المجلة الدولية للدراسات التربوية والنفسية
e-ISSN 2520-4149 , p-ISSN 2520-4130
المجلد٦- العدد١- ٢٠١٩، ص ٦٦-٨٨
https://doi.org/DOI:10.31559/EPS2019.6.1.5

رفاد للدراسات والأبحاث
www.refaad.com

أثر موقع البديل الصحيح في اختبار اختيار من متعدد على دقة تقدير معالم الفقرة وفق النموذج ثنائي المعلمة لنظرية الاستجابة للفقرة

أشواق ضيف الله سليم المرواني
ماجستير في القياس والتقويم - جامعة تبوك- المملكة العربية السعودية
A.ashwaqalmarwani@hotmail.com

شاهر خالد سليمان
أستاذ مشارك في القياس والتقويم - جامعة تبوك - المملكة العربية السعودية
s.suliman@ut.edu.sa

الملخص:

هدف البحث الحالي إلى الكشف عن أثر موقع البديل الصحيح في اختبار اختيار من متعدد ذي أربعة بدائل على دقة تقدير معلمة الصعوبة والتمييز للفقرات، وعلى تقديرات معاملات الثبات التجريبي ومعاملات ثبات الأفراد، وفق النموذج ثنائي المعلمة لنظرية الاستجابة للفقرة. ولتحقيق أهداف البحث تم بناء اختبار تحصيلي من نوع اختيار من متعدد ذي أربعة بدائل في مادة العلوم مكون في صورته النهائية من (٣١) فقرة، ومن ثلاثة نماذج لها نفس المحتوى والمتغير هو موقع البديل الصحيح، حيث النموذج الأول يكون موقع البديل الصحيح على البديلين (أ، ب)، والنموذج الثاني يكون موقع البديل الصحيح على البديلين (ج، د)، والنموذج الثالث يتوزع البديل الصحيح بشكل عشوائي على جميع البدائل الأربعة. وطبقت النماذج الثلاثة على عينة عشوائية عنقودية بلغت (٥٠٠) طالبة من طالبات الصف الثالث متوسط في مدينة ينبع. وتم استخدام البرنامج الإحصائي(SPSS) وبرنامج (Bilog-Mg3) لتحليل البيانات. وأظهرت النتائج عدم وجود فروق دالة إحصائيًا بين متوسطات دقة تقدير معلمة الصعوبة والتمييز للفقرات تعزى لموقع البديل الصحيح. وكذلك عدم وجود فروق ذات دلالة إحصائية بين متوسطات تقديرات معاملات الثبات التجريبي تعزى لموقع البديل الصحيح، بينما توجد فروق ذات دلالة إحصائية بين متوسطات تقديرات معاملات ثبات الأفراد تعزى لموقع البديل الصحيح. وفي ضوء النتائج تم تقديم عدد من التوصيات من أبرزها توعية المعلمين والمعلمات إلى ضرورة العناية بفقرات اختبار الاختيار من متعدد عند إعدادها وكذلك ضرورة العناية بتوزيع الإجابات الصحيحة على جميع البدائل بشكل عشوائي.

الكلمات المفتاحية: موقع البديل الصحيح، اختبار اختيار من متعدد، دقة تقدير معالم الفقرة، النموذج ثنائي المعلمة، نظرية الاستجابة للفقرة.

المقدمة:

تُعد الاختبارات بجميع أنواعها من الأدوات المعتمدة لقياس التحصيل المعرفي لدى الطلبة في جميع المراحل الدراسية المختلفة، و بناءً على نتائجها يتم اتخاذ العديد من القرارات التي تصُب في مصلحة الطلبة، من حيث الانتقال من مستوى تعليمي إلى أخر، أو تُستخدم لأغراض الانتقاء والتصنيف كتحديد مسار معين للطلبة بما يتناسب مع قدراتهم ومهاراتهم وغير ذلك من قرارات.

والاختبارات التحصيلية تُعد من أنسب الأدوات لتقويم نتاجات التعلم المعرفية، حيث يُعتمد عليها في تقويم تقدم الطلبة وانتقالهم من صف دراسي إلى آخر، ومن مرحلة تعليمية إلى أخرى، ومن هنا جاء الاهتمام بكيفية إعدادها وأهم الخطوات التي يمكن اتباعها للوصول إلى اختبار تحصيلي جيد (هاشم، والخليفة، ٢٠١٥، ص٩٩).

ويعرف (عودة، ٢٠١٠ ، ص٦٣) الاختبار التحصيلي بأنه طريقة منظمة لتحديد مستوى تحصيل الطلبة لمعلومات ومهارات في مادة دراسية معينة كان قد تم تعلمها مسبقا بصفة رسمية وذلك من خلال إجاباتهم على عينة من الأسئلة (الفقرات) التي تمثل محتوى المادة الدراسية.

وتتنوع طرق وضع الاختبارات بتنوع الأهداف المراد تحقيقها منها ولكنها غالبًا ما تنطلق من مخطط تفصيلي أو جدول مواصفات يحدد محتوى الاختبار، ثم يتم اتخاذ قرارات متعلقة بطبيعة الأسئلة التي سيتضمنها الاختبار: أسئلة موضوعية (اختيار من متعدد أو أسئلة الصواب والخطأ أو أسئلة الإجابات القصيرة أو أسئلة المطابقة) أو أسئلة ذات إجابات مفتوحة (مقالية) (Aiken, 1998/2007, p.37).

وتعد الاختبارات التحصيلية من نوع الاختيار من متعدد أكثر أشكال التقويم انتشارًا في التربية ومما زاد من انتشار هذا النوع من الاختبارات وتفوقها على كافة أشكال الفقرات الموضوعية الأخرى وكذلك على الأسئلة المقالية، كفاءتها وتعدد استعمالاتها، إذ يمكن بواسطتها قياس أهداف بسيطة وأخرى مركبة في مختلف المواد الدراسية، وعلى اختلاف المراحل التعليمية فهي أقل تأثرًا من غيرها من الفقرات الموضوعية بعامل التخمين، وكذلك إلى الاستعدادات الخاصة للاستجابة بطرق معينة (أبو فودة، وبني يونس، ٢٠١٢، ص١٥).

ويشير (عبابنة، ٢٠٠٩، ص٦٨) إلى أن الفقرات من نوع الاختيار من متعدد تتألف من متن (stem) أو مقدمة الفقرة وبدائل (alternative) ويطلب من الطالب اختيار الجواب الصحيح أو الجواب الأقرب إلى الصحة، وتعتبر فقرات الاختيار من متعدد من أفضل أنواع الفقرات وأكثرها مرونة فهي لا تقتصر على قياس مستوى تذكر المعلومات بل يمكن أن توظف لقياس قدرات عقلية عليا مثل التطبيق والتحليل والتركيب.

ويؤكد (عبد الرحمن،٢٠١١، ص٢١٦) بأنه لا بد من تنويع البدائل الصحيحة بين الفقرات في محاولة لمساعدة الطالب على اختيار البديل الصحيح معتمدًا على الفهم والتمييز وليس على التخمين. ويلاحظ أن بعض المعلمين يتخذون نمطًا معينًا للبدائل الصحيحة من أجل تسهيل عملية التصحيح وفي ذلك إشارة للطالب الذكي والنبيه لمعرفة الإجابات الصحيحة لبقية الفقرات، وبالتالي لن تعكس علامته حقيقة مستوى قدراته وتكون عملية القياس في هذه الحالة غير دقيقة. ولذلك لا بد من الاختيار العشوائي للبدائل الصحيحة وعدم اتخاذ نمط واحد لها إلا إذا اقتضى الأمر ذلك لأسباب منطقية.

ولقد أسفرت جهود العلماء في مجال القياس والتقويم إلى ظهور بعض الاتجاهات الحديثة، ومن هذه الاتجاهات نظرية الاستجابة للفقرة Item Response Theory (IRT) أو نظرية السمات الكامنة Latent Trait Theory (LTT) أو نظرية منحنى خصائص الفقرة Item Characteristic Curve Theory (ICCT). وقد أطلق على كل هذه الاتجاهات نظرية الاختبارات المعاصرة Modern Test Theory (MTT) (الفقي، ٢٠١٤، ص ٢٤٣).

ونظرية الاستجابة للفقرة تفترض تفسير أداء الطلبة تبعًا للسمات التي لا تعتمد على المجموعة وإنما تعتمد على استجابات الطالب لفقرات الاختبار، وتُقاس السمة بطريقة غير مباشرة من خلال إجابة الطالب على فقرات الاختبار. وهذه النظرية تسمح بقيم مختلفة لخصائص الفقرة من أن توضع وحدات القياس على متصل القدرة موضع القياس. فمثلًا يمكن للفقرات أن توضع تبعًا لخاصية الصعوبة أو التمييز أو الاثنين معًا (الصراف، ٢٠١٤، ص٤٠٧).

ويذكر بيكر (Baker, 2001/2010) أن نظرية الاستجابة للفقرة تنطلق منها ثلاثة نماذج رياضية لمنحنى خصائص الفقرة، وهذه النماذج تزودنا بمعادلة رياضية للعلاقة بين احتمالية الاستجابة الصحيحة للفقرة ودرجات القدرة. وكل نموذج منها يستخدم معلم Parameter أو أكثر. وهذه النماذج هي: النموذج ثلاثي المعلمة ويعتمد على معلمة الصعوبة والتمييز، والتخمين. والنموذج ثنائي المعلمة ويعتمد على معلمة الصعوبة والتمييز. بالإضافة إلى النموذج أحادي المعلمة والذي يعتمد على معلمة الصعوبة فقط .

ويعتمد مستوى أداء الطالب في الاختبار على خصائص الاختبار، وخصائص الطالب، ومن خصائص فقرة الاختيار من متعدد المؤثرة في مستوى أداء الطالب، صعوبة الفقرة وقدرتها التمييزية وجاذبية المموهات، ومدى تأثرها بعامل التخمين. (Plake, Thompson & Lowry, 1981, P. 214)

ويذكر نوادنيجو ونايبي (Nwadinigwe & Naibi, 2013) أن صعوبة الفقرة تعتمد على موقع البديل الصحيح بالنسبة لبقية البدائل في اختبار الاختيار من متعدد.

وتتأثر القدرة اللازمة للإجابة على الفقرة، بعوامل فنية تعتمد على درجة الوضوح في صياغة المتن، ومدى تجانس البدائل، ودرجة علاقتها مع متن السؤال، ومدى تحرر الفقرة بشكل عام من أي خلل في صياغتها (Lane & Bull, 1987, P. 865-879).

وأجرى كيزيك (Cizek, 1994) دراسة هدفت إلى معرفة ما إذا كان هناك أثر لإعادة ترتيب بدائل الإجابات في فقرات اختبار اختيار من متعدد على مستوى صعوبة الفقرة، وطُبقت الدراسة على عينة مكونة من (٧٥٩)من المختبرين في المجلس الطبي التخصصي بالولايات المتحدة الأمريكية، وتكون الاختبار من (٢٠)فقرة من نوع اختيار من متعدد في اختيار التأهيل الطبي. وتوصلت نتائج الدراسة إلى أن إعادة ترتيب بدائل الإجابة له أثر دال إحصائيًا على صعوبة الفقرة، ولكن من غير الممكن التنبؤ بالأداء على الفقرة من خلال إعادة ترتيب بدائل الإجابة.

كما قام ويفرك (Wevrick, 2001) بدراسة بعنوان الاستعداد الخاص للاستجابة في اختبار اختيار من متعدد إذا كان البديل الصحيح يحتل موقعًا متحيزًا (البديل الأوسط)، وطُبقت الدراسة على عينة عشوائية مكونة من (٤٦٧) طالبًا جامعيًا، وتم بناء ثلاثة نماذج من اختبار من نوع اختيار من متعدد ذي ثلاثة بدائل ويتكون من (١٠٠)فقرة، وتختلف النماذج الثلاثة بالتالي: في النموذج الأول يكون موقع البديل الصحيح في الوسط، والنموذج الثاني البديل الوسط يتضمن مموه للإجابة، والثالث موقع البديل الصحيح موزع توزيعًا عشوائيًا، وأشارت نتائج الدراسة إلى أنه إذا كان البديل الصحيح يقع في الوسط، فإن درجة الطالب على الاختبار عالية أي أن معامل صعوبة الفقرة يزداد، أما إذا كان البديل الأوسط يحتله مموه من مموهات الإجابة فإن معامل صعوبة الفقرة يزداد، أما إذا كان البديل الصحيح موزعًا عشوائيًا فإنه لا تأثير يُذكر للاستعداد الخاص للطالب.

وأجرى كل من كولاولي واولاروتيمي (Kolawole & Olarotimi, 2008) دراسة بهدف التحقق من مدى التحيز نحو الرموز الوسطى والمتطرفة في الاختبارات من أربعة خيارات، وقد استخدم الباحثين المنهج التحليلي، حيث قاما بتحليل اختبارات مجلس امتحانات أفريقيا الغربية لمدة خمس سنوات من (٢٠٠٤ – ٢٠٠٨)، وتوصلت نتائج الدراسة إلى أن واضعي الاختبارات يميلون إلى وضع البديل الصحيح في الوسط، مما يُضعف جودة الاختبارات.

كما أجرت (الصبح، ٢٠١١) دراسة هدفت إلى بيان أثر بنية فقرات الاختيار من متعدد ومستوى القدرة لدى الأفراد على دقة التقديرات لمعالم الفقرات والأفراد وفق نظرية الاستجابة للفقرة. ولتحقيق هدف الدراسة تم بناء اختبار تحصيلي من نوع الاختيار من متعدد ذي أربعة بدائل في مادة الرياضيات للصف العاشر الأساسي، وتكون من (٣٥) فقرة محكمة البناء، ثم جرى صياغة فقرات الاختبار في نموذج أخر مخالف في بنائه لقواعد صياغة الأسئلة، ومن ضمن هذه المخالفات موقع البديل المميز (البديل الأول) . وتكونت عينة الدراسة من (١٠٥٧) طالب وطالبة من طلبة الصف العاشر، وكشفت نتائج الدراسة عن وجود فروق ذات دلالة إحصائية بين متوسطات الأخطاء المعيارية في تقديرات معالم الصعوبة للفقرات وفق النموذجين الثنائي والثلاثي المعلمة، وعدم وجود فروق ذات دلالة إحصائية بين متوسطات الأخطاء المعيارية في تقديرات معالم التمييز للفقرات.

كما قام (الكناني وفيضي،٢٠١٢) بدراسة بعنوان فاعلية استخدام موقع البديل الصحيح عند ابعاد البديل الخاطئ في صعوبة الفقرات لأسئلة الاختيار من متعدد وتمييزها، وتكونت عينة الدراسة من (٦٠) طالبًا من طلبة المرحلة الثالثة لكلية التربية الرياضية، ولتحقيق هدف الدراسة تم بناء مقياس البديل الصحيح في أسئلة الاختيار من متعدد لمادة طرائق تدريس التربية الرياضية، وسعى الباحثان إلى الكشف عن أثر تغير موقع البديل الصحيح عند استخدام تقنية إبعاد البدائل الخاطئة على تحصيل الطلبة في أسئلة الاختيار من متعدد، حيث تم تقسيم عينة الدراسة إلى مجموعتين متساويتين وأجاب (٣٠) طالبًا وطالبة عن أسئلة النموذج الأول، كما أجاب (٣٠) طالبًا وطالبة عن أسئلة النموذج الثاني وأشارت النتائج إلى عدم وجود فروق معنوية بين نموذجي الاختبار من حيث الصعوبة والتمييز.

وأجرت (قاسم، ٢٠١٣) دراسة هدفت إلى الكشف عن أثر ترتيب مواقع بدائل لأسئلة الاختيار من متعدد على نسبة الإجابات الصحيحة المخمنة، ولتحقيق أهداف الدراسة تم بناء اختبار في اللغة العربية للصف الأول ثانوي، يتكون من (٥٠) فقرة منها (١٠) فقرات بأربعة بدائل بدون إجابة صحيحة، والأخرى بأربعة بدائل واحد منها صحيحة، بحيث توزع الإجابة الصحيحة على جميع المواقع بالتساوي، وطُبقت الدراسة على عينة مكونة من (٤٢٩) طالبًا وطالبة من طلبة الصف الأول ثانوي تم اختيارهم بطريقة العينة الطبقية العنقودية العشوائية، وتوصلت نتائج الدراسة إلى وجود فروق دالة إحصائيًا في نسبة تخمين البدائل كإجابات صحيحة باختلاف موقع الإجابة الصحيحة، ولصالح البديلين (ب ،ج)، وتشير هذه النتيجة إلى أن الطلبة يتجهون إلى اختيار البدائل المتوسطة، وتبين كذلك أن توزيع الإجابات الصحيحة للمفحوصين على فقرات اختبار الاختيار من متعدد غير متساو على رموز الفقرات المستخدمة.

وأجرى (الأمير، ٢٠١٦) دراسة بعنوان الأخطاء الشائعة في بناء فقرات اختبار الاختيار من متعدد وتأثيرها على الخصائص السيكومترية للاختبارات التحصيلية لدى معلمي الرياضيات بالمرحلة المتوسطة بمدينة مكة المكرمة. وتم استخدام عينة مكونه من (٥١) نموذجًا من أسئلة معلمي الرياضيات للفصل الدراسي الأول للصف الثاني متوسط، ولتحقيق هدف الدراسة قام الباحث ببناء قائمة تدقيق الأخطاء تحتوي على جميع الأخطاء في بناء فقرات اختبار الاختيار من متعدد وعددها (٢٢) خطأ، وذلك لحصر جميع الأخطاء في بناء فقرات اختبار الاختيار من متعدد من خلال نماذج أسئلة المعلمين، ثم تحديد الأخطاء الشائعة فيها والتي تكررت بنسبة ٢٠% فأكثر، وتوصلت النتائج إلى تحديد الأخطاء الشائعة في بناء فقرات الاختيار من متعدد لمعلمي الرياضيات بالصف الثاني متوسط إلى خمسة أخطاء شائعة وهي: خطأين في مدخل الفقرة، وخطأين في البدائل، وخطأً واحدًا في الإجابة الصحيحة وهو عدم توزيع الإجابة الصحيحة توزيعًا عشوائيًا.

وأجرى (سليمان والصالح، ٢٠١٧) دراسة بعنوان أثر موقع البديل الصحيح في اختبار اختيار من متعدد على تقديرات معالم الفقرات والقدرة وفق النموذج اللوجستي ثلاثي المعلمة، ولتحقيق أهداف الدراسة تم بناء اختبار تحصيلي من نوع اختيار من متعدد من أربعة بدائل مكون من (٣٥) فقرة في القياس والتقويم التربوي المقرر لطلبة دبلوم التربية العام، ولقد أشتمل على ثلاثة نماذج لها نفس المحتوى والمتغير هو موقع البديل الصحيح، بحيث إن النموذج الأول يكون فيه موقع البديل الصحيح على البديلين أ ، ب والنموذج الثاني يكون فيه موقع البديل الصحيح على البديلين ج ، د والنموذج الثالث يكون توزيع البديل الصحيح عشوائيًا على مختلف البدائل الأربعة، وطُبقت النماذج على عينة عشوائية مكونة من (١٥٤١) طالبًا وطالبة من جامعة تبوك، وتوصلت النتائج إلى وجود فروق دالة إحصائيًا بين معاملات تمييز وتخمين الفقرات والثبات التجريبي للاختبار تُعزى لموقع البديل الصحيح ولا توجد فروق دالة إحصائيًا بين معاملات الصعوبة وتقديرات معلمة القدرة تُعزى لموقع البديل الصحيح.

كما أجرى (الصالح، ٢٠١٧) دراسة بعنوان مقارنة تأثيرات تغيير موقع البديل الصحيح في اختبار اختيار من متعدد على تقديرات الصعوبة للفقرات وتقديرات القدرة لدى الأفراد باستخدام نموذج راش ولتحقيق أهداف الدراسة تم بناء اختبار تحصيلي من نوع اختيار من متعدد بأربعة بدائل في القياس والتقويم التربوي المقرر لطلبة دبلوم التربية العام مكون من (٣٤) فقرة، ولقد اشتمل على ثلاثة نماذج لها نفس المحتوى والمتغير هو موقع البديل الصحيح، بحيث إن النموذج الأول يكون فيه موقع البديل الصحيح على البديلين أ ، ب والنموذج الثاني يكون فيه موقع البديل الصحيح على

البديلين ج ، د والنموذج الثالث يكون توزيع البديل الصحيح عشوائيًا على مختلف البدائل الأربعة، وطُبقت النماذج على عينة عشوائية مكونة من (١٥٤١) طالبًا وطالبة من جامعة تبوك، ودلت نتائج الدراسة على عدم وجود فروق دالة إحصائيًا بين المتوسطات الحسابية لمعاملات صعوبة الفقرات تُعزى لموقع البديل الصحيح.

ويتضح مما سبق أهمية موقع البديل الصحيح في اختبار الاختيار من متعدد، وتأثيره على تقديرات معالم الفقرة وخصائص الاختبار، ويتميز البحث الحالي باستخدام النموذج ثنائي المعلمة على العكس من دراسة (سليمان والصالح، ٢٠١٧) والتي استخدمت النموذج ثلاثي المعلمة، وكذلك دراسة (الصالح، ٢٠١٧) والتي استخدمت النموذج راش الأحادي المعلمة. ولقد توضح لدى الباحثين عدم وجود دراسات سابقة تناولت أثر تغيير موقع البديل الصحيح وفق النموذج ثنائي المعلمة وهذا يدل على الحاجة الملحة لإجراء هذا البحث. وكذلك يتميز البحث الحالي عن الدراسات السابقة في الكشف عن أثر موقع البديل الصحيح في اختبار اختيار متعدد على دقة تقدير معالم الفقرة باستخدام الخطأ المعياري لمعالم الفقرة وهو أكثر دقة في تقدير معالم الفقرة.

مشكلة البحث:

يسعى علماء القياس والمختصين بالعلوم التربوية والنفسية للتوصل الى الموضوعية في قياس سلوك الافراد واستجاباتهم، فقد بات بناء الوسيلة التي تمكن الباحث من الوصول الى التقدير الموضوعي للسلوك من الأهداف المهمة التي ينشدها علماء النفس والتربويين .

لذا فقد نشأت فكرة الاختبارات التحصيلية نتيجة رغبة المتخصصين في التعرف على أدوات جديدة تمكنهم من الوصول الى معلومات أكثر عن مستوى أداء الطلبة موضوع القياس، تفيد في اتخاذ القرارات المتنوعة، وكان من أسباب الاهتمام بهذا النوع من الاختبارات أن معظم متخصصي القياس والتقويم التربوي كانوا يركزون في الماضي على عدد قليل من ادوات القياس، ولم يحظ بالاهتمام ما يحتاجه المعلم من ادوات قياس لتشخيص حاجات الطلبة الفردية او تقويم البرامج التعليمية الخاصة وفق النظريات الحديثة في القياس.

والتعليم بشكل عام، بحاجة الى توافر اختبارات تحصيلية تتمتع بدرجة مقبولة من الصدق والثبات، يمكن من خلالها قياس تحصيل الطلبة والتحقق من مدى اتقان الطلبة للمادة العلمية. وتمثل هذه الاختبارات أدوات مناسبة لتشخيص ومعرفة الأهداف التربوية التي تحققت، والتي لم تتحقق واقتراح الأساليب التدريسية في ضوء نتائج عملية التشخيص.

والواقع التعليمي الحالي، يعتمد بدرجة كبيرة على اختبارات التحصيل كوسيلة أساسية في تقويم الطلبة في المدارس، وهي غالباً اختبارات تقليدية يشوبها الكثير من السلبيات والأخطاء في ضوء معايير القياس والتقويم الحديث، وتبتعد من ناحية ثانية عن رسم الصورة الصحيحة لحقيقة قدرات واستعدادات ومستويات تحصيل الطالب المراد تقويمها تربوياً.

ولقد أجريت بعض الأبحاث التي تناولت أثر موقع البديل الصحيح على خصائص الاختبار وفقراته. حيث أوصت جميع هذه الأبحاث الى ضرورة إجراء المزيد من الدراسات البحثية المعمقة في هذا الخصوص وفق النظريات الحديثة في القياس، كون هذا الميدان ما زال بحاجة الى المزيد من البحث والتمحيص، وأن خصائص الاختبار نفسه وخصائص فقراته تتأثر بنوع الفقرات في الاختبار وموقع البديل الصحيح .

ومن هنا، فقد جاء البحث الحالي لمحاولة الكشف عن أثر موقع البديل الصحيح في اختبار اختيار من متعدد على دقة تقدير معالم الفقرة وفق النموذج ثنائي المعلمة لنظرية الاستجابة للفقرة ليتناول أثر موقع البديل الصحيح والذي يأخذ بالاعتبار اختلاف معالم الصعوبة والتمييز للفقرات، ولما تتمتع به هذه الاختبارات من دقة في قياس خصائص الأفراد التربوية والنفسية، والتي افتقرت إليها الأبحاث السابقة باستخدام أحد نماذج نظرية الاستجابة للفقرة وهو النموذج اللوجستي ثنائي المعلم، حيث يسمح هذا النموذج باختلاف معالم الفقرة ، إذا تختلف هذه المنحنيات في ميلها (معلمة التمييز وموقعها على متصل السمة (معلمة الصعوبة)، (Hambelton& Swaminathan,1985, Lord,1980:241)

أسئلة البحث:

١. هل توجد فروق دالة إحصائيا بين متوسطات الأخطاء المعيارية في دقة تقدير معلمة الصعوبة للفقرات في اختبار اختيار من متعدد وفق نظرية الاستجابة للفقرة تُعزى إلى اختلاف موقع البديل الصحيح؟
٢. هل توجد فروق دالة إحصائيا بين متوسطات الأخطاء المعيارية في دقة تقدير معلمة التمييز للفقرات في اختبار اختيار من متعدد وفق نظرية الاستجابة للفقرة تُعزى إلى اختلاف موقع البديل الصحيح؟
٣. هل توجد فروق ذات دلالة إحصائية بين متوسطات تقديرات معاملات الثبات التجريبي في اختبار اختيار من متعدد وفق نظرية الاستجابة للفقرة تُعزى إلى اختلاف موقع البديل الصحيح؟
٤. هل توجد فروق ذات دلالة إحصائية بين متوسطات تقديرات معاملات ثبات الأفراد في اختبار اختيار من متعدد وفق نظرية الاستجابة للفقرة تُعزى إلى اختلاف موقع البديل الصحيح؟

أهداف البحث:

يسعى البحث الحالي إلى تحقيق الأهداف التالية:

١. الكشف عن أثر موقع البديل الصحيح على دقة تقدير معلمة الصعوبة للفقرات في اختبار اختيار من متعدد وفق نظرية الاستجابة للفقرة.
٢. الكشف عن أثر موقع البديل الصحيح على دقة تقدير معلمة التمييز للفقرات في اختبار اختيار من متعدد وفق نظرية الاستجابة للفقرة.
٣. التحقق من أثر موقع البديل الصحيح على تقديرات معاملات الثبات التجريبي في اختبار اختيار من متعدد وفق نظرية الاستجابة للفقرة.
٤. التحقق من أثر موقع البديل الصحيح على تقديرات معاملات ثبات الأفراد في اختبار اختيار من متعدد وفق نظرية الاستجابة للفقرة.

أهمية البحث:

تأتي أهمية البحث الحالي من الآتي:

أ- **الأهمية النظرية:**

١. أن البحث الحالي يسعى إلى الكشف عن تأثير اختلاف موقع البديل الصحيح في اختبار اختيار من متعدد وفق النموذج ثنائي المعلمة على دقة تقديرات معالم الفقرات وكذلك الثبات التجريبي، وثبات الأفراد.
٢. يُعد هذا البحث إضافة علمية جيدة، ومحاولة الوصول إلى إطار نظري في هذا المجال حيثُ إنه لا يوجد في البحوث والدراسات السابقة -على حد علم الباحثة- دراسة عن أثر موقع البديل الصحيح على دقة تقدير معالم الفقرة وفق النموذج ثنائي المعلمة.

ب- **الأهمية التطبيقية:**

١. يوفر البحث الحالي أداة قياس يمكن الاعتماد عليها في قياس تحصيل الطالبات في مادة العلوم للصف الثالث متوسط من قبل المعلمات.
٢. مساعدة الجهات المختصة في الميدان التربوي ببناء وتطوير الاختبارات التحصيلية وتحليل وتفسير نتائجها.
٣. قد تُسهم نتائج هذا البحث في مواصلة الباحثين بطريقة التراكم المعرفي في هذا المجال من خلال القيام بالمزيد من البحوث والدراسات المشابهة على نطاق أوسع وفي مواد دراسية أخرى.

مصطلحات البحث:

- **اختبار الاختيار من متعدد:** " يتكون اختبار الاختيار من متعدد من المتن (Stem) أو النص أو الجذر حيث يشرح فيه المعلم المشكلة، ويتبعه ثلاثة بدائل (Alternatives) أو أكثر، أحد هذه البدائل هو الإجابة الصحيحة (أو الأصح أحيانًا)، وباقي البدائل عبارة عن مموهات (Distractors) يطلب من الطالب أن يبين الإجابة الصحيحة، وقد يظهر المطلوب على شكل سؤال أو على شكل جملة مفتوحة تكتمل بالإجابة الصحيحة " (عودة، ٢٠١٠، ص١٥٩).

التعريف الإجرائي لاختبار الاختيار من متعدد: هو الاختبار الذي يتكون من جملة تقريرية أو سؤال، في مادة العلوم للصف الثالث متوسط، ومتبوع بعدد من البدائل؛ حيث يختار الطالب الإجابة الصحيحة من بين هذه البدائل.

التعريف الإجرائي لموقع البديل الصحيح: هو موقع رمز الإجابة الصحيحة والتي يختارها الطلبة ذوو الأداء المرتفع أكثر من الطلبة ذوي الأداء المنخفض.

- **نظرية الاستجابة للفقرة:** العلاقة الوظيفية النظرية بين سلسلة للقدرات الكامنة عند الفرد وبين الإجابات على فقرات الاختبار (Aiken, 1998/2007, p.95).

التعريف الإجرائي لنظرية الاستجابة للفقرة: هي نظرية حديثة في مجال القياس، تعتمد على نماذج رياضية، توضح العلاقة بين القدرة و احتمال استجابة الطالب الذي يمتلك قدرة معينة استجابة صحيحة على الفقرة.

- **النموذج ثنائي المعلمة:** هو الفرق بين القدرة (θ) التي يمتلكها الطالب في الصفة التي يراد تقديرها، ودرجة صعوبة الفقرة التي يرغب الطالب في الإجابة عنها والتي يمثلها الرمز (b_i) مضروباً في درجة التمييز التي تتصف فيها الفقرة والتي يرمز لها بالرمز (a_i) (التقي، ٢٠١٣، ص٢٢).

التعريف الإجرائي للنموذج ثنائي المعلمة: هو النموذج اللوجستي الرياضي، الذي يعتمد في معادلته الرياضية على معلمين هما: معلمة الصعوبة، ومعلمة التمييز في حالة عدم وجود معلمة التخمين.

- **معلمة صعوبة الفقرة:** " نقطة على مقياس القدرة عندما يكون احتمال الاستجابة الصحيحة مساويًا لـ ٠,٥ " (Baker, 2001/2010, p.25).

التعريف الإجرائي لمعلمة صعوبة الفقرة: هي معلمة الفقرة التي يرمز لها بالرمز (b_i) وتقع على نفس مقياس القدرة؛ لذلك يُطلق عليها ايضاً معلمة الموقع Location Parameter.

- **معلمة تمييز الفقرة:** هي ميل المنحنى المميز للفقرة عند نقطة الانقلاب (علام، ٢٠١٥، ص٦٩٨).

التعريف الإجرائي لمعلمة تمييز الفقرة: هي معلمة الفقرة التي يرمز لها بالرمز (a_i) وهي قدرة الفقرة على التمييز بين الطلبة ذي القدرة المرتفعة وذي القدرة المنخفضة.

- **تقدير معالم الفقرة:** هو التقدير باستخدام دالة الأرجحية العظمى (Ayala, 2009/2017, p.47).

التعريف الإجرائي لتقدير معالم الفقرة: هو تقدير رقمي لمعلمة الصعوبة، ومعلمة التمييز في النموذج ثنائي المعلمة، عند افتراض معلومية تقدير معلمة القدرة.

- **دقة تقدير معالم الفقرة:** تعبير يشير إلى جودة الفقرات التي يميزها الاحتمالية الكبيرة في أن التقدير قريب من القيمة الحقيقية باستخدام الأخطاء المعيارية في التقدير (Baker, 2001).

التعريف الإجرائي لدقة تقدير معالم الفقرة: هو التقدير الذي يتم باختيار تقدير الخطأ المعياري الأقل.

حدود البحث:

يقتصر البحث الحالي على الحدود التالية:

- حدود موضوعية: اقتصر البحث الحالي على دقة تقدير معالم الفقرة وفق النموذج ثنائي المعلمة لاختبار من نوع اختيار من متعدد ذي أربعة بدائل، لوحدتين في مادة العلوم الوحدة الأولى (طبيعة العلم وتغيرات الأرض)، والوحدة الثانية (كيمياء المادة) للصف الثالث متوسط.
- حدود ديموغرافية: اقتصر البحث على طالبات الصف الثالث متوسط المنتظمات في المدارس الحكومية، والتابعة لإدارة التعليم في مدينة ينبع.
- حدود مكانية: تم تطبيق البحث على المدارس الحكومية للبنات، للمرحلة المتوسطة والتابعة لإدارة التعليم في مدينة ينبع.
- حدود زمانية: تم تطبيق البحث خلال الفصل الدراسي الأول من العام الدراسي ١٤٣٩هـ/١٤٤٠هـ .

الطريقة والإجراءات:

مجتمع البحث:

يتكون مجتمع البحث من جميع طالبات الصف الثالث متوسط المنتظمات في المدارس الحكومية التابعة لإدارة التعليم في مدينة ينبع للعام الدراسي ١٤٣٩هـ/١٤٤٠هـ للفصل الدراسي الأول، وذلك بعد استبعاد مدارس تحفيظ القرآن الكريم. وبلغ عدد مجتمع البحث (٢٣١١) طالبة موزعين على (٤٥) مدرسة، وذلك وفقًا للتقارير الإحصائية، والتي تم الحصول عليها من قسم التخطيط والتطوير بإدارة التعليم في مدينة ينبع خلال العام الدراسي ١٤٣٩هـ/١٤٤٠هـ.

عينة البحث:

تم اختيار عينة البحث بالطريقة العشوائية العنقودية، حيث تم اختيار عدد من المدارس عشوائيًا ثم فصل من كل مدرسة ثم شعبة من كل فصل. وبلغ عدد أفراد العينة (٥٠٠) طالبة، تم اختيارهم من (١٤) مدرسة عشوائية من مدارس البنات في مدينة ينبع كما هو موضح في الجدول رقم (١).

جدول(١): توزيع عدد أفراد عينة البحث من حيث المدارس وعدد الطالبات

م	اسم المدرسة	عدد الطالبات	النسبة المئوية من العينة الكلية
١	المتوسطة الثانية	٣٦	%٧,٢
٢	المتوسطة الرابعة	٤٠	%٨
٣	المتوسطة الخامسة	٤٤	%٨,٨
٤	المتوسطة السادسة	٤١	%٨,٢
٥	المتوسطة السابعة	٢٧	%٥,٤
٦	المتوسطة الثامنة	٤٦	%٩,٢
٧	المتوسطة التاسعة	٥٨	%١١,٦
٨	المتوسطة العاشرة	٣٤	%٦,٨
٩	المتوسطة الحادية عشر	١٧	%٣,٤
١٠	المتوسطة الثالثة عشر	٤٧	%٩,٤
١١	المتوسطة الخامسة عشر	٣٥	%٧

١٢	متوسطة أم المؤمنين عائشة رضي الله عنها	٣١	%٦,٢
١٣	متوسطة الخنساء	١١	%٢,٢
١٤	متوسطة حفصة بنت عمر	٣٣	%٦,٦
	المجموع	٥٠٠	%١٠٠

أداة البحث:

لتحقيق أهداف البحث تم بناء اختبار تحصيلي من نوع اختيار من متعدد ذي أربعة بدائل في مادة العلوم للصف الثالث متوسط في الوحدتين التاليتين: الوحدة الأولى (طبيعة العلم وتغيرات الأرض)، والوحدة الثانية (كيمياء المادة)، وقد تم تتبع الخطوات والإجراءات التالية في بناء الاختبار:

١. تحديد الغرض من الاختبار: تم بناء الاختبار لقياس تحصيل طالبات الصف الثالث متوسط ومعرفة مدى اكتسابهن للمعلومات والمهارات المطلوبة في مادة العلوم، وبالتحديد في الوحدة الأولى (طبيعة العلم وتغيرات الأرض)، والوحدة الثانية (كيمياء المادة).
٢. تحليل المحتوى الدراسي: تم تحليل المحتوى لمادة العلوم للوحدتين الأولى، والثانية، وتم صياغة الأهداف السلوكية لكل موضوع على حده، وذلك بعد الرجوع إلى كتاب العلوم للصف الثالث متوسط، وكذلك كتاب دليل المعلم، وتم صياغة (٤٢) هدفًا سلوكيًا.
٣. بناء جدول المواصفات للاختبار: حيث تم إعداد جدول المواصفات وتم فيه ربط مستويات الأهداف بمحتوى المادة الدراسية موضوع الاختبار.
٤. صياغة فقرات الاختبار: تم صياغة (٥٢) فقرة من نوع اختيار من متعدد ذي أربعة بدائل، وتم مراعاة الأسس الفنية في كتابة هذا النوع من الفقرات ومطابقتها للهدف الذي تقيسه من حيث المحتوى ومستوى الهدف.
٥. التأكد من صدق المحتوى للاختبار: وللتحقق من صدق المحتوى لأداة البحث، تم عرض الاختبار بصورته الأولية، وكذلك تحليل المحتوى، والأهداف السلوكية وجدول المواصفات، على عدد من المحكمين من أصحاب الخبرة والكفاءة، وعددهم (١٥) محكمًا من معلمات ومشرفات لمادة العلوم، وكذلك من المختصين في بناء الاختبار من مجال القياس والتقويم وغيرهم؛ وذلك من أجل أخذ وجهة نظرهم حول مدى صدق فقرات الاختبار في قياس الهدف السلوكي المحدد وكذلك يقترحون ما يرونه مناسبًا لتعديل الفقرات أو الأهداف أو الأهداف والفقرات معًا. وتم إبداء المقترحات في نموذج تحكيم الأداة، وبناءً على الاقتراحات تم إجراء تعديل في صياغة الفقرة رقم (٤١). ولقد أجمع المحكمون على أن فقرات الاختبار مناسبة وتؤدي الغرض من الاختبار، وممثلة للمجال السلوكي الذي يتم قياسه.
٦. إعداد الاختبار في صورته الأولية: تم إعداد الاختبار المكون من (٥٢) فقرة، وكذلك ورقة تعليمات الاختبار، والنموذج المخصص للإجابة.
٧. التجريب الأولي للاختبار: تم تطبيق الأداة في صورتها الأولية والمكونة من (٥٢) فقرة على عينة استطلاعية من طالبات الصف الثالث متوسط والمقدر عددهم بـ (٥٠) طالبة، حيث تم اختيار المدرسة عشوائيًا من خارج عينة البحث، وهي المدرسة الثالثة للبنات في مدية ينبع وكان الهدف من التطبيق هو التأكد من وضوح الفقرات، وكذلك للكشف عن معامل ثبات الأداة ومعاملات تمييز وصعوبة الفقرات وفق النظرية التقليدية في القياس، حيث تم ذلك بإدخال البيانات في برنامج (ITEMAN) وذلك للحصول على معامل التمييز لكل فقرة من خلال معامل الارتباط الثنائي Point Biserial (Correlation) بين درجة الفقرة والدرجة الكلية على الاختبار وإيجاد معامل الصعوبة لكل فقرة، وذلك من خلال نسبة الطالبات اللاتي أجبن عن الفقرة إجابة صحيحة إلى عدد اللاتي حاولن الإجابة عن الفقرة (أي عدد الطالبات الكلي) والنتائج تتضح كما في الجدول (٢).

جدول (٢): معاملات الصعوبة والتمييز لفقرات الاختبار للعينة الاستطلاعية وفق النظرية التقليدية

رقم الفقرة	معامل الصعوبة	معامل التمييز	رقم الفقرة	معامل الصعوبة	معامل التمييز	رقم الفقرة	معامل الصعوبة	معامل التمييز
١	٠,٥٨	٠,٢٣	١٨	٠,٣٨	٠,٢١	٣٥	٠,٦٦	٠,٠٠
٢	٠,٥٠	٠,٣٧	١٩	٠,٥٤	٠,٠٤-	٣٦	٠,٥٦	٠,٤٩
٣	٠,٨٨	٠,٢٢	٢٠	٠,٦٠	٠,٢٢	٣٧	٠,٢٠	٠,٣٣-
٤	٠,٩٦	٠,٢٤	٢١	٠,٥٤	٠,٦٩	٣٨	٠,٢٠	٠,١١-
٥	٠,٧٦	٠,٣٨	٢٢	٠,٧٦	٠,٢٤	٣٩	٠,٦٦	٠,٦١
٦	٠,٧٠	٠,٣٩	٢٣	٠,٥٠	٠,٥٧	٤٠	٠,٧٤	٠,٤٩
٧	٠,٨٠	٠,٣٤	٢٤	٠,٥٤	٠,١٥	٤١	٠,٥٤	٠,٠٨
٨	٠,٧٠	٠,٣٢	٢٥	٠,٤٠	٠,٠٢-	٤٢	٠,٥٢	٠,٣٨
٩	٠,٦٨	٠,٠٥	٢٦	٠,٦٠	٠,٣٥	٤٣	٠,٣٦	٠,٣١
١٠	٠,٢٢	٠,٢٤	٢٧	٠,٥٤	٠,٣١	٤٤	٠,٦٨	٠,٥٤
١١	٠,٣٨	٠,٠٢	٢٨	٠,٥٠	٠,١١	٤٥	٠,٤٢	٠,١٠

١٢	٠,٦٢	٠,٣٩	٢٩	٠,٦٠	٠,٥٦	٤٦	٠,٧٤	٠,٢٥
١٣	٠,٤٢	٠,٢١	٣٠	٠,٢٠	-٠,١٩	٤٧	٠,٢٠	-٠,١٠
١٤	٠,٤٤	٠,٠٢	٣١	٠,٤٦	٠,٣٤	٤٨	٠,٢٤	٠,١٠
١٥	٠,٧٤	٠,٤٤	٣٢	٠,٧٢	٠,٣٧	٤٩	٠,٤٢	-٠,١٣
١٦	٠,٦٨	-٠,٠١	٣٣	٠,٦٨	٠,٣٨	٥٠	٠,٢٤	٠,٠٨
١٧	٠,٣٦	٠,٢٣	٣٤	٠,٤٨	٠,٢٧	٥١	٠,٥٦	٠,٣٩
						٥٢	٠,٢٦	-٠,٢٥
المتوسط الحسابي لمعاملات الصعوبة			٠,٥٣		المتوسط الحسابي لمعاملات التمييز			٠,٢٢

يتضح من الجدول رقم (٢) أن معاملات الصعوبة لفقرات الاختبار قد تراوحت بين (٠,٢٠) و (٠,٩٦) بمتوسط حسابي قدره (٠,٥٣). بينما معاملات التمييز فقد تراوحت بين (-٠,٣٣) و (٠,٦٩) وبمتوسط حسابي قدره (٠,٢٢). ومن المعلوم بأن معاملات الصعوبة والتي تتراوح قيمتها بين (٠,٢٠ – ٠,٨٠) وبمتوسط قدره (٠,٥٠) يمكن أن تكون مقبولة (عوده، ٢٠١٠، ص٢٨٧). ويذكر علام (٢٠١٦، ص١١٦) بأنه إذا كانت قيمة معاملات التمييز تتراوح بين (٠,٢٠- ٠,٤٠)، فإن التمييز يكون لا بأس به. وإذا قلت عن (٠,٢٠) يكون التمييز ضعيفًا أما إذا كان قيمة معامل التمييز صفر فهذا يعني أن الفقرة غاية في السهولة أو الصعوبة، ولا تمييز بين مرتفعي ومنخفضي التحصيل.

ويؤكد النبهان (٢٠٠٤، ص٢١٦) بأن مؤشرات تحليل الفقرة هي معلومات أولية وليست مطلقة حيث يفترض بعض العاملين في مجال الاختبارات بأن معاملات الصعوبة والتمييز ثابتة أو نهائية حيث من المعلوم بأن معلومات تحليل الفقرة تتأثر بطبيعة المجموعة الخاضعة للاختبار، أو عدد الطلبة الذين طُبق عليهم الاختبار، والأسلوب التعليمي المتبع في تدريسهم، وأخطاء الصدفة. ويجب أن يكون حكمنا على جودة الفقرة مبنيًا على مدى قياسها أو عدم قياسها لهدف تعليمي ذي قيمة، وليس على حجم معاملات صعوبتها وتمييزها. ومن خلال ما تقدم وبعد مراجعة قيم معاملات صعوبة وتمييز الفقرات، يتضح لنا ضرورة حذف الفقرات التالية:

- الفقرة رقم (٤،٣) وذلك لارتفاع معامل صعوبتها، ويدل ذلك على أن هذه الفقرات غاية في السهولة.
- الفقرة رقم (٩، ١١، ١٤، ٢٤، ٢٨، ٤١، ٤٥، ٤٨، ٥٠) وذلك لانخفاض معامل تمييزها وذلك يدل على انخفاض قدرة هذه الفقرات على التمييز بين الطالبات.
- الفقرة رقم (٣٥) وذلك لأن قيمة معامل تمييزها يساوي صفر، وذلك يدل على أن الفقرة لا تُميز بين الطالبات ذوات التحصيل المرتفع والمنخفض.
- الفقرة رقم (١٦، ١٩، ٢٥، ٣٠، ٣٧، ٣٨، ٤٧، ٤٩، ٥٢) وذلك لأنها ذات قيم تمييز سالبة ويدل ذلك على أن الطالبات ذوات التحصيل المرتفع لم يجيبن على هذه الفقرات إجابة صحيحة، بينما أجاب عليها نسبة كبيرة من الطالبات ذوات التحصيل المنخفض.

لذلك تم حذف هذه الفقرات والبالغ عددها (٢١) فقرة، والاحتفاظ بالفقرات المتبقية وعددها (٣١) فقرة. أما ما يتعلق بمعامل ثبات الاختبار فقد تم الحصول عليه أيضاً من برنامج (ITEMAN) حيث تم حساب معامل الاتساق الداخلي للاختبار باستخدام معادلة ألفا لكرونباخ، وكذلك بطريقة التجزئة النصفية من خلال معادلة سبيرمان براون بالطرق الثلاثة (الفردي والزوجي، النصف الأول والنصف الثاني، التجزئة العشوائية) ويتضح ذلك من الجدول رقم (٣):

جدول (٣): قيم معاملات الثبات للاختبار للعينة الاستطلاعية

الطريقة		قيمة معامل الثبات
معامل ألفا لكرونباخ		٠,٧٧
التجزئة النصفية	الفردي والزوجي S-B Odd-Even	٠,٧٨
	النصف الأول والنصف الثاني S-B First-Last	٠,٧٦
	التجزئة العشوائية S-B Random	٠,٦٨

ويتضح من الجدول السابق، بأن قيم معاملات الثبات بالطرق المختلفة تُعد مناسبة ومقبولة وتخدم أغراض البحث.

٨. تطبيق الاختبار بصورته النهائية: بعد تحليل فقرات الاختبار للعينة الاستطلاعية وحذف الفقرات غير المناسبة، تم إعداد الاختبار في صورته النهائية والمكون من (٣١) فقرة، حيث تم إعداد ثلاثة نماذج من الاختبار لها نفس المحتوى والمتغير هو موقع البديل الصحيح، بحيث إن النموذج الأول يكون فيه موقع البديل الصحيح على البديلين (أ ، ب)، والنموذج الثاني يكون فيه موقع البديل الصحيح على البديلين (ج ، د) والنموذج الثالث يكون توزيع البديل الصحيح عشوائيًا على جميع البدائل الأربعة. وطُبقت على العينة الأساسية والبالغ عددها (٥٠٠) طالبة. وتم تصحيح نماذج الاختبار

باستخدام برنامج (Zip Grade) وتم إدخال بيانات كل نموذج على حده إلى ذاكرة الحاسوب وذلك لاستخدامها في البرمجية الإحصائية (SPSS)، وكذلك برنامج (Bilog-Mg3) من أجل تحليل بيانات البحث.

متغيرات البحث:

- المتغير المستقل: موقع البديل الصحيح وله ثلاثة مستويات.
- المتغير التابع: دقة تقديرات معالم الفقرة، والثبات التجريبي، وثبات الأفراد.

الأساليب الإحصائية المستخدمة:

- تم استخدام برنامج (ITEMAN) لحساب ما يلي:
  ١. استخراج معامل صعوبة الفقرة بالنظرية التقليدية للقياس .
  ٢. معامل الارتباط الثنائي (Point Biserial Correlation) بين الدرجة على الفقرة والدرجة الكلية للاختبار لإيجاد معامل تميز الفقرة بالنظرية التقليدية للقياس.
  ٣. معامل ثبات الاختبار بالطرق المختلفة مثل طريقة ألفا لكرونباخ (Alpha)، وطريقة التجزئة النصفية (Split-Half).
- وتم استخدام البرمجية الإحصائية (SPSS) لحساب ما يلي:
  ١. التحليل العاملي الاستكشافي للتحقق من افتراض أحادية البعد للاختبار.
  ٢. اختبار تحليل التباين الأحادي (One Way ANOVA) للكشف عن دلالة الفروق بين المتوسطات الحسابية لدقة تقدير معالم الفقرة باختلاف موقع البديل الصحيح.
- استخدام البرنامج الإحصائي (Bilog-Mg3) لحساب التالي:
  ١. معاملات صعوبة وتميز الفقرة وفق النموذج الثنائي المعلمة لنظرية الاستجابة للفقرة.
  ٢. إيجاد دقة تقدير معالم الفقرة (الصعوبة، والتمييز) وفق النموذج الثنائي المعلمة لنظرية الاستجابة للفقرة.
  ٣. اختبار مربع كاي لحسن المطابقة للكشف عن مدى مطابقة الفقرات للنموذج اللوجستي ثنائي المعلمة.

عرض ومناقشة نتائج البحث:

عرض نتائج البحث:

أولاً: النتائج المتعلقة بالتحقق من افتراضات نظرية الاستجابة للفقرة (النموذج ثنائي المعلمة).

١. أحادية البعد:

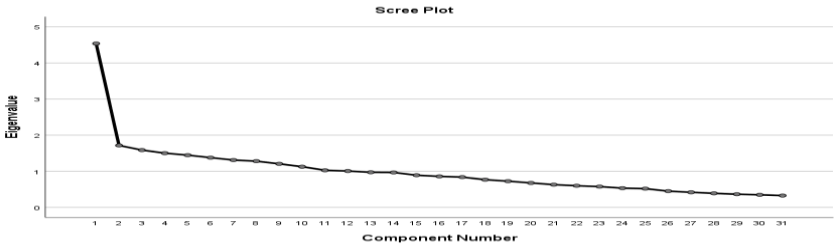
من افتراضات نظرية الاستجابة للفقرة، أن تكون استجابة الفرد على الفقرة تقيس عاملًا واحدًا وليس عدة عوامل، أي تقيس سمة معينة بذاتها. وتم التحقق من افتراض أحادية البعد باستخدام برنامج (SPSS)، من خلال التحليل العاملي الاستكشافي (Factor Analysis) بطريقة المكونات الرئيسية (Principle Component) لنماذج الاختبار الثلاثة. ويظهر في الجدول رقم (٤) قيمة الجذر الكامن (Eigen Value) ونسبة التباين المفسر (Explained Variance) للعامل الأول والثاني، وكذلك ناتج قسمة الجذر الكامن للعامل الأول على العامل الثاني، وذلك للنماذج الثلاثة.

جدول (٤): قيم الجذور الكامنة ونسب التباين المفسروالتراكمي لنماذج الاختبار الثلاثة

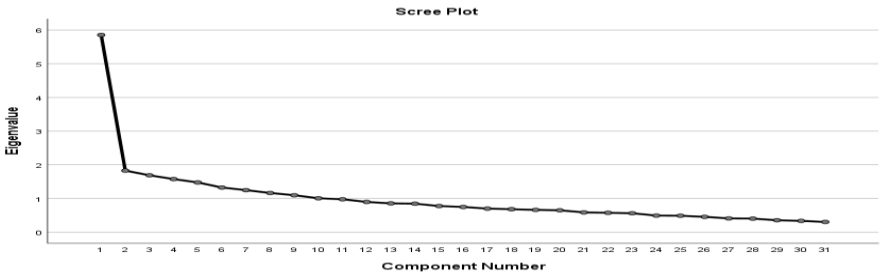
النموذج	العامل	الجذر الكامن	التباين المفسر	التباين المفسر التراكمي	الجذر الكامن للعامل الأول / الجذر الكامن للعامل الثاني
الأول (البديل الصحيح على أ و ب)	الأول	٤,٥٤	١٤,٦٣	١٣,٦٣	٢,٦٤
	الثاني	١,٧٢	٥,٥٤	٢٠,١٧	
الثاني (البديل الصحيح على ج و د)	الأول	٥,٨٥	١٨,٨٨	١٨,٨٨	٣,٢٠
	الثاني	١,٨٣	٥,٨٩	٢٤,٧٨	
الثالث (البديل الصحيح موزع عشوائياً)	الأول	٤,٩٥	١٥,٩٥	١٥,٩٥	٢,٤٤
	الثاني	٢,٠٣	٦,٥٤	٢٢,٤٩	

يتبين من الجدول رقم (٤) أن ناتج قسمة الجذر الكامن للعامل الأول على الثاني في النماذج الثلاثة كانت أكبر من (٢)، مما يدل على وجود سمة سائدة للاختبار، وبالتالي يمكن القول بتحقق افتراض أحادية البعد للاختبار (Hatti, 1984).

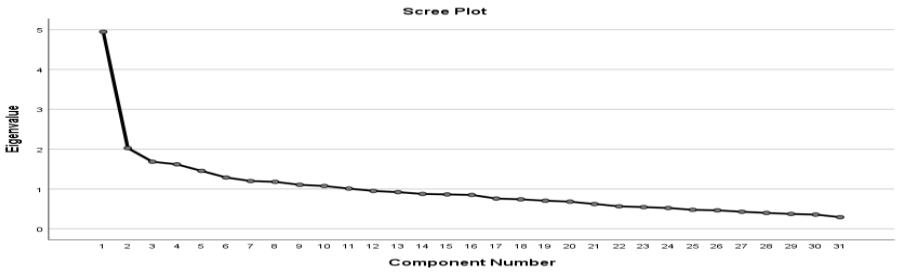
ومن خلال مخطط سكري (Scree Plot) للعوامل مع الجذور الكامنة، والذي يتضح في الأشكال رقم (١)، (٢)، (٣) يتبين أن الجذر الكامن للعامل الأول يتميز بشكل كبير عن الجذور الكامنة لبقية العوامل وذلك للنماذج الثلاثة، وهذا مؤشر ايضًا على أحادية البعد لبيانات الاختبار



شكل (١): مخطط سكري لقيم الجذور الكامنة للعوامل في النموذج الأول



شكل (٢): مخطط سكري لقيم الجذور الكامنة للعوامل في النموذج الثاني



شكل (٣): مخطط سكري لقيم الجذور الكامنة للعوامل في النموذج الثالث

ويتضح من الأشكال السابقة رقم (١)، (٢)، (٣) أن افتراض أحادية البعد قد تحقق للنماذج الثلاثة حيث يتبين من خلال الرسم البياني باستخدام (Scree Plot) أن قيمة الجذر الكامن للعامل الأول كانت مرتفعة نسبيًا عن بقية العوامل والتي تبدو متقاربة، ويتضح ذلك بشكل أكبر في النموذج الثاني للاختبار، ثم في النموذج الثالث، ثم في النموذج الأول.

٢. الاستقلال الموضعي:

ويقصد بذلك أن إجابة الطالب على الفقرة لا تتأثر بإجابته عن فقرة أخرى والعامل الوحيد المؤثر على احتمال الاجابة عن الفقرة هو مقدار القدرة(Ɵ) وخصائص الفقرة. وهذا الافتراض مكافئ لافتراض أحادية البعد، حيث يذكر هامبلتون وسوامينثان(Hambleton & Swaminathan,1985) أنه إذا تحقق افتراض أحادية البعد يعني تحقق افتراض الاستقلال الموضعي. ويتضح من خلال النتائج السابقة تحقق افتراض أحادية البعد وبالتالي تحقق افتراض الاستقلال الموضعي.

٣. التحرر من السرعة في الأداء:

ويقصد بذلك إعطاء الطلبة الوقت الكافي للإجابة على فقرات الاختبار. ولقد تم تحقق هذا الافتراض بحيث تم الانتهاء من الاختبار لجميع الطالبات وعلى كل الفقرات في الوقت المحدد وذلك للنماذج الثلاثة للاختبار، وهذا يدل على أن الوقت المحدد كان مناسبًا للإجابة على فقرات الاختبار .

ثانياً: التحقق من مطابقة الأفراد والفقرات للنموذج ثنائي المعلمة.

١. التحقق من مطابقة الأفراد للنموذج ثنائي المعلمة:

والمقصود بذلك تحديد ما إذا كان الطالب يتصرف بطريقة متوافقة مع النموذج أو لا. وبشكل أكثر تحديدًا فالأفراد غير المتطرفين لا يستبعدون من النموذج، حيث إن المفحوص المتطرف يكون قد حصل على الدرجة الخام صفر، أو الدرجة الكاملة وبالتالي يُستبعد من النماذج ويُعد غير مطابق للنموذج (Ayala,2009 /2017, p.62).

وبهدف التحقق من درجة مطابقة الأفراد للنموذج ثنائي المعلمة، تم إدخال البيانات الخاصة باستجابة أفراد عينة البحث (٥٠٠) طالبة إلى برنامج (Bilog-Mg3) لتحليل البيانات للنماذج الثلاثة للاختبار، وأظهرت نتائج التحليل عدم مطابقة طالبة واحدة فقط في النموذج الأول للاختبار (موقع البديل الصحيح على البديلين أ و ب)، وكذلك عدم مطابقة طالبة واحدة فقط في النموذج الثاني للاختبار (موقع البديل الصحيح على البديلين ج و د)، بينما تبين عدم مطابقة طالبتين في النموذج الثالث للاختبار (موقع البديل الصحيح عشوائيًا). وبالتالي تم استبعاد استجابات الطالبات غير المطابقة للنموذج ثنائي المعلمة والإبقاء على استجابات الطالبات المطابقات للنموذج وعددهم (٤٩٦) طالبة.

٢. التحقق من مطابقة الفقرات للنموذج ثنائي المعلمة:

تم فحص مطابقة الفقرات للنموذج ثنائي المعلمة باستخدام برنامج (Bilog-Mg3) باستخدام اختبار مربع كاي (x^2) عند مستوى الدلالة (٠,٠١) وذلك للنماذج الثلاثة للاختبار وكانت النتائج كالتالي:

جدول (٥): نتائج مطابقة الفقرات للنموذج ثنائي المعلمة في النموذج الأول للاختبار

رقم الفقرة	قيمة كاي تربيع	مستوى دلالة كاي تربيع	رقم الفقرة	قيمة كاي تربيع	مستوى دلالة كاي تربيع	رقم الفقرة	قيمة كاي تربيع	مستوى دلالة كاي تربيع
١	٢,٢	٠,٨٢	١١	١٤,٤	٠,٠٧	٢١	١١,١	٠,١٤
٢	٣,٥	٠,٨٤	١٢	٧,٣	٠,٣٩	٢٢	١٣,١	٠,٠٧
٣	٣,٩	٠,٧٩	١٣	١٤,٥	٠,٠٤	٢٣	١٧,٤	٠,٠٣
٤	١,٥	٠,٩١	١٤	٢,٥	٠,٦٤	٢٤	١,٦	٠,٩٥
٥	٤,٠	٠,٥٥	١٥	٤,٧	٠,٤٥	٢٥	٢,٤	٠,٨٨
٦	٣,٤	٠,٣٤	١٦	٦,٨	٠,٤٥	٢٦	١,٣	٠,٨٦
٧	٩,٧	٠,٢٩	١٧	٢,٠	٠,٩٦	٢٧	٤,١	٠,٧٧
٨	٦,٩	٠,٣٣	١٨	٣,٢	٠,٩٢	٢٨	١٠,٥	٠,١٦
٩	١٢,٦	٠,٠٨	١٩	٤,٥	٠,٧٢	٢٩	٥,٠	٠,٦٦
١٠	٢,١	٠,٨٤	٢٠	٦,٥	٠,٤٨	٣٠	٨,٢	٠,٣٢
						٣١	٢,٦	٠,٨٥

يتضح من الجدول رقم (٥) مطابقة جميع فقرات الاختبار في النموذج الأول للاختبار (موقع البديل الصحيح على البديلين أ و ب) للنموذج ثنائي المعلمة، حيث كانت قيمة مستوى دلالة مربع كاي أكبر من (٠,٠١) لجميع الفقرات.

جدول (٦): نتائج مطابقة الفقرات للنموذج ثنائي المعلمة في النموذج الثاني للاختبار

رقم الفقرة	قيمة كاي تربيع	مستوى دلالة كاي تربيع	رقم الفقرة	قيمة كاي تربيع	مستوى دلالة كاي تربيع	رقم الفقرة	قيمة كاي تربيع	مستوى دلالة كاي تربيع
١	١,٩	٠,٩٣	١١	٤,١	٠,٨٥	٢١	٧,١	٠,٤٢
٢	١,٧	٠,٩٧	١٢	٦,٣	٠,٥٠	٢٢	١,٤	٠,٩٩
٣	١,٤	٠,٩٦	١٣	١٢,٠	٠,١٠	٢٣	٦,٠	٠,٦٥
٤	٣,٩	٠,٥٧	١٤	١,٨	٠,٩٣	٢٤	١,٦	٠,٩٥
٥	٦,١	٠,٢٩	١٥	٢,٩	٠,٩٠	٢٥	٤,٨	٠,٦٨
٦	٢,٦	٠,٧٦	١٦	٢,١	٠,٩٥	٢٦	٥,٤	٠,٢٥
٧	٣,٥	٠,٩٠	١٧	٥,٢	٠,٦٤	٢٧	٨,٥	٠,٤٨
٨	٥,٥	٠,٥٩	١٨	٦,١	٠,٦٣	٢٨	٥,٥	٠,٦٠
٩	٨,٩	٠,٢٦	١٩	١,٩	٠,٩٣	٢٩	٢,٦	٠,٧٦
١٠	١,٢	٠,٩٨	٢٠	٧,٦	٠,٤٧	٣٠	٤,٨	٠,٥٨
						٣١	٤,١	٠,٥٤

يتضح من الجدول رقم (٦) مطابقة جميع فقرات الاختبار في النموذج الثاني للاختبار (موقع البديل الصحيح على البديلين ج و د) للنموذج ثنائي المعلمة، حيث كانت قيمة مستوى دلالة مربع كاي أكبر من (٠,٠١) لجميع الفقرات.

جدول(٧): نتائج مطابقة الفقرات للنموذج ثنائي المعلمة في النموذج الثالث للاختبار

رقم الفقرة	قيمة كاي تربيع	مستوى دلالة كاي تربيع	رقم الفقرة	قيمة كاي تربيع	مستوى دلالة كاي تربيع	رقم الفقرة	قيمة كاي تربيع	مستوى دلالة كاي تربيع
١	١,١	٠,٩٥	١١	٧,٠	٠,٤٣	٢١	٣,٤	٠,٨٤
٢	٦,٩	٠,٣٣	١٢	٢,٠	٠,٩٦	٢٢	٤,٠	٠,٥٥
٣	١,٨	٠,٧٦	١٣	١٠,٣	٠,١٧	٢٣	٣,٧	٠,٨١
٤	١,٢	٠,٧٦	١٤	٢,٤	٠,٧٩	٢٤	٩,٣	٠,٠٥
٥	٢,٣	٠,٥١	١٥	١,٢	٠,٩٤	٢٥	٦,٢	٠,١٩
٦	٠,٩	٠,٩٣	١٦	١٤,٣	٠,٠٥	٢٦	٤,٧	٠,٢٠
٧	٣,١	٠,٨٧	١٧	٠,٥	٠,٩٩	٢٧	١,٣	٠,٩٩
٨	٣,١	٠,٦٨	١٨	٢,٠	٠,٩٦	٢٨	٥,٥	٠,٦٠
٩	٥,٢	٠,٣٩	١٩	٩,٣	٠,١٠	٢٩	٤,٠	٠,٥٥
١٠	٣,٣	٠,٦٦	٢٠	١٠,٤	٠,١١	٣٠	١,٦	٠,٨٠
						٣١	٨,٤	٠,١٤

يتضح من الجدول رقم (٧) مطابقة جميع فقرات الاختبار في النموذج الثالث للاختبار (موقع البديل الصحيح موزع عشوائيًا) للنموذج ثنائي المعلمة، حيث كانت قيمة مستوى دلالة مربع كاي أكبر من (٠,٠١) لجميع الفقرات.

ثالثاً: النتائج المتعلقة بالإجابة عن اسئلة البحث.

السؤال الأول: هل توجد فروق دالة إحصائيا بين متوسطات الأخطاء المعيارية في دقة تقدير معلمة الصعوبة للفقرات في اختبار اختيار من متعدد وفق نظرية الاستجابة للفقرة تُعزى إلى اختلاف موقع البديل الصحيح؟

للإجابة عن هذا السؤال تم استخدام برنامج (Bilog-Mg3) لحساب الخطأ المعياري في دقة تقدير معلمة الصعوبة لفقرات الاختبار وذلك للنماذج الثلاثة للاختبار، كما هو موضح في الجدول رقم (٨) والذي يوضح كذلك تقديرات معلمة الصعوبة. وحساب المتوسط الحسابي والانحراف المعياري لدقة تقدير معلمة الصعوبة وذلك للنماذج الثلاثة للاختبار، كما هو موضح في الجدول رقم (٩) والذي يوضح أيضًا المتوسط الحسابي والانحراف المعياري لتقديرات معلمة الصعوبة.

جدول (٨): تقديرات قيم معلمة الصعوبة والأخطاء المعيارية في دقة تقديرها لفقرات الاختبار للنماذج الثلاثة

رقم الفقرة	النموذج الأول (البديل الصحيح على البديلين أ و ب)		النموذج الثاني (البديل الصحيح على البديلين ج و د)		النموذج الثالث (البديل الصحيح موزع عشوائياً)	
	معلمة الصعوبة	الخطأ المعياري	معلمة الصعوبة	الخطأ المعياري	معلمة الصعوبة	الخطأ المعياري
١	٠,٩٢-	٠,١٩٩	٠,٠٣	٠,١٩٣	٠,٧٨-	٠,٢٣١
٢	٠,٠٣	٠,٢٢٦	٠,٩٥	٠,٢٣١	٠,٣٥	٠,٢٢٢
٣	١,١٩-	٠,٣٥٨	٠,٢٠	٠,١٤٥	٠,٧٤-	٠,١٩٦
٤	١,٣٤-	٠,٣٣٢	٠,٠٩-	٠,١٢١	٠,٦٧-	٠,١٢٠
٥	٠,٩٤-	٠,٢٠٥	٠,٨٨-	٠,١٧٤	١,٢٨-	٠,٢٥٤
٦	٠,٧١-	٠,١٣٢	٠,٣٣-	٠,١٢٠	١,٠٦-	٠,٢٧٧
٧	١,٢٣	٠,٦٤١	١,٤٥	٠,٥٠٥	١,٥٤	٠,٤٨٤
٨	٠,٨٠-	٠,١٩٠	٠,٢٣-	٠,١٦٨	٠,٥٦-	٠,١٨٧
٩	٠,٠٥-	٠,١٩٥	٠,٣٩	٠,٢٢٠	٠,٢٥	٠,١٩٤
١٠	٠,٨٤-	٠,٢٠١	٠,٧١-	٠,١٧٠	٠,٩٦-	٠,٢٥٢
١١	٠,٦١	٠,٣٤٥	١,٣١	٠,٣٢٧	٠,٩٢	٠,٣٣٨
١٢	٠,٢٤-	٠,٣٠٠	١,٣٥	٠,٤٢٦	٠,٣٨	٠,٢٢٨
١٣	٠,٤١	٠,٣٨٩	١,٣٧	٠,٣٧٤	٠,٩٢	٠,٣٥٤
١٤	٠,٩٠-	٠,١٧٠	٠,٧٦-	٠,١٦٩	٠,٣٨-	٠,١٢٩
١٥	١,١١-	٠,٢٤٤	١,٢٤-	٠,٣١٥	١,١٢-	٠,٢٨٤
١٦	٠,١٥-	٠,٢٢٣	٠,٠٩-	٠,٢٤٥	٠,٠٤-	٠,٣٢٢
١٧	٠,٥٦-	٠,٢٣٢	٠,٠٩-	٠,٢٢٣	٠,٨٣-	٠,٢٢٠
١٨	٠,٢١-	٠,٣٥٧	١,١٦	٠,٤٦٤	٠,٨٧	٠,٣٤٥
١٩	٠,٠٨-	٠,٢٦٠	٠,٥٤	٠,١٨٩	٠,١٥	٠,١٧١
٢٠	٠,١٩	٠,٢٢٣	٠,٣١	٠,٢٣٤	٠,٤٦	٠,٢٠٩
٢١	٠,٥٥-	٠,٣٢٤	٠,٣٩	٠,١٧٧	٠,٤٤	٠,٢٥٣
٢٢	١,٠٨-	٠,٤٠٨	٠,٢١-	٠,١٧٠	٠,٦٧-	٠,١٩٢
٢٣	٠,٩٣	٠,٥٠٥	١,٩٠	٠,٥٩٢	٠,٥٩	٠,٣٤٣
٢٤	٠,٦٢-	٠,١٩٥	٠,١٠	٠,١٤٠	٠,٠٥-	٠,١٣٧
٢٥	١,١١-	٠,٢٧٠	٠,٥٠-	٠,١٦٥	٠,٣١-	٠,١٣٣
٢٦	٠,٩٩-	٠,١٧٧	٠,٥٤-	٠,١١٠	٠,٨٢-	٠,١٣٦
٢٧	٠,٣١-	٠,٢٨٨	٠,٩٩	٠,٥٠٢	٠,٣٨	٠,٢٩٣
٢٨	٠,١٩	٠,٢١٦	١,٣٠	٠,٤٣٧	٠,٦٦	٠,٢٨٦
٢٩	٠,٥٧-	٠,٢٣٣	٠,٣٠-	٠,١٤٩	٠,٠٤-	٠,١٣٥
٣٠	٠,٣٨-	٠,٢٨٢	٠,٢١	٠,١٦٨	٠,٧٢-	٠,١٦٧
٣١	٠,٢٨-	٠,١٢٧	٠,٤٩-	٠,١٤٨	٠,٢٧-	٠,١٥٧

جدول (٩): المتوسط الحسابي والانحراف المعياري لمعلمة الصعوبة والأخطاء المعيارية في دقة تقديرها لفقرات الاختبار للنماذج الثلاثة

النموذج	المتوسط الحسابي		الانحراف المعياري	
	معلمة الصعوبة	الخطأ المعياري	معلمة الصعوبة	الخطأ المعياري
النموذج الأول (البديل الصحيح على البديلين أ و ب)	٠,٤٠-	٠,٢٧٢	٠,٦٤	٠,١٠٩
النموذج الثاني (البديل الصحيح على البديلين ج و د)	٠,٢٤	٠,٢٥١	٠,٨١	٠,١٣٤
النموذج الثالث (البديل الصحيح موزع عشوائيًا)	٠,١١-	٠,٢٣٤	٠,٧٣	٠,٠٨٥

لا بد أن نوضح أولًا أن صعوبة الفقرة تحدد مكان فعالية الفقرة على مقياس القدرة فعلى سبيل المثال الفقرة السهلة تظهر فعاليتها بين المفحوصين منخفضي القدرة، بينما الفقرة الصعبة تظهر فعاليتها بين المفحوصين مرتفعي القدرة لذا فالصعوبة دليل لموضع الفقرة على المنحنى المميز لها. ويتضح من الجدول السابق رقم (٨)، والجدول رقم (٩)، أن قيم معلمة الصعوبة للنموذج الأول للاختبار تتراوح بين(-١,٣٤ إلى ١,٢٣) بمتوسط حسابي قدره (-٠,٤٠)، بينما قيم معلمة الصعوبة للنموذج الثاني للاختبار تتراوح بين (-١,٢٤ إلى ١,٩٠) بمتوسط حسابي قدره (٠,٢٤)، وقيم معلمة الصعوبة للنموذج الثالث للاختبار تتراوح بين (-١,٢٨ إلى ١,٥٤) بمتوسط حسابي قدره (-٠,١١). ويذكر ايالا (Ayala, 2009/2017, p.17) أن الفقرات التي تتراوح قيمتها تحت الصفر و-٢ توصف بأنها فقرات سهلة، والفقرات التي فوق الصفر فقرات صعبة. أما الفقرات التي قيمتها حوالي الصفر تعتبر

متوسطة الصعوبة. وبناءً على ما تقدم نلاحظ بأن المتوسط الحساب لقيم معلمة الصعوبة للنماذج الثلاثة تقترب من الصفر وهذا يدل على أن فقرات الاختبار متوسطة الصعوبة، ويتضح بأن النموذج الثاني للاختبار هو الأصعب ثم النموذج الثالث ثم النموذج الأول.

أما ما يتعلق بالخطأ المعياري لتقدير معلمة الصعوبة فهو دلالة على دقة تقدير معلمة الصعوبة وهو أدق من تقدير معلمة الصعوبة نفسها. ويتبين من الجدول السابق رقم (٨) بأن الفقرة رقم (٣١) في النموذج الأول للاختبار تبلغ فيها قيمة الخطأ المعياري لتقدير معلمة الصعوبة (٠,١٢٧) وهي أقل قيمة في الجدول بالنسبة للنموذج الأول للاختبار، وبالتالي تُعد هذه الفقرة أكثر دقة في تقدير معلمة الصعوبة. والفقرة رقم (٢٦) في النموذج الثاني للاختبار تبلغ فيها قيمة الخطأ المعياري لتقدير معلمة الصعوبة (٠,١١٠) وهي أقل قيمة بالنسبة للنموذج الثاني، وبالتالي تُعد هذه الفقرة أكثر دقة في تقدير معلمة الصعوبة. بينما في النموذج الثالث، فالفقرة رقم (٤) تبلغ فيها قيمة الخطأ المعياري لتقدير معلمة الصعوبة (٠,١٢٠) وهي أقل قيمة بالنسبة للنموذج الثالث، وبالتالي تُعد هذه الفقرة أكثر دقة في تقدير معلمة الصعوبة في النموذج الثالث للاختبار. ويتضح من الجدول رقم (٩) بأن المتوسط الحسابي للأخطاء المعيارية في دقة تقدير معلمة الصعوبة للنموذج الثالث هو الأقل خطاءً مقارنة ببقية النماذج، ثم النموذج الثاني وأخيرًا النموذج الأول.

وللكشف عن الدلالة الإحصائية للفروق بين المتوسطات الحسابية للأخطاء المعيارية في دقة تقدير معلمة الصعوبة للنماذج الثلاثة للاختبار باختلاف موقع البديل الصحيح، تم استخدام تحليل التباين الأحادي، وتظهر النتائج في الجدول رقم (١٠).

جدول (١٠): نتائج تحليل التباين الأحادي للكشف عن الفروق بين المتوسطات الحسابية للأخطاء المعيارية في دقة تقدير معلمة الصعوبة باختلاف موقع البديل الصحيح

مصدر التباين	مجموع المربعات	الدرجات الحرية	متوسط المربعات	قيمة F	الدلالة الإحصائية
بين المجموعات (موقع البديل الصحيح)	٠,٠٢٣	٢	٠,٠١٢	٠,٩٤٠	٠,٣٩٤
داخل المجموعات (الخطأ)	١,١١٤	٩٠	٠,٠١٢		
الكلي	١,١٣٨	٩٢			

ويتبين من الجدول رقم (١٠) عدم وجود فروق دالة إحصائيًا عند مستوى الدلالة (٠,٠٥) بين المتوسطات الحسابية للأخطاء المعيارية في دقة تقدير معلمة الصعوبة للنماذج الثلاثة للاختبار تُعزى إلى اختلاف موقع البديل الصحيح.

السؤال الثاني: هل توجد فروق دالة إحصائيا بين متوسطات الأخطاء المعيارية في دقة تقدير معلمة التمييز للفقرات في اختبار اختيار من متعدد وفق نظرية الاستجابة للفقرة تُعزى إلى اختلاف موقع البديل الصحيح؟

للإجابة عن هذا السؤال تم استخدام برنامج (Bilog-Mg3) لحساب الخطأ المعياري في دقة تقدير معلمة التمييز لفقرات الاختبار وذلك للنماذج الثلاثة للاختبار، كما هو موضح في الجدول رقم (١١) والذي يوضح كذلك تقديرات معلمة التمييز. وحساب المتوسط الحسابي والانحراف المعياري لدقة تقدير معلمة التمييز وذلك للنماذج الثلاثة للاختبار، كما هو موضح في الجدول رقم (١٢) والذي يوضح أيضاً المتوسط الحسابي والانحراف المعياري لتقديرات معلمة التمييز.

جدول (١١): تقديرات قيم معلمة التمييز والأخطاء المعيارية في دقة تقديرها لفقرات الاختبار للنماذج الثلاثة

رقم الفقرة	النموذج الأول (البديل الصحيح على البديلين أ و ب)		النموذج الثاني (البديل الصحيح على البديلين ج و د)		النموذج الثالث (البديل الصحيح موزع عشوائيًا)	
	معلمة التمييز	الخطأ المعياري	معلمة التمييز	الخطأ المعياري	معلمة التمييز	الخطأ المعياري
١	١,٢٢	٠,٣٢٣	٠,٨٨	٠,١٩٦	٠,٩٠	٠,٢٤٠
٢	٠,٧٣	٠,١٧١	١,٠٥	٠,٢١٣	٠,٨٦	٠,١٩٧
٣	٠,٦٩	٠,١٨٦	١,٤٢	٠,٢٧٢	١,٠٥	٠,٢٥٤
٤	٠,٨٦	٠,٢٢٧	١,٦٤	٠,٣٤٦	١,٨١	٠,٥١١
٥	١,٢٠	٠,٢٨٣	١,٤٥	٠,٣٣٩	١,٥٢	٠,٤٥٧
٦	١,٧٠	٠,٤٥١	١,٦٥	٠,٣٤٨	٠,٩٧	٠,٣٠٠
٧	٠,٣٠	٠,٠٨٨	٠,٤٨	٠,١٢٦	٠,٥٣	٠,١٣٣
٨	١,١٩	٠,٢٥١	١,٠٤	٠,٢١٥	٠,٩٩	٠,٢٥٥
٩	٠,٨٨	٠,٢٠٧	٠,٨٧	٠,١٨٥	١,٠٢	٠,٢٣٦
١٠	١,١٤	٠,٣٠٠	١,٢٤	٠,٣١٧	٠,٩٧	٠,٢٦٤
١١	٠,٥٤	٠,١٤٣	٠,٨٢	٠,١٧٣	٠,٦٩	٠,١٦٥
١٢	٠,٥٣	٠,١٤٧	٠,٥٩	٠,١٤٨	٠,٧٩	٠,١٧٢
١٣	٠,٤٢	٠,١٢٣	٠,٦٦	٠,١٥٤	٠,٥٩	٠,١٤٢
١٤	١,٥٣	٠,٣٤٠	١,٢٦	٠,٣١٤	١,٤٢	٠,٣١٩

١٥	١,١٩	٠,٣٣٢	٠,٨٦	٠,٢١٠	٠,٨٩	٠,٢٢١
١٦	٠,٧٤	٠,١٨٩	٠,٦٦	٠,١٥٨	٠,٤٨	٠,١٢٦
١٧	٠,٧٩	٠,١٩٥	٠,٧٣	٠,١٧٦	٠,٩٩	٠,٢٦٧
١٨	٠,٤٤	٠,١٢٣	٠,٤٥	٠,١٢٣	٠,٥٩	٠,١٤٤
١٩	٠,٦٢	٠,١٥٥	١,٢١	٠,٢٤٢	١,١٩	٠,٢٩٩
٢٠	٠,٧٩	٠,١٩٥	٠,٧٣	٠,١٦٤	٠,٩٦	٠,٢٠٢
٢١	٠,٥٥	٠,١٤١	١,٠٦	٠,٢٠٠	٠,٧٤	٠,١٧٨
٢٢	٠,٥٤	٠,١٤٣	١,٠٣	٠,٢٠٧	١,٠٢	٠,٢٥١
٢٣	٠,٣٦	٠,١٠٦	٠,٥٠	٠,١٣٦	٠,٥١	٠,١٢٧
٢٤	٠,٩٧	٠,٢٥٣	١,٤٤	٠,٣٠٤	١,٤٨	٠,٣٦٥
٢٥	٠,٩٥	٠,٢٧٢	١,١٣	٠,٢٣٥	١,٣٧	٠,٣٤٧
٢٦	١,٥٧	٠,٤١٧	١,٩٧	٠,٤٥٠	١,٧٨	٠,٥٣٨
٢٧	٠,٥٧	٠,١٥٠	٠,٣٧	٠,١٠٣	٠,٥٩	٠,١٤٨
٢٨	٠,٨٠	٠,١٨٧	٠,٥٤	٠,١٤٠	٠,٧٠	٠,١٧٣
٢٩	٠,٧٩	٠,١٨٧	١,٢٤	٠,٢٣٣	١,٣٧	٠,٢٦٥
٣٠	٠,٦٠	٠,١٥٤	١,٠٩	٠,٢١٨	١,٢٨	٠,٣٣٥
٣١	١,٥٢	٠,٣٥٦	١,٢٩	٠,٢٨٤	١,١٠	٠,٢٥٦

جدول (١٢): المتوسط الحسابي والانحراف المعياري لمعلمة التمييز والأخطاء المعيارية في دقة تقديرها لفقرات الاختبار للنماذج الثلاثة

النموذج	المتوسط الحسابي		الانحراف المعياري	
	معلمة التمييز	الخطأ المعياري	معلمة التمييز	الخطأ المعياري
النموذج الأول (البديل الصحيح على البديلين أ و ب)	٠,٨٦	٠,٢١٩	٠,٣٨	٠,٠٩٣
النموذج الثاني (البديل الصحيح على البديلين ج و د)	١,٠١	٠,٢٢٤	٠,٤٠	٠,٠٨٢
النموذج الثالث (البديل الصحيح موزع عشوائيًا)	١,٠٠	٠,٢٥٤	٠,٣٦	٠,١٠٧

يتضح من الجدول رقم (١١)، والجدول رقم (١٢)، أن قيم معلمة التمييز للنموذج الأول للاختبار تتراوح بين(٠,٣٠ إلى ١,٧٠) بمتوسط حسابي قدره (٠,٨٦)، بينما قيم معلمة التمييز للنموذج الثاني للاختبار تتراوح بين (٠,٣٧ إلى ١,٩٧) بمتوسط حسابي قدره (١,٠١)، وقيم معلمة التمييز للنموذج الثالث للاختبار تتراوح بين (٠,٤٨ إلى ١,٨١) بمتوسط حسابي قدره (١,٠٠). ونجد بأن متوسط معلمة التمييز للنموذج الثاني هو الأعلى ثم النموذج الثالث ويليه النموذج الأول.

ويتضح من الجدول السابق رقم (١١) بأن الفقرة رقم (٧) في النموذج الأول للاختبار تبلغ فيها قيمة الخطأ المعياري لتقدير معلمة التمييز (٠,٠٨٨) وهي أقل قيمة في الجدول بالنسبة للنموذج الأول للاختبار، وبالتالي تُعد هذه الفقرة أكثر دقة في تقدير معلمة التمييز. والفقرة رقم (١٨) في النموذج الثاني للاختبار تبلغ فيها قيمة الخطأ المعياري لتقدير معلمة التمييز (٠,١٢٣) وهي أقل قيمة بالنسبة للنموذج الثاني، وبالتالي تُعد هذه الفقرة أكثر دقة في تقدير معلمة التمييز. بينما في النموذج الثالث فالفقرة رقم (١٦) تبلغ فيها قيمة الخطأ المعياري لتقدير معلمة التمييز (٠,١٢٦) وهي أقل قيمة بالنسبة للنموذج الثالث، وبالتالي تُعد هذه الفقرة أكثر دقة في تقدير معلمة التمييز في النموذج الثالث للاختبار. ويتضح من الجدول رقم (١٢) بأن المتوسط الحسابي للأخطاء المعيارية في دقة تقدير معلمة التمييز للنموذج الأول هو الأقل خطاءً مقارنة ببقية النماذج، ثم النموذج الثاني وأخيرًا النموذج الثالث.

وللكشف عن الدلالة الإحصائية للفروق بين المتوسطات الحسابية للأخطاء المعيارية في دقة تقدير معلمة التمييز للنماذج الثلاثة للاختبار باختلاف موقع البديل الصحيح، تم استخدام تحليل التباين الأحادي، وتظهر النتائج في الجدول رقم (١٣).

جدول (١٣): نتائج تحليل التباين الأحادي للكشف عن الفروق بين المتوسطات الحسابية للأخطاء المعيارية في دقة تقدير معلمة التمييز باختلاف موقع البديل الصحيح

مصدر التباين	مجموع المربعات	الدرجات الحرية	متوسط المربعات	قيمة F	الدلالة الإحصائية
بين المجموعات (موقع البديل الصحيح)	٠,٠٢٣	٢	٠,٠١١	١,٢٨٥	٠,٢٨٢
داخل المجموعات (الخطأ)	٠,٧٩٩	٩٠	٠,٠٠٩		
الكلي	٠,٨٢٢	٩٢			

ويتبين من الجدول رقم (١٣) عدم وجود فروق دالة إحصائيًا عند مستوى الدلالة (٠,٠٥) بين المتوسطات الحسابية للأخطاء المعيارية في دقة تقدير معلمة التمييز للنماذج الثلاثة للاختبار تُعزى إلى اختلاف موقع البديل الصحيح.

السؤال الثالث: هل توجد فروق ذات دلالة إحصائية بين متوسطات تقديرات معاملات الثبات التجريبي في اختبار اختيار من متعدد وفق نظرية الاستجابة للفقرة تُعزى إلى اختلاف موقع البديل الصحيح؟

للإجابة عن هذا السؤال تم استخدام برنامج (Bilog-Mg3) لحساب معاملات الثبات التجريبي لكل نموذج من النماذج الثلاثة للاختبار، والجدول رقم (١٤) يوضح هذه النتائج.

جدول (١٤): نتائج قيم معاملات الثبات التجريبي لنماذج الاختبار الثلاثة

النموذج	معامل الثبات التجريبي
النموذج الأول (موقع البديل الصحيح على البديلين أ و ب)	٠,٨٥
النموذج الثاني (موقع البديل الصحيح على البديلين ج و د)	٠,٨٨
النموذج الثالث (موقع البديل الصحيح موزع عشوائيًا)	٠,٨٧

يتضح من الجدول رقم (١٤) أن قيمة معامل الثبات التجريبي للنموذج الثاني (موقع البديل الصحيح على البديلين ج و د) هي الأعلى قيمة، وبالتالي هي الأكثر ثباتًا من النماذج الأخرى للاختبار ويليها النموذج الثالث (موقع البديل الصحيح موزع عشوائيًا)، واخيرًا النموذج الأول (موقع البديل الصحيح على البديلين أ و ب).

ولمعرفة دلالة الفروق في معاملات الثبات التجريبي بين نماذج الاختبار الثلاثة، فقد تم استخدام الاختبار الإحصائي (M) والذي يتبع توزيع مربع كاي بدرجات حرية (عدد المعاملات – ١) وتتضح المعادلة الرياضية لهذا الاختبار بالصورة التالية:

$$M = \frac{J-1}{18\,J}\left\{\sum_{K=1}^{K} B_K - \frac{\left[\sum_{K=1}^{K} B_K\,(1-r_k)^{\frac{-1}{3}}\right]^2}{\sum_{K}^{K} B_K (1-r_k)^{\frac{-2}{3}}}\right\}$$

حيث تُحسب B_K من المعادلة التالية:

$$B_K = \frac{(9n_k - 11)^2}{n_k - 1}$$

حيث إن ...

J: عدد فقرات الاختبار ، r_k: معامل الثبات التجريبي ، n_k: عدد أفراد العينة ، K : رقم نموذج الاختبار حسب موقع البديل الصحيح.

وتم حساب قيمة الإحصائي (M) لكل مقارنة بين نماذج الاختبار الثلاثة لحساب معامل الثبات التجريبي كما تظهر في الجدول رقم (١٥).

جدول(١٥): قيم الإحصائي (M) للمقارنات الثنائية بين معاملات الثبات التجريبي لنماذج الاختبار الثلاثة

المقارنات الثنائية	قيم الإحصائي (M)
النموذج الأول*النموذج الثاني	١,٩٦
النموذج الأول*النموذج الثالث	٠,٨١
النموذج الثاني*النموذج الثالث	٠,٢٥

من الجدول رقم (١٥) يتضح أن جميع قيمة (M) المحسوبة أقل من القيمة الحرجة لتوزيع كاي تربيع بدرجة حرية (١) وعند مستوى دلالة (٠,٠٥) والتي تساوي (٣,٨٤)، والذي يدل على عدم وجود فروق ذات دلالة إحصائية بين قيم معاملات الثبات التجريبي لنماذج الاختبار الثلاثة، وهذا يعني أن موقع البديل الصحيح ليس له تأثير ذو دلالة إحصائية على معاملات الثبات التجريبي.

السؤال الرابع: هل توجد فروق ذات دلالة إحصائية بين متوسطات تقديرات معاملات ثبات الأفراد في اختبار اختيار من متعدد وفق نظرية الاستجابة للفقرة تُعزى إلى اختلاف موقع البديل الصحيح؟

للإجابة عن هذا السؤال تم حساب معامل الفصل بين الأفراد للنماذج الثلاثة للاختبار من خلال المعادلة التالية:

$$\text{معامل الفصل للأفراد (G)} = \frac{\text{الانحراف المعياري لقيم معلمة الصعوبة للفقرات}}{\text{الوسط الحسابي للخطأ المعياري في تقدير معلمة الصعوبة}}$$

ويتم استخراج هذه القيم من خلال نتائج برنامج (Bilog-Mg3).

أ- معامل ثبات الأفراد للنموذج الأول:

في النموذج الأول للاختبار تبلغ قيمة الانحراف المعياري لقيم معلمة الصعوبة للفقرات (٠,٦٤) وقيمة الوسط الحسابي للخطأ المعياري في تقدير معلمة الصعوبة (٠,٢٧٢)، وعند التعويض في المعادلة السابقة نحصل على قيمة معامل الفصل للأفراد كالتالي: (G) = $\frac{٠,٦٤}{٠,٢٧٢}$ = ٢,٣٥ ومن خلال قيمة معامل الفصل للأفراد يتم حساب معامل ثبات الأفراد من المعادلة التالية:

معامل ثبات الأفراد= $\frac{G^2}{1+G^2}$ وبالتعويض في هذه المعادلة ينتج لنا التالي ...

معامل ثبات الأفراد للنموذج الأول: $\frac{٢,٣٥^2}{1+٢,٣٥^2}$ = ٠,٨٥

ب- معامل ثبات الأفراد للنموذج الثاني:

ما يتعلق بالنموذج الثاني للاختبار فتبلغ قيمة الانحراف المعياري لقيم معلمة الصعوبة للفقرات (٠,٨١) وقيمة الوسط الحسابي للخطأ المعياري في تقدير معلمة الصعوبة (٠,٢٥١) وعند التعويض في معادلة معامل الفصل للأفراد نحصل على القيمة التالي: (G) = $\frac{٠,٨١}{٠,٢٥١}$ = ٣,٢٣ ومن خلال قيمة معامل الفصل للأفراد يتم حساب معامل ثبات الأفراد من المعادلة التالية:

معامل ثبات الأفراد للنموذج الثاني = $\frac{G^2}{1+G^2}$ = ٠,٩١

ج- معامل ثبات الأفراد للنموذج الثالث:

في النموذج الثالث للاختبار تبلغ قيمة الانحراف المعياري لقيم معلمة الصعوبة للفقرات (٠,٧٣) وقيمة الوسط الحسابي للخطأ المعياري في تقدير معلمة الصعوبة (٠,٢٣٤) وعند التعويض في معادلة معامل الفصل للأفراد نحصل على التالي: (G) = $\frac{٠,٧٣}{٠,٢٣٤}$ = ٣,١٢ ومن خلال قيمة معامل الفصل للأفراد يتم حساب معامل ثبات الأفراد من المعادلة التالية:

معامل ثبات الأفراد للنموذج الثالث = $\frac{G^2}{1+G^2}$ = ٠,٩١

ومن خلال الجدول رقم (١٦) يتضح قيم معاملات ثبات الأفراد للنماذج الثلاثة للاختبار.

جدول (١٦): نتائج قيم معاملات ثبات الأفراد لنماذج الاختبار الثلاثة

النموذج	معامل ثبات الأفراد
النموذج الأول (موقع البديل الصحيح على البديلين أ و ب)	٠,٨٥
النموذج الثاني (موقع البديل الصحيح على البديلين ج و د)	٠,٩١
النموذج الثالث (موقع البديل الصحيح موزع عشوائيًا)	٠,٩١

يتضح من الجدول رقم (١٦) بأن النموذج الثاني (موقع البديل الصحيح على البديلين ج و د) والنموذج الثالث (موقع البديل الصحيح موزع عشوائيًا) هما الأعلى ثباتًا حيث يتضح تساوي قيمة معامل ثبات الأفراد للنموذجين الثاني والثالث وتبلغ القيمة (٠,٩١) وهي قيمة مرتفعة للثبات، بينما النموذج الأول (موقع البديل الصحيح على البديلين أ و ب) هو الأقل ثباتًا من النموذجين الثاني والثالث.

ولمعرفة دلالة الفروق في معاملات ثبات الأفراد بين نماذج الاختبار الثلاثة، فقد تم استخدام الاختبار الإحصائي (M) والذي يتبع توزيع مربع كاي بدرجات حرية (عدد المعاملات – ١) وتتضح المعادلة الرياضية لهذا الاختبار بالصورة التالية:

$$M = \frac{J-1}{18\,J}\left\{\sum_{K=1}^{K} B_K - \frac{\left[\sum_{K=1}^{K} B_K\,(1-r_k)^{\frac{-1}{3}}\right]^2}{\sum_{K}^{K} B_K (1-r_k)^{\frac{-2}{3}}}\right\}$$

حيث تُحسب B_K من المعادلة التالية :

$$B_K = \frac{(9n_k - 11)^2}{n_k - 1}$$

حيث إن ...

J: عدد فقرات الاختبار ، r_k: معامل الثبات التجريبي ، n_k: عدد أفراد العينة ، K : رقم نموذج الاختبار حسب موقع البديل الصحيح.

وتم حساب قيمة الإحصائي (M) لكل مقارنة بين نماذج الاختبار الثلاثة لحساب معامل ثبات الأفراد كما تظهر في الجدول رقم (١٧).

جدول (١٧) : قيم الإحصائي (M) للمقارنات الثنائية بين معاملات ثبات الأفراد لنماذج الاختبار الثلاثة

المقارنات الثنائية	قيم الإحصائي (M)
النموذج الأول*النموذج الثاني	١٠,١٨
النموذج الأول*النموذج الثالث	١٠,١٨
النموذج الثاني*النموذج الثالث	٠,٠٠

من الجدول رقم (١٧) يتضح أن بعض قيم (M) المحسوبة أعلى من القيمة الحرجة لتوزيع كاي تربيع بدرجة حرية (١) وعند مستوى دلالة (٠,٠٥) والتي تساوي (٣,٨٤)، والذي يدل على وجود فروق ذات دلالة إحصائية بين قيم معاملات ثبات الأفراد، وهذا يعني أن موقع البديل الصحيح له تأثير ذو دلالة إحصائية على معاملات ثبات الأفراد. حيث يوجد فروق ذات دلالة إحصائية بين النموذج الأول (موقع البديل الصحيح على البديلين أ و ب)، والنموذج الثاني (موقع البديل الصحيح على البديلين ج و د)، لصالح النموذج الثاني. ووجود فروق ذات دلالة إحصائية بين النموذج الأول (موقع البديل الصحيح على البديلين أ و ب)، والنموذج الثالث (موقع البديل الصحيح موزع عشوائيًا) لصالح النموذج الثالث. وعدم وجود فروق ذات دلالة إحصائية بين النموذج الثاني (موقع البديل الصحيح على البديلين ج و د)، والنموذج الثالث (موقع البديل الصحيح موزع عشوائيًا)، حيث هنا تساوت قيم معامل ثبات الأفراد بين النموذجين الثاني، والثالث وتبلغ (٠,٩١) لذلك لا توجد فروق ذات دلالة إحصائية بين النموذجين.

مناقشة نتائج البحث:

أظهرت النتائج المتعلقة بتحقق افتراضات نظرية الاستجابة للفقرة للنموذج ثنائي المعلمة وبالتحديد افتراض أحادية البعد بأن قيمة الجذر الكامن للعامل الأول إلى قيمة الجذر الكامن للعامل الثاني كانت أعلى من (٢) للنماذج الثلاثة للاختبار، وهذا يدل على وجود سمة سائدة للاختبار، وهذه القيمة تكون أعلى في النموذج الأول (البديل الصحيح على أ و ب)، ثم النموذج الثالث (البديل الصحيح موزع عشوائيًا)، ويليه النموذج الثاني (البديل الصحيح على ج و د)، وتتفق هذه النتائج مع دراسة سليمان والصالح (٢٠١٧) في تحقق افتراض أحادية البعد وتختلف معها في أن قيمة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني تكون أعلى في النموذج الثالث (البديل الصحيح موزع عشوائيًا)، ثم النموذج الأول (البديل الصحيح على أ و ب)، ويليه النموذج الثاني (البديل الصحيح على ج و د).

أما ما يتعلق بمناقشة نتائج مطابقة الأفراد للنموذج ثنائي المعلمة وفق نظرية الاستجابة للفقرة فلقد تم استبعاد استجابات طالبة واحدة من النموذجين الأول، والثاني للاختبار، ويُفسر عدم مطابقتهم للنموذج إلى عدم جدية بعض الطالبات في الإجابة على فقرات الاختبار، وميلهن إلى الإجابة على فقرات الاختبار بشكل عشوائي. وكذلك اعتقادهم باستحالة أن تكون الإجابات الصحيحة مركزه على بدائل معينة واعتقادهم بأنها مرتبة بطريقة عشوائية على البدائل، وبالتالي حصولهم على درجات متدنية في الاختبار. وكذلك تم استبعاد استجابات طالبتين من النموذج الثالث للاختبار، وذلك بسبب حصولهم على درجات مرتفعة. وتتفق هذه النتائج مع دراسة سليمان والصالح (٢٠١٧) في أن النموذج الثاني للاختبار تم فيه استبعاد استجابات طالب واحد غير مطابق للنموذج ثلاثي المعلمة وتختلف عنها في استبعاد استجابات ثلاثة طلبة من النموذج الأول، واستبعاد استجابات طالب واحد من النموذج الثالث.

أما ما يتعلق بمناقشة نتائج أسئلة البحث، فتم ذكرها كالتالي ...

أولاً: النتائج المتعلقة بالإجابة عن السؤال الأول والذي ينص على: هل توجد فروق دالة إحصائيا بين متوسطات الأخطاء المعيارية في دقة تقدير معلمة الصعوبة للفقرات في اختبار اختيار من متعدد وفق نظرية الاستجابة للفقرة تُعزى إلى اختلاف موقع البديل الصحيح؟

أظهرت النتائج عدم وجود فروق دالة إحصائيًا عند مستوى الدلالة (٠,٠٥) بين متوسطات الأخطاء المعيارية في دقة تقدير معلمة الصعوبة للفقرة تُعزى إلى اختلاف موقع البديل الصحيح. ولقد اتضح أن النموذج الثالث (البديل الصحيح موزع عشوائيًا) هو الأكثر دقة في تقدير معلمة الصعوبة للفقرة من النماذج الأخرى، وتختلف هذه النتيجة مع دراسة الصبح (٢٠١١) والتي كشفت عن وجود فروق دالة إحصائيًا بين متوسطات الأخطاء المعيارية في دقة تقدير معلمة الصعوبة للفقرة وفق موقع البديل المميز (البديل الأول) للنموذجين الثنائي والثلاثي المعلمة. وقد يعود سبب عدم وجود فروق دالة إحصائيًا بين متوسطات الأخطاء المعيارية في دقة تقدير معلمة الصعوبة للفقرة تُعزى إلى اختلاف موقع البديل الصحيح؛ إلى أن محتوى الاختبار واحد لكل النماذج الثلاثة، وربما يكون السبب الأخر هو أن الطالبات لا يمتلكن مهارة حكمة الاختبار.

ثانياً: النتائج المتعلقة بالإجابة عن السؤال الثاني والذي ينص على: هل توجد فروق دالة إحصائيا بين متوسطات الأخطاء المعيارية في دقة تقدير معلمة التمييز للفقرات في اختبار اختيار من متعدد وفق نظرية الاستجابة للفقرة تُعزى إلى اختلاف موقع البديل الصحيح؟

أظهرت النتائج عدم وجود فروق دالة إحصائيًا عند مستوى الدلالة (٠,٠٥) بين متوسطات الأخطاء المعيارية في دقة تقدير معلمة التمييز للفقرة تُعزى إلى اختلاف موقع البديل الصحيح. ولقد اتضح أن النموذج الأول (البديل الصحيح على أ و ب) هو الأكثر دقة في تقدير معلمة التمييز للفقرة من النماذج الأخرى، وتتفق هذه النتيجة مع دراسة الصبح (٢٠١١) والتي كشفت عن عدم وجود فروق دالة إحصائيًا بين متوسطات الأخطاء المعيارية في دقة تقدير معلمة التمييز للفقرة وفق موقع البديل المميز (البديل الأول) للنموذجين الثنائي والثلاثي المعلمة. وقد يعود سبب عدم وجود فروق دالة إحصائيًا بين متوسطات الأخطاء المعيارية في دقة تقدير معلمة التمييز للفقرة تُعزى إلى اختلاف موقع البديل الصحيح إلى أحد السببين حيث السبب الأول قد يتمحور حول أن محتوى الاختبار واحد لكل النماذج الثلاثة للاختبار، والسبب الثاني ربما يعود إلى عدم امتلاك الطالبات لمهارة حكمة الاختبار.

ثالثاً: النتائج المتعلقة بالإجابة عن السؤال الثالث والذي ينص على: هل توجد فروق ذات دلالة إحصائية بين متوسطات تقديرات معاملات الثبات التجريبي في اختبار اختيار من متعدد وفق نظرية الاستجابة للفقرة تُعزى إلى اختلاف موقع البديل الصحيح؟

أظهرت النتائج عدم وجود فروق ذات دلالة إحصائية عند مستوى الدلالة (٠,٠٥) بين متوسطات تقديرات معاملات الثبات التجريبي تُعزى إلى اختلاف موقع البديل الصحيح. واختلفت هذه النتيجة مع دراسة سليمان والصالح (٢٠١٧)، والتي توصلت إلى وجود فروق ذات دلالة إحصائية بين متوسطات تقديرات معاملات الثبات التجريبي تُعزى إلى اختلاف موقع البديل الصحيح.

رابعاً: النتائج المتعلقة بالإجابة عن السؤال الرابع والذي ينص على: هل توجد فروق ذات دلالة إحصائية بين متوسطات تقديرات معاملات ثبات الأفراد في اختبار اختيار من متعدد وفق نظرية الاستجابة للفقرة تُعزى إلى اختلاف موقع البديل الصحيح؟

أظهرت النتائج وجود فروق ذات دلالة إحصائية عند مستوى الدلالة (٠,٠٥) بين متوسطات تقديرات معاملات ثبات الأفراد تُعزى إلى اختلاف موقع البديل الصحيح، وذلك لصالح النموذجين الثاني (البديل الصحيح على ج و د)، والثالث (البديل الصحيح موزع عشوائيًا).

التوصيات:

بناءً على ما تم التوصل إليه من نتائج، يمكن الخروج بالتوصيات التالية:

١. إقامة الدورات التدريبية وورش العمل للمعلمين والمعلمات في مجال قواعد صياغة اختبارات الاختيار من متعدد.
٢. توعية المعلمين والمعلمات إلى ضرورة العناية بفقرات اختبار الاختيار من متعدد عند إعدادها وكذلك ضرورة العناية بتوزيع الإجابات الصحيحة على جميع البدائل بشكل عشوائي، دون إتباع نمط معين في ترتيب الإجابات الصحيحة.
٣. إجراء المزيد من البحوث والدراسات حول أثر موقع البديل الصحيح في اختبار اختيار من متعدد في مواد دراسية مختلفة، وفي مراحل دراسية مختلفة، وكذلك في وجود متغيرات أخرى كالجنس والمرحلة الدراسية، وعدد البدائل وغيرها.

المقترحات:

١. الكشف عن أثر موقع البديل الصحيح في اختبار اختيار من متعدد على تحصيل الطلبة.
٢. دراسة عن أثر موقع البديل الصحيح في اختبار اختيار من متعدد ذي خمسة بدائل على معالم الفقرة والثبات التجريبي وفق نظرية الاستجابة للفقرة.
٣. دراسة مقارنة بين نموذج اختبار اختيار من متعدد ذي أربعة بدائل ونموذج اختبار اختيار من متعدد ذي ثلاثة بدائل في ضوء موقع البديل الصحيح وفق نظرية الاستجابة للفقرة.
٤. دراسة مقارنة بين النظرية التقليدية في القياس، ونظرية الاستجابة للفقرة في ضوء موقع البديل الصحيح وأثره على معلمة الصعوبة والتمييز.

المراجع:

أولاً: المراجع العربية:

١. إيالا، دي، (٢٠١٧) النظرية والتطبيق في نظرية الاستجابة للفقرة، (ط١)، ترجمة: عبدالله الكيلاني، إسماعيل البرصان. الرياض: دار جامعة الملك سعود للنشر.
٢. أبو فودة، باسل خميس & بني يونس، نجاتي أحمد، (٢٠١٢) الاختبارات التحصيلية، (ط١)، عمان: دار المسيرة للنشر والتوزيع.
٣. الأمير، بندر زيد سالم، (٢٠١٦) الأخطاء الشائعة في بناء مفردات اختبار الاختيار من متعدد وتأثيرها على الخصائص السيكومترية للاختبارات التحصيلية لدى معلمي الرياضيات بالمرحلة المتوسطة بمدينة مكة المكرمة، المجلة العربية للعلوم الاجتماعية، ٥(٩)،١٣٧-١٧٥، https://doi.org/10.12816/0033017
٤. أيكين، لويس ر، (٢٠٠٧) الاختبارات والامتحانات قياس القدرات والأداء، (ط١)، ترجمة: فرح السراج. الرياض : شركة العبيكان للأبحاث والتطوير.
٥. بيكر، فرانك ب، (٢٠١٠) أسس نظرية الاستجابة للمفردة، (ط١)، ترجمة: عبد الرحمن الطريري، السيد أبو هاشم، سوسن شلبي. الرياض: النشر العلمي والمطابع – جامعة الملك سعود.
٦. التقي، أحمد محمد، (٢٠١٣) النظرية الحديثة في القياس، (ط٢)، عمان: دار المسيرة للنشر والتوزيع.
٧. سليمان، شاهر خالد & الصالح، علي محمد خالد، (٢٠١٧) أثر موقع البديل الصحيح في اختبار اختيار من متعدد على تقديرات معالم الفقرات والقدرة وفق النموذج اللوجستي ثلاثي المعلمة، دراسات عربية في التربية وعلم النفس، ٩٠، ٩٨-١٢٠، https://doi.org/10.12816/0041974
٨. الصالح، علي محمد خالد، (٢٠١٧) مقارنة تأثيرات تغيير موقع البديل الصحيح في اختبار اختيار من متعدد على تقديرات الصعوبة للفقرات وتقديرات القدرة لدى الأفراد باستخدام نموذج راش، المجلة التربوية الدولية المتخصصة، ٦ (٩)، ١٢٤-١٣٣.

٩. الصبح، رانيا محمد خير حسين، (٢٠١١) أثر بنية فقرات الاختيار من متعدد ومستوى القدرة لدى الأفراد على دقة التقديرات لمعالم الفقرات والأفراد وفق نظرية الاستجابة للفقرة، مجلة جامعة أم القرى للعلوم التربوية والنفسية، ٣(٢)، ٤٥-١١٠.
١٠. الصراف، قاسم علي، (٢٠١٤) القياس والتقويم في التربية والتعليم، (ط٢)، القاهرة : دار الكتاب الحديث.
١١. عبابنة، عماد غصاب، (٢٠٠٩) الاختبارات محكية المرجع فلسفتها وأسس تطويرها، (ط١)، عمان: دار المسيرة للنشر والتوزيع.
١٢. عبد الرحمن، أحمد محمد، (٢٠١١) تصميم الاختبارات، (ط١)، عمان: دار أسامة للنشر والتوزيع.
١٣. علام، صلاح الدين محمود، (٢٠١٥) القياس والتقويم التربوي والنفسي أساسياتُه وتطبيقاتُه وتوجهاتُه المعاصرة، (ط٦)، القاهرة: دار الفكر العربي.
١٤. علام، صلاح الدين محمود، (٢٠١٦) الاختبارات والمقاييس التربوية والنفسية، (ط٥)، عمان: دار الفكر.
١٥. عودة، أحمد سليمان، (٢٠١٠) القياس والتقويم في العملية التدريسية، (ط٤)، إربد : دار الأمل للنشر والتوزيع.
١٦. الفقي، اسماعيل، (٢٠١٤) التقويم والقياس النفسي والتربوي، (ط٣)، الرياض : مكتبة الرشد.
١٧. قاسم، ياسمين محمود ضيف الله، (٢٠١٣) أثر موقع البديل في فقرات اختبارات الاختيار من متعدد على نسبة اختيارها بالتخمين كإجابات صحيحة للفقرة، رسالة ماجستير غير منشورة، جامعة مؤتة، كلية العلوم التربوية، الأردن.
١٨. الكناني، عايد كريم & فيضي، فاضل عبد، (٢٠١٢) فاعلية استخدام موقع البديل الفعال عند إبعاد البديل الخاطئ في صعوبة فقرات الأسئلة من متعدد وتمييزها، مجلة القادسية لعلوم التربية الرياضية، ١٢ (١)، ٢٠٧-٢٢٦.
١٩. النبهان، موسى، (٢٠٠٤) أساسيات القياس في العلوم السلوكية، (ط١)، عمان : دار الشروق للنشر والتوزيع.
٢٠. هاشم، كمال الدين محمد & الخليفة، حسن جعفر، (٢٠١٥) التقويم التربوي مفهومه أساليبه مجالاته توجهاته الحديثة، (ط٥)، الرياض : مكتبة الرشد ناشرون.

ثانياً: المراجع الأجنبية:

[1] Baker. F., The Basic of item Response Theory, (2th Ed), USA: Eric Clearing House on Assessment and Evaluation, (2001)

[2] Cizek. G. J., The Effect of Altering the Position of Options in a Multiple-Choice Examination, Educational and Psychological Measurement, 54(1)(1994),8-20, https://doi.org/10.1177/0013164494054001002

[3] Hambleton. R., & Swaminathan. H., Item Response Theory: Principles and applications, Boston: Kluwer Nijhoff Publishing, (1985)

[4] Hatti. J., An empirical study of various indices for determining unidimensionality, Multivariate Behavioral Research, 19(1)(1984), 49-78, https://doi.org/10.1207/s15327906mbr1901_3

[5] Kolawole. E. B. & Olarotimi. J. T., A survey of the Anchor Bias in Mathematics Objective Tests in West African Examination Council (WAEC), Journal of Emerging Trends in Educational Research and Policy Studies (JETERAPS), 2(3)(2008),171-173.

[6] Lane. D. S. & Bull. K. S., The effects of Knowledge of item arrangement Gender And Statistical and cognitive item Difficulty on Test Performance, Educational and Psychological Measurement, 47(4)(1987), 865– 879, https://doi.org/10.1177/0013164487474002

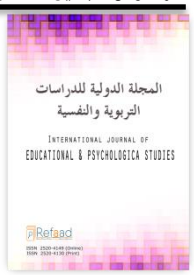
[7] Nwadinigwe. P., I. & Naibi. L., The Number of Options in a Multiple-Choice Test Item and the Psychometric Characteristics, Journal of Education and Practice, Nigeria, 4(28)(2013), 189-196.

[8] Plake. B. Thompson. A. & Lowry. S., Effects of item arrangement and test anxiety on two scoring methods, The Journal of Experimental Education, 49(4)(1984) 214– 219, https://doi.org/10.1080/00220973.1981.11011786

[9] Wevric. L, Response set in a multiple-choice test, Educational and psychological Measurement, (2001) (33).

Refaad
Research, Enlightment, Findings Accelerated Applications Development
رفاد للدراسات والأبحاث
www.refaad.com

المجلة الدولية للدراسات التربوية والنفسية
e-ISSN 2520-4149 , p-ISSN 2520-4130
المجلد٦- العدد١- ٢٠١٩، ص ٦٦-٨٨
https://doi.org/DOI:10.31559/EPS2019.6.1.5



The Impact of the Correct Alternative Position in Multiple Choice Test on the Precision of Item Parameters Estimation According to Two Parameter Model of the Item Response Theory

Ashwaq Daifallah Selim Al-Marwani
M.Sc. in Measurement and Evaluation- University of Tabuk- KSA
A.ashwaqalmarwani@hotmail.com

Shaher Khalid Suleiman
Associate Professor of Measurement and Evaluation- University of Tabuk- KSA
s.suliman@ut.edu.sa

Abstract: This study aimed to detect the impact of the correct alternative position in multiple choice test with four alternatives on the precision of estimation difficulty parameter and item discrimination, and on the estimation of experiential reliability coefficients and individual reliability coefficients according to two parameter model of the item response theory.

To achieve research objectives, a multiple choice test with four alternatives in science was built consisted of (31) items and of three patterns with the same content and variable which is the correct alternative position where the first pattern is the correct alternative position on the two alternatives (A & B) and the second pattern on the alternatives (C & D) and the third pattern randomly distributes the correct alternative position on the four alternatives.

The three patterns were applied to a random clustered sample consisting of (500) third intermediate female students in Yanbu city. (SPSS) statistical program package and (Bilog-Mg 3) were used for data analysis .

The results of the study indicated that there are no statistically significant differences between the means of precision of estimation difficulty parameter and item discrimination attributed to correct alternative position, and there are no statistically significant differences between the means of experiential reliability coefficients ratings attributed to correct alternative position attributed to correct alternative position , while there are no statistically significant differences between the means of individual reliability coefficients ratings attributed to correct alternative position.

In light of these results a number of recommendations were presented, the most notably are the male and female teacher's awareness of the need to take care of the multiple choice test items preparation, as well as the need to care in the distribution of the correct answers randomly to all alternatives .

Keywords: *Correct Alternative Position, Multiple Choice Test, Precision of Item Parameters Estimation, Two Parameter Model, Item Response Theory.*

References:

[1] ʿbạbnh. ʿmạd Gẖsạb, Ạlạkẖtbạrạt Mḥkyẗ Ạlmrjʿ Flsfthạ Wảss Tṭwyrhạ, (Ṭ1), ʿmạn: Dạr Ạlmsyrh Llnsẖr Wạltwzyʿ, (2009)

[2] ʿbd Ạlrḥmn. Ảḥmd Mḥmd, Tṣmym Ạlạkẖtbạrạt, (Ṭ1), ʿmạn: Dạr osạma Llnsẖr Wạltwzyʿ, (2011)

[3] ʿlam. Ṣlaḥ Aldyn Mḥmwd, Alqyas Waltqwym Altrbwy Walnfsy Ȧsasyatuh Wttbyqatuh Wtwjhatuh Almʿaṣrh, (T6), Alqahrh: Dar Alfkr Alʿrby, (2015)

[4] ʿlam. Ṣlaḥ Aldyn Mḥmwd, Alakẖtbarat Walmqayys Altrbwyh Walnfsyh, (T5), ʿman: Dar Alfkr, (2016)

[5] ʿwdh. Ȧḥmd Slyman, Alqyas Waltqwym Fy Alʿmlyh Altdrysyh, (T4), Ạrbd: Dar Alȧml Llnsẖr Waltwzyʿ, (2010)

[6] Ȧbw Fwdh. Basl Kẖmys & Bny Ywns. Njaty Ȧḥmd, Alakẖtbarat Althṣylyh, (T1), ʿman: Dar Almsyrh Llnsẖr Waltwzyʿ, (2012)

[7] Alȧmyr. Bndr Zyd Salm, Alȧkẖtaʾ Alsẖaẏʿh Fy Bnaʾ Mfrdat Akẖtbar Alakẖtyar Mn Mtʿdd Wtȧthyrha ʿla Alkẖṣaẏṣ Alsykwmtryh Llakẖtbarat Althṣylyh Lda Mʿlmy Alryaḍyat Balmrḥlh Almtwsth Bmdynẗ Mkh Almkrmh, Almjlh Alʿrbyh Llʿlwm Alajtmaʿyh, 5(9)(2016),137-175, https://doi.org/10.12816/0033017

[8] Alfqy. Asmaʿyl, Altqwym Walqyas Alnfsy Waltrbwy, (T3), Alryaḍ : Mktbẗ Alrsẖd, (2014)

[9] Alknany. ʿayd Krym & Fydy. Faḍl ʿbd, Faʿlyẗ Astkẖdam Mwqʿ Albdyl Alfʿal ʿnd ABʿad Albdyl Alkẖatẏ Fy Sʿwbẗ Fqrat Alȧsẏlh Mn Mtʿdd Wtmyyzha, Mjlẗ Alqadsyh Lʿlwm Altrbyh Ạlryaḍyh, 12(1)(2012), 207-226.

[10] Alnbhan. Mwsa, Ȧsasyat Alqyas Fy Alʿlwm Alslwkyh, (T1), ʿman: Dar Alsẖrwq Llnsẖr Waltwzyʿ, (2004)

[11] Alsalḥ. ʿly Mḥmd Kẖald, Mqarnẗ Tȧthyrat Tgẖyyr Mwqʿ Albdyl Alsḥyḥ Fy Akẖtbar Akẖtyar Mn Mtʿdd ʿla Tqdyrat Alsʿwbh Llfqrat Wtqdyrat Alqdrh Lda Alȧfrad Bastkẖdam Nmwdẖj Rasẖ, Almjlh Altrbwyh Aldwlyh Almtkẖṣṣh, 6(9) (2017), 124-133.

[12] Alsbḥ. Ranya Mḥmd Kẖyr Ḥsyn, Ȧthr Bnyẗ Fqrat Alakẖtyar Mn Mtʿdd Wmstwy Alqdrh Lda Alȧfrad ʿla Dqẗ Altqdyrat Lmʿalm Alfqrat Walȧfrad Wfq Nẓryẗ Alastjabh Llfqrh, Mjlẗ Jamʿẗ Ȧm Alqra Llʿlwm Altrbwyh Walnfsyh, 3(2)(2011), 45-110.

[13] Alsraf. Qasm ʿly, Alqyas Waltqwym Fy Altrbyh Waltʿlym, (T2), Alqahrh: Dar Alktab Alḥdyth, (2014)

[14] Altqy. Ȧḥmd Mḥmd, Alnẓryh Alḥdyth Fy Alqyas, (T2), ʿman: Dar Almsyrh Llnsẖr Waltwzyʿ, (2013)

[15] Ayala. Dy, Alnẓryh Waltṭbyq Fy Nẓryẗ Alastjabh Llfqrh, (T1), Trjmẗ: ʿbdallh Alkylany, Asmaʿyl Albrsan. Alryaḍ: Dar Jamʿẗ Almlk Sʿwd Llnsẖr, (2017)

[16] Ȧykyn. Lwys R, Alakẖtbarat Walamthanat Qyas Alqdrat Walȧdaʾ, (T1), Trjmẗ: Frḥ Alsraj. Alryaḍ: Sẖrkẗ Alʿbykan Llȧbḥath Waltṭwyr, (2007)

[17] Baker. F., The Basic of item Response Theory, (2th Ed), USA: Eric Clearing House on Assessment and Evaluation, (2001)

[18] Bykr. Frank B, Ȧss Nẓryẗ Alastjabh Llmfrdh, (T1), Trjmẗ: ʿbd Alrḥmn Altryry, Alsyd Ȧbw Hasẖm, Swsn Sẖlby. Alryaḍ: Alnsẖr Alʿlmy Walmṭabʿ – Jamʿẗ Almlk Sʿwd, (2010)

[19] Cizek. G. J., The Effect of Altering the Position of Options in a Multiple-Choice Examination, Educational and Psychological Measurement, 54(1)(1994),8-20, https://doi.org/10.1177/0013164494054001002

[20] Hambleton. R., & Swaminathan. H., Item Response Theory: Principles and applications, Boston: Kluwer Nijhoff Publishing, (1985)

[21] Hasẖm. Kmal Aldyn Mḥmd & Alkẖlyfh. Ḥsn Jʿfr, Altqwym Altrbwy Mfhwmh Ȧsalybh Mjalath Twjhath Alḥdyth, (T5), Alryaḍ: Mktbẗ Alrsẖd Nasẖrwn, (2015)

[22] Hatti. J., An empirical study of various indices for determining unidimensionality, Multivariate Behavioral Research, 19(1)(1984), 49-78, https://doi.org/10.1207/s15327906mbr1901_3

[23] Kolawole. E. B. & Olarotimi. J. T., A survey of the Anchor Bias in Mathematics Objective Tests in West African Examination Council (WAEC), Journal of Emerging Trends in Educational Research and Policy Studies (JETERAPS), 2(3)(2008),171-173.

[24] Lane. D. S. & Bull. K. S., The effects of Knowledge of item arrangement Gender And Statistical and cognitive item Difficulty on Test Performance, Educational and Psychological Measurement, 47(4)(1987), 865– 879, https://doi.org/10.1177/0013164487474002

[25] Nwadinigwe. P., I. & Naibi. L., The Number of Options in a Multiple-Choice Test Item and the Psychometric Characteristics, Journal of Education and Practice, Nigeria, 4(28)(2013), 189-196.

[26] Plake. B. Thompson. A. & Lowry. S., Effects of item arrangement and test anxiety on two scoring methods, The Journal of Experimental Education, 49(4)(1984) 214– 219, https://doi.org/10.1080/00220973.1981.11011786

[27] Qạsm. Yạsmyn Mḥmwd Ḍyf Ạllh, Ậṯḫr Mwqʿ Ạlbdyl Fy Fqrạt Ạkḫtbạrạt Ạlạkḫtyạr Mn Mtʿdd ʿla Nsbẗ Ạkḫtyạrha Bạltkḫmyn Kạjạbạt Ṣḥyḥh Llfqrh, Rsạlẗ Mạjstyr Gḫyr Mnsḫwrh, Jạmʿẗ Mẁth, Klyẗ Ạlʿlwm Ạltrbwyh, Ạlậrdn, (2013)

[28] Slymạn. Sḫạhr Kḫạld & Ạlṣạlḥ. ʿly Mḥmd Kḫạld, Ậṯḫr Mwqʿ Ạlbdyl Ạlṣḥyḥ Fy Ạkḫtbạr Ạkḫtyạr Mn Mtʿdd ʿla Tqdyrạt Mʿạlm Ạlfqrạt Wạlqdrh Wfq Ạlnmwdḫj Ạllwjsty Tḫlạtḫy Ạlmʿlmh, Drạsạt ʿrbyh Fy Ạltrbyh Wʿlm Ạlnfs, 90(2017), 98-120, https://doi.org/10.12816/0041974

[29] Wevric. L, Response set in a multiple-choice test, Educational and psychological Measurement, (2001) (33).