المجلة الأردنية في العلوم التربوية، مجلد 9، عدد 3، 2013، 319 - 333

أثر عدد البدائل وتغيير موقع المموه القوي في فقرات اختبار الاختيار من متعدد على معالم الفقرات وقدرة الفرد ودالة المعلومات

زايد صالح بني عطا* ابراهيم محمد الرباعي**

تاريخ تسلم البحث 2012/10/3 تاريخ قبوله 2013/3/14

ملخص: هدفت الدراسة إلى التحقق من أثر عدد بدائل الفقرة وموقع المموه القوي في فقرات اختبار الاختيار من متعدد على الخصائص السيكومترية للاختبار ومعالم الفقرات وقدرة الأفراد. ولتحقيق الغرض من الدراسة تم بناء اختبار تحصيلي من نوع الاختيار من متعدد في الرياضيات لطلبة الصف العاشر مكون من 41 فقرة بصورته النهائية, وقد اشتمل على أربعة نماذج حسب عدد البدائل وموقع المموه القوي. وباستخدام برنامج Bilog- Mg3 تم تحليل استجابات المفحوصين البالغ عددهم 2111 طالبا وطالبة لجميع نماذج الاختبار الأربعة وفق النموذج اللوجستي ثلاثي المعلمة. كشفت نتائج تحليل التباين الثنائي بأنه لا يوجد فروق ذات دلالة إحصائية بين متوسطات معالم الصعوبة للفقرات تعزى لعدد بدائل الفقرة وموقع المموه القوي، وكذلك أظهرت النتائج أيضا بأنه لا يوجد فروق ذات دلالة إحصائية بين متوسطات معالم التخمين للفقرات تعزى لمتغيري الدراسة، في حين كشفت النتائج عن فروق ذات دلالة إحصائية بين متوسطات معالم التمييز تعزى لموقع المموه القوي وقد جاءت تلك الفروق لصالح النموذج الثاني (5- بدائل موقع المموه القوي بعيد). وبينت نتائج التحليل كذلك بأنه لا يوجد فروق دالة إحصائيا بين متوسطات معلمة القدرة للأفراد تعزى لعدد بدائل الفقرة وموقع المموه القوي. وكشفت النتائج كذلك بان دالة المعلومات للاختبار قد تباينت بتباين نماذج الاختبار، وكانت هناك فروق دالة إحصائيا لقيم معاملات الثبات الامبريقي حيث جاءت لصالح النموذج الثاني. (**الكلمات المفتاحية:** المموه القوي، عدد البدائل، النموذج اللوجستي ثلاثي المعلمة، اختبار الاختيار من متعدد، دالة المعلومات).

The Effect of Alternatives Number and Changing the Position of the Strong Distractor on Items Parameters, Person Ability and Information Function

Zaid Bani Ata, *Faculty of Education, Yarmouk University, Irbid, Jordan.*
Ibraheem Al-Rabbaie, *Customs Department Irbid, Jordan.*

Abstract*: This study aimed at verifying the effect of the number of alternatives and changing the strong distractor position on the items parameters, person's ability and information function. To achieve this aim, a 41 item multiple – choice achievement test in 10th grade mathematics was constructed.The test had four different forms according to number of alternative and the strong distractor position. The responses of 2111 examinees on the four forms were analyzed by Bilog–Mg3 programs according to the three parameter logistic model. The results of two-way ANOVA revealed no significant statistical differences between the means of item difficulty and guessing parameters. Differences were found, however, between the means of item discrimination, due to the position of the strong distractor; the differences were in favor of the second form. No differences were found in the means person's ability parameter, but the information function of the test varied due to the test forms. Significant differences were noticed in the values of empirical reliability coefficients in favor of the second form. (**Key words:** Strong Distractor, Number of Alternatives, Three- Parameter Logistic Model, Multiple-Choice Test, Information Function).*

مقدمة: إن محور اهتمام المربين في كثير من دول العالم هو التحصيل الدراسي من حيث اكتساب الطلبة وتعلمهم محتوى دراسي معين , ويمكن تعريف التحصيل الدراسي بأنه مقدار ما تحقق عند الفرد (المتعلم) من نتاجات التعلم المنشودة نتيجة مروره بخبرة تعليمية مقصودة ومنظمة, وتستخدم الاختبارات التحصيلية لقياسه التي لعبت دورا رئيسا في تقويم تعلم الطلاب, حتى أن جرونلد (Gronlund, 1982) اعتبرها الأداة الوحيدة أحيانا في تحديد ما اكتسبه الطالب من معارف ومهارات, فهي توفر مقاييس جيدة لكثير من مخرجات التعلم كما أنها تساعد على قياس كثير من التغيرات المرغوبة في سلوك الطالب (حمدان, 1986؛ مراد وسليمان , 2002).

ويمكن تصنيف الاختبارات التحصيلية وفق معايير مختلفة, احدها أن تصنف حسب طبيعة الاجابة المطلوبة فتصنف الى: الاختبارات ذات الإجابة المصاغة والاختبارات ذات الإجابة المنتقاة. ويتوقف استخدام أحد هذه الأنواع من الفقرات على مجموعة من العوامل منها: مستوى الهدف المراد قياسه، وعمر المفحوص، وغرض الاختبار، وعدد الطلبة، وظروف التطبيق، ومهارة المعلم في صياغة الأسئلة (عودة، 2010). ويرى بروان (Brown, 1983) أن اختيار شكل الفقرات المناسب يعتمد على خصائص الفقرة ومهارة المعلم في كتابة الفقرات , وفلسفة التربية.

* قسم علم النفس الإرشادي والتربوي، جامعة اليرموك، إربد، الأردن.
** دائرة الجمارك العامة، إربد، الأردن.
© حقوق الطبع محفوظة لجامعة اليرموك، اربد، الأردن.

ويعد الاختيار من متعدد من أكثر أشكال الفقرات شيوعا واستخداما في الاختبارات التحصيلة, وذلك بسبب كفاءتها فهي تستطيع قياس التحصيل الدراسي بشكل أكثر ثباتا وصدقا(Aiken, 1987; Gay, 1980; Frisbie & Sweeney, 1982). وتتكون فقرة اختبار الاختيار من متعدد من جزئين رئيسيين: الجزء الأول الأرومة (الجذر) وهو عبارة عن سؤال أو جملة إنشائية تطرح سؤالاً أو قضية بحاجة لإجابة، أما الجزء الثاني فهو مجموعة البدائل المحتملة كاجابة للسؤال المطروح، ويكون أحدها هو الاجابة الصحيحة للسؤال والبدائل الاخرى مموهات يفترض أن تكون ذات صلة بالمشكلة التي يطرحها السؤال وبذا يفترض ان تكون جميعها جذابة وبنفس الدرجة للأفراد الذين لا يعرفون الاجابة الصحيحة للسؤال يتراوح عددها من أثنين إلى خمسة بدائل (Nitko, 2001).

وعلى الرغم من وجود العديد من الميزات لفقرات اختبار الاختيار من متعدد، إلا أن هنالك عدداً من الانتقادات قد وجهت إليها منها: مشكلة عدد البدائل، ومشكلة التحيز لموقع البدائل، وموقع الإجابة الصحيحة بالنسبة لموقع المموه القوي وباقي المموهات، والتي بدورها تؤثر على استجابة المفحوصين وعلى فقرة الاختبار. وتجدر الإشارة كذلك إلى أنّ اختيار المفحوص للبديل الصحيح يعتمد على موقع البديل الصحيح بالنسبة للبدائل الأخرى، وعلى محتوى فقرات الاختبار(Blunch, 1984).

لقد اعتبرت قضية عدد بدائل الفقرة من القضايا السيكومترية التي اهتم بها البحث السيكومتري لمعرفة الأثر الناتج منها على الخصائص السيكومترية للفقرات وللاختبار, حيث أشارت نتائج دراسة كريهان وهالدينا وبريور(Crehan, Haladyna & Brewer, 1993) أن الفقرات ذات البدائل الثلاثة كانت أصعب من الفقرات ذات البدائل الأربعة, في حين لم يكن هناك فروق ذات دلالة لعدد البدائل على تمييز الفقرات. أما دراسة فقوسة (1995) فقد أشارت النتائج إلى وجود فروق ذات دلالة إحصائية بين متوسطات معاملات الصعوبة للفقرات ذات الثلاثة بدائل والفقرات ذات الأربعة والخمسة بدائل، ولصالح الفقرات ذات الثلاثة بدائل. أما بالنسبة لمعاملات التمييز فقد أظهرت النتائج وجود فروق ذات دلالة إحصائية بين متوسطات معاملات تمييز الفقرات ذات الخمسة بدائل والفقرات ذات الأربعة بدائل ولصالح الفقرات ذات الخمسة بدائل. كما أظهرت النتائج وجود فروق ذات دلالة إحصائية بين متوسطات معاملات تمييز الفقرات ذات الثلاثة بدائل وذات البديلين، لصالح الفقرات ذات الثلاثة بدائل. أما بالنسبة لمعاملات الثبات فقد أظهرت النتائج وجود فروق ذات دلالة إحصائية بين معامل ثبات النموذج الذي لفقراته خمسة بدائل والنموذج الذي لفقراته بديلين، ولصالح النموذج الذي لفقراته خمسة بدائل، بينما للنماذج الأخرى لم يكن هناك فروق ذات دلالة إحصائية في معاملات الثبات.

وقام رودريكز (Rodriguez, 2005) بدراسة هدفت إلى الكشف فيما إذا كان الاختبار الذي يحتوي على ثلاثة بدائل هو الاختبار الأفضل في اختبارات الاختيار من متعدد. وقد قام الباحث بتحليل مجموعة من البحوث والدراسات التربوية والتي أجريت خلال (80) سنة، وذلك للتعرف على أثر عدد البدائل ودور المموهات في اختبارات الاختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته أظهرت نتائج الدراسة أن الاختبار ذا الثلاثة بدائل كان الأفضل في اختبارات الاختيار من متعدد في معظم الدراسات، وأن التغير من الخمسة بدائل إلى الأربعة بدائل يعمل على انخفاض معامل صعوبة الفقرة بمعدل (0.02)، ومعامل التمييز بمعدل (0.04)، ومعامل الثبات بمعدل (0.035). وبينت نتائج دراسة كل من شيزيوكا وتاكيوشي وياشيما ويوشيزاوا (Shizuka, Takeuchi, Yashima, & Yoshizawa, 2006) بأن الاختبار ذا الثلاثة بدائل يؤدي الهدف أو الغرض منه كما في الاختبار ذي الأربعة بدائل، حيث أن زيادة عدد بدائل الفقرات يقلل من وقت الاختبار، ويزيد من فرصة عامل التخمين.

وبينت نتائج دراسة الشريفين وطعامنة (2009) التي هدفت إلى الكشف عن أثر عدد البدائل في اختبار الاختيار المتعدد في تقديرات القدرة للأفراد ومعالم صعوبة الفقرات بأنه لا يوجد فروق دالة إحصائيا بين متوسطات الأخطاء المعيارية لصعوبة الفقرات تعزى لعدد بدائل الفقرة, في حين كانت هناك فروق ذات دلالة إحصائية بين متوسطات الأخطاء المعيارية في تقديرات معلمة القدرة للأفراد تعزى لعدد البدائل حيث جاءت دقة التقدير لصالح الاختبار ذو الثلاثة بدائل. بينما توصلت دراسة يامان (Yamman,2011) التي هدفت لتحديد العدد الأمثل للبدائل في اختبارات الاختيار من متعدد، وذلك بهدف مقارنة الخصائص السيكومترية لاختبارات الاختيار من متعدد. أشارت نتائج الدراسة بأن مستوى ثبات فقرات الاختبار المكون من ثلاثة وخمسة بدائل أعلى منها في الفقرات المكونة من أربعة بدائل. كما أشارت أيضاً إلى عدم وجود فروق ذات دلالة إحصائية بين الاختبارات الثلاثة في الخصائص السيكومترية للفقرات الخاصة بمعلمة الصعوبة والتمييز لكل منها. وقد أوصت الدراسة بأن فقرات الاختبار المكون من ثلاثة بدائل هي الأفضل؛ ويعود السبب في ذلك إلى أن هذا النوع من الاختبارات يتم تطويره وتحليله بطريقة أسهل.

ولم يقتصر البحث السيكومتري على تناول أثر عدد البدائل في الخصائص السيكومترية للفقرات والاختبار بل تعدى ذلك إلى دراسة اثر موقع البديل الصحيح بالنسبة لأحد المموهات التي يدعى المموه القوي (Strong Distractor)، فوجود المموه القوي بجانب البديل الصحيح، أو بعيداً عنه ببديلين أو أكثر ربما يؤثر في استجابة المفحوص، وبالتالي يؤثر على الخصائص السيكومترية للفقرة. فمن الدراسات التي اهتمت بذلك دراسة أمبوسعيدي (2003) التي هدفت إلى تقصي أثر تغيير موقع المتشتت القوي عند استخدام تقنية إبعاد البدائل الخطأ في أسئلة الاختيار من متعدد على التحصيل الدراسي (المتوسط الحسابي)، وعلى بعض الخصائص السيكومترية مثل: معاملات الصعوبة ومعاملات التمييز

لفقرات الاختبار. ولتحقيق الهدف من الدراسة تم استخدام اختبار في مادة الفيزياء للصف العاشر من التعليم العام مكوناً من (24) فقرة من نوع الاختيار من متعدد. وتم تصميم نموذجين متشابهين في كل شيء من هذا الاختبار ما عدا موقع المتشتت القوي بالنسبة إلى البديل الصحيح، كما تم استخدام تقنية إبعاد البدائل الخطأ من الورقة الامتحانية في النموذجين. أظهرت نتائج الدراسة عدم وجود فروق ذات دلالة إحصائية بين المتوسطات الحسابية لدرجات الطالبات في نموذجي الاختبار، وأشارت نتائج التحليل إلى عدم وجود فروق ذات دلالة إحصائية بين المتوسطات الحسابية لمعاملات الصعوبة، ومعاملات التمييز لنموذجي الاختبار.

وفي دراسة أخرى مماثلة قام بها أمبوسعيد والعفيفي (2004) هدفت إلى تقصي أثر تغيير موقع المشتت القوي في أسئلة الاختيار من متعدد في تحصيل الطلاب، وفي بعض الخصائص السيكومترية لفقرات الاختبار مثل: معاملات الصعوبة، ومعاملات تمييزها. ولتحقيق الهدف من الدراسة تم بناء اختبار تحصيلي في مبحث الفيزياء للصف الأول الثانوي مكون من (30) فقرة في صورته النهائية، وقد جرى إعداد نموذجين متشابهين في كل شيء ما عدا موقع المشتت القوي بالنسبة للبديل الصحيح, وقد تم تطبيق النموذجين على عينة مكونة من

من (197) طالبا. أظهرت نتائج الدراسة وجود فروق ذات دلالة إحصائية بين المتوسطات الحسابية لدرجات الطلبة في نموذجي الاختبار، وأشارت نتائج الدراسة إلى وجود فروق ذات دلالة إحصائية بين المتوسطات الحسابية لمعاملات الصعوبة، ومعاملات التمييز لنموذجي الاختبار.

وقام الكعابنة (2009) بدراسة هدفت إلى الكشف عن أثر طريقة اختيار المموهات في فقرات الاختيار من متعدد على القيم المقدرة لمؤشرات معالم الفقرة والقدرة، وفق النظرية الكلاسيكية ونظرية الاستجابة للفقرة, ولتحقيق الهدف من الدراسة تم إعداد اختبار تحصيلي في مبحث الرياضيات لطلبة الصف التاسع الأساسي، وقد تكون الإختبار بصورته النهائية من (41) فقرة، وجرى بناء مموهات تلك الفقرات في ثلاث نماذج (A, B, C) وفق ثلاث طرق مختلفة لاختيار المموهات، فالطريقة الأولى كانت قائمة على اختبار المعلمين، بينما كانت الطريقة الثانية قائمة على أسس إحصائية، في حين كانت الطريقة الثالثة قائمة على الإجابات المفتوحة للطلبة عن فقرات الاختبار. أشارت النتائج إلى وجود فروق ذات دلالة إحصائية بين متوسطات الأخطاء المعيارية لتقديرات معلمة القدرة، ووجود فروق ذات دلالة إحصائية بين قدرات المفحوصين المقدرة بالعلامة الخام، تعزى إلى الطريقة المتبعة في بناء مموهات فقرات الاختبار، كما أشارت نتائج التحليل إلى وجود فروق ذات دلالة إحصائية بين التقديرات الخاصة بمعلمة الصعوبة، والتقديرات الخاصة بمعلمة التمييز، وكذلك الأخطاء المعيارية في تقديرها. وبينت نتائج الدراسة عدم وجود فروق ذات دلالة إحصائية بين التقديرات الخاصة بمعلمة التخمين، بينما أظهرت النتائج وجود فروق بين الأخطاء المعيارية لتقديرات معلمة التخمين، وكذلك وجود فروق ذات دلالة إحصائية بين قيم معاملات الصعوبة وقيم معاملات التمييز وفق النظرية الكلاسيكية، في حين لم تظهر النتائج وجود فروق ذات دلالة إحصائية بين الأخطاء المعيارية لتقديرات معلمة الصعوبة.

وأظهرت النتائج الخاصة بقيم معامل ثبات كرونباخ ألفا عدم وجود فروق ذات دلالة إحصائية بين معاملات الثبات، ووجود فروق ذات دلالة إحصائية بين معاملات الثبات التجريبي، وفيما يتعلق بدالة معلومات الاختبار، فقد أشارت النتائج إلى عدم وجود فروق ذات دلالة إحصائية بين قيم دالة المعلومات تعزى للطريقة المتبعة في اختيار المموهات.

وأجرت تارانت وواير ومحمد (Tarrant, ware & mohammed, 2009) دراسة هدفت إلى معرفة فاعلية وعدم فاعلية المموهات في اختبارات الاختيار من متعدد معتمدين على نتائج التحليل الوصفي للبيانات. ومن أجل تحقيق هدف الدراسة قام الباحثون بالبحث والكشف عن المموهات غير الفاعلة في اختبارات الاختيار من متعدد، والمطورة من قبل المعلم، والمقدمة في تخصص التمريض في إحدى الجامعات في هونج، حيث قام الباحثون ببناء سبعة اختبارات تحصيلية من نوع الاختيار من متعدد في تخصص التمريض، وقد قام الباحثون بإجراء الاختبار وتوزيع نماذجه على طلبة الكلية، بهدف تحليل نتائج الدراسة معتمدين في ذلك على دالة المعلومات، وذلك لتقييم المموهات غير الفاعلة في الاختبارات.أشارت نتائج الدراسة أنّ نسبة المموهات الفاعلة في فقرات الاختبارات قد بلغت 52.2 %، وقد كانت فاعلة بدرجة جيدة في الاختبارات، وأشارت نتائج الدراسة إلى أن 10.2% من المموهات كانت نسبة اختيارها في الإجابة صفر، وأنّ الفقرات ذات المموهات الأكثر فاعلية كانت تتصف بدرجة أكبر من الصعوبة والتمييز مقارنة مع الفقرات الأخرى.

وفي دراسة أخرى أجرت تارنت وواير (Tarrant & Ware, 2010)دراسة لمقارنة الخصائص السيكومترية لاختبارات الاختيار من متعدد الذي تحوي فقراته على ثلاثة بدائل أو أربعة بدائل والمستخدمة في عملية تقييم طلبة تخصص التمريض. ولتحقيق أهداف الدراسة قام الباحثان بتطبيق اختبار الاختيار من متعدد ذو الأربعة بدائل على عينة استطلاعية، وذلك لفحص ومقارنة الخصائص السيكومترية لفقرات الاختبار، وباستخدام بيانات تحليل الفقرة، تم معرفة المموه الضعيف في عملية الإجابة، ليتم بعدها إعادة كتابة نفس الفقرات بعد حذفه. وقام الباحثان بإعداد الصورة النهائية للاختبار، بحيث يتكون الاختبار من (41) فقرة في كل نموذج، فالنموذج الأول يحوي ثلاثة بدائل، أما النموذج الثاني فيحوي أربعة بدائل. وقد قام طلبة تخصص التمريض بالاستجابة عن فقرات الاختبار.وبعد تحليل استجابات الطلبة، أشارت نتائج الدراسة أن الاختبارات التي تحتوي على ثلاثة بدائل كانت أكثر فاعلية، على الرغم من قلة المموهات؛ وذلك بسبب قوة هذه المموهات. وأشارت نتائج الدراسة بأن المموهات المستخدمة في

اختبار الاختيار من متعدد، تصبح ذات درجة عالية من التمييز حين يتم حذف المموهات التي لا يتم اختيارها بشكل متكرر في عملية الإجابة.

إن المتصفح لنتائج الدراسات السابقة يجد التباين في نتائجها حول العدد الأمثل لعدد فقرات الاختبار من اختيار متعدد, وبنفس الوقت تباينت نتائج الدراسات حول اثر عدد البدائل في الخصائص السيكومترية للفقرات والاختبار ككل, وان هناك أثرا لموقع المموه القوي على خصائص الفقرات. وتجدر الإشارة كذلك بان غالبية هذه الدراسات اعتمدت على النظرية التقليدية في القياس في تحليل النتائج باستثناء دراسة الشريفين وطعامنة (2009) ودراسة الكعابنة(2005) التي اعتمدت نظرية الاستجابة للفقرة (Item Response Theory) والتي تعد ثورة في القياس النفسي والتربوي, حيث ساعدت على تقديم الكثير من الحلول، لمشاكل تتعلق ببناء الاختبارات وتطويرها، خاصة فيما يتعلق بتكافؤ الاختبارات ومعادلتها، وبناء الاختبارات المحكية المرجع، والاختبارات التكيفية، وبناء بنوك الأسئلة، والكشف عن تحيز الفقرات ونحو ذلك والميزة الكبرى لنظرية استجابة الفقرة، أنها تقود إلى فقرات تتسم معالمها باللاتغير عند تغيير العينة (Embretson & Reise, 2000).

لذا جاءت الحاجة لإجراء هذه الدراسة من خلال تناولها أثر عدد البدائل وموقع المموه القوي على معالم الفقرات وقدرة الأفراد معا, والتي افتقرت إليه الدراسات السابقة باستخدام احد نماذج نظرية الاستجابة للفقرة وهو النموذج اللوجستي ثلاثي المعلمة (Three – Parameter Logistic Model), حيث يسمح هذا النموذج باختلاف معالم الفقرة الثلاثة، إذ تختلف هذه المنحنيات في ميلها (معلمة التمييز)، وموقعها على متصل السمة (معلمة الصعوبة)، وخط التقارب السفلي لمنحنى خصائص الفقرة (معلمة التخمين)، أو كما أطلق عليه لورد معلم مستوى الصدفة Pseudo Chance Level، وتمثل هذه المعلمة احتمال توصل الفرد ذو القدرة المنخفضة للإجابة الصحيحة عن الفقرة بالتخمين (Hambelton & Swaminathan, 1985 ; Lord, 1980).

مشكلة الدراسة وأسئلتها

نظراً للانتشار الواسع لاستخدام اختبارات الاختيار من متعدد، وخاصة في المجال التعليمي؛ لما تتمتع به هذه الاختبارات من ميزات كثيرة؛ إذ إن هذه الاختبارات لديها القدرة على شمول المحتوى بشكل جيد، وكذلك سهولة التطبيق، وموضوعية التصحيح، مما جعلها تمتع بدرجة عالية من الصدق والثبات، حيث يتم استخدامها في الاختبارات المقننة بدرجة كبيرة نظراً لدرجة الثقة التي تتمتع بها في تقدير قدرات المفحوصين، وبالرغم من ذلك إلا أن هذا النوع من الاختبارات له عيوب من أبرزها أمكانية الوصول إلى الإجابة الصحيحة عن طريق التخمين، وحتى يتم التخفيف من هذا الأثر في تقدير قدرات المفحوصين يُنصح بانئي الاختبارات عادة بزيادة عدد بدائل الفقرة لتقليل فرصة التخمين. لذا اهتمت الدراسات التربوية بالبحث عن العدد المثالي للبدائل، حيث لم تتفق نتائج الدراسات على عدد البدائل الأفضل، فقد أشارت معظمها إلى أن عدد البدائل الأفضل يتباين بتباين قدرات المفحوصين.

كما أن هناك دراسات اهتمت بموقع المموه القوي بين بدائل فقرة الاختيار من متعدد، وبالمثل فقد تباينت نتائج الدراسات حول موقع المموه القوي الذي يؤثر في الخصائص السيكومترية للاختبار أو في تقديرات القدرة للمفحوصين، ومعظم الدراسات التي بحثت في أثر عدد البدائل وموقع المموه استخدمت في عملية التحليل لبياناتها النظرية التقليدية في القياس رغم أوجه القصور التي تعاني منها هذه النظرية أبرزها اختلاف معالم الفقرات مثل الصعوبة والتمييز باختلاف عينة المفحوصين، ومعظم الدراسات اعتمدت نموذج راش للمعلمة الواحدة، والذي يفترض تكافؤ الفقرات في التمييز والتخمين، لذا جاءت هذه الدراسة لتناول تفاعل هذين العاملين وهما: عدد بدائل فقرات الاختيار من متعدد وموقع المموه القوي في التأثير على الخصائص السيكومترية للاختبار ومعالم الفقرات وقدرات الأفراد باستخدام نظرية الاستجابة للفقرة لتلافي عيوب النظرية الكلاسيكية وكذلك باستخدام النموذج اللوجستي ثلاثي المعلمة والذي يأخذ بالاعتبار اختلاف معالم التخمين، والصعوبة، والتمييز للفقرات، وبالتحديد، فإن هذه الدراسة سعت للوصول إلى إجابة حول التساؤلات الآتية:

1. *هل تختلف تقديرات معالم الفقرات (الصعوبة، التمييز، التخمين) باختلاف عدد بدائل الفقرة (3 أو 5 بدائل) وموقع المموه القوي (قريب أو بعيد) من الإجابة الصحيحة؟*
2. *هل تختلف تقديرات معلمة القدرة للأفراد (θ) باختلاف عدد بدائل الفقرة (3 أو 5 بدائل) وموقع المموه القوي (قريب أو بعيد) من الإجابة الصحيحة؟*
3. *هل تختلف دالة المعلومات للاختبار باختلاف عدد بدائل الفقرة (3 أو 5 بدائل) وموقع المموه القوي (قريب أو بعيد) من الإجابة الصحيحة؟*

أهمية الدراسة:

تكمن الأهمية النظرية لهذه الدراسة في توضيح أثر عدد البدائل وتغير موقع المموه القوي في اختبار الاختيار من متعدد على التقديرات المختلفة لنظرية استجابة الفقرة. ولعل دراسة أثر عدد البدائل وتغير موقع المموه القوي تعدّ من الأمور المهمة نظرا للاستخدام الواسع لاختبارات الاختيار من متعدد في تقييم التحصيل الدراسي, وذلك للوقوف على تأثير ذلك على خصائص الفقرة الإحصائية وقدرات المفحوصين. وتأتي أهمية الدراسة كذلك الإسهام العلمي المتواضع في التراكم المعرفي في هذا المجال حيث ستسهم نتائج الدراسة قيام بعض الباحثين بالمزيد من الدراسات المشابهة على نطاق أوسع.

وتكمن الأهمية التطبيقية لهذه الدراسة في توفيرها لأداة قياس في مبحث الرياضيات يتحقق من خلالها مؤشرات إحصائية جيدة، يمكن الاعتماد عليها مستقبلاً في قياس تحصيل الطلبة في مبحث الرياضيات، وهو ما يمكن الاستفادة منه من قبل المعلمين والمعلمات في الميدان التربوي، إضافةً إلى ما ستوفره هذه الدراسة من نتائج، يمكن الاعتماد عليها في بناء الاختبارات.

مصطلحات الدراسة

- **المموه القوي:** ويقصد به البديل غير الصحيح الذي اختاره العدد الأكبر من المفحوصين بعد اختيارهم للبديل الصحيح.
- **موقع المموه القوي:** ويقصد به موقع المموه القوي سواء أكان قريب (بجانب البديل الصحيح) أم بعيد (ببديل أو بديلين) من موقع الإجابة الصحيحة لفقرة الاختبار.
- **عدد البدائل:** ويقصد بذلك عدد بدائل فقرة الاختيار من متعدد، وهي في هذه الدراسة ثلاثة أو خمسة بدائل للفقرة الواحدة.
- **اختبار الاختيار من متعدد:** ويتكون اختبار الاختيار من متعدد في هذه الدراسة من أربعة نماذج، بحيث يكون المتن لكل أسئلة الاختبار واحد، ويكون الاختلاف بين هذه النماذج فقط في عدد البدائل لكل نموذج، سواء أكان عدد بدائل الفقرة ثلاثة بدائل أم خمسة بدائل، وكذلك الاختلاف في موقع المموه القوي سواء أكان قريب أو بعيد عن الإجابة الصحيحة.
- **النموذج ثلاثي المعلمة:** هو أحد نماذج نظرية الاستجابة للفقرة ثنائية التدريج، حيث يمكن لهذا النموذج تقدير ثلاث معلمات هي: معامل صعوبة الفقرة ومعامل التمييز ومعامل تخمين الفقرة.
- **معالم الفقرات:** وهي معالم الصعوبة، التمييز، التخمين المنبثقة عن النموذج اللوجستي ثلاثي المعلمة.

الطريقة والإجراءات

مجتمع الدراسة وعينتها

تكون مجتمع الدراسة من جميع طلاب وطالبات الصف العاشر الأساسي في المدارس الحكومية التابعة لمديرتي إربد الأولى ولواء الرمثا في محافظة اربد خلال العام الدراسي 2012/2011م والبالغ عددهم 8954 طالبا وطالبة. منهم (6225) طالباً وطالبة من مديرية تربية اربد الأولى، و(2729) طالبا وطالبة من مديرية تربية لواء الرمثا, وقد تم الحصول على أعداد الطلبة وفق التقرير الإحصائي التربوي السنوي في المديريتين خلال العام الدراسي2012/2011م.

وبعد الرجوع للتقرير الإحصائي في المديرتين تم حصر المدارس، إذ اختيرت عينة من المدارس بطريقة عشوائية بلغ عددها (19) مدرسة، منها (11) مدرسة للإناث و(8) مدارس للذكور، ومن ثمّ جرى اختيار من كل مدرسة وبطريقة عشوائية عنقودية مجموعة من الشعب بلغ عددها(60) شعبة، منها(30) شعبة إناث، و(30) شعبة للذكور؛ وهذا الاختيار العشوائي لضمان عدم التحيز, وبذلك يكون عدد افراد عينة الدراسة (2123) طالبا وطالبة.

أداة الدراسة

لتحقيق أهداف الدراسة قام الباحثان بإعداد اختبار تحصيلي في مبحث الرياضيات، لطلبة الصف العاشر الأساسي حيث تم الاعتماد على الإجراءات التي أوردها غرونلند ولين (Gronlund & Linn, 1990) في تصميم اختبارات التحصيل الصفية, وفيما يلي الإجراءات التي اتبعها الباحثان:-

أولا: تحديد الغرض من الاختبار التحصيلي وهو قياس مقدار ما اكتسبه الطالب من معارف ومهارات وقوانين في وحدتي الاقترانات، والدائرة والممارسات والأشكال الرباعية الدائرية لمبحث الرياضيات للصف العاشر الأساسي.

ثانيا: تم تحليل محتوى كل وحدة من الوحدات المعنية من كتاب الرياضيات للصف العاشر الأساسي بالاستفادة من منهاج الرياضيات وخطوطه العريضة إلى مفاهيم، ومصطلحات، ورموز، وتعميمات، ومهارات، وتطبيقات.

ثالثا: تحديد نواتج التعلم التي يقيسها الاختبار حيث تم صياغة النتاجات التعليمية للوحدات الدراسية سابقة الذكر بالاعتماد على دليل المعلم لمبحث الرياضيات واستشارة عدد من ذوي الخبرة التربوية بوصفهم محكمين للأهداف، للتحقق من صحة صياغة الأهداف, ومدى ملاءمتها للمقرر الدراسي. وقد تم إعداد جدول المواصفات (انظر ملحق 1)، إذ جرى ربط مستويات النتاجات التعليمية (الأهداف السلوكية)، بمحتوى الوحدات الدراسية موضع الاختبار, حيث تم توزيع النسب المئوية لكل وحدة وذلك بالاعتماد على عدد الحصص التي يقضيها المعلم في التدريس، كما هو موضح في دليل المعلم لمادة الرياضيات للصف العاشر الأساسي للعام 2011-2012م.

رابعا: كتابة فقرات الاختبار وتعد هذه الخطوة من أهم مراحل بناء الاختبار التحصيلي، حيث تم كتابة 77 فقرة من نوع الاختيار من متعدد لكل فقرة منها خمسة بدائل، وقد روعي عند كتابتها الأسس الفنية في كتابة هذا النوع من الفقرات.بحيث تتلاءم مع المستوى المعرفي للهدف الذي تقيسه، ومن الأمثلة على ذلك:

مثال (1): إذا علمت أن ق(س)= 5 فان مجال هذا الاقتران يقع في الفترة:

ا) [5, ∞) ب) [0, ∞) جـ (-∞ , 5]

د) (-3, ∞) هـ) [-3, -∞)

مثال (2): مجموع قياس كل زاويتين متقابلتين في الشكل الرباعي الدائري يساوي:

أ) ه-180 ب) 90 جـ) 180 د) 360 هـ) ه+180

خامسا: بعد الانتهاء من كتابة فقرات الاختبار تم عرض الاختبار بصورته الأولية (77 فقرة) على 20 محكما من أهل الخبرة والاختصاص من أساتذة الجامعات في تخصصي أساليب تدريس الرياضيات والقياس والتقويم, بالإضافة إلى معلمي ومشرفي مبحث الرياضيات في وزارة التربية والتعليم في الأردن, وطلب إليهم إبداء الرأي حول وضوح الفقرات ومدى ملاءمتها لقياس النتاجات التعليمة واقتراح ما يرونه مناسبا من خلال الإجابة عن استبيان معد لهذا الغرض, وبعد استرجاع نسخ المحكمين تم إجراء التعديلات اللازمة على فقرات الاختبار، وحذف بعض الفقرات التي أجمع المحكمون على أنها فقرات تؤدي إلى الهدف نفسه، وقد بلغ عددها (14) فقرة، كما قدّم المحكمون اقتراحات قيّمة متعلقة بصياغة بعض الفقرات واقتراح بدائل أكثر قوة ساعدت الباحثان في مراجعة عدد الفقرات التي أعيدت صياغتها والنظر ببعض البدائل المقترحة، وقد أصبح الاختبار مؤلف من (63) فقرة.

سادسا: تم تطبيق الاختبار على عينة استطلاعية مكونة من (120) طالبا وطالبة من طلبة المدرسة النموذجية في جامعة اليرموك موزعين على أربع شعب؛ شعبتان للذكور وشعبتان للإناث تم اختيارهما بصورة عشوائية من شعب الصف العاشر الأساسي في هذه المدرسة. وجرى إبلاغ الطلبة من خلال معلميهم بموعد تطبيق الاختبار، والمادة التي سيطبق عليها الاختبار. وبالاعتماد على التجريب الأولي للاختبار، تم تحديد الزمن المناسب للإجابة عن فقرات الاختبار بواقع ساعة وربع، وتم حساب صعوبة الفقرات وذلك بإيجاد نسبة الطلبة الذين أجابوا عن الفقرة إجابة صحيحة من بين الطلبة الذين حاولوا الإجابة عن هذه الفقرة، كما جرى حساب معامل التمييز للفقرات، وكذلك حساب فاعلية المموهات لكل فقرة، وذلك لتحديد المموه القوي لكل فقرة من فقرات الاختبار. ويبين الجدول1 قيم معاملات الصعوبة والتمييز وكذلك فعالية المموهات لكل فقرة من فقرات الاختبار بصورته الأولية.

الجدول 1: قيم معاملات الصعوبة والتمييز وفاعلية المموهات مرتبه تنازلياً حسب قوتها.

رقم الفقرة	معامل الصعوبة	معامل التمييز	فاعلية المموهات			
			الاول	الثاني	الثالث	الرابع
1	*0.49*	*0.74*	*0.22*	*0.18*	*0.07*	*0.04*
2	*0.62*	*1.00*	*0.17*	*0.09*	*0.07*	*0.05*
3	*0.62*	*0.91*	*0.25*	*0.08*	*0.03*	*0.02*
4	*0.73*	*0.63*	*0.17*	*0.05*	*0.03*	*0.02*
5	*0.28*	*0.45*	*0.36*	*0.23*	*0.10*	*0.03*
6	*0.69*	*0.75*	*0.21*	*0.05*	*0.04*	*0.01*
7	*0.77*	*0.53*	*0.16*	*0.04*	*0.03*	*0.01*
8	*0.67*	*0.45*	*0.14*	*0.09*	*0.08*	*0.02*
9	*0.44*	*0.21*	*0.21*	*0.19*	*0.13*	*0.03*
10	*0.20*	*0.46*	*0.29*	*0.27*	*0.16*	*0.08*
11	*0.20*	*0.46*	*0.44*	*0.12*	*0.09*	*0.05*
12	*0.68*	*0.63*	*0.15*	*0.09*	*0.05*	*0.03*
13	*0.61*	*0.70*	*0.22*	*0.15*	*0.02*	*0.01*
14	*0.32*	*0.02*	*0.48*	*0.14*	*0.03*	*0.03*
15	*0.38*	*0.82*	*0.36*	*0.14*	*0.07*	*0.05*
16	*0.51*	*0.79*	*0.26*	*0.16*	*0.04*	*0.03*
17	*0.58*	*0.82*	*0.25*	*0.09*	*0.06*	*0.02*
18	*0.66*	*0.56*	*0.18*	*0.12*	*0.03*	*0.01*
19	*0.14*	*0.37*	*0.39*	*0.26*	*0.17*	*0.04*
20	*0.66*	*0.91*	*0.19*	*0.09*	*0.04*	*0.02*
21	*0.26*	*0.45*	*0.28*	*0.19*	*0.15*	*0.12*
22	*0.43*	*0.08*	*0.27*	*0.14*	*0.10*	*0.06*
23	*0.18*	*0.46*	*0.28*	*0.26*	*0.18*	*0.10*
24	*0.37*	*0.73*	*0.25*	*0.19*	*0.16*	*0.03*
25	*0.39*	*0.50*	*0.23*	*0.16*	*0.13*	*0.09*
26	*0.36*	*0.54*	*0.32*	*0.23*	*0.06*	*0.03*
27	*0.50*	*1.00*	*0.26*	*0.10*	*0.08*	*0.06*
28	*0.41*	*0.66*	*0.30*	*0.19*	*0.06*	*0.04*
29	*0.41*	*0.66*	*0.30*	*0.19*	*0.06*	*0.04*
30	*0.49*	*0.74*	*0.22*	*0.18*	*0.07*	*0.04*
31	*0.62*	*1.00*	*0.17*	*0.09*	*0.07*	*0.05*
32	*0.62*	*0.91*	*0.25*	*0.08*	*0.03*	*0.02*
33	*0.32*	*0.30*	*0.33*	*0.15*	*0.12*	*0.08*
34	*0.26*	*0.16*	*0.32*	*0.26*	*0.08*	*0.08*
35	*0.36*	*0.11*	*0.28*	*0.19*	*0.10*	*0.07*
36	*0.35*	*0.45*	*0.25*	*0.19*	*0.13*	*0.08*
37	*0.41*	*0.70*	*0.25*	*0.20*	*0.13*	*0.01*
38	*0.49*	*0.48*	*0.23*	*0.16*	*0.07*	*0.05*
39	*0.13*	*0.29*	*0.29*	*0.25*	*0.20*	*0.13*
40	*0.44*	*0.03*	*0.25*	*0.15*	*0.12*	*0.04*
41	*0.34*	*0.13*	*0.28*	*0.22*	*0.11*	*0.05*
42	*0.32*	*-0.38*	*0.31*	*0.14*	*0.13*	*0.10*
43	*0.21*	*-0.02*	*0.33*	*0.28*	*0.11*	*0.07*
44	*0.24*	*0.01*	*0.31*	*0.26*	*0.14*	*0.05*
45	*0.25*	*-0.18*	*0.28*	*0.22*	*0.14*	*0.11*
46	*0.23*	*0.28*	*0.28*	*0.23*	*0.17*	*0.09*
47	*0.23*	*0.26*	*0.32*	*0.28*	*0.14*	*0.03*
48	*0.30*	*-0.22*	*0.29*	*0.17*	*0.16*	*0.08*
49	*0.34*	*0.34*	*0.20*	*0.18*	*0.17*	*0.11*
50	*0.17*	*0.17*	*0.32*	*0.24*	*0.19*	*0.08*
51	*0.19*	*0.37*	*0.34*	*0.16*	*0.15*	*0.06*
52	*0.23*	*0.17*	*0.28*	*0.25*	*0.14*	*0.10*
53	*0.34*	*0.46*	*0.23*	*0.22*	*0.15*	*0.06*
54	*0.32*	*-0.31*	*0.29*	*0.22*	*0.13*	*0.04*
55	*0.12*	*-0.09*	*0.29*	*0.25*	*0.20*	*0.14*
56	*0.32*	*0.51*	*0.27*	*0.16*	*0.15*	*0.10*
57	*0.20*	*0.11*	*0.34*	*0.25*	*0.13*	*0.08*
58	*0.10*	*0.11*	*0.37*	*0.26*	*0.14*	*0.13*
59	*0.10*	*-0.02*	*0.39*	*0.28*	*0.16*	*0.07*
60	*0.14*	*0.26*	*0.28*	*0.24*	*0.20*	*0.14*

0.12	*0.21*	*0.24*	*0.25*	*0.49*	*0.18*	**61**	*0.04*	*0.11*	*0.26*	*0.36*	*-0.04*	*0.14*	**29**
0.03	*0.23*	*0.28*	*0.43*	*-0.09*	*0.03*	**62**	*0.04*	*0.07*	*0.20*	*0.22*	*0.65*	*0.47*	**30**
0.15	*0.20*	*0.21*	*0.26*	*0.26*	*0.18*	**63**	*0.09*	*0.13*	*0.18*	*0.31*	*0.26*	*0.29*	**31**
							0.04	*0.05*	*0.19*	*0.29*	*1.00*	*0.43*	**32**

يتضح من نتائج الجدول 1 أنّ أعلى معامل صعوبة يساوي (0.77) للفقرة رقم (7)، بينما كان أقل معامل صعوبة محسوب يساوي (0.03) للفقرة رقم (62)، أما أعلى معامل تمييز فكان يساوي (1.00) صحيح للفقرات رقم (27،2، 32)، بينما كان أقل معامل تمييز يساوي (-0.38) وذلك للفقرة رقم(43). وقد كان أكبر قيمة للممو القوي تساوي (0.48) للفقرة (14)، وأقل قيمة للممو القوي تساوي (0.12) للفقرة(4). وقد تم حذف (9) فقرات نظراً لتدني معاملات تمييزها (تميزها سالب) ليصبح الاختبار بصورته النهائية مكون من (54) فقرة.

سابعا: بعد الخروج بالصورة النهائية للاختبار. تم إعداد نماذج الاختبار الأربعة: نموذج الاختبار الأول والذي تحوي كل فقرة من فقراته خمسة بدائل، بحيث يكون موقع الممو القوي قريب من الإجابة الصحيحة(النموذج الأول)، أما النموذج الثاني فيحوي خمسة بدائل يكون موقع الممو القوي بعيداً عن الإجابة الصحيحة (النموذج الثاني). ومن ثمّ قام الباحثان بحذف أضعف مموهين من خلال نتائج فاعلية المموهات المبينة في الجدول 1 لكل فقرة من فقرات الاختبار، ليتم بذلك تشكيل النموذج الثالث والرابع، فالنموذج الثالث يحوي ثلاثة بدائل بحيث يكون الممو القوي قريب من الإجابة الصحيحة (النموذج الثالث)، بينما يحوي النموذج الرابع لكل فقرة من فقراته ثلاثة بدائل بحيث يكون الممو القوي بعيداً عن الإجابة الصحيحة (النموذج الرابع).

ثامناً: تم إعداد ورقة تعليمات الاختبار، وكذلك ورقة الإجابة النموذجية لكل نموذج من نماذج الاختبار.

جمع وتحليل البيانات

لتحقيق الهدف من الدراسة، قام الباحثان بتطبيق نماذج الاختبار عشوائيا على أفراد عينة الدراسة والبالغ عددهم 2123 طالبا وطالبة, تغيب منهم 11 طالبا وطالبة عن الاختبار بواقع نموذج واحد لكل طالب. وبعد تطبيق نماذج الاختبار على أفراد عينة الدراسة، تمّ فرز كل نموذج على حده، وتصحيحه، ثمّ أدخلت البيانات إلى ذاكرة الحاسوب، واستخدم البرنامج الإحصائي (*SPSS*)، وبرنامج (*Bilog- Mg3*) لإجراء التحليلات الإحصائية اللازمة للإجابة عن أسئلة الدراسة، وقد تمت عملية التحليل وفق المراحل التالية:-

أولا: التحقق من افتراضات النموذج

تم التحقق من افتراضات النموذج اللوجستي ثلاثي المعلمة وهو أحادية البعد لما له من أثر على دقة التقديرات، وذلك باستخدام التحليل العاملي لكل نموذج من نماذج الاختبار الأربعة باستخدام طريقة المكونات الرئيسة(*Principle Component Analysis*). ويظهر الجدول 2 قيم الجذر الكامن(*Eigen Values*) ونسب التباين المفسر(*Explained Variance*) للعامل الأول والثاني وناتج قسمة الجذر الأول على العامل الثاني في نماذج الاختبار الأربعة.

الجدول 2: قيم الجذر الكامن ونسب التباين المفسر للعامل الأول والثاني وناتج قسمة قيمة الجذر الأول على العامل الثاني في التوزيعات المختلفة

نماذج الاختبار		العامل الأول	العامل الثاني	ناتج القسمة
الأول	**الجذر الكامن**	*10.512*	*2.033*	5.171
	التباين المفسر	*20.215*	*3.910*	
الثاني	**الجذر الكامن**	*10.519*	*2.260*	4.654
	التباين المفسر	*20.230*	*4.347*	
الثالث	**الجذر الكامن**	*10.443*	*2.377*	4.393
	التباين المفسر	*20.083*	*4.570*	
الرابع	**الجذر الكامن**	*10.433*	*2.313*	4.511
	التباين المفسر	*20.063*	*20.063*	

يتضح من النتائج الواردة في الجدول 2 بأن جميع نسب التباين المفسر من العامل الأول كانت أعلى من 20 لكل نماذج الاختبار، وأن ناتج قسمة الجذر الكامن للعامل الأول على العامل الثاني اكبر من 2، مما يشير إلى وجود سمة سائدة للاختبار؛ وهذا

يعني تحقق افتراض أحادية البعد للاختبار (*Hatti, 1984 ; Reckase, 1997*).

ثانيا: مطابقة البيانات للنموذج

بهدف التحقق من درجة مطابقة الأفراد والفقرات للنموذج اللوجستي ثلاثي المعلمة، تم استخدام برنامج Bilog-Mg 3 لتحليل البيانات لكل نموذج من نماذج الاختبار الأربعة، حيث أظهرت نتائج التحليل مطابقة جميع الطلاب في نموذجي الاختبار الأول والثالث (5 بدائل- المموه القوي قريب، 3 بدائل- المموه القوي بعيد) والبالغ عددهم على الترتيب (534، 543) طالباً وطالبة، وعدم مطابقة فردين من طلاب عينة الدراسة ضمن نموذج الاختبار الرابع (3 بدائل- المموه القوي بعيد) والبالغ عددهم (519) طالباً وطالبة، وكذلك عدم مطابقة طالباً واحداً من طلاب عينة النموذج الثاني (5 بدائل- المموه القوي بعيد) والبالغ عددهم (516) طالباً.

وفيما يتعلق بفحص مطابقة الفقرات للنموذج المستخدم، أعيد التحليل باستخدام برنامج Bilog-Mg 3 بعد حذف استجابة الطالب غير المطابقة للنموذج، وبناءً على مؤشر مربع كآي عند مستوى الدلالة($\alpha = 0.01$)، الذي يعتبر الفقرة غير مطابقة للنموذج إذا كان مقدار الاحتمالية لهذه الفقرة أقل من (0.01)، حيث أظهرت نتائج التحليل عدم مطابقة الفقرات ذات الأرقام 5، 14، 20، 21 في نموذج الاختبار الثالث (3 بدائل- المموه القوي قريب)، وعدم مطابقة الفقرات ذات الأرقام 1، 21، 23، 47، 53 في نموذج الاختبار الأول (5 بدائل- المموه القوي قريب)، وعدم مطابقة الفقرات ذات الأرقام 8، 16،13،24، 44، 47، 53 في نموذج الاختبار الرابع (3 بدائل- المموه القوي بعيد)، وأخيراً عدم مطابقة الفقرات ذات الأرقام 1، 21، 23، 44، 47،53 في نموذج الاختبار الأول (5 بدائل- المموه القوي بعيد).وبعد حذف جميع الفقرات المشتركة في جميع نماذج الاختبارات الأربعة والبالغ عددها (13) فقرة، أصبح الاختبار بنماذجه الأربعة مكون من (41) فقرة.

المعالجة الإحصائية

من أجل الإجابة عن أسئلة الدراسة، تم استخدام المعالجات الإحصائية التالية:-

- إيجاد معاملات الصعوبة، التمييز، التخمين لكل نموذج من نماذج الاختبار الأربعة حسب متغيري: عدد بدائل الفقرة وموقع المموه القوي باستخدام النموذج ثلاثي المعلمة.
- استخدام تحليل التباين الثنائي. واختبار M للكشف عن الفروق في معاملات الثبات الامبريقي لنماذج الاختبار.

عرض النتائج ومناقشتها

للإجابة عن السؤال الأول تم تقدير قيم معالم الفقرات (الصعوبة، التمييز، التخمين) باستخدام النموذج اللوجستي ثلاثي المعلمة لكل نموذج من نماذج الاختبار الأربعة حسب متغيري عدد البدائل وموقع المموه القوي, وتم أيضا حساب المتوسطات الحسابية والانحرافات المعيارية لقيم هذه المعالم. ويوضح الجدول 3 قيم معالم الفقرات والمتوسطات الحسابية لكل نموذج من نماذج الاختبار.

الجدول 3: قيم معالم الفقرات (الصعوبة، التمييز، التخمين) والمتوسطات الحسابية والانحرافات المعيارية لقيم معالم الفقرات لكل نموذج من نماذج الاختبار حسب متغيري عدد البدائل وموقع المموه القوي

رقم الفقرة	نموذج الاختبار											
	النموذج الأول 5-بدائل المموه القوي قريب			النموذج الثاني 5-بدائل المموه القوي بعيد			النموذج الثالث 3-بدائل المموه القوي قريب			النموذج الرابع 3-بدائل المموه القوي بعيد		
	الصعوبة	التمييز	التخمين	الصعوبة	التمييز	التخمين	الصعوبة	التمييز	التخمين	الصعوبة	التمييز	التخمين
1	0.148	3.070	0.189	0.151	2.472	0.182	0.290	1.269	0.486	0.229	1.890	0.409
2	0.077	2.345	0.184	-0.041	2.020	0.375	0.228	1.712	0.273	0.186	2.267	0.205
3	-0.110	1.770	0.380	-0.789	1.471	0.473	-0.859	1.281	0.491	0.244	1.901	0.268
4	0.213	2.054	0.267	-0.131	4.829	0.500	1.517	3.264	0.164	-0.136	2.622	0.358
5	0.763	1.855	0.263	-2.043	1.767	0.488	-2.328	1.450	0.495	-0.016	3.268	0.324
6	0.661	2.746	0.194	0.813	2.388	0.244	-0.064	2.134	0.297	1.316	3.153	0.133
7	0.066	1.945	0.367	-1.196	2.210	0.430	1.297	1.266	0.412	-1.797	2.500	0.467
8	0.024	3.930	0.352	-0.667	1.738	0.471	-1.101	2.611	0.437	0.191	1.789	0.391
9	-0.312	4.235	0.500	-0.848	3.277	0.309	-0.589	1.480	0.476	0.858	2.482	0.250
10	1.271	3.189	0.139	0.200	3.026	0.408	-0.881	2.716	0.412	-0.187	2.302	0.336
11	-0.153	3.312	0.196	-0.868	1.105	0.446	0.611	2.976	0.242	-0.498	2.932	0.289
12	0.236	1.942	0.407	0.184	1.314	0.302	0.318	2.956	0.396	-1.201	2.248	0.447
13	0.829	2.431	0.250	-0.304	2.859	0.309	-0.912	1.325	0.458	-0.787	1.728	0.452
14	0.772	1.308	0.376	1.056	1.747	0.296	0.606	1.553	0.479	-0.918	3.210	0.339
15	-1.251	2.194	0.396	1.324	2.266	0.261	-0.312	3.013	0.297	0.689	4.205	0.263
16	0.057	2.962	0.199	-0.645	2.163	0.429	1.070	1.735	0.323	0.222	2.769	0.424

رقم الفقرة	نموذج الاختبار											
	النموذج الأول 5-بدائل المموه القوي قريب			النموذج الثاني 5-بدائل المموه القوي بعيد			النموذج الثالث 3-بدائل المموه القوي قريب			النموذج الرابع 3-بدائل المموه القوي بعيد		
	الصعوبة	التمييز	التخمين	الصعوبة	التمييز	التخمين	الصعوبة	التمييز	التخمين	الصعوبة	التمييز	التخمين
17	-0.844	1.664	0.433	-0.304	3.383	0.239	0.102	1.846	0.388	-1.000	1.031	0.428
18	-0.906	3.246	0.314	0.360	3.182	0.270	-0.336	1.740	0.349	0.567	1.298	0.395
19	0.589	3.731	0.250	-1.697	1.712	0.500	1.638	1.347	0.366	1.311	2.449	0.260
20	0.106	2.520	0.408	0.017	1.164	0.468	-1.797	1.259	0.500	0.327	1.816	0.418
21	0.153	1.336	0.310	-1.614	2.596	0.436	-0.688	2.435	0.394	0.393	2.985	0.278
22	0.580	1.497	0.402	1.229	3.818	0.181	1.514	2.750	0.248	1.281	1.269	0.290
23	1.070	1.534	0.294	0.665	3.206	0.250	1.519	2.427	0.144	-0.728	3.398	0.381
24	0.120	1.643	0.362	1.474	1.117	0.262	0.249	1.241	0.469	1.250	4.748	0.183
25	-0.746	2.228	0.411	-1.154	2.544	0.362	-0.949	2.549	0.423	1.150	2.544	0.149
26	-0.692	1.106	0.494	1.349	4.545	0.169	0.461	0.981	0.469	1.404	1.157	0.256
27	-0.658	3.881	0.500	0.169	2.377	0.409	0.190	2.557	0.347	-1.129	3.162	0.376
28	-1.620	2.951	0.431	0.492	2.105	0.277	0.069	2.761	0.324	0.227	2.456	0.280
29	0.631	2.908	0.250	0.821	2.606	0.262	1.448	2.106	0.187	0.107	2.414	0.401
30	0.906	2.036	0.339	0.337	1.072	0.441	0.388	3.557	0.218	0.494	1.803	0.269
31	0.200	2.108	0.302	-0.142	2.264	0.260	-0.535	3.154	0.300	0.924	2.102	0.268
32	1.061	3.227	0.159	0.096	4.032	0.286	1.139	1.254	0.335	0.263	0.921	0.427
33	-0.957	3.717	0.405	1.167	1.950	0.188	0.414	3.115	0.262	-0.116	2.266	0.287
34	0.186	2.482	0.285	1.410	3.304	0.206	0.301	1.385	0.375	0.072	4.455	0.290
35	1.397	4.097	0.171	0.550	2.603	0.196	1.110	2.102	0.233	1.135	2.228	0.188
36	0.037	2.392	0.390	-0.491	3.323	0.207	1.769	1.301	0.337	0.013	1.419	0.367
37	0.880	2.347	0.285	0.429	2.963	0.190	-1.322	0.803	0.500	0.426	2.978	0.196
38	-0.134	2.431	0.287	0.951	1.832	0.229	1.312	2.513	0.239	0.370	3.802	0.216
39	0.064	5.600	0.304	0.826	3.267	0.195	1.423	1.968	0.301	0.786	3.010	0.148
40	0.386	1.303	0.462	-0.034	4.436	0.500	1.632	3.646	0.182	1.401	2.596	0.129
41	1.121	1.792	0.171	0.710	3.343	0.138	-1.323	1.481	0.495	-1.153	2.562	0.401
الوسط الحسابي	**0.152**	**2.563**	**0.319**	**0.093**	**2.571**	**0.318**	**0.210**	**2.074**	**0.354**	**0.199**	**2.491**	**0.308**
الانحراف المعياري	**0.704**	**0.969**	**0.101**	**0.897**	**0.959**	**0.113**	**1.054**	**0.776**	**0.106**	**0.812**	**0.881**	**0.097**

يتضح من النتائج الواردة في الجدول 3 تباينا ملحوظا في مدى صعوبة الفقرات، حيث كان المدى في صعوبة الفقرات هو الأعلى في النموذج الثالث (3-بدائل المموه القوي قريب) الذي تراوحت قيمه من 2.328 – إلى 1.769، في حين كان الأقل في النموذج الأول (5- بدائل المموه القوي قريب) حيث تراوحت قيمه من 1.620 – إلى 1.397. بينما أشارت نتائج الجدول نفسه بأن أعلى مدى لتمييز الفقرات كان في النموذج الأول (5-بدائل المموه القوي قريب) والأقل في النموذج الثالث (3-بدائل المموه القوي قريب). وبينت النتائج كذلك بأن مدى معلمة التخمين كانت متقاربة لجميع نماذج الاختبار حسب متغيري: عدد بدائل الفقرة وموقع المموه القوي, وقد جاءت متفقة مع الاتجاه العام لصعوبة الفقرة حيث كانت معلمة التخمين للفقرة الصعبة عالية، مما يعني لجوء الأفراد ذوي القدرة المتدنية للإجابة عن مثل هذه الفقرة إلى التخمين.

إن التباين في مدى قيم معالم الفقرات اظهر تباينا ملحوظا في الأوساط الحسابية لقيم هذه المعالم , ولمعرفة الاختلافات بين معالم الفقرات حسب متغيري عدد بدائل الفقرة وموقع المموه القوي. فقد تم استخدام تحليل التباين الثنائي (*two way ANOVA*)، ويوضح الجدول 4 نتائج تحليل التباين الثنائي لمعرفة دلالة الفروق في معالم الفقرات لكل نموذج من نماذج الاختبار تبعا لمتغيري: عدد بدائل الفقرة وموقع المموه القوي

الجدول 4: نتائج تحليل التباين الثنائي (Two Way ANOVA) للفروق بين المتوسطات الحسابية لتقديرات معالم الفقرات حسب متغيري: عدد بدائل الفقرة، وموقع المموه القوي والتفاعل بينهما.

معالم الفقرات	مصدر التباين	مجموع المربعات	درجات الحرية	وسط المربعات	قيمة *f*	قيمة الاحتمال
الصعوبة	عدد بدائل الفقرة	.0500	1	.0500	.0650	.7990
	موقع المموه القوي	.2790	1	.2790	.3640	.5470

	التفاعل	.0230	1	.0230	.0310	.8610
	الخطأ	122.741	160	.7670		
	الكلي	2123.09	163			
التمييز	عدد بدائل الفقرة	71.85	1	71.85	2.2945	2.130
	موقع المموه القوي	3.312	1	83.312	4.0933	.044*0
	التفاعل	81.71	1	81.71	2.1229	.1470
	الخطأ	3129.47	160	.80920		
	الكلي	60136.3	163			
التخمين	عدد بدائل الفقرة	3.020	1	3.020	2.069	.1520
	موقع المموه القوي	7.000	1	7.000	7.590	.4410
	التفاعل	1.020	1	1.020	71.91	8.160
	الخطأ	61.74	160	1.010		
	الكلي	71.79	163			

* دال إحصائيا عند مستوى الدلالة $\alpha = 0.05$

تشير النتائج الواردة في الجدول 4 إلى أنه لا توجد فروق ذات دلالة إحصائية عند مستوى الدلالة ($\alpha = 0.05$) بين متوسطات معالم الصعوبة للفقرات تعزى لعدد بدائل الفقرة وموقع المموه القوي أو التفاعل بينهما. ويمكن تفسير عدم وجود الفروق إلى أنّ طبيعة التلاعب في المموه القوي، وعدد بدائل فقرة الاختبار، وكذلك التفاعل بينهما، لم تكن بمثابة عامل كافي لجعل معلمة الصعوبة لفقرات الاختبار تختلف وفق نماذج الاختبار، كون متون هذه الفقرات واحدة في كل النماذج، والتلاعب فقط في موقع المموه القوي وعدد بدائل الفقرة، وبالتالي فالمعرفة واحدة بطبيعة المشكلة التي يطرحها متن فقرة الاختبار في هذه النماذج، بمعنى أنّ إجابات الطلبة على هذه النماذج كانت متقاربة، ناهيك عن استغلال بعض الطلبة لبعض المؤشرات أو المفاتيح في بنية الفقرة، أدى إلى عدم وجود الفروق الدالة إحصائية، على الرغم من جهل بعض المفحوصين لمحتوى الفقرة، أو افتقارهم للمعرفة اللازمة للإجابة عنها، مما جعل تأثير موقع المموه القوي سواء كان قريب أو بعيد عن الإجابة الصحيحة وكذلك عدد بدائل فقرة الاختبار تعطي تقديرات متقاربة لمعلمة الصعوبة، بمعنى أنه لم يلعب موقع المموه القوي وعدد بدائل فقرة الاختبار دوراً بارزاً في اختلاف تقديرات معلمة الصعوبة لفقرات نماذج الاختبار.

كما يعزو الباحثان عدم وجود الفروق الدالة إحصائيا بين تقديرات معلمة الصعوبة تبعاً لاختلاف موقع المموه القوي وعدد بدائل فقرة الاختبار، إلى أنّ الطلبة في مادة الرياضيات يحاولون الوصول إلى الإجابة الصحيحة بطريقة علمية مبنية على خطوات متسلسلة في حل المسألة، دون النظر إلى عدد البدائل وموقع المموه القوي لهذه الفقرة.

وقد اتفقت نتائج هذه الدراسة الخاصة بموقع المموه القوي مع نتائج دراسة امبوسعيدي (2003) التي أظهرت وفق منظور النظرية التقليدية في القياس عدم وجود فروق ذات دلالة إحصائية لموقع المموه القوي على معلمة الصعوبة، بينما تعارضت هذه النتائج مع نتائج دراسة امبوسعيدي والعفيفي(2004)، والتي أظهرت وجود فروق ذات دلالة إحصائية لموقع المموه القوي على معلمة الصعوبة.

وتتفق نتائج هذه الدراسة فيما يتعلق بعدد البدائل مع ما توصلت إليه دراسة الشريفين وطعامنة (2009)، ودراسة شيزيزكا وتاكيوشي وياشيما ويوشيزاوا (shizuka, takeuchi, yashima & yoshizawa; 2006)، إذ أشارت نتائج هذه الدراسات وفق منظور النظرية التقليدية في القياس إلى عدم وجود فروق دلالة إحصائية في تقديرات معلمة الصعوبة لفقرات الاختبار سواء كان الاختبار يحوي ثلاثة بدائل أو خمسة بدائل، بينما تعارضت نتائج هذه الدراسة مع ما توصلت إليه نتائج دراسة الكعابنة (2009)، إذ أشارت نتائج هذه الدراسة إلى وجود فروق ذات دلالة إحصائية في تقديرات معلمة الصعوبة تعزى إلى عدد البدائل، سواء كانت ثلاثة بدائل أو خمسة بدائل.

أما بالنسبة لمعاملات التمييز، فقد أشارت النتائج الخاصة بمعلمة التمييز عدم وجود فروق ذات دلالة إحصائية في تقديرات معلمة التمييز لفقرات الاختبار تعزى إلى موقع المموه القوي سواء كان قريب أو بعيد عن الإجابة الصحيحة، بينما أظهرت نتائج التحليل وجود فروق تعزى إلى عدد البدائل، لصالح نموذج الاختبار ذي الخمسة بدائل، إذ كانت تقديرات معلمة التمييز لفقرات هذا النموذج أعلى مقارنة بنموذج الاختبار ذي الثلاثة بدائل. ولم تظهر نتائج التحليل وجود فروق ذات دلالة إحصائية تعزى إلى التفاعل بين موقع المموه القوي وعدد البدائل لفقرة الاختبار.

وقد اتفقت نتائج هذه الدراسة مع ما توصلت إليه نتائج دراسة الكعابنة (2009 إذ أشارت نتائج هذه الدراسة إلى وجود فروق ذات دلالة إحصائية لتقديرات معلمة التمييز تعزى إلى عدد البدائل، لصالح النموذج ذي الثلاثة بدائل أو الخمسة بدائل. واختلفت نتائج هذه الدراسة فيما يتعلق بعدد البدائل مع ما توصلت إليه نتائج بعض الدراسات من منظور النظرية التقليدية في القياس مثل نتائج دراسة الشريفين وطعامنة (2009)، ودراسة شيزيزكا وتاكيوشي وياشيما ويوشيزاوا (shizuka, takeuchi , yashima &

(yoshizawa;2006)، إذ أشارت نتائج هذه الدراسات إلى عدم وجود فروق ذات دلالة إحصائية بين تقديرات معلمة التمييز لفقرات الاختبار تعزى إلى عدد البدائل سوء كانت بدائل فقرات الاختبار مكونة من ثلاثة بدائل أو من خمسة بدائل.

أما بالنسبة لموقع المموه القوي، فقد اتفقت نتائج هذه الدراسة مع ما توصلت إليه دراسة امبوسعيدي (2003) التي أظهرت عدم وجود فروق ذات دلالة إحصائية في تقديرات معلمة التمييز تعزى إلى موقع المموه القوي. بينما تعارضت نتائج هذه الدراسة مع نتائج دراسة امبوسعيدي والعفيفي (2004) إذ أظهرت نتائج هذه الدراسة وجود فروق ذات دلالة إحصائية في تقديرات معلمة التمييز تعزى إلى موقع المموه القوي.

ويعزو الباحثان عدم وجود الفروق الدالة إحصائياً تبعاً لاختلاف موقع المموه القوي، إلى أنّ معلمة التمييز تقسم المفحوصين إلى متمكنين وغير متمكنين، وتعتمد إجابة الطالب على فقرة الاختبار في مادة الرياضيات على قدرة المفحوص، فالمفحوص المتمكن من محتوى الفقرة، يستطيع إتباع خطوات حل علمية بصورة متسلسلة دون الوقوع في أخطاء، بعكس المفحوص غير المتمكن، كما أنّ محتوى متن الفقرة واحد عند جميع المفحوصين، وبالتالي فمهما كان موقع المموه القوي فلن يؤثر ذلك على تقديرات معلمة التمييز لفقرات نموذج الاختبار، أي لم تتأثر تقديرات معلمة التمييز باختلاف موقع المموه القوي، بينما كان الاختلاف في عدد بدائل فقرة الاختبار، ولصالح نموذج الاختبار ذي الخمسة بدائل.

وأظهرت النتائج المتعلقة بمعلمة التخمين عدم وجود فروق ذات دلالة إحصائية في تقديرات معلمة التخمين لفقرات الاختبار تعزى إلى اختلاف موقع المموه القوي سواء كان قريب أو بعيد عن الإجابة الصحيحة، وكذلك عدم وجود فروق ذات دلالة إحصائية تعزى إلى عدد البدائل سواء كان عدد البدائل ثلاثة أو خمسة بدائل، وعدم وجود فروق دالة إحصائيا تعزى إلى التفاعل بينهما. وتشير في هذا المجال دراسة شيزيزكا وتاكيوشي وياشيما ويوشيزاوا (shizuka, takeuchi, yashima & yoshizawa; 2006) إلى أن فرصة التخمين تزيد بزيادة عدد البدائل، وذلك لعدم توفر الوقت، بينما الدراسات التي بحثت في موقع المموه القوي فلم تشر أية دراسة إلى تقديرات معلمة التخمين لفقرات الاختبار.

ويمكن تفسير تقارب تقديرات قيم معلمة التخمين، وبالتالي عدم وجود الفروق الدالة إحصائياً تبعاً إلى اختلاف موقع المموه القوي، أو عدد بدائل الفقرة، أو التفاعل بينهما، من أنّ نموذج الاختبار المطبق لم يتغير، فمتون فقرات الاختبار واحدة في كل نماذج الاختبار، وإنما تغيرت فقط طبيعة ترتيب مموهات فقرة الاختبار، وكذلك اختلاف عدد بدائل فقرة الاختبار، وهذا الأمر لم يظهر فروق دالة إحصائيا في تقديرات معلمة التخمين، فالطالب قد يتعرض إلى مواقف اختباريه متنوعة، وممارسة وتدريب، تجعل إمكانية تخمينه لإجابة فقرة الاختبار أقل ما يمكن، بغض النظر عن موقع المموه القوي أو عدد بدائل فقرة الاختبار.

ولدى تتبع قيم الخطأ المعياري كمؤشر إحصائي لتقييم دقة تقديرات معالم الفقرات, وعلى الرغم من أن النتائج أظهرت بعدم وجود فروق دالة إحصائيا بين معالم الفقرات تعزى لعدد بدائل الفقرة وموقع المموه القوي إلا أن المتوسط الحسابي لأخطاء التقدير في معالم الفقرات أظهر بأن نموذج الاختبار ذي الخمسة بدائل وموقع المموه البعيد عن الإجابة الصحيحة كان الأكثر دقة في تقدير معالم الفقرات مقارنة مع بقية النماذج الأخرى. مثل هذه النتيجة تستدعي إجراء دراسة تتناول دراسة اثر عدد بدائل الفقرة وموقع المموه القوي على دقة تقدير معالم الفقرات وفق نظرية الاستجابة للفقرة.

أما بالنسبة للسؤال الثاني والمتعلق بأثر عدد بدائل الفقرة وموقع المموه القوي على تقدير معلمة القدرة للأفراد (θ), فقد تم إيجاد الأوساط الحسابية والانحرافات المعيارية لتقديرات معلمة القدرة للأفراد حسب متغيري: عدد بدائل الفقرة، وموقع المموه القوي، ويبين الجدول 5 هذه القيم.

الجدول 5: الأوساط الحسابية والانحرافات المعيارية لتقديرات معلمة القدرة للأفراد(θ) حسب متغيري: عدد بدائل الفقرة، وموقع المموه القوي لكل نموذج من نماذج الاختبار

نماذج الاختبار	العدد	الوسط الحسابي	الانحراف المعياري
الأول: 5-بدائل المموه القوي قريب	534	.0040	0.970
الثاني: 5-بدائل المموه القوي بعيد	515	.0080	0.963
الثالث: 3-بدائل المموه القوي قريب	543	.0100	0.961
الرابع: 3-بدائل المموه القوي بعيد	517	- 0.008	.9980

تشير النتائج الواردة في الجدول 5 وجود فروق ظاهرية بين المتوسطات الحسابية لتقديرات معلمة القدرة للأفراد حسب متغيري: عدد بدائل الفقرة، وموقع المموه القوي، ولمعرفة الدلالة الإحصائية لتلك الفروق الظاهرية تبعاً لمتغيري: عدد بدائل الفقرة، وموقع المموه القوي والتفاعل بينهما؛ تم استخدام تحليل التباين الثنائي (Two-Way ANOVA)، ويبين الجدول 6 نتائج تحليل التباين الثنائي.

الجدول 6: نتائج تحليل التباين الثنائي (Two-Way ANOVA) للمتوسطات الحسابية لتقديرات معلمة القدرة للأفراد حسب متغيري: عدد بدائل الفقرة، وموقع المموه القوي والتفاعل بينهما

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة f	الدلالة الإحصائية
موقع المموه القوي	0.023	1	0.023	0.026	0.872
عدد بدائل الفقرة	0.015	1	0.015	0.016	0.899
التفاعل	0.058	1	0.058	0.061	0.805
الخطأ	1994.962	2105	0.948		
الكلي	1995.058	2108			

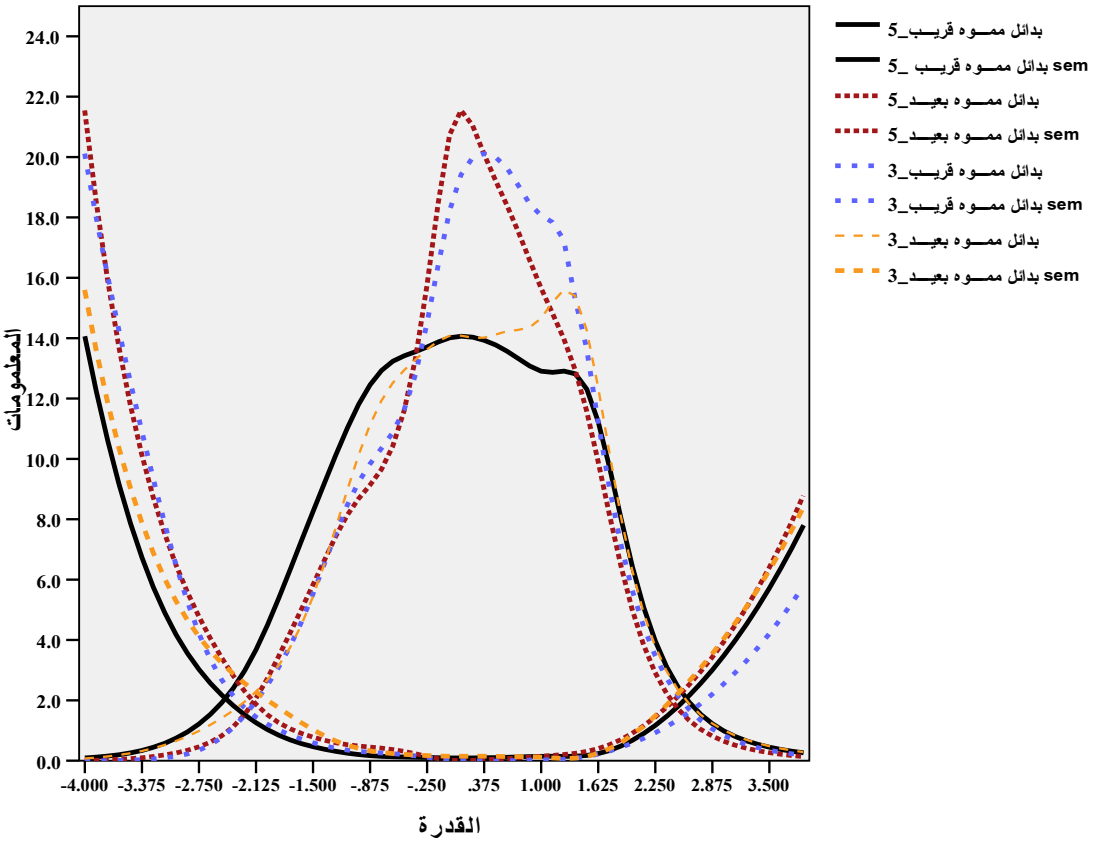
يتضح من النتائج الواردة في الجدول 6 بأنه لا يوجد فروق ذات دلالة إحصائية عند مستوى الدلالة الإحصائية ($\alpha = 0.05$) بين المتوسطات الحسابية لتقديرات معلمة القدرة للأفراد يُعزى لمتغير موقع المموه القوي (بعيد، وقريب.ولمتغير عدد بدائل الفقرة (3 بدائل، و5 بدائل). ويمكن تفسير النتائج السابقة كون نماذج الاختبار المطبقة على الطلبة هي نفسها، إذ لم تتغير متون هذه الفقرات أو حتى عدد فقرات الاختبار، وإنما كان التغيير فقط في موقع المموه القوي سواء كان قريب أو بعيد عن الإجابة الصحيحة، كما كان التغيير فقط في عدد بدائل فقرة الاختبار، وهذه التغيرات لم تؤثر على تقديرات معلمة القدرات للمفحوصين، فجميع الطلبة درسوا نفس محتوى مادة الاختبار، وبأساليب تدريس متنوعة، وأنشطة إثرائية كذلك، مما جعل تقديرات قدراتهم لا تختلف باختلاف متغيري الدراسة: موقع المموه القوي وعدد البدائل.

أما بالنسبة للسؤال الثالث والمتعلق بأثر عدد بائل الفقرة وموقع المموه القوي على دالة المعلومات والخطأ المعياري للفقرات والمقياس ككل، حيث تعتبر دالة معلومات المقياس (test information function) من المؤشرات التي يستدل منها على معامل الثبات للمقياس في نظرية الاستجابة للفقرة (Hambleton, 1994). فقد تم استخدام منحنيات دوال المعلومات للاختبار والخطأ المعياري للقياس لكل نموذج من نماذج الاختبار حسب عدد بائل الفقرة وموقع المموه الاقوى من برنامج BILOG-MG 3، ولتسهيل عملية المقارنة بين منحنيات دوال المعلومات وضعت تلك المنحنيات في شكل واحد كما يظهر في الشكل 1.

يتضح من الشكل *1* بأن أكبر كمية لدالة المعلومات كانت للنموذج الثاني (5-بدائل مموه قوي بعيد) حيث أعطى اكبر كمية للمعلومات في مدى القدرة *-0.625* , ولغاية *1.500* في حين كانت اقل كمية لدالة المعلومات للنموذج الأول (5- بدائل مموه قوي قريب) والنموذج الرابع (3-بدائل مموه قوي بعيد) مقارنة مع النموذج الثالث (3 – بدائل مموه قوي قريب). ومع النموذج الثاني (5-بدائل مموه بعيد)،وربما يعود ذلك إلى معلمة التمييز للفقرات، حيث كانت قيم التمييز للفقرات في النموذج الثاني (5-بدائل مموه قوي بعيد) هي الأعلى بينما كانت الأقل لفقرات النموذج الثالث (3-بدائل مموه قوي قريب)، فكلما زاد معلم التمييز للفقرة زادت كمية المعلومات التي تسهم بها الفقرة حول معلمة الصعوبة.

ويظهر الشكل 1 كذلك تناقص قيم منحنى الخطأ المعياري للنموذج الأول (5- بدائل مموه قوي قريب) في مدى القدرة - *1.875* و*2.125*، بينما تناقصت قيم منحنى الخطأ المعياري للنموذج الثاني في المدى *-1.625* و*2.00* ويبين الشكل كذلك بأن قيم الخطأ المعياري للنموذج الثالث (3- بدائل مموه قوي قريب) تناقصت في المدى *1.75-* و*2.25*, في حين تناقصت قيم الخطأ المعياري للنموذج الرابع (3_بدائل مموه قوي بعيد) في مدى القدرة *1.375* – و*2.00*. ومما تجدر الإشارة إليه بأن القيم الدنيا للخطأ المعياري للاختبار كانت في مدى القدرة *1.375* – و*2.00* لجميع نماذج الاختبار, وبالتالي فان فقرات الاختبار أكثر دقة في هذا المدى بغض النظر عن عدد بدائل الفقرة وموقع المموه القوي، وبذلك فهي تعطي أقصى كمية من المعلومات في هذا المدى لارتباط الخطأ المعياري ارتباطا عكسيا مع دالة المعلومات كما أشار بيكر(*Baker, 2001*).

ولأن مفهوم الثبات في نظرية الاستجابة للفقرة يرتبط بدالة معلومات الفقرة والاختبار، فقد تم إيجاد معاملات الثبات الامبريقي المنبثق من نظرية الاستجابة للفقرة لكل نموذج من نماذج الاختبار، ويوضح الجدول 7 هذه القيم.



الشكل 1: دالة معلومات الاختبار والخطأ المعياري للقياس لكل نموذج من نماذج الاختبار حسب عدد بدائل الفقرة وموقع المموه القوي

الجدول 7: قيم معاملات الثبات الامبريقي لكل نموذج من نماذج الاختبار

نموذج الاختبار	الثبات الامبريقي
الأول (5-بدائل مموه قوي قريب)	0.93
الثاني (5-بدائل مموه قوي بعيد)	0.95
الثالث (3-بدائل مموه قوي قريب)	0.94
الرابع (3-بدائل مموه قوي بعيد)	0.93

يتضح من النتائج الواردة في الجدول 7 تباين قيم معاملات الثبات الامبريقي كانت متماثلة بالنسبة للنموذجين الثاني (5-بدائل مموه قوي بعيد) والنموذج الثالث (3-بدائل مموه قوي قريب) وبالنظر إلى الشكل 1 نجد بأن كمية المعلومات للنموذجين متماثلة وهي الأكبر, بينما جاءت قيم معاملات الثبات الامبريقي للنموذجين الأول (5-بدائل مموه قوي بعيد) والرابع (3- بدائل مموه قوي بعيد) أيضا متماثلة.

ولمعرفة دلالة الفروق في معاملات الثبات الامبريقي المحسوبة من نظرية الاستجابة للفقرة، فقد تم استخدام اختبار (M) المقترح من قبل هاكستين وولن (*Hakstin & Whalen, 1976*) والذي يتبع توزيع مربع كاي بدرجات حرية (عدد المعاملات – 1) حيث كانت قيمة الإحصائي M المحسوبة تساوي *19.873* وهي أعلى من القيمة الحرجة لكاي تربيع بدرجات حرية (3)، والذي يشير إلى وجود فروق ذات دلالة إحصائية بين قيم معاملات الثبات الامبريقي،وقد جاءت تلك الفروق دالة لصالح النموذج الثاني (5-بدائل مموه قوي بعيد). ويعزو الباحثان السبب في ذلك ربما إلى تباين قدرة المفحوصين واختلاف تمييز الفقرات التي جاءت لصالح المموه القوي البعيد. وهي بذلك انسجمت مع نتائج الشكل 1 وتحليل التباين الثنائي الوارد في الجدول 4.

لقد تبين من النتائج التي توصلت إليها الدراسة الحالية بأنه لا يوجد أثر لعدد بدائل الفقرة على تقديرات معالم الفقرة في حين كان هناك اثر لموقع المموه القوي على معلمة التمييز للفقرة وعلى دالة المعلومات, وفي ضوء تلك النتائج فان الباحثان يوصيان بإجراء دراسة مماثلة بحيث يتم فيها تثبيت موقع المموه القوي أما بعد أو قبل الإجابة الصحيحة.ويوصي الباحثان كذلك بإجراء دراسات تتناول أثر عدد البدائل وموقع المموه القوي على دقة التقديرات لمعالم الفقرات وقدرة الأفراد من خلال الخطأ المعياري للتقدير.

المراجع

أمبوسعيدي، عبد الله والعفيفي، منى. (2004). أثر تغيير موقع المشتت القوي (Strong Distractor) على بعض الخصائص السيكومترية لأسئلة الاختيار من متعدد في مادة الفيزياء. **مجلة العلوم التربوية والنفسية**، جامعة البحرين، 5(2)، 169-191.

أمبوسعيدي، عبد الله. (2003). أثر تغيير موقع المشتت القوي (Strong Distractor) استخدام تقنية إبعاد البدائل الخاطئة في صعوبة الفقرات لأسئلة الاختيار من متعدد في مادة الفيزياء وتميزها. **مجلة العلوم التربوية والنفسية**، جامعة البحرين، 6(3)، 250-273.

الكعابنه، نايل. (2009). **أثر طريقة اختيار المموهات لفقرات اختبار الاختيار من متعدد في خصائصها السيكومترية وفق النظرية الكلاسيكية والنظرية الحديثة في القياس**. رسالة ماجستير غير منشورة. جامعة اليرموك، اربد، الأردن.

حمدان، محمد زياد. (1986). **تقييم التحصيل اختباراته وعملياته وتوجيهه للتربية المدرسية**. عمان: دار التربية الحديثة.

الشريفين، نضال وطعامنه، ايمان. (2009). أثر عدد البدائل في اختبار الاختيار من متعدد على تقديرات القدرة للأفراد والخصائص السيكومترية للفقرات والاختبار وفق نموذج راش في نظرية الاستجابة للفقرة. **المجلة الأردنية في العلوم التربوية**، 5(4) ,309-335.

عودة، أحمد. (2010). **القياس والتقويم في العملية التدريسية**، اربد، الأردن: دار الأمل.

فقوسه، محمد. (1994). **أثر عدد البدائل في اختبار الاختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته**. رسالة ماجستير غير منشورة. الجامعة الأردنية، الأردن.

مراد، صلاح احمد وسليمان، امين علي. (2002). **الاختبارات والمقاييس في العلوم النفسية والتربوية خطوات إعدادها وخصائصها**. القاهره: دار الكتاب الحديث

Aiken, J. (1987). Testing with Multiple-Choice Items: *Journal of* Development *in education,20,*44 - 57.

Baker, F.B.(2000). *The basics of item response theory(2nd ed).* College Park, MD: ERIC Clearing House on Assessment and Evaluation.

Blunch, N. J. (1984). Positional Bias in Multiple-Choice Questions. *Journal of Marketing Research*, *21*, 216-220.

Brown, Frederick G.(1983). *Principles of educational and psychological testing (3ed)ed)* , New York: Holt, Rinehart and Winston.

Crehan, K., Haladyna, T & Brewer, B.W.(1993). "Use of An Inclusive Option and the Optimal Number of Options for Multiple-Choice Items". *Educational and Psychological Measurement, 53*, 241-247.

Embretson , S. E. & Reise, S. P.(2000). *Item Response Theory for Psychologists.* New Jersey: Lawrence Erlbaum Associates, Publishers.

Frisbie.,D. A., and Sweeney, D.S. (1982). The relative merits of Multiple true-False achievement test. *Journal of Educational measurement. 19*, 29-35.

Gay,L.(1980). *Educational evaluation and measurement.*Ohio chales: E.Merrill.

Gronlund, N. (1982). *Constructing Achievement Test (3ed ed.)*. New Jersey: Prentice – Hall.

Gronlund, N., and Linn, R. (1990). *Measurement and Evaluation in Teaching.* New York: Macmillan publishing Co., Inc.

Hakstain, A.R & Whalen, T.E.(1976). A K-sample significance test for independent alpha coefficients. *Pyschometrika, 41,*219-231

Hambleton, R.,K.(1994). Item Response Theory: A broad psychometric frame work for measurement advances. *Psicothema,6,*535-556.

Hambleton, R. K., & Swaminathan, H. (1985). *Item Response Theory Principles and Application.* Boston: Kluwer, Nijh off Publishing.

Hattie, J.(1985).Methodology Review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement.9* ,139-164.

Lord, F. M. (1980). *Application of Item Response Theory to Practical Testing Problems*. Hillsdale, New Jersey: Lawrence Erlbaum Associates

Nitko, A. J. (2001). *Educational Assessment of Students* (3rd ed.). New Jersey: Prentice-Hall, Inc.

Reckase, M.D.(1997). The past and future of multidimensional item response theory. *Applied Psychological Measurement,1,* 25-36.

Rodriguez, M. C. (2005). Three Options Are Optimal for Multiple-Choice Items: Ameta-Analysis of 80 years of Research. *Educational Measurement, 24* , 3-11.

Shizuka, T., Takeuchi, O., Yashima, Y., & Yoshizawa, K. (2006). A Comparison of Three- and Four-Option English Tests for University Entrance

functioning distracters in multiple-choice questions: A descriptive analysis. *BMC Medical Education*, *9,* 2-11.

Yaman , S.(2011). The optimal number of choices in multiple –choice tests: Some evidence for science and technology education. *New educational review,23,227-241*

Selection Purposes in Japan. *Language Testing, 23* , 35-57.

Tarrant, M. & Ware, J., (2010). A comparison of the psychometric of three and four- option multiple choice questions in nursing assessment. *Nurse education today, 30,539*-543.

Tarrant, Marie; Ware, James; Mohammed, Ahmed M.(2009). An assessment of functioning and Non-

الملحق 1: جدول المواصفات

الوحدة	الموضوع (النسبة)	مستويات الأهداف (النسبة)				عدد الفقرات
		معرفة (15%)	استيعاب (30%)	تطبيق (40%)	تحليل (15%)	
الاقترانات	الاقتران الحقيق (9%)	1	2	3	1	**7**
	الاقتران المتشعب (9%)	1	2	3	1	**7**
	اقتران القيمة المطلقة (9%)	1	2	3	1	**7**
	الاقتران الدرجي واقتران اكبر عدد صحيح (9%)	1	2	3	1	**7**
	تركيب الاقترانات (9 %)	1	2	3	1	**7**
	الاقتران (9%)	1	2	3	1	**7**
الدائرة والممارسات والأشكال الرباعية الدائرية (46%)	أوتار الدائرة 9(%)	1	2	3	1	**7**
	الزاويا المركزية والزاوية المحيطية (9%)	1	2	3	1	**7**
	مماسات الدائرة (9%)	1	2	3	1	**7**
	الزاوية المماسية (9%)	1	2	3	1	**7**
	الشكل الرباعي والزاوية الخارجة عنه (9%)	1	2	3	1	**7**
عدد الفقرات		**11**	**22**	**33**	**11**	**77**