد (1997) والتي أظهرت وجود فروق بين متوسطات الصعوبة تُعزى لطريقة تصفية المموهات.

الجدول(13) يتضمن المتوسطات الحسابية والانحرافات المعيارية لمعاملات التمييز لنماذج تصفية المموهات.

كما تم استخدام أسلوب تحليل التباين الأحادي للكشف عن مدى الفروق بين المتوسطات الحسابية لمعاملات التمييز، للبيانات في الجدول (13)، والجدول (14) يبين نتائج تحليل التباين الأحادي للتمييز لجميع نماذج تصفية المموهات.

الجدول (11) المتوسطات الحسابية والانحرافات المعيارية لمعاملات الصعوبة

نموذج تصفية المموهات	عدد الفقرات	متوسط معامل الصعوبة	الانحراف المعياري
العشوائي	26	0.584	0.13
التجريبي	28	0.655	0.11
التحكيم	26	0.600	0.11
المنطقي	26	0.596	0.13

الجدول (12) تحليل التباين الأحادي للمتوسطات الحسابية لمعاملات الصعوبة لنماذج تصفية المموهات

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط مجموع المربعات	قيمة ف المحسوبة	الدلالة الإحصائية
بين نماذج تصفية المموهات	0.081	3	0.027	1.914	0.132
داخل الأفراد	1.444	102	0.014		
الكلي	1.525	105			

الجدول (13) المتوسطات الحسابية والانحرافات المعيارية لمعاملات التمييز لنماذج تصفية المموهات

نموذج تصفية المموهات	عدد الفقرات	متوسط معامل التميز	الانحراف المعياري
العشوائي	26	0.417	0.08
التجريبي	28	0.456	0.08
التحكيم	26	0.358	0.07
المنطقي	26	0.437	0.10

الجدول (14) تحليل التباين الأحادي للمتوسطات الحسابية لمعاملات التمييز لنماذج تصفية المموهات

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط مجموع المربعات	قيمة ف المحسوبة	الدلالة الإحصائية
بين نماذج تصفية المموهات	0.143	3	0.048	7.129	0.000
داخل الأفراد	0.683	102	0.007		
الكلي	0.826	105			

يتضح من خلال الجدول (14) وجود فروق ذات دلالة إحصائية بين المتوسطات الحسابية لمعاملات التمييز للنماذج المختلفة عند مستوى الدلالة ($\alpha = 0.05$) تُعزى لطريقة تصفية المموهات، وهذه النتيجة تتفق مع نتيجة ظاظا (2000)، والتي أظهرت وجود فروق دالة إحصائياً بين متوسطات معاملات التمييز تُعزى لطريقة تصفية المموهات، بينما تتعارض هذه النتيجة مع ما جاء في كل من سيزيك وأودي (Cizek & O'day, 1994)، وأبو جراد (1997) وعيلبوني (2007)، حيث لم تُظهر نتائجهم أية فروق دالة إحصائياً بين متوسطات معاملات التمييز تُعزى لطريقة تصفية المموهات، وهذا الفرق الدال إحصائياً لمتوسطات معاملات التمييز في هذه الدراسة مؤشر إضافي على ضرورة معادلة درجات المفحوصين على نماذج الاختبار التي تم تصفية مموهاتها بطرق مختلفة.

ومن خلال الاطلاع على الإحصائيات المتحققة حسب النظرية الكلاسيكية في القياس، يتضح أن الطريقة التي تم بها تصفية المموهات وربما أثرت بشكل واضح على درجات المفحوصين، وعلى بعض خصائص الفقرات لنماذج الاختبار المكونة بطرق تصفية المموهات، مما قد يؤثر في عدالة هذه النماذج من الاختبارات المختلفة بطريقة تصفية مموهاتها، وخاصة أنها تعد ضمن صور مختلفة للاختبار نفسه. وهذا يقودنا إلى التأكيد على ضرورة إجراء المعادلة لدرجات نماذج الاختبار المختلفة بطريقة تصفية مموهاتها.

ونظراً لما تم التوصل إليه من نتائج في تأثير طريقة تصفية المموهات على درجات الأفراد وبعض خصائص الفقرات والاختبار من خلال نتائج تحليل التباين الأحادي التي أظهرت وجود فروق دالة إحصائياً بين المتوسطات الحسابية لدرجات الأفراد وبعض خصائص الفقرات، كان من الضروري إجراء المعادلة لدرجات الأفراد على نماذج تصفية المموهات لتعديل هذه الفروق بين الدرجات على النماذج المختلفة، وقد استخدمت الدراسة طريقة المعادلة وفقا النموذج أحادي المعلم (راش)، وبالتالي كان لا بد من من اجراء مطابقة للفقرات والافراد وفق نموذج راش لإتمام عملية اجراء المعادلة.

ثالثاً: النتائج المتعلقة بالسؤال الفرعي الثالث (ما مؤشرات مطابقة الفقرات والافراد لنماذج الاختبار المكونة بطرق تصفية المموهات وفق نموذج راش؟) تم فحص مطابقة الفقرات والأفراد باستخدام برنامج (Bilog-Mg) وتم أيضاً نقل البيانات إلى ملف من نوع (Text Document) لكل نموذج من نماذج تصفية المموهات بشكل مستقل ومطابقة البيانات للنموذج أحادي المعلم، ومن ثم تقدير معالم الفقرات المتمثلة بالصعوبة، وتقديرات القدرة للأفراد ولكل نموذج على حدا، باستخدام تقديرات الأرجحية العظمى (Maximum Likelihood Estimation)، كما تم استخراج قيم مربع كاي وما يقابلها من قيم احتمالية لتحديد الفقرات والأفراد المطابقين للنموذج حيث مثلت القيم الاحتمالية التي تقل عن (0.01) الفقرات والأفراد غير المطابقين للنموذج، والجدول (15) يبين أرقام وأعداد الفقرات والأفراد غير المطابقين للنموذج أحادي المعلم ولكل نموذج من نماذج تصفية المموهات:

الجدول (15) أرقام وأعداد الفقرات والأفراد غير المطابقين للنموذج أحادي المعلم للنماذج المختلفة

نموذج تصفية المموهات	المؤشر	العدد	الأرقام
العشوائي	عدم مطابقة الأفراد	2	105,180
	عدم مطابقة الفقرات	5	13، 15، 24، 29، 30
التجريبي	عدم مطابقة الأفراد	10	122,102,96,86,76,39,34,22,20,10
	عدم مطابقة الفقرات	3	3، 24، 30
التحكيم	عدم مطابقة الأفراد	4	16, 153, 200, 217
	عدم مطابقة الفقرات	3	7، 22، 30
المنطقي	عدم مطابقة الأفراد	12	218,197,191,155,131,127,124,119,118,113,78,22
	عدم مطابقة الفقرات	11	12, 13, 14, 21, 24, 25, 26, 28, 30, 31, 32

ويلاحظ من خلال الجدول (15) أن نموذج التصفية المنطقية كانت بياناته الأقل مطابقة لنموذج أحادي المعلم إذ يلاحظ عدم مطابقة (11) فقرة و(12) فرد، ثم يليه النموذج العشوائي بعدم مطابقة (5) فقرات وفردين، وأما النموذج التجريبي يظهر عدم مطابقة (3) فقرات و(10) أفراد للنموذج، والأفضل مطابقة كان لنموذج التصفية بالتحكيم حيث أظهرت بياناته عدم مطابقة (3) فقرات و(4) أفراد للنموذج.

ويوضح الجدول (16) عدد الفقرات والأفراد المطابقة للنموذج أحادي المعلم (راش) لكل نموذج من نماذج تصفية المموهات.

بعد استبعاد الفقرات والأفراد غير المطابقين لنموذج أحادي المعلم تم استخراج تقديرات معلمة الصعوبة والخطأ المعياري لتقديرات معلمة الصعوبة ولكل نموذج على حدا، والجدول (17) يبين تقديرات معلمة الصعوبة والخطأ المعياري لتقديرات معلمة الصعوبة بعد حذف الفقرات غير المطابقة للنموذج (حيث تمثل الخلايا الفارغة الفقرات المحذوفة).

الجدول (16) عدد الفقرات والأفراد المطابقة لنموذج أحادي المعلم للنماذج المختلفة

نموذج تصفية المموهات	المؤشر	العدد
العشوائي	مطابقة الأفراد	218
	مطابقة الفقرات	29
التجريبي	مطابقة الأفراد	210
	مطابقة الفقرات	31
التحكيم	مطابقة الأفراد	216
	مطابقة الفقرات	31
المنطقي	مطابقة الأفراد	208
	مطابقة الفقرات	23

الجدول (17) تقديرات معلمة الصعوبة والخطأ المعياري للتقديرات لنماذج تصفية المموهات

رقم الفقرة	نموذج تصفية المموهات							
	عشوائي		تجريبي		تحكيم		منطقي	
	معلمة الصعوبة	الخطأ المعياري لمعلمة الصعوبة	معلمة الصعوبة	الخطأ المعياري لمعلمة الصعوبة	معلمة الصعوبة	الخطأ المعياري لمعلمة الصعوبة	معلمة الصعوبة	الخطأ المعياري لمعلمة الصعوبة
1	-0.23	0.16	-0.08	0.13	-0.48	0.17	-0.46	0.13
2	-0.87	0.17	-0.50	0.13	-1.20	0.21	-1.34	0.15
3	0.96	0.16			1.41	0.16	0.86	0.12
4	1.03	0.17	1.07	0.12	1.13	0.16	1.42	0.13
5	-0.27	0.16	0.09	0.12	-0.16	0.16	-0.11	0.12
6	-0.21	0.16	-0.28	0.13	-0.23	0.17	-0.30	0.13
7	-0.43	0.16	-0.37	0.13			-0.71	0.13
8	-0.36	0.16	-0.44	0.13	-0.74	0.18	-0.06	0.12
9	-0.46	0.16	-0.74	0.14	0.37	0.16	-0.77	0.14
10	0.94	0.16	1.12	0.12	0.62	0.15	-0.04	0.12
11	-0.30	0.16	0.01	0.12	-0.03	0.16	-0.24	0.13
12	-1.14	0.18	-1.34	0.16	-1.41	0.22		
13			-0.70	0.14	-0.48	0.17		
14	0.67	0.16	0.82	0.12	0.91	0.15		
15			0.41	0.12	1.29	0.16	0.86	0.12
16	-1.32	0.19	-1.21	0.15	-1.24	0.21	-1.19	0.15
17	0.67	0.16	0.90	0.12	0.33	0.16	0.34	0.12
18	0.42	0.16	0.44	0.12	0.89	0.15	0.11	0.12

رقم الفقرة	نموذج تصفية المموهات							
	عشوائي		تجريبي		تحكيم		منطقي	
	معلمة الصعوبة	الخطأ المعياري لمعلمة الصعوبة	معلمة الصعوبة	الخطأ المعياري لمعلمة الصعوبة	معلمة الصعوبة	الخطأ المعياري لمعلمة الصعوبة	معلمة الصعوبة	الخطأ المعياري لمعلمة الصعوبة
19	1.73	0.19	0.57	0.12	0.75	0.15	2.20	0.14
20	0.12	0.16	0.80	0.12	0.82	0.15	0.21	0.12
21	-0.84	0.17	-0.67	0.14	-0.91	0.19		
22	0.60	0.16	0.12	0.12			-0.48	0.13
23	-0.50	0.16	-0.60	0.14	0.07	0.16	-0.30	0.13
24					-0.05	0.16		
25	0.10	0.16	0.17	0.12	0.14	0.16		
26	0.99	0.17	0.70	0.12	0.73	0.15		
27	0.56	0.16	1.82	0.12	0.95	0.15	0.97	0.13
28	-0.16	0.16	-0.10	0.13	-0.94	0.19		
29			-0.34	0.13	0.33	0.16	-0.40	0.13
30								
31	0.34	0.16	0.15	0.12	-0.23	0.17		
32	-0.84	0.17	-1.16	0.15	-1.20	0.21		
33	-1.58	0.20	-1.16	0.15	-1.70	0.24	-1.46	0.16
34	0.40	0.16	0.49	0.12	0.23	0.16	0.89	0.12

يتضح من الجدول (17) أن قيم معلمة الصعوبة لنموذج العشوائي تراوحت بين (1.73 – 1.58−) ولنموذج التجريبي تراوحت بين (1.82 – 1.34−) ولنموذج التصفية بالتحكيم تراوحت بين (1.41−1.70−) ولنموذج المنطقي تراوحت بين (2.20 – 1.46−) وتعتبر هذه القيم لمعلمة الصعوبة قيم مقبولة عملياً حسب ما ذكره بيكر (Baker, 2001). كما ويلاحظ أن قيم الخطأ المعياري لتقديرات معلمة الصعوبة كانت متدنية ومقبولة. كما تم استخراج تقديرات القدرات للأفراد والخطأ المعياري لتقديرات قدرات الأفراد لكل نموذج من نماذج تصفية المموهات على حدا، والجدول الآتي يبين تقديرات القدرات للأفراد والخطأ المعياري في تقدير القدرة ونسبة الإجابة الصحيحة وتكرارها.

ويتضح من خلال الجدول (18) أن نموذج التصفية المنطقية كان الأقل انتشاراً على متصل القدرة ومرد ذلك أن بيانات نموذج التصفية المنطقية كانت الأقل مطابقة لنموذج أحادي المعلم، إلا أنه يمكن القول أن الأفراد يتوزعون بشكل مناسب على متصل السمة ولجميع النماذج. ويوضح الجدول (19) المتوسطات الحسابية والانحرافات المعيارية لقدرات لأفراد المقدرة.

تم تحليل البيانات السابقة باستخدام تحليل التباين الأحادي للكشف عن مدى الفروق في المتوسطات الحسابية لتقديرات قدرات الأفراد بين نماذج تصفية المموهات، لضمان إجراء خطوات المعادلة ضمن النظرية الحديثة في القياس، والجدول (20) يبين عرضاً تفصيلياً لنتائج تحليل التباين الأحادي للكشف عن الفروق بين المتوسطات الحسابية السابقة لنماذج تصفية المموهات.

ومن خلال النظر إلى الجدول (20) يتبين وجود فروق ذات دلالة إحصائية بين المتوسطات الحسابية لتقديرات قدرات الأفراد عند مستوى الدلالة ($\alpha = 0.05$) تُعزى لطريقة تصفية المموهات، وهذا يُؤكد على ضرورة معادلة قدرات الأفراد على نماذج تصفية المموهات.

التحقق من الافتراضات الأساسية لنموذج أحادي المعلم:

أولاً: التحقق من افتراض أحادية البعد (Unidimensionality)

للكشف عن أحادية البعد تم استخدام التحليل العاملي من خلال دراسة المؤشرات التي اعتمدت على التحليل العاملي للمكونات الأساسية (Principal Components Analysis) وذلك باستخدام برنامج (SPSS)، ولكل نموذج على حدا، والجدول (21) يبين النتائج التي تم الوصول إليها من إجراء التحليل العاملي على نماذج تصفية المموهات.

الجدول (18): تقدير قدرات الأفراد والأخطاء المعيارية في قياسها لنماذج تصفية المموهات

نموذج تصفية المموهات															
العشوائي				التجريبي				التحكيم				المنطقي			
نسبة الإجابة الصحيحة	التكرار	القدرة	الخطأ المعياري في تقدير القدرة	نسبة الإجابة الصحيحة	التكرار	القدرة	الخطأ المعياري في تقدير القدرة	نسبة الإجابة الصحيحة	التكرار	القدرة	الخطأ المعياري في تقدير القدرة	نسبة الإجابة الصحيحة	التكرار	القدرة	الخطأ المعياري في تقدير القدرة
10.34	1	-2.39	0.63	12.90	3	-2.110	0.552	12.90	1	-2.168	0.559	13.04	1	-2.145	0.640
13.79	2	-2.04	0.56	16.13	1	-1.832	0.506	19.35	2	-1.636	0.481	21.74	2	-1.476	0.531
17.24	2	-1.76	0.51	19.35	3	-1.594	0.473	22.58	3	-1.416	0.457	26.09	5	-1.211	0.502
20.69	7	-1.51	0.48	22.58	3	-1.381	0.449	25.81	1	-1.216	0.439	30.43	9	-0.970	0.481
24.14	4	-1.29	0.46	25.81	3	-1.188	0.431	29.03	3	-1.030	0.425	34.78	10	-0.746	0.467
27.59	11	-1.09	0.44	29.03	3	-1.009	0.417	32.26	4	-0.855	0.414	39.13	14	-0.532	0.458
31.03	6	-0.91	0.43	32.26	6	-0.840	0.406	35.48	6	-0.687	0.405	43.48	16	-0.326	0.452
34.48	10	-0.73	0.42	35.48	2	-0.679	0.397	38.71	9	-0.526	0.398	47.83	17	-0.123	0.450
37.93	10	-0.56	0.41	38.71	5	-0.524	0.391	41.94	7	-0.369	0.394	52.17	16	0.080	0.451
41.38	13	-0.40	0.40	41.94	3	-0.373	0.387	45.16	4	-0.216	0.390	56.52	14	0.285	0.455
44.83	10	-0.24	0.40	45.16	9	-0.224	0.384	48.39	5	-0.064	0.389	60.87	19	0.494	0.462
48.28	17	-0.08	0.40	48.39	8	-0.078	0.383	51.61	8	0.087	0.388	65.22	7	0.713	0.473
51.72	12	0.08	0.40	51.61	11	0.069	0.383	54.84	18	0.238	0.389	69.57	4	0.944	0.489
55.17	14	0.24	0.40	54.84	8	0.216	0.384	58.06	12	0.390	0.392	73.91	13	1.194	0.511
58.62	13	0.40	0.40	58.06	9	0.364	0.387	61.29	18	0.545	0.396	78.26	5	1.471	0.542
62.07	12	0.56	0.41	61.29	5	0.516	0.392	64.52	8	0.704	0.402	82.61	9	1.788	0.587
65.52	13	0.73	0.41	64.52	15	0.672	0.398	67.74	13	0.868	0.410	86.96	16	2.170	0.655
68.97	9	0.91	0.43	67.74	7	0.834	0.407	70.97	15	1.040	0.420	91.30	20	2.672	0.773
72.41	10	1.09	0.44	70.97	11	1.004	0.418	74.19	15	1.222	0.434	95.65	11	3.465	1.049
75.86	15	1.29	0.46	74.19	10	1.184	0.432	77.42	4	1.418	0.452				
79.31	9	1.51	0.48	77.42	6	1.379	0.451	80.65	12	1.632	0.475				
82.76	5	1.76	0.51	80.65	10	1.593	0.475	83.87	14	1.873	0.507				
86.21	6	2.04	0.56	83.87	14	1.834	0.508	87.10	12	2.153	0.553				
89.66	1	2.39	0.63	87.10	13	2.115	0.555	90.32	12	2.496	0.623				
93.10	2	2.85	0.75	90.32	14	2.460	0.625	93.55	4	2.955	0.744				
96.55	4	3.61	1.03	93.55	16	2.923	0.746	96.77	6	3.700	1.025				
				96.77	12	3.672	1.028								

الجدول (19) المتوسطات الحسابية والانحرافات المعيارية لقدرات لأفراد المقدرة

القدرة	العدد	المتوسط الحسابي	الانحراف المعياري
العشوائي	218	0.258	1.07
التجريبي	210	1.016	1.37
التحكيم	216	0.841	1.13
المنطقي	208	0.705	1.31

الجدول (20) تحليل التباين الأحادي للمتوسطات الحسابية لتقديرات قدرات الأفراد على نماذج تصفية المموهات

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط مجموع المربعات	قيمة ف المحسوبة	الدلالة الإحصائية
بين نماذج تصفية المموهات	67.778	3	22.593	14.983	0.000
داخل الأفراد	1272.531	808.002*	1.501		
الكلي	1340.308	851			
	*وفقاً لطريقة (Brown-Forthyse)				

الجدول (21) العوامل الثلاثة الأولى وجذورها الكامنة ونسب تباينها ومؤشرات الحكم على أحادية البعد للنماذج المختلفة

نموذج تصفية المموهات	رقم العامل	الجذور الكامنة الاستهلالية			الجذر الكامن الأول	التباين المفسر التراكمي بعد استخلاص مجموع مربعات التشبعات
		الكلي	التباين المفسر%	التباين المفسر التراكمي%	الجذر الكامن الثاني	
العشوائي	1	5.394	18.599	18.599	2.829	60.510
	2	1.907	6.575	25.174		
	3	1.557	5.369	30.543		
التجريبي	1	8.070	26.032	26.032	4.859	56.924
	2	1.661	5.358	31.390		
	3	1.607	5.184	36.574		
التحكيم	1	5.885	18.984	18.984	3.384	56.164
	2	1.739	5.609	24.593		
	3	1.698	5.477	30.070		
المنطقي	1	5.344	23.233	23.233	3.592	55.852
	2	1.487	6.467	29.700		
	3	1.378	5.990	35.690		

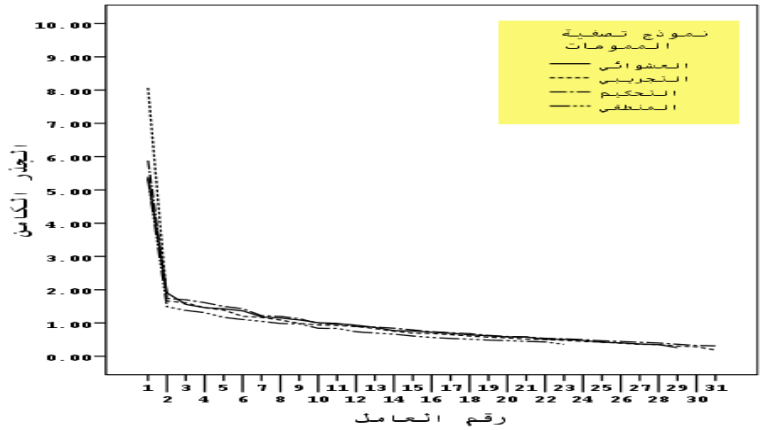
يتضح من خلال الجدول (21) التأكيد على أحادية البعد ولكل نموذج من نماذج تصفية المموهات من خلال:

أ. نسبة ما يفسره العامل الأول من التباين الكلي أكبر من (20%) بالنسبة لنموذج التصفية التجريبية ونموذج التصفية المنطقية، أما بالنسبة لنموذج التصفية العشوائية والتصفية بالتحكيم كان يفسر العامل الأول ما نسبته (18%) تقريباً وهي نسبة عالية وقريبة من (20%).

ب. نسبة ما يفسره العامل الأول إلى العامل الثاني كانت أكبر من (2) ولجميع النماذج.

ج. التباين المفسر التراكمي بعد استخلاص مجموع مربعات التشبعات كان أكبر من (54) ولجميع النماذج.

وتم رسم الجذور الكامنة بيانيا من خلال ما يعرف ب (Scree Plot) حيث يمثل المحور الأفقي العوامل ويمثل المحور العمودي مقدار الجذر الكامن والذي أكد أحادية البعد من خلال تَميُز العامل الأول عن باقي العوامل ولجميع النماذج وهذا ما يؤكده الشكل البياني (1).



الشكل (1): رسم بياني يمثل الجذور الكامنة للعوامل

ثانياً: الاستقلال الموضعي (Local Independence)

اعتبر الباحث شرط الاستقلال الموضعي متحقق من خلال التحقق من أحادية البعد وهو ما أكده كل من هامبلتون وسواميناثان (Hambleton & Swaminathan, 1985)، إذ تم اعتبار أن افتراض الاستقلال الموضعي يكافئ افتراض أحادية البعد.

رابعاً: النتائج المتعلقة بالسؤال الفرعي الرابع (هل تختلف دقة وفاعلية المعادلة لنماذج الاختبار (من خلال قيم محك الخطأ المعياري للمعادلة ومحك الصدق التقاطعي) باختلاف طريقة تصفية المموهات؟)

تمت معادلة تقديرات قدرات الأفراد، من خلال إيجاد قيم القدرة المتعادلة لنماذج تصفية المموهات عند الرتب المئينية وعند كل مستوى من مستويات القدرة. كما تم تقدير دالة المعلومات لجميع النماذج عند كل مستوى من مستويات القدرة باستخدام المعادلة $I(\Theta) = \frac{1}{SE(\Theta)^2}$، وذلك لإيجاد الكفاءة النسبية لكل نموذج من نماذج تصفية المموهات باستخدام المعادلة $RE(\Theta) = \frac{I_A(\Theta)}{I_B(\Theta)}$، لمقارنة دوال المعلومات بين النماذج المختلفة، (على اعتبار أن النموذج العشوائي هو النموذج الأساس الذي تم معادلة الدرجات والمقارنة لدوال المعلومات على أساسه بالنسبة للنماذج المختلفة وذلك لاعتبارات عدة منها شيوع استخدام هذه الطريقة لسهولتها) ومن ثم التعرف إلى حجم المعلومات التي يزود بها كل نموذج بالمقارنة بالنموذج العشوائي. ويبين الجدول (22) تقديرات القدرة المتعادلة عند الرتب المئينية وما يناظرها من قيم لدالة المعلومات لنماذج تصفية المموهات، والكفاءة النسبية للنماذج المتعادلة.

يتضح من خلال الجدول (22) أنه يمكن تعيين تقديرات القدرة المتعادلة على النماذج المختلفة تبعاً للمئين المقابل لها، فمثلا المئين (45) يقابل قدرة مقدارها (0.08) لوجيت على النموذج العشوائي، وقدرة مقدارها (0.67) على النموذج التجريبي، وقدرة مقدارها (0.54) لوجيت على نموذج التصفية بالتحكيم، وقدرة مقدارها (0.28) لوجيت على نموذج التصفية المنطقية، وعليه تكون تقديرات القدرة المقابلة لهذا المئين من خلال النماذج متعادلة.

ويلاحظ من خلال قيم الكفاءة النسبية للمقارنة بين دوال المعلومات لنموذج التصفية التجريبية ونموذج التصفية العشوائية يلاحظ أن نموذج التصفية التجريبية يقدم معلومات أكثر من نموذج التصفية العشوائية عند مستويات القدرة المنخفضة، وفي متوسط القدرات كان النموذجان يقدمان نفس المعلومات تقريباً، وأما عند مستويات القدرة المرتفعة كان النموذج العشوائي يقدم معلومات أكثر من نموذج التصفية التجريبية، وهذه نتيجة منطقية؛ لأن متوسط قدرات الأفراد لنموذج التصفية التجريبية مرتفع إذ بلغ (1.02) مقارنة بالعشوائي إذ بلغ (0.26)، هذا يعني أن نموذج التصفية التجريبية يناسب الأفراد ذوي القدرات المنخفضة أكثر من نموذج التصفية العشوائية، أما بمقارنة دوال المعلومات لنموذج التصفية بالتحكيم مقارنة بنموذج التصفية العشوائية، يلاحظ كذلك أن نموذج التصفية بالتحكيم يقدم معلومات أكثر من نموذج التصفية العشوائية عند مستويات القدرة المنخفضة، وعند المستويات المرتفعة للقدرة فقد قدم نموذج التصفية العشوائية معلومات أكثر من نموذج التحكيم، وفي متوسط متصل السمة يقدم النموذجان نفس المعلومات تقريباً، وهذا يعني أن نموذج التصفية بالتحكيم يناسب مستويات القدرة المنخفضة أكثر من نموذج التصفية العشوائية، أما عند المقارنة بين نموذج التصفية المنطقية ونموذج التصفية العشوائية نلاحظ أن نموذج التصفية العشوائية قد قدم معلومات أكثر من نموذج التصفية المنطقية عند جميع مستويات القدرة تقريباً.

وبشكل عام يلاحظ أن نموذج التصفية العشوائية يقدم معلومات عند الأفراد ذوي القدرات المرتفعة أكثر من غيره من النماذج، أي أنه يُناسب الأفراد ذوي القدرات المرتفعة أكثر من باقي النماذج، ويمكن أن يُرد ذلك إلى أن نموذج التصفية العشوائية لم يُتخذ في حقه أي إجراء عملي لتصفية المموهات إذ كانت آلية التصفية بشكل عشوائي مما أُبقي على عدد من المموهات كانت تسبب تشتت في الإجابة عند الطلبة مما زاد من صعوبة الاختبار، أما باقي النماذج أُتخذت في حقها بعض الإجراءات لتصفية المموهات فتم إزالة معظم المموهات التي كانت تسبب تشتت في إجابة الطالب الأمر الذي زاد من سهولة الاختبار وبالتالي قلة قدرة الاختبار على التمييز بين الأفراد ذوي القدرة المرتفعة بشكل جيد.

الجدول (22) تقديرات القدرة المتعادلة لنماذج تصفية المموهات عند الرتب المئينية وما يناظرها من قيم لدالة المعلومات والكفاءة النسبية

المئين	نموذج تصفية المموهات															
	العشوائي			التجريبي				التحكيم				المنطقي				
	قدرة الأفراد	الخطأ المعياري في تقدير القدرة	دالة المعلومات	قدرة الأفراد المعادلة	الخطأ المعياري	دالة المعلومات	الكفاءة النسبية	قدرة الأفراد المعادلة	الخطأ المعياري	دالة المعلومات	الكفاءة النسبية	قدرة الأفراد المعادلة	الخطأ المعياري	دالة المعلومات	الكفاءة النسبية
1	-2.04	0.56	3.22	-2.11	0.55	3.29	1.02	-1.64	0.48	4.33	1.34	-1.48	0.53	3.55	1.10
2	-1.76	0.51	3.81	-1.78	0.51	3.91	1.03	-1.42	0.46	4.79	1.26	-1.21	0.50	3.97	1.04
3	-1.51	0.48	4.34	-1.59	0.47	4.47	1.03	-1.31	0.46	4.79	1.10	-1.21	0.50	3.97	0.92
4	-1.51	0.48	4.34	-1.38	0.47	4.47	1.03	-1.03	0.42	5.55	1.28	-1.12	0.50	3.97	0.92
5	-1.51	0.48	4.34	-1.28	0.43	5.39	1.24	-0.88	0.41	5.85	1.35	-0.97	0.48	4.32	1.00
6	-1.29	0.46	4.80	-1.19	0.43	5.39	1.12	-0.85	0.41	5.85	1.22	-0.97	0.48	4.32	0.90
7	-1.29	0.46	4.80	-1.01	0.42	5.76	1.20	-0.69	0.40	6.10	1.27	-0.97	0.48	4.32	0.90
8	-1.09	0.44	5.19	-0.86	0.41	6.07	1.17	-0.69	0.40	6.10	1.17	-0.97	0.48	4.32	0.83
9	-1.09	0.44	5.19	-0.84	0.41	6.07	1.17	-0.69	0.40	6.10	1.17	-0.75	0.47	4.58	0.88
10	-1.09	0.44	5.19	-0.84	0.41	6.07	1.17	-0.53	0.40	6.30	1.21	-0.75	0.47	4.58	0.88
11	-1.09	0.44	5.19	-0.68	0.40	6.33	1.22	-0.53	0.40	6.30	1.21	-0.75	0.47	4.58	0.88
12	-1.09	0.44	5.19	-0.52	0.39	6.54	1.26	-0.53	0.40	6.30	1.21	-0.75	0.47	4.58	0.88
13	-0.91	0.43	5.53	-0.52	0.39	6.54	1.18	-0.53	0.40	6.30	1.14	-0.71	0.47	4.58	0.83
14	-0.91	0.43	5.53	-0.44	0.39	6.69	1.21	-0.37	0.39	6.45	1.17	-0.53	0.46	4.78	0.86
15	-0.91	0.43	5.53	-0.37	0.39	6.69	1.21	-0.37	0.39	6.45	1.17	-0.53	0.46	4.78	0.86
16	-0.73	0.42	5.80	-0.22	0.38	6.79	1.17	-0.37	0.39	6.45	1.11	-0.53	0.46	4.78	0.82
17	-0.73	0.42	5.80	-0.22	0.38	6.79	1.17	-0.23	0.39	6.56	1.13	-0.53	0.46	4.78	0.82
18	-0.73	0.42	5.80	-0.22	0.38	6.79	1.17	-0.22	0.39	6.56	1.13	-0.53	0.46	4.78	0.82
19	-0.73	0.42	5.80	-0.22	0.38	6.79	1.17	-0.06	0.39	6.62	1.14	-0.53	0.46	4.78	0.82
20	-0.59	0.41	6.02	-0.08	0.38	6.83	1.13	-0.06	0.39	6.62	1.10	-0.37	0.45	4.90	0.81
21	-0.56	0.41	6.02	-0.08	0.38	6.83	1.13	0.02	0.39	6.63	1.10	-0.33	0.45	4.90	0.81
22	-0.56	0.41	6.02	-0.08	0.38	6.83	1.13	0.09	0.39	6.63	1.10	-0.33	0.45	4.90	0.81
23	-0.56	0.41	6.02	-0.08	0.38	6.83	1.13	0.09	0.39	6.63	1.10	-0.33	0.45	4.90	0.81
24	-0.56	0.41	6.02	0.07	0.38	6.83	1.13	0.09	0.39	6.63	1.10	-0.33	0.45	4.90	0.81
25	-0.40	0.40	6.19	0.07	0.38	6.83	1.10	0.24	0.39	6.60	1.07	-0.33	0.45	4.90	0.79
26	-0.40	0.40	6.19	0.07	0.38	6.83	1.10	0.24	0.39	6.60	1.07	-0.33	0.45	4.90	0.79
27	-0.40	0.40	6.19	0.07	0.38	6.83	1.10	0.24	0.39	6.60	1.07	-0.33	0.45	4.90	0.79
28	-0.40	0.40	6.19	0.07	0.38	6.83	1.10	0.24	0.39	6.60	1.07	-0.12	0.45	4.94	0.80
29	-0.40	0.40	6.19	0.22	0.38	6.78	1.10	0.24	0.39	6.60	1.07	-0.12	0.45	4.94	0.80
30	-0.40	0.40	6.19	0.22	0.38	6.78	1.10	0.24	0.39	6.60	1.07	-0.12	0.45	4.94	0.80
31	-0.24	0.40	6.30	0.22	0.38	6.78	1.08	0.24	0.39	6.60	1.05	-0.12	0.45	4.94	0.79
32	-0.24	0.40	6.30	0.22	0.38	6.78	1.08	0.24	0.39	6.60	1.05	-0.12	0.45	4.94	0.79
33	-0.24	0.40	6.30	0.36	0.39	6.67	1.06	0.33	0.39	6.52	1.03	-0.12	0.45	4.94	0.79
34	-0.24	0.40	6.30	0.36	0.39	6.67	1.06	0.39	0.39	6.52	1.03	-0.12	0.45	4.94	0.79
35	-0.13	0.40	6.35	0.36	0.39	6.67	1.05	0.39	0.39	6.52	1.03	-0.12	0.45	4.94	0.78
36	-0.08	0.40	6.35	0.36	0.39	6.67	1.05	0.39	0.39	6.52	1.03	0.08	0.45	4.92	0.78
37	-0.08	0.40	6.35	0.52	0.39	6.51	1.03	0.39	0.39	6.52	1.03	0.08	0.45	4.92	0.78
38	-0.08	0.40	6.35	0.52	0.39	6.51	1.03	0.39	0.39	6.52	1.03	0.08	0.45	4.92	0.78
39	-0.08	0.40	6.35	0.56	0.40	6.30	0.99	0.54	0.40	6.38	1.01	0.08	0.45	4.92	0.78
40	-0.08	0.40	6.35	0.67	0.40	6.30	0.99	0.54	0.40	6.38	1.01	0.08	0.45	4.92	0.78

المئين	نموذج تصفية المموهات															
	العشوائي			التجريبي				التحكيم				المنطقي				
	قدرة الأفراد	الخطأ المعياري في تقدير القدرة	دالة المعلومات	قدرة الأفراد المعادلة	الخطأ المعياري	دالة المعلومات	الكفاءة النسبية	قدرة الأفراد المعادلة	الخطأ المعياري	دالة المعلومات	الكفاءة النسبية	قدرة الأفراد المعادلة	الخطأ المعياري	دالة المعلومات	الكفاءة النسبية
41	-0.08	0.40	6.35	0.67	0.40	6.30	0.99	0.54	0.40	6.38	1.01	0.08	0.45	4.92	0.78
42	-0.08	0.40	6.35	0.67	0.40	6.30	0.99	0.54	0.40	6.38	1.01	0.08	0.45	4.92	0.78
43	0.08	0.40	6.35	0.67	0.40	6.30	0.99	0.54	0.40	6.38	1.00	0.08	0.45	4.92	0.78
44	0.08	0.40	6.35	0.67	0.40	6.30	0.99	0.54	0.40	6.38	1.00	0.28	0.45	4.84	0.76
45	0.08	0.40	6.35	0.67	0.40	6.30	0.99	0.54	0.40	6.38	1.00	0.28	0.45	4.84	0.76
46	0.08	0.40	6.35	0.68	0.40	6.30	0.99	0.54	0.40	6.38	1.00	0.28	0.45	4.84	0.76
47	0.08	0.40	6.35	0.83	0.41	6.04	0.95	0.70	0.40	6.20	0.98	0.28	0.45	4.84	0.76
48	0.10	0.43	5.53	0.83	0.41	6.04	1.09	0.70	0.40	6.20	1.12	0.28	0.45	4.84	0.87
49	0.24	0.40	6.30	0.83	0.41	6.04	0.96	0.70	0.40	6.20	0.98	0.28	0.45	4.84	0.77
50	0.24	0.40	6.30	1.00	0.42	5.72	0.91	0.70	0.40	6.20	0.98	0.39	0.45	4.84	0.77
51	0.24	0.40	6.30	1.00	0.42	5.72	0.91	0.87	0.41	5.96	0.95	0.49	0.46	4.68	0.74
52	0.24	0.40	6.30	1.00	0.42	5.72	0.91	0.87	0.41	5.96	0.95	0.49	0.46	4.68	0.74
53	0.24	0.40	6.30	1.00	0.42	5.72	0.91	0.87	0.41	5.96	0.95	0.49	0.46	4.68	0.74
54	0.24	0.40	6.30	1.00	0.42	5.72	0.91	0.87	0.41	5.96	0.95	0.49	0.46	4.68	0.74
55	0.40	0.40	6.19	1.18	0.43	5.35	0.86	0.87	0.41	5.96	0.96	0.49	0.46	4.68	0.76
56	0.40	0.40	6.19	1.18	0.43	5.35	0.86	0.87	0.41	5.96	0.96	0.49	0.46	4.68	0.76
57	0.40	0.40	6.19	1.18	0.43	5.35	0.86	1.04	0.42	5.66	0.92	0.49	0.46	4.68	0.76
58	0.40	0.40	6.19	1.18	0.43	5.35	0.86	1.04	0.42	5.66	0.92	0.49	0.46	4.68	0.76
59	0.40	0.40	6.19	1.18	0.43	5.35	0.86	1.04	0.42	5.66	0.92	0.56	0.46	4.68	0.76
60	0.40	0.40	6.19	1.38	0.45	4.92	0.79	1.04	0.42	5.66	0.92	0.71	0.47	4.46	0.72
61	0.56	0.41	6.03	1.38	0.45	4.92	0.82	1.04	0.42	5.66	0.94	0.71	0.47	4.46	0.74
62	0.56	0.41	6.03	1.38	0.45	4.92	0.82	1.04	0.42	5.66	0.94	0.71	0.47	4.46	0.74
63	0.56	0.41	6.03	1.59	0.48	4.42	0.73	1.04	0.42	5.66	0.94	0.94	0.49	4.18	0.69
64	0.56	0.41	6.03	1.59	0.48	4.42	0.73	1.22	0.43	5.31	0.88	0.94	0.49	4.18	0.69
65	0.56	0.41	6.03	1.59	0.48	4.42	0.73	1.22	0.43	5.31	0.88	1.19	0.51	3.82	0.63
66	0.65	0.41	5.81	1.59	0.48	4.42	0.76	1.22	0.43	5.31	0.91	1.19	0.51	3.82	0.66
67	0.73	0.41	5.81	1.68	0.48	4.42	0.76	1.22	0.43	5.31	0.91	1.19	0.51	3.82	0.66
68	0.73	0.41	5.81	1.83	0.51	3.87	0.67	1.22	0.43	5.31	0.91	1.19	0.51	3.82	0.66
69	0.73	0.41	5.81	1.83	0.51	3.87	0.67	1.22	0.43	5.31	0.91	1.19	0.51	3.82	0.66
70	0.73	0.41	5.81	1.83	0.51	3.87	0.67	1.22	0.43	5.31	0.91	1.19	0.51	3.82	0.66
71	0.73	0.41	5.81	1.83	0.51	3.87	0.67	1.42	0.45	4.90	0.84	1.47	0.54	3.40	0.59
72	0.85	0.43	5.53	1.83	0.51	3.87	0.70	1.47	0.45	4.90	0.89	1.47	0.54	3.40	0.61
73	0.91	0.43	5.53	1.83	0.51	3.87	0.70	1.63	0.48	4.43	0.80	1.65	0.59	2.91	0.53
74	0.91	0.43	5.53	2.12	0.55	3.25	0.59	1.63	0.48	4.43	0.80	1.79	0.59	2.91	0.53
75	0.91	0.43	5.53	2.12	0.55	3.25	0.59	1.63	0.48	4.43	0.80	1.79	0.59	2.91	0.53
76	0.99	0.43	5.53	2.12	0.55	3.25	0.59	1.63	0.48	4.43	0.80	1.79	0.59	2.91	0.53
77	1.09	0.44	5.20	2.12	0.55	3.25	0.63	1.63	0.48	4.43	0.85	1.79	0.59	2.91	0.56
78	1.09	0.44	5.20	2.12	0.55	3.25	0.63	1.87	0.51	3.88	0.75	2.17	0.65	2.33	0.45
79	1.09	0.44	5.20	2.12	0.55	3.25	0.63	1.87	0.51	3.88	0.75	2.17	0.65	2.33	0.45
80	1.09	0.44	5.20	2.39	0.62	2.56	0.49	1.87	0.51	3.88	0.75	2.17	0.65	2.33	0.45
81	1.29	0.44	5.20	2.46	0.62	2.56	0.49	1.87	0.51	3.88	0.75	2.17	0.65	2.33	0.45
82	1.29	0.44	5.20	2.46	0.62	2.56	0.49	1.87	0.51	3.88	0.75	2.17	0.65	2.33	0.45

المئين	نموذج تصفية المموهات															
	العشوائي			التجريبي				التحكيم				المنطقي				
	قدرة الأفراد	الخطأ المعياري في تقدير القدرة	دالة المعلومات	قدرة الأفراد المعادلة	الخطأ المعياري	دالة المعلومات	الكفاءة النسبية	قدرة الأفراد المعادلة	الخطأ المعياري	دالة المعلومات	الكفاءة النسبية	قدرة الأفراد المعادلة	الخطأ المعياري	دالة المعلومات	الكفاءة النسبية
83	1.29	0.44	5.20	2.46	0.62	2.56	0.49	1.87	0.51	3.88	0.75	2.17	0.65	2.33	0.45
84	1.29	0.44	5.20	2.46	0.62	2.56	0.49	1.95	0.51	3.88	0.75	2.17	0.65	2.33	0.45
85	1.29	0.44	5.20	2.46	0.62	2.56	0.49	2.15	0.55	3.27	0.63	2.50	0.77	1.67	0.32
86	1.29	0.44	5.20	2.46	0.62	2.56	0.49	2.15	0.55	3.27	0.63	2.67	0.77	1.67	0.32
87	1.29	0.44	5.20	2.92	0.75	1.80	0.35	2.15	0.55	3.27	0.63	2.67	0.77	1.67	0.32
88	1.51	0.48	4.34	2.92	0.75	1.80	0.41	2.15	0.55	3.27	0.75	2.67	0.77	1.67	0.39
89	1.51	0.48	4.34	2.92	0.75	1.80	0.41	2.15	0.55	3.27	0.75	2.67	0.77	1.67	0.39
90	1.51	0.48	4.34	2.92	0.75	1.80	0.41	2.50	0.62	2.58	0.59	2.67	0.77	1.67	0.39
91	1.51	0.48	4.34	2.92	0.75	1.80	0.41	2.50	0.62	2.58	0.59	2.67	0.77	1.67	0.39
92	1.76	0.51	3.81	2.92	0.75	1.80	0.47	2.50	0.62	2.58	0.68	2.67	0.77	1.67	0.44
93	1.76	0.51	3.81	2.92	0.75	1.80	0.47	2.50	0.62	2.58	0.68	2.67	0.77	1.67	0.44
94	2.00	0.56	3.22	3.18	0.75	1.80	0.56	2.50	0.62	2.58	0.80	2.67	0.77	1.67	0.52
95	2.04	0.56	3.22	3.67	1.03	0.95	0.29	2.56	0.62	2.58	0.80	3.46	1.05	0.91	0.28
96	2.04	0.56	3.22	3.67	1.03	0.95	0.29	2.95	0.74	1.81	0.56	3.46	1.05	0.91	0.28
97	2.59	0.63	2.54	3.67	1.03	0.95	0.37	3.32	0.74	1.81	0.71	3.46	1.05	0.91	0.36
98	3.32	0.75	1.79	3.67	1.03	0.95	0.53	3.70	1.03	0.95	0.53	3.46	1.05	0.91	0.51
99	3.61	1.03	0.94	3.67	1.03	0.95	1.00	3.70	1.03	0.95	1.01	3.46	1.05	0.91	0.96
القيمة الصغرى	-2.04	0.40	0.94	-2.11	0.38	0.95	0.29	-1.64	0.39	0.95	0.53	-1.48	0.45	0.91	0.28
القيمة العظمى	3.61	1.03	6.35	3.67	1.03	6.83	1.26	3.70	1.03	6.63	1.35	3.46	1.05	4.94	1.10
المتوسط الحسابي	0.26	0.44	5.46	1.02	0.50	4.85	**0.87**	0.84	0.46	5.29	**0.96**	0.71	0.55	3.86	**0.70**
الانحراف المعياري	1.06	0.08	1.06	1.36	0.16	1.87	0.27	1.11	0.11	1.46	0.18	1.30	0.15	1.29	0.18

وللتحقق من فاعلية ودقة المعادلة وفق النموذج أحادي المعلم (راش)، تم التحقق من فاعلية ودقة القدرات المتعادلة بطريقة الرتب المئينية باستخدام الخطأ المعياري للمعادلة. وفق المعادلة الآتية (kolen & Brennan, 2004):

$$\hat{se}[\hat{e}_Y(x_i)] = \sqrt{\frac{1}{R-1}\sum_{r=1}^{R}[\hat{e}_{Y.r}(x_i) - \hat{e}_{Y.}(x_i)]^2}$$

حيث R: عدد العينات العشوائية المختارة من الصورة Y.

$\hat{e}_{Y\ .r}(x_i)$: العلامة المعادلة على العينة r من الصورة Y للعلامة الخام xi.

$\hat{e}_{Y\ .}(x_i)$: المتوسط الحسابي للعلامات المعادلة على جميع العينات من الصورة Y للعلامة الخام xi ويتم إيجاده من خلال المعادلة الآتية:

$$\hat{e}_{Y.}(x_i) = \frac{\sum_{r=1}^{R}\hat{e}_{Y.r}(x_i)}{R}$$

وأظهرت النتائج أن قيم الخطأ المعياري لمعادلة القدرة بطريقة الرتب المئينية لنموذج التصفية التجريبية تراوحت بين(0.5153 – 0.000) وبمتوسط حسابي مقداره (0.1608)، وتراوحت لنموذج التصفية بالتحكيم بين (0.8518 – 0.000) وبمتوسط حسابي مقداره (0.1615)، ولنموذج التصفية المنطقية تراوحت بين (0.9153 – 0.000) وبمتوسط حسابي مقداره (0.2441)، وتبين هذه النتائج أن القدرات المتعادلة بطريقة الرتب المئينية لنموذج التصفية التجريبية كانت الأكثر دقة وفاعلية عند معظم القدرات المتعادلة، ثم يليها دقة وفاعلية

القدرات المتعادلة لنموذج التصفية بالتحكيم، أما القدرات المتعادلة لنموذج التصفية المنطقية كانت الأقل دقة وفاعلية من باقي النماذج، ومدلول ذلك أن بيانات نموذج التصفية المنطقية كانت الأقل مطابقة للنموذج أحادي المعلم (راش)، إذ يتبين أن أكثر من (11) فقرة و(12) فرداً لم تطابق النموذج، ويبين هذا أن نموذج راش قدم نتائج أفضل عندما كانت البيانات متطابقة مع النموذج، وأنه في حال مطابقة البيانات لنموذج تكون نتائج المعادلة لصالحه.

وللتحقق من فاعلية معادلة القدرة بطريقة الرتب المئينية لنماذج تصفية المموهات باستخدام محك الصدق التقاطعي، تم استخدام المعادلة التالية (Peterson, et al., 1989):

$$C = \sum_i \frac{(Yi - Y'i)^2}{NK}$$

حيث Yi: العلامات المعادلة في العينة الأولى.

$Y'i$: العلامات المعادلة في العينة الثانية.

N: عدد العلامات الملاحظة في توزيع الصدق التقاطعي.

K: عدد الفقرات في الاختبار المتعادل.

والجدول (23) يبين قيم الصدق التقاطعي التي تم الحصول عليها ولكل نموذج من نماذج تصفية المموهات:

الجدول (23): قيم الصدق التقاطعي لنماذج تصفية المموهات لمعادلة القدرة بطريقة الرتب المئينية

نموذج تصفية المموهات	الصدق التقاطعي
التجريبي	0.0031
التحكيم	0.0053
المنطقي	0.0079

ويتبين من خلال الجدول (23) أن أصغر قيمة للصدق التقاطعي كانت عند نموذج التصفية التجريبية (0.0031)، ثم عند نموذج التصفية بالتحكيم (0.0053)، وأكبرها كانت عند النموذج المنطقي (0.0079)، أي أن أكثر المعادلات دقة وفاعلية عندما تمت معادلة القدرة للنموذج التجريبي بما يعادلها من قدرات على النموذج العشوائي، وأقلها دقة وفاعلية عندما تمت معادلة القدرة للنموذج المنطقي بما يعادلها من قدرات على النموذج العشوائي، والسبب يعود إلى أن بيانات نموذج التصفية المنطقية كانت الأقل مطابقة من بين النماذج الأخرى لنموذج أحادي المعلم، وتتفق هذه النتيجة مع ما تم التوصل إليه عند حساب الخطأ المعياري لمعادلة القدرة. ويلاحظ أن قيم الصدق التقاطعي وقيم الخطأ المعياري للمعادلة وفق نموذج راش كانت صغيرة، وهذا يبين دقة وفاعلية المعادلة وفق نموذج راش.

استنتاجات الدراسة:

ذكرت العديد من الدراسات أن الطريقة التي يتم بها تصفية المموهات تؤثر على خصائص الاختبار وفقراته وعلى درجات الأفراد، الأمر الذي يؤثر في عدالة القياس لنماذج الاختبار المكونة بالطرق المختلفة لتصفية المموهات بسبب الاختلاف في خصائص الاختبار وفقراته، وقد توصلت الدراسة إلى ما يلي:

أ. التأكيد على ضرورة معادلة الدرجات للاختبارات المكونة من طرق تصفية المموهات، لما أظهرته النتائج من تأثر خصائص الاختبار وفقراته ودرجات الأفراد بطريقة تصفية المموهات، ولما أظهرته النتائج من اختلاف واضح في دقة المعادلة وفاعليتها يُعزى إلى طريقة تصفية المموهات.

ب. تتأثر دقة وفاعلية المعادلة باختلاف الطريقة التي يتم بها تصفية المموهات، وفق طريق المعادلة المستخدمة في الدراسة، كما يتضح ذلك من قيم الخطأ المعياري للمعادلة وقيم الصدق التقاطعي.

ج. أشارت النتائج إلى مقدرة طريقة المعادلة وفق نموذج راش في تحقيق أهدافها.

د. تقل فاعلية المعادلة وفق نموذج راش عندما لا تكون البيانات متطابقة مع النموذج، وتزداد دقة وفاعلية المعادلة وفق نموذج راش في حال مطابقة البيانات لافتراضات النموذج.

توصيات الدراسة

في ضوء ما توصلت إليه نتائج الدراسة فإن الباحث يوصي بما يلي:

أ. دراسة أثر طريقة تصفية المموهات على معادلة الاختبار باستخدام طرق معادلة جديدة غير مستخدمة في الدراسة الحالية مثل طريقة تكر الخطية وطريقة ليفن الخطية القائمة على النظرية الكلاسيكية، وطريقة معادلة الدرجات الحقيقية القائمة على النظرية الحديثة في القياس. وباستخدام معايير جديدة للحكم على دقة وفاعلية المعادلة غير المستخدمة في الدراسة الحالية كمعيار الجذر التربيعي لمتوسط مربعات الخطأ في المعادلة، ومعيار تحيز المعادلة.

ب. دراسة أثر طريقة تصفية المموهات على معادلة الاختبار مع الأخذ بالاعتبار التصحيح لأثر التخمين للدرجات

ولمعاملات الصعوبة للفقرات، إذ لم تراعي الدراسة الحالية التخمين، وافترضت بأن التخمين تأثيره متساوٍ.

ج. إجراء مقارنة بين دقة وفاعلية المعادلة باستخدام النماذج المختلفة لنظرية الاستجابة للفقرة (النموذج ثلاثي المعلم، والنموذج ثنائي المعلم، بالإضافة إلى النموذج أحادي المعلم (راش) نظراً لاختلاف التشدد في افتراضات كل نموذج من نماذج نظرية الاستجابة للفقرة.

المراجع

أبو جراد، ح. (1997) أثر الحذف التجريبي والعشوائي للمموهات في فقرات اختبار التحصيل من نوع الاختيار من متعدد على خصائصه السيكومترية، رسالة ماجستير غير منشورة، جامعة اليرموك: اربد، الأردن.

ثورندايك، وهيجن.(1986) القياس والتقويم في علم النفس والتربية، (ترجمة: عبد الله زيد الكيلاني، وعبد الرحمن عدس)، عمان: مركز الكتب الأردني.

الحياصات، خ. (2011) مدى تحقق معايير الفاعلية في معادلة نموذجي اختبار مع بقاء الفقرات ذات الأداء التفاضلي للجنس وحذفها، رسالة دكتوراه غير منشورة، جامعة اليرموك: اربد، الأردن.

الحواري، أ. (2008) الخصائص السيكومترية لصور مختارة من اختبارات الرخصة الدولية لقيادة الحاسوب في الأردن ومعادلة درجاتها، رسالة ماجستير غير منشورة، جامعة اليرموك: اربد، الأردن.

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.

ظاظا، ح. (2000) المقارنة بين أثر أربع طرق لحذف المموهات في فقرات اختبار الاختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته، رسالة ماجستير غير منشورة، الجامعة الأردنية: عمان، الأردن.

عودة، أ. (2010) *القياس والتقويم في العملية التدريسية*، اربد: دار الأمل للنشر والتوزيع.

عيلبوني، س. (2007) أثر عدد البدائل وتمييز المموهات في فقرات الاختيار من متعدد على التوافق مع نموذج ثلاثي المعلم، رسالة دكتوراه غير منشورة، جامعة اليرموك: اربد، الأردن.

المدانات، ص. (2004) تطوير اختبار رياضيات متعدد المستويات للصفوف الأساسية (3–6) بفقرات متعددة التدريج، رسالة دكتوراه غير منشورة، جامعة عمان العربية: عمان، الأردن.

Aiken, L. (2003) Psychological testing and assessment. Boston: Pearson Education Group.

Baker, F. (2001) The basics of item response theory. USA: ERIC Clearinghouse on Assessment & Evaluation.

Cizek, G. & O'day, D. (1994). Further investigation of nonfunctiong options in multiple-choice test items. Educational and Psychological Measurement, 54(4), 861-872.

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

Hakstian, A. & Whalen, T. (1976) A k-samples significance test for independent alpha coefficients. Psychometrika, 41(2), 219-231.

Heh, V. (2007). Equating accuracy using small samples in the random groups design. (Doctoral dissertation, Ohio University) http://etd.ohiolink.edu/view.cgi?acc_num=ohiou1178299995.

Kim, S. & Hanson, B. (2002) Test equating under the multiple – choice model. Applied Psychological Measurement, 26(3), 255 – 270.

Kolen, M.& Brennan, R. (2004) Test Equating, scaling, and linking methods and linking. New York: Springer Science and Business Media, LLC.

Peterson, N., Kolen, M. & Hoover, H. (1989) Scaling norming and equating. Educational Measurement, (ED). By Linn, R. L.: (3rd ed). Washington, D.C.: American Council on Education, 241-262.

Suen, H. (1990). Principles of test theory. Hillsdale, NJ: Lawrence Elbaum.

The Effect of Filtering Distractors Method in Multiple Choice Items on Test Equating

*Mousa Mohmmad Mahmood Kofahi, Reema Fayez Ali Zawahrah **

ABSTRACT

This study aims at investigating the effect of filtering distractors method in multiple-choice items on test equating. To achieve this purpose, the researcher adopted a math test consisted of (40) multiple-choice items with six Alternatives each. Completing the first form of the test, the researcher administered it to a sample of (130) students. Then the difficulty and discrimination coefficients were calculated. The final form of the test consisted of (34) multiple-choice items with six Alternatives each. Upon deleting two distractors from the Alternatives of the final test, formulated four test forms with four Alternatives for each item. These forms come as follows: a form according to the random filtering method of distractors, a form according to the experimental filtering method, a form according to the Judicial filtering method, and a form according to the logical filtering method.

The four test forms were applied to a sample of (880) students from the seventh grade, During the period from10/4/2013 up to 16/4/2013.used The Rach method according to the modern theory in measurement. Then, the researcher, in equating procedures, adopted the Random Group Design. Upon analyzing the data, the results of the study emphasized that the scores of the distractors filtering forms shall necessarily be equated. Furthermore, the results showed that the accuracy and efficiency of the equating can be affected by the different methods of distractors filtering used. according to the results, the effectiveness of equation decreases in accordance with Rash Model when the data are not in compatible with the model, and the effectiveness and punctuality of the equation increases, in accordance with Rash Model in case the data are not compatible with the model assumptions.

Keywords: Filtering Distrasctors, Test Equating, Multiple Choice Items.

* The Unit of Measurement and Evaluation, Deanship of Quality and Accreditation, Islamic University in Madinah, Saudi Arabia; and Competitive Examination, Civil Service Bureau, Amman, Jordan. Received on 14/10/2014 and Accepted for Publication on 15/3/2015.