" أثر موقع البديل الصحيح في اختبار اختيار من متعدد على تقديرات معالم الفقرات والقدرة وفق النموذج اللوجستي ثلاثي المعلمة "

د/ شـــاهر خالـــد سليمـــان د/ علـــي محمـــد الصـــالـــح

• **المستخلص** :

هدفت الدراسة إلى الكشف عن أثر اختلاف موقع البديل الصحيح في اختبار اختيار من متعدد أربعة بدائل على الخصائص السيكومترية للفقرة وفق النموذج اللوجستي ثلاثي المعلمة (الصعوبة ، والتمييز ، والتخمين) ، وعلى تقديرات معلمة القدرة للأفراد ودالة معلومات الاختبار، ولتحقيق أهدف الدراسة تم بناء اختبار تحصيلي اختيار من متعدد أربعة بدائل في القياس والتقويم التربوي المقرر لطلاب دبلوم التربية العام مكون من (35) فقرة بصورته النهائية ، وقد أشتمل على ثلاثة نماذج لها نفس المحتوى والمتغير هو موقع البديل الصحيح ، النموذج الاول يكون فيه موقع البديل الصحيح على البديلين أ، ب ، والنموذج الثاني يكون فيه موقع البديل الصحيح على البديلين ج ، د، والنموذج الثالث توزيع البديل الصحيح عشوائيا على مختلف البدائل الأربعة، وطبقت النماذج الثلاثة على عينة عشوائية من طلاب وطالبات جامعة تبوك عدد أفرادها (1541) طالبا وطالبة ، استخدم برنامج Bilog-Mg3لتحليل البياناتٍ لكل نموذج من نماذج الاختبار الثلاثة. دلت نتائج الدراسة على وجود فروق دالة احصائيا بين المتوسطات الحسابية لمعاملات تمييز وتخمين الفقرات والثبات الامبريقي للاختبار تعزى لموقع البديل الصحيح ، ولا يوجد فروق دالة احصائيا بين المتوسطات الحسابية لمعاملات صعوبة الفقرات وتقديرات معلمة القدرة للأفراد تعزى لموقع البديل الصحيح ، كما يوجد تباين في كمية دالة معلومات الاختبار تبعا لتغير موقع البديل الصحيح .

الكلمات المفتاحية : موقع البديل الصحيح ، اختبار اختيار من متعدد ، معالم الفقرات ، قدرة الأفراد ، النموذج اللوجستي ثلاثي المعلمة .

The Impact of Correct Alternative Position in the Multiple-Choice Test on Items Parameters and the Ability According to the Three-Parameter Logistic Model

Dr. Shaher Khalid Suleiman

Dr .Ali Mohammad Al-Saleh

Abstract

The aim of the study is to detect the effect of differing correct alternative positions in multiple-choice test of four alternatives on the psychometric properties of an item according to the three-parameter logistic model (difficulty, discrimination and guessing), and on estimates of persons ability parameter and test information function. To achieve the aims of the study, it was built a multiple-choice achievement test of four alternatives in educational measurement and evaluation which scheduled for students of General Education Diploma consisted of (35) item in its final image. It included the three forms which have the same content and the variable is the

correct alternative position. The first form consists of the correct alternative position focusing on the two alternatives a and b. The second form consists of the correct alternative position focusing on the two alternatives c and d. The third form distributed correct alternative randomly at various four alternatives. The three forms have been applied on a random sample of students at the University of Tabuk which includes (1541) students. The Bilog-Mg3 program was used to analyze the data for each of the three tests forms. Results of the study indicated that there were statistically significant differences between the arithmetic averages of the items guess and recognition factors and consistency empirical testing due to the correct alternative position. Also, there was no statistically significant differences between the arithmetic averages of items difficult factors and estimates of persons ability parameter due to the correct alternative position. In addition, there was a variation in the quantitative information function test depending on the change of position of the correct alternative.

Key words: correct alternative position, multiple-choice test, item parameters, person's ability, logistic model three-parameter.

- **المقدمة** :

تعد اختبارات الاختيار من متعدد من أكثر الاختبارات الاختيارية شيوعًا في المؤسسات التربوية فالفقرة المحكمة البناء تتميز بكفاءتها ، وتعدد استعمالاتها ، بالإضافة إلى إمكانية استخدامها في قياس العمليات العقلية العليا بالرغم من صعوبة قياسها لمستويي التركيب والتقويم (Hoepfi,1994 أمبو سعيدي ، ٢٠٠٥).

وفي المقابل ؛ هناك جوانب ضعف لفقرة الاختيار من متعدد، فقد تستغرق الفقرة المحكمة البناء وقتا لإعدادها ، كما أنها تتصف بصعوبة توفير البدائل الجذابة ، وعدم إمكانية استخدامها في الغالب في قياس القدرة على التعبير عن الأفكار وتنظيمها (Gronlund,1993) .

وتتكون فقرة الاختيار من متعدد من مكونين : الأول على شكل سؤال ، والثاني على شكل عبارة ناقصة (Cantor,1995)، أما المكون الثاني للفقرة ؛ فهو البدائل Alternative وهي بمنزلة إجابات محتملة للسؤال الوارد في المتن Stem ، وأحد هذه البدائل هو الإجابة الصحيحة في حين تُعْرف بقية البدائل بالموهات (Distractors Crocker &Algina,1996) .

وحظي بناء اختبارات الاختيار من متعدد باهتمام المختصين في القياس منذ بداية القرن الماضي (Rodriguez,1997) إذ كثر الحديث عن قواعد صياغتها وبات من المهم تطوير قواعد لصياغتها تستند إلى التجريب أكثر من كونها عُرفاً (Klimko,1982) وأعتبرها رودجيوز (Rodriguez,1997) فنا وثمة من أعتبرها فنا وعلما (Withers,2005) ، كما أشار كل من جليسون وألي وبيكر (Gleason,Alley &Baker,2010) إلى ضرورة فحص كل فقرة والإجابة عنها لتحديد مدى التزام واضع الفقرة بقواعد صياغتها .

ويرى هامبلتون وسوامنياثان (Hambleton & swaminathan,1995) أن مستوى الطالب في الاختبار يتوقف على خصائص كل من الفقرة أو الاختبار وعلى خصائص الطالب الذي يتعرض للاختبار ، ويتأثر مستوى أداء الطالب في الاختبار بخصائص الفقرة من حيث الصعوبة والتمييز وفاعلية الموهات .

أما بالنسبة لصعوبة الفقرة يرى علام (٢٠٠٨ : ٥٦) أن من العوامل المهمة المؤثرة على صعوبة فقرات الاختيار من متعدد إضافة إلى قدرة الطالب على الإجابة ، عوامل فنية تتعلق بقواعد صياغة فقرات الاختيار من متعدد مثل درجة تجانس البدائل ، وسلامة الصياغة اللغوية ، والتوزيع العشوائي للبدائل الصحيحة.

أما فيما يتعلق بخصائص الطالب التي تؤثر في أدائه على الاختبار ، مستوى قدرة الطالب على السمة المقاسة ، وحكمته في الإجابة (Test Wiseness)، وقد عُرّفت حكمة الاختبار بأنها قدرة الفرد على استغلال خصائص الاختبار وطريقة بنائه في الحصول على درجات مرتفعة في حالة عدم معرفته بمحتوى فقرات الاختبار ، ومن الاستراتيجيات المتعلقة بحكمة الاختبار استراتيجيات تتعلق بقواعد بناء الاختبار مثل الاستفادة من خبرة الطالب ببناي الاختبار في تمييز البديل الصحيح كميل واضع الاختبار لوضع البديل الصحيح في موقع مميز بين بدائل الفقرة (سليمان ، ٢٠١٤) .

ويرى ويتن (Weiten ,1984) أن قواعد صياغة فقرات الاختبار التحصيلية تمثل نتاج خبرة عملية ، وأحكام خبراء مبنية على المنطق في هذا المجال والخلل أو العيوب التي قد يقع بها واضع الاختبار ، وأشار المالكي (٢٠١٥) أن معلمي العلوم في المرحلة المتوسطة يلتزمون ب ٦٠ ٪ من قواعد بناء أسئلة الاختيار من متعدد أربعة بدائل ، وقد يعزى ذلك إلى أن معظم المدرسين مضغوطين بالوقت ومتطلبات وظيفتهم ، إضافة إلى عدم معرفة البعض بكيفية صياغة الأسئلة ، وبالنتيجة فإن اختباراتهم قد تكون عرضة للتلوث بفقرات ضعيفة الصياغة ، وإن وجود هذه العيوب تكون مؤشرات لما يعرف بالحكمة في الإجابة لدى بعض الطلاب ، مما قد يؤثر سلباً على خصائص الفقرة والاختبار بالإضافة إلى تضخيم درجات الاختبار بشكل مضلل ، لأن فقرات الاختبار بدلاً من أن تقيس السمة المراد قياسها أصبحت تقيس سمة أخرى دخيلة ، وهي حكمة الاختبار أو المثابرة واستغلال الخلل للحصول على درجات مرتفعة على الاختبار (Sarnaki,1997).

ويرى ننلي (Nunnally, 1982) أن الشخص الذي يعرف الإجابة فإن بديلاً واحداً من بدائل الفقرة يشده ، وهو المقبول لديه وهو البديل الصحيح ، بينما الشخص الذي لا يعرف الإجابة ، فإن كل البدائل تبدو له متساوية الأهمية

وبدائل ذات احتمال معقولة له ، وبالتالي فإن احتمال اختيار بديل احتمال واحد لجميع البدائل ، وإن أي عيب في أي من البدائل بطريقة ما تجعله يشد انتباه بعض الطلبة ممن يجهلون الإجابة عن الفقرة ، مما يغير نسبة احتمال الإجابة عن ذلك البديل .

وتعد نظرية استجابة الفقرة (Item Response Theory) من التطورات الحديثة في مجال القياس والتقويم التربوي والنفسي بسبب ما قدمته من طرق سيكومترية ذات فاعلية كبيرة في بناء المقاييس التربوية والنفسية وطريقة تفسير الدرجات على هذه المقاييس مقارنة بالنظرية التقليدية في القياس وتقوم نظرية استجابة الفقرة على افتراض وجود متصل على السمة بحيث يمكن تقدير احتمال الإجابة الصحيحة للفرد عن فقرة إذا علم موقعه على هذا المتصل (Warm ,2008:56)

ولنظرية استجابة الفقرة مزايا منها : يكون تقدير قدرة الفرد مستقلاً عن عينة الفقرات التي تطبق عليها أي أن تقدير قدرات الأفراد متحررة من الفقرات وتقديرات معالم الفقرات مستقلة عن عينة الأفراد التي استخدمت في تقدير هذه الخصائص ، وأنه يمكن الحصول على إحصائي (الخطأ المعياري في التقدير) لتقدير درجة الدقة في قياس القدرة لكل فرد ويختلف الإحصائي من فرد لآخر(Hambleton and swaminathan& Rogers,1999:11). وتعد استقلالية القياس بمثابة الفرق الجوهري بين النظرية التقليدية ونظرية استجابة الفقرة وهو ما عبّر عنه لورد (Lord,1980) بخاصية اللاتغاير Invariance (Rup&Zumbo,2006) ، هذا وقد بنيت هذه النظرية على افتراضات ينبغي تحقيقها في البيانات لكي تؤدي إلى نتائج يمكن الوثوق بها ، ومن أهم الافتراضات :

- **أحادية البعد (Unidimensionality)** : يعني أن هناك قدرة واحدة تفسر أداء الفرد في الاختبار لذا تسمى بالنماذج أحادية البعد ، أم النماذج التي تفترض وجود أكثر من قدرة واحدة تكمن وراء هذا الأداء تسمى متعددة الأبعاد .
- **الاستقلال الموضعي(Local Independence)**: أي أنٍ تكون استجابات الفرد على الفقرات المختلفة في الاختبار مستقلة استقلالا احصائياً عند مستوىٍ معينٍ من القدرة أي أن استجابة الفرد عن فقرة ما يجب أن لا تؤثر سلباً أوإيجاباً على استجابته لفقرة أخرى (Croker&Algina,1986) .
- **منحنى خصائص الفقرة Item Characteristic Curve (ICC)**: وهو اقتران رياضي يربط بين احتمال استجابة الفرد استجابة صحيحة على فقرة الاختبار، وبين القدرة التي تقيسها مجموعة الفقرات التي تحتوي على تلك الفقرة، ويشير بوضوح إلى أن احتمال إجابة الفقرة إجابة صحيحة يزداد بازدياد قدرة المفحوص، بسبب أن المنحنى تراكمي صاعد، وتوصف هذه

المنحنيات في نماذج الاختبارات المصممة لقياس سمة واحدة (أحادية البعد) بدلالة بارامتر واحد، أو بارامترين، أو ثلاثة بارامترات.

- **السرعة في الأداء (Speediness)**: تفترض أن عامل السرعة ليس له دور في الإجابة عن الفقرة ، أي أن خصائص الفرد في الإجابة عن فقرات الاختبار يرجع إلى انخفاض قدرته وليس إلى تأثير عامل السرعة في إجابته.(Allen & Yen ,1989: 97)

ويوجد مجموعة من الافتراضات خاصة بكل نموذج من نماذج استجابة الفقرة البارامترية، حيث إن النموذج الثلاثي يعتبر الحالة العامة للنماذج، وذلك لأنه يتم فيه تقدير بارامتر الصعوبة (b)، ومعلمة التمييز (a)، ومعلمة التخمين (c).

- والنموذج اللوجستي ذو الثلاثة معالم (الصعوبة ، والتمييز ، التخمين) ، وهذا النموذج هو المستخدم في الدراسة الحالية ، و يمثل بالعلاقة الآتية:

$$pi(\theta) = ci + (1 - ci) \quad \frac{e^{Dai(\theta - bi)}}{1 + e^{Dai(\theta - bi)}}$$

PI (θ) : احتمال إجابة المفحوص الذي اختير عشوائيا من مستوى القدرة (θ) على الفقرة (i) اجابة صحيحة، bi: معلم الصعوبة ، ai: معلم التمييز، ci معلم التخمين، (θ) معلم القدرة ، a : المستوى المشترك من التمييز لكافة الفقرات ، D : تمثل عامل التدريج (Scaling Factor) وقيمته(١.٧) .

وتلعب دالة معلومات الاختبار دوراً مهماً في نظرية استجابة الفقرة إذ يمكن من خلالها تحديد الخطأ المعياري في التقدير وتتميز دالة معلومات الاختبار بأنها تمثل مجموع دوال معلومات الفقرات عند مستوى معين من القدرة ، ودالة معلومات الاختبار مستقلة عن عينة المفحوصين وبذلك تقدم نظرية الاستجابة للفقرة مميزات إضافية فيما يتعلق بزيادة القدرة على تقدير أخطاء القياس Adedoyin, , Nenty & Chilisa,. (2008) والخطأ في التقدير يرتبط عكسيا مع الجذر التربيعي لدالة معلومات الاختبار وفق العلاقة التالية:

$$SE = \frac{1}{\sqrt{I(\theta)}}$$

حيث إن :

I (θ) : تمثل دالة معلومات الاختبار

SE (θ) : تمثل الخطأ المعياري في التقدير

وهذا يعني أن الخطأ المعياري في التقدير يكون أقل ما يمكن عند مستويات القدرة التي تناظر أقصى معلومات . ومن الاستخدامات المفيدة لدالة معلومات

الفقرة إمكانية معرفة مدى مساهمة كل فقرة في دالة معلومات الاختبار بشكل مستقل عن الفقرات الأخرى للاختبار ، فإذا علمنا قدرات مجموعة من المفحوصين فيمكن انتقاء فقرات الاختبار التي تعمل على زيادة المعلومات التي يقدمها الاختبار في المدى الذي تتوزع فيه قدرات المفحوصين في الاختبار (Brannick,2003).

وفي جامعة أريزونا في الولايات المتحدة قامت كارتر(Carter,1986) الأمريكية بدراسة بعنوان" الحنكة في الإجابة لكل من الطلبة والمعلمين " ، إذ جاءت الدراسة نتيجة انتشار برامج تعلم الطلبة مهارة أخذ الاختبار .وجرى إعداد اختباراً مؤلفا من خمس فقرات قام بكتابتها مجموعة من الخبراء عالجوا بها خمس معضلات اكتشفوها نتيجة مراجعة اختبارات الاختيار من متعدد التي يقوم المعلمون بصياغتها، وقد صيغت الفقرات الخمس على نص أدبي في مستوى الصف السابع، ولا يستطيع الطالب الإجابة عن هذه الأسئلة دون قراءة النص الأدبي، أو استغلال المؤشرات الثانوية الموجودة في كل فقرة من الفقرات الخمس، والدالة على الإجابة الصحيحة.

وجرى تطبيق الفقرات على عينة مؤلفة من(٣١٢) طالباً من طلبة الصف السابع في أريزونا، ثم عملت مقابلة مع هؤلاء الطلبة ليقرروا ما الاستراتيجيات التي تم اتباعها في الإجابة عن الفقرات الخمس. ومن أهم نتائج الدراسة أن نسبة الطلبة الذين اختاروا البديل المتوسط كانت ٦٩.٢٣ ٪، وفي مقابلة مع الطلبة أكدوا أن البديل(ج)غالبا ما يكون صحيحا، وهذا ما تعودوه من مدرسيهم، كما وجد أن ٨٦٪من الطلبة قد اختاروا البديل الأطول بين بدائل الفقرة، لذا جعل الاختيار للبديل الأطول أحد عناصر حنكة الطالب في الإجابة حتى يتسنى على الطالب اختيار البديل الأطول، عندما لا يعرف الإجابة عن الفقرة .

أجرى ويفرك (Wevrick,2001) دراسة بعنوان الاستعداد الخاص للإستجابة في اختبار اختيار من متعدد إذا كان البديل الصحيح يحتل موقعا متحيزا (البديل الأوسط) لتحقيق هدف الدراسة صمم الباحث اختبار اختيار من متعدد ثلاثة بدائل يتكون (١٠٠) فقرة يشمل ثلاثة نماذج: الأول يكون موقع البديل الصحيح في الوسط والثاني البديل الوسط يتضمن مموه للإجابة والنموذج الثالث موقع البديل الصحيح موزع توزيعا عشوائيا طبقت نماذج الاختبار عينة عشوائية تكونت (٤٦٧) طالبا جامعيا . دلت النتائج 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 موزعا عشوائيا فإنه لا تأثير يذكر لما يسمى باستعداد الطالب الخاص للإجابة عن الفقرات بطريقة معينة .

وفي دراسة قام بها السويلم (٢٠٠٣) بعنوان القواعد العامة في صياغة أسئلة الاختيار من متعدد ومدى التزام أساتذة كلية التربية بجامعة الملك سعود بهذه القواعد، وهدفت الدراسة إلى تقديم دليل توضيحي يساعد المعلم والتربوي في صياغة وبناء أسئلة الاختيار من متعدد ذات كفاءة عالية في دقة التشخيص وسلامة التقويم، مع التعرف على ممارسات أعضاء هيئة التدريس في كلية التربية نحو أسئلة الاختبارات بشكل عام ، وعلى درجة الدقة وسلامة الصياغة لأسئلة الاختيار من متعدد في الاختبارات التي يستخدمها أعضاء هيئة التدريس في كلية التربية ،وذلك بناء على درجة التزامهم بالقواعد العامة المذكورة في الدليل .تكونت عينة الدراسة من (٩٧) عضوا من أعضاء هيئة التدريس بكلية التربية ، كما اشتملت عينة تحليل محتوى الأسئلة على(٤٨٩) سؤالا . دلت نتائج الدراسة أن أكثر الأنواع الاسئلة استخداما في الاختبارات الأكاديمية هو أسئلة الاختيار من متعدد ،إلا أنها كشفت عن قصور وضعف صدقها المحكي حيث إن ٣.٦٥ ٪ من الأسئلة المستخدمة يحتوي على مخالفات للشروط السيكومترية الواجب توافرها في سؤال الاختيار من متعدد وأوصت الدراسة بضرورة الاعتناء بكفاءة أسئلة الاختبارات، وأهمية تطوير قدرات أعضاء هيئة التدريس لبناء اختبارات صادقة نظرا لضخامة القرارات المرتبطة بنتائج تلك الاختبارات .

وأجرت الحارثي(٢٠٠٧) دراسة هدفت إلى تقويم كفايات بناء الاختبارات التحصيلية لدى معلمات المرحلة الثانوية تبعا لمتغيرات الدراسة (التخصص والخبرة ، التدريب) وقد استخدمت الباحثة المنهج الوصفي في دراستها من خلال بطاقة تحليل الأسئلة المعدة من قبل عينة من معلمات المرحلة الثانوية التي بلغت (٢٥١) معلمة بمحافظة جدة . دلت النتائج على تدني درجة معرفة وممارسة معلمات العلوم لكفايات بناء الاختبار التحصيلية ، وعلى وجود فروق ذات دلالة إحصائية في متغير التدريب لصالح المتدربات .

كما أجرى سليمان (٢٠٠٨) دراسة هدفت إلى استقصاء أثر مخالفة قواعد صياغة فقرات اختبار الاختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته ، ولتحقيق هذا الهدف تم إعداد اختبارا تحصيليا في مادة علم النفس التربوي ، وهو أحد المواد التي تطرح في الجامعة ، وتكون الاختبار من (٤٠) فقرة مخالفة من نوع اختيار من متعدد أربعة بدائل بواقع (١٠) فقرات لكل نوع من المخالفات الأربع الآتية :طول البديل الصحيح ، موقع البديل الصحيح البديل جميع ما ذكر صحيح ، ، البديل كلاهما صحيحان ، ويناظره اختبارا مكونا من (٤٠) فقرة متحررة من المخالفات ولها نفس المحتوى ، طبق الاختبار على عينة عشوائية مكونة من (٤٨٢) طالبا من جامعة تبوك . دلت النتائج أن الفقرات المخالفة لقواعد الصياغة ، أسهل من نظيرتها الفقرات المتحررة

من المخالفات وبفارق دال احصائيا ، ولا توجد فروق دالة احصائياً بين معاملات تمييز الفقرات المخالفة الصياغة ونظيرتها المتحررة من المخالفات ، كما يوجد فروق دالة احصائيا بين معاملات صدق وثبات الاختبار الذي فقراته مخالفة لقواعد صياغة الفقرات والاختبار الذي فقرات متحررة من المخالفات والفروق لصالح الاختبار الذي فقراته متحررة من المخالفات .

وأجرى الكناني وفيضي (٢٠١٢) دراسة بعنوان فاعلية استخدام موقع البديل الصحيح عند ابعاد البديل الخاطئ في صعوبة فقرات لأسئلة اختيار من متعدد وتمييزها . هدفت الدراسة إلى بناء مقياس البديل الصحيح في أسئلة الاختيار من متعدد لمادة طرائق تدريس التربية الرياضية ، ومن خلال سعي الباحثان إلى الكشف عن أثر تغير موقع البديل الصحيح عند استخدام تقنية إبعاد البدائل الخاطئة على تحصيل الطلبة في أسئلة الاختيار من متعدد ، اشتملت عينة الدراسة على (٦٠) طالبا من طلبة المرحلة الثالثة لكلية التربية الرياضية في جامعة المثنى في العراق ، حيث تم تقسيم عينة الدراسة إلى مجموعتين متساويتين حيث أجاب (١٣٠) طالبا وطالبة عن أسئلة النموذج الأول كما أجاب (١٣٠) طالبا عن أسئلة النموذج الثاني . دلت النتائج على عدم وجود فروق معنوية بين نموذجي الاختبار من حيث الصعوبة والتمييز.

كماأجرى يعقوب ، وأبو فودة (٢٠١٢) دراسة هدفت إلى معرفة أثر مخالفة قواعد صياغة فقرات الاختيار من متعدد في الخصائص السيكومترية للفقرات (الصعوبة، والتمييز)، والخصائص السيكومترية للاختبار (الثبات، والصدق). ولتحقيق ذلك تمّ بناء اختبار من نوع الاختيار من متعدد في مبحث الرياضيات تكون من ٤٠ فقرة ذات أربعة بدائل، ثم جرى صياغة فقرات الاختبار في نموذج آخر مخالف لقواعد الصياغة. جرى تطبيق نموذجي الاختبار على عينة مكونة من ٣٠٠ طالب، بواقع ١٥٠ طالبا لكل نموذج، وفي ظروف صفية عادية. وبعد إجراءات التطبيق حسبت قيم معاملات صعوبة الفقرات وتمييزها في نموذجي الاختبار، وكذلك معاملات صدقها وثباتها. فأظهرت النتائج أن الفرق كان دالاً إحصائياً بين صعوبة وتمييز الفقرات التي لم تحتوي المخالفات في قواعد الصياغة، أو التي احتوتها، وبينت النتائج أن الفرق بين قيم معاملي ثبات الاتساق الداخلي في نموذجي الاختبار كان دالاً إحصائياً، وكان الفرق بين قيم معاملي صدقها دالاً أيضاً. الكلمات المفتاحية: مخالفة قواعد صياغة فقرات الاختيار من متعدد، صعوبة الفقرة، تمييز الفقرة، معامل ثبات الاتساق الداخلي، معامل صدق المحك التلازمي.

وأجرى ظاظا (٢٠١٢) دراسة تهدف إلى الكشف عن مدى انتهاك قواعد صياغة فقرات الاختيار من متعدد في أسئلة امتحان شهادة الدراسة الثانوية العامة في الأردن في ضوء مراجعة (٢٩١٠) فقرات للأعوام ٢٠٠٧ ــ ٢٠١٠م باستخدام (١٨)

قاعدة للكشف عن مخافة الفقرة لقواعد صياغتها . أظهرت النتائج أن (٤١٦) فقرة أي ما نسبته ١٢.٦ ٪ تضمنت أكثر من مخالفة ، وبشكل اجمالي وجدت ٧٨٢ فقرة أي ما نسبته٢٦.٩٪فقرة مخالفة ، وتبين أن أكثر من قاعدتين انتهاكاً كانتا : ترتيب البدائل عمودياً بنسبة ٦٧.٧ ٪ وترتيب البدائل منطقيا أو عدديا أو أبجديا بنسبة ٩.٨ ٪ في حين حافظت الفقرات على عدم انتهاك ست قواعد كما توصلت الدراسة إلى أن امتحان الثانوية العامة في الأردن بشكل عام يراعي قواعد صياغة فقرات الاختيار من متعدد .

وأجرى بني عطا ، و الرباعي (٢٠١٣) دراسة هدفت إلى التحقق من أثر عدد بدائل الفقرة وموقع المموه القوي في فقرات اختبار اختيار 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 (٤١) فقرة بصورته النهائية ، وقد اشتمل على أربعة نماذج حسب عدد البدائل ، وكان عدد أفراد عينة الدراسة (٢١١١) طالبا وطالبة لجميع نماذج الاختبارات الأربعة ، استخدم الباحثان النموذج اللوجستي ثلاثي المعلمة .كشف النتائج أنه لا توجد فروق ذات دلالة احصائية بين متوسطات معالم الصعوبة والتمييز والتخمين تعزى لعدد البدائل أو موقع المموه.

وتأتي هذه الدراسة للكشف عن أثر موقع البديل الصحيح في اختبار اختيار من متعدد أربعة بدائل على معالم الفقرة : الصعوبة والتمييز والتخمين باستخدام نظرية استجابة الفقرة .

- **أهداف الدراسة** :

تهدف الدراسة الحالية إلى الآتي :

- الكشف عن أثر تغير موقع البديل الصحيح في اختبار اختيار من متعدد على خصائص الفقرة وفق نظرية استجابة الفقرة .
- الكشف عن أثر تغير موقع البديل الصحيح في اختبار اختيار من متعدد على الخصائص السيكومترية للاختبار وفق نظرية استجابة الفقرة .
- التعرف على أثر تغير موقع البديل الصحيح على قيمة معلمة القدرة للأفراد على الاختبار .
- التعرف على أثر تغير موقع البديل الصحيح على قيمة دالة المعلومات للاختبار .

- **أسئلة الدراسة** :

- السؤال الأول : هل تختلف تقديرات معالم الفقرات (الصعوبة ، التمييز التخمين) باختلاف موقع البديل الصحيح في اختبار اختيار من متعدد أربعة بدائل ؟

- السؤال الثاني : هل تختلف تقديرات معلمة القدرة للأفراد (θ) باختلاف موقع البديل الصحيح في اختبار اختيار من متعدد أربعة بدائل ؟
- السؤال الثالث : هل تختلف دالة المعلومات للاختبار باختلاف موقع البديل الصحيح في اختبار اختيار من متعدد أربعة بدائل ؟

- **حدود الدراسة** :

- اقتصرت عينة الدراسة على طلاب وطالبات جامعة تبوك تخصص دبلوم عام في التربية للعام الدراسي ١٤٣٤ ه/ ١٤٣٥ هـ.
- اقتصرت الدراسة على ثلاثة مواقع للبديل الصحيح النصف الأول ، النصف الثاني، والعشوائي في اختبار اختيار من متعدد أربعة بدائل.
- اقتصرت الدراسة على استخدام النموذج اللوجستي ثلاثي المعلمة كأحد نماذج نظرية استجابة الفقرة للتأكد من صدق وثبات للاختبار وخصائص الفقرات من حيث مطابقتها للنموذج.

- **مصطلحات الدراسة** :

- **النموذج اللوجستي ثلاثي المعْلم** :

هو أحد نماذج نظرية الاستجابة للفقرة ثنائية التدريج حيث يمكن لهذا النموذج تقدير ثلاثة معالم للفقرة هي : الصعوبة ، التمييز ، التخمين .

- **معالم الفقرة** :

هي معالم الصعوبة والتمييز والتخمين المنبثقة عن النموذج اللوجستي ثلاثي المعلمة .

- **معْلمة الصعوبة** :

هي نقطة على متصل القدرة تقابل احتمال (1+ci/2) للإجابة عن الفقرة (i)إجابة صحيحة حيث يشير الرمز(ci) إلى معلمة التخمين.

- **معْلمة التمييز** :

هي ميل منحنى خصائص الفقرة الذي يحدث عنده تغيير في اتجاه المنحنى (Inflexion) الذي يقابل الصعوبة على متصل القدرة .

- **معْلمة التخمين** :

هي احتمال إجابة فرد ذو قدرة متدنية على فقرة ما إجابة صحيحة .

- **الطريقة والاجراءات** :

- **عينة الدراسة** :

تكونت من جميع أفراد مجتمع الدراسة وهم طلاب وطالبات جامعة تبوك تخصص الدبلوم العام في العام الدراسي ١٤٣٤ / ١٤٣٥ هـ ، وبلغ عددهم (١٥٤٥) طالبا وطالبة .

- **أداة الدراسة** :

لقد مر بناء اداة الدراسة بالمراحل الآتي :

- تحديد الغرض من الاختبار وهو قياس التحصيل الدراسي للطلاب في مقرر القياس والتقويم ومدى إتقانه لمهارات القياس .
- تحديد محتوى الاختبار وذلك بالرجوع الى وصف المقرر والأدب النظري المتعلق بموضوع القياس التربوي وقد انحصر المحتوى بالموضوعات الآتية : مفاهيم أساسية في القياس والتقويم ، أنواع الاختبارات ، بناء الاختبارات تحليل نتائج الاختبار .
- صياغة الاهداف السلوكية التي تناولتها الموضوعات والمتوقع إتقانها من قبل الطلبة.

قام الباحثان بصياغة (٤٠) فقرة بصورة أولية من نوع الاختيار من متعدد أربعة بدائل أحدها يمثل الإجابة الصحيحة لقياس كل هدف من الأهداف السلوكية ، وقد تطلب قياس بعضها أكثر من فقرة ، وقد روعيت المبادئ الأساسية التي تحكم صياغة هذا النوع من الفقرات ومطابقتها للهدف السلوكي الذي تقيسه من حيث المحتوى والمستوى المعرفي. كما أعد الباحثان ثلاثة نماذج من الاختبار جميعهم يتكون من نفس المحتوى والفقرات ، النموذج الأول موقع البديل الصحيح في النصف الأول (البديل أ ، أو البديل ب) والنموذج الثاني نفس محتوى النموذج الأول لكن موقع البديل الصحيح في النصف الثاني من البدائل (البديل الصحيح ج ، أو البديل د)، والنموذج الثالث يتكون من نفس الفقرات باستثناء موقع البديل الصحيح موزع بشكل عشوائي وتم تطبيق النماذج على عينة مكونة من(١٥٤٥) طالبا وطالبة تخصص الدبلوم العام في التربية .

- **صدق الاداة**:

وللتأكد من صدق المحتوى ، عرضت فقرات الاختبار بصيغتها الأولية وعددها (٤٠) فقرة على عشرة محكمين من أعضاء هيئة التدريس في الجامعات الذين يدرسون مقرر القياس والتقويم التربوي وطلب من كل منهم أن يبدي رأيه في : درجة الاتفاق بين الأهداف السلوكية والهدف الأساسي للاختبار وتمثيلها له ، ومدى قياس الفقرات للهدف السلوكي المتعلق بها ، ومدى اتساق الاختبار ككل مع الهدف الأساسي منه ، وبناء على ملاحظاتهم لم يتم حذف فقرات ، وإنما جرى تعديل على بعض الفقرات ، و أجمع المحكمون أن هناك فقرات تؤدي غرضها ، وأنها عينة ممثلة للمجال السلوكي الذي يقيسه.

- **جمع وتحليل البيانات** :

لتحقيق الهدف من الدراسة، قام الباحثان بتطبيق نماذج الاختبار عشوائياًعلى أفراد عينة الدراسة والبالغ عددهم (١٥٤٥)طالباًوطالبة، وبعد تطبيق نماذج الاختبار على أفراد عينة الدراسة، تم فرز كل نموذج على حده وتصحيحه، ثم أدخلت البيانات إلى ذاكرة الحاسوب، واستخدمت البرامج

الإحصائية لإجراء التحليلات الإحصائية وهي: (Bilog- Mg3) ، وبرنامج (SPSS-20) للإجابة عن أسئلة الدراسة، وقد تمت عملية التحليل وفق المراحل التالية:

- **أولا :التحقق من افتراضات النموذج** :

تم التحقق من افتراضات النموذج اللوجستي ثلاثي المعلمة وهي أحادية البعد لما له من أثر على دقة التقديرات، وذلك باستخدام التحليل العاملي لكل نموذج من نماذج الاختبار الثلاثة باستخدام طريقة المكونات الرئيسة (Principle Component Analysis) ويظهر الجدول (١) قيم الجذر الكامن (Eigen Values) ونسب التباين المفسر (Explained Variance) للعامل الأول والثاني وناتج قسمة الجذر الأول على العامل الثاني في نماذج الاختبار الثلاثة .

جدول (١) 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	الجذر الكامن للعامل الاول / الجذر الكامن للعامل الثاني
الاول	الجذر الكامن	٩.٣٤	٣.٧٩	٢.٤٦
	التباين المفسر	٢٣.٥	٩.٤٨	
الثاني	الجذر الكامن	١١.٤٣	٥.٠١	٢.٢٨
	التباين المفسر	٢٤.١١	١٠.٤٢	
الثالث	الجذر الكامن	١٠.٢	٤.٠٢	٢.٥٤
	التباين المفسر	٢٣.٧٩	٩.٤٦	

يتضح من الجدول (١) بأن جميع نسب التباين المفسر من العامل الأول كانت أعلى من(٢٠) لكل نماذج الاختبار، وأن ناتج قسمة الجذر الكامن للعامل الأول على العامل الثاني أكبر من (٢)، مما يشير إلى وجود سمة سائدة للاختبار؛ وهذا يعني تحقق افتراض أحادية البعد للاختبار (Hatti، 1984,87) .

- **ثانيا :مطابقة البيانات للنموذج** :

بهدف التحقق من درجة مطابقة الأفراد والفقرات للنموذج اللوجستي ثلاثي المعلمة، تم استخدام برنامج:Bilog-Mg3 لتحليل البيانات لكل نموذج من نماذج الاختبار الثلاثة، حيث أظهرت نتائج التحليل عدم مطابقة طالبا واحدا فقط في نموذجي الاختبار الثاني والثالث (النموذج الثاني ـ مواقع البديل الصحيح توزعت في ج ، د ، النموذج الثالث ـ مواقع البديل الصحيح توزعت عشوائيا) والبالغ عددهم في الأصل (٥١٠ ، ٥١٣) على الترتيب ، وكذلك عدم مطابقة ثلاثة طلاب من طلاب نموذج الاختبار الأول (النموذج الأول ـ مواقع البديل الصحيح توزعت في أ ، ب) والبالغ عددهم في الأصل (٥٢٢) طالبا. لذا أصبحت عينة الدراسة بصورتها النهائية بعد استبعاد الطلبة غير المطابقين للنموذج (١٥٤١) طالبا وطالبة .

وفيما يتعلق بفحص مطابقة الفقرات للنموذج المستخدم أُعيد التحليل باستخدام برنامج Bilog-Mg بعد حذف استجابة الطلبة غير المطابقين

للنموذج وبناءً على مؤشر مربع كاي عند مستوى الدلالة α = ٠.٠٥ الذي يعتبر الفقرة غير مطابقة للنموذج إذا كان مقدار الاحتمالية لهذه الفقرة أقل من ٠.٠٥حيث أظهرت نتائج التحليل عدم مطابقة الفقرات ذات الأرقام ٤٠ ، ٣٩ ، ٣٧ ، ٦ ، ٥ للنموذج الأول ، و عدم مطابقة الفقرات ذات الأرقام ٤٠ ، ٣٧ ، ٦ ، ٥ للنموذجين الثاني والثالث .

وبعد حذف جميع الفقرات المشتركة في جميع نماذج الاختبارات الثلاثة والبالغ عددها (٥) فقرات أصبح الاختبار بنماذجه الثلاثة يشتمل على(٣٥)فقرة

- **المعالجات الإحصائية** :

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

- إيجاد معاملات الصعوبة، التمييز، التخمين لكل فقرة في كل نموذج من نماذج الاختبار الثلاثة حسب متغير موقع البديل الصحيح باستخدام النموذج ثلاثي المعلمة .
- استخدام تحليل التباين متعدد المتغيرات MANOVA (Multivariate Analysis of Variance)
- استخدام اختبار M للكشف عن الفروق في معاملات الثبات الامبريقي لنماذج الاختبار.
- استخدام تحليل التباين الأحادي لاختبار أثر موقع البديل الصحيح على تقديرات معلمة القدرة لدى الأفراد .

- **النتائج ومناقشتها** :

للإجابة عن السؤال الأول " **هل تختلف تقديرات معالم الفقرات (الصعوبة التمييز، التخمين) باختلاف موقع البديل الصحيح ؟** "

تم تقدير قيم معالم الفقرات الصعوبة، التمييز، التخمين باستخدام النموذج اللوجستي ثلاثي المعلمة لكل نموذج من نماذج الاختبار الثلاثة حسب متغير موقع البديل الصحيح ، وتم أيضاحساب المتوسطات الحسابية والانحرافات المعيارية لقيم هذه المعالم .ويوضح جدول (٢) قيم معالم الفقرات والمتوسطات الحسابية لكل نموذج من نماذج الاختبار.

يتضح من الجدول (٢) تبايناً ظاهرياً في مدى تمييز الفقرات وقد كان المدى في تمييز الفقرات هو الأعلى في النموذج الثالث (توزيع البديل الصحيح عشوائيا) الذي تراوحت قيمه من ٠.٤٦ الى ٢.٧٨ بمتوسط حسابي مقداره ١.٣٢ في حين كان الأقل في النموذج الأول (توزيع البديل الصحيح على أ، ب) الذي تراوحت قيم معاملات تمييز فقراته من ٠.٢٣ الى ٢.٢٨ بمتوسط حسابي مقداره ٠.٨٧

جدول (٢) قيم معالم الفقرات(الصعوبة، التمييز، التخمين)والمتوسطات الحسابية والانحرافات المعيارية لقيم معالم الفقرات لكل نموذج من نماذج الاختبار حسب متغير موقع البديل الصحيح .

رقم الفقرة	نموذج الاختبار								
	الأول :توزيع البديل الصحيح على أ ، ب			الثاني : توزيع البديل الصحيح على ج ، د			الثالث : توزيع البديل الصحيح عشوائيا		
	التمييز	الصعوبة	التخمين	التمييز	الصعوبة	التخمين	التمييز	الصعوبة	التخمين
١	٠.٦٣	٢.٤٧	٠.١٥	٢.٠٩	١.٠٢	٠.١١	٢.٢٢	١.٧٣	٠.١٩
٢	٠.٧٣	٢.٩٧	٠.١٠	١.٢٩	٢.٦٠	٠.١٨	٠.٨٧	٣	٠.٠٦
٣	١.٧٩	٠.٥٥	٠.١٢	١.١٦	١.٠٩	٠.١٧	١.٠٧	١	٠.١٥
٤	١.٤٥	١.٦٢	٠.٢٩	٢.٠٥	٢.٢٧	٠.٢٤	٠.٧٧	٢.١٤	٠.٢٩
٥	٠.٧٧	٢.٤٧	٠.٢١	١.٦٩	٠.٩٤	٠.٠٨	١.١٨	١.٦٤	٠.٤
٦	٠.٣٣	٢.٩٨	٠.١٣	٢.٥٧	٠.٨٥	٠.١١	٢.٢٩	١.٣٣	٠.٣٥
٧	٠.٧٥	١.٥٨	٠.٣٥	٢.٣٨	١.٠٢	٠.٣	٢.٣٠	١.٨٦	٠.٣
٨	٠.٣٤	٢.٩٦	٠.٣٢	٢.٦٠	١.٤٥	٠.١٧	٢.٤٥	١.٧٠	٠.١٩
٩	١.٧٢	٠.٦٤	٠.٣٥	١.٥٨	١.٢١	٠.٣١	٠.٨٣	١.٤٧	٠.٢٩
١٠	٠.٧٣	١.٠٧	٠.٢٩	١.٢٧	٢.٨٤	٠.١٢	٠.٤٧	٢.٨٨	٠.٣٤
١١	٠.٣٩	٢.٧٥	٠.٣٠	١.٣٤	٢.١٥	٠.٠٩	٠.٥٨	١.٣٨	٠.٢٩
١٢	٠.٣٦	٠.١٣	٠.٣٤	١.٣٤	٠.٩٤	٠.٠٩	٠.٨٥	١.٠٣	٠.٢٨
١٣	٠.٧٦	٠.٣٤	٠.٢٣	٢.٧٥	١.٢٢	٠.٣٤	٢.١٤	١.٨٥	٠.٣٦
١٤	٠.٥٥	٠.٢٧	٠.٣٢	٠.٧١	٢.٥٦	٠.١٢	٠.٤٧	٢	٠.٢٩
١٥	٠.٢٣	٢.٧٢	٠.٣١	١.٨٤	٢.٤٢	٠.١٨	٠.٦٥	٢.٤٩	٠.٢٦
١٦	١.١٢	-٠.٦٩	٠.٢٨	٠.٩٦	٠.٤٠	٠.١١	٠.٦٦	٠.٥٨	٠.٢٥
١٧	٠.٩٧	٠.٧٥	٠.٢٩	٢.٤٣	١.٤٦	٠.٢٧	٠.٩٦	٢.٤٠	٠.٣٥
١٨	١.٣٦	-٠.٥٥	٠.٤١	٢.٦٧	١.٩٠	٠.٢٥	١.٥٨	٢.٩٥	٠.٢١
١٩	٠.٤٨	-٠.٢٢	٠.٢٦	٠.٧٥	٢.٢٩	٠.١٣	٠.٤٩	٢.٥٢	٠.٣٨
٢٠	٠.٤٦	-١.٣٤	٠.٣١	١.١١	-٠.٥٣	٠.٠٧	٠.٩٨	-٠.٥٨	٠.٢٧
٢١	٠.٦٥	٠.١٤	٠.٢٥	٢.٤٣	١.١١	٠.١١	١.٢٥	١.٣٢	٠.٣٨
٢٢	٠.٥٨	١.٥٥	٠.٢١	٢.٥٠	١.٧٢	٠.٠٩	٠.٤٦	٢.٠٧	٠.٣٩
٢٣	٠.٣٨	٢.٤٦	٠.٢٤	١.٠٧	-١.٤٤	٠.٢٨	٠.٩١	-٠.٥٥	٠.٣٢
٢٤	٢.٢٨	-٠.٠٧	٠.٤٨	٢.١٥	٠.١١	٠.١٣	٠.٨٤	-٠.٥١	٠.٣٧
٢٥	٠.٤٧	٢.٥٢	٠.٢٦	١.٩٤	-٠.٠٨	٠.٢٣	١.٢٣	٠.٠٧	٠.٢٥
٢٦	٠.٤٩	٢	٠.٢٧	١.٢٤	١.٤٠	٠.٣٣	٠.٧٤	١.٥٩	٠.٢٨
٢٧	١.١٤	-٠.٦٧	٠.٢٠	٢.٥٣	٠.٥٧	٠.٥	٢.١٢	١.٤٠	٠.٤١
٢٨	٢.٢٤	٠.٥٥	٠.١٠	٢.٣٣	-٠.٢٦	٠.٢٢	٢.٥٥	-٠.٠١	٠.٣٦
٢٩	٠.٤٨	١.٩٣	٠.٢٤	٢.٩٩	٠.٤٧	٠.٢٧	٢.٢٤	٠.٩٣	٠.٣
٣٠	٠.٥٤	٢.٢٨	٠.٠١	١.٢٣	١.٠٢	٠.١	٠.٨٨	١.٢٩	٠.٣٤
٣١	١.٠١	٠.٢٧	٠.٢٩	٢.٢٦	٠.٧٩	٠.١٢	٠.٧٨	١.٣٦	٠.٥
٣٢	٠.٤٥	١.٥٥	٠.٠١	١.٨٦	-٠.٢٥	٠.٣	١.٣٩	-٠.٤٧	٠.٢٣
٣٣	٠.٧٩	٠.٤٦	٠.٢٣	٢.٣٨	٠.٢١	٠.١١	٢.٧٨	٠.٥٥	٠.٢١
٣٤	٢.٠١	٠.١٢	٠.١٤	٢.٦٨	١.١٤	٠.١١	٢.٧٨	٠.٥٧	٠.٢٢
٣٥	١.٠٢	-٠.٤٦	٠.٠٩	٢.٤٩	٠.٦٩	٠.١٥	١.٣٤	١.١٠	٠.٢٤
المتوسط الحسابي	٠.٨٧	١.٠٩	٠.٢٤	١.٩٠	١.٠٧	٠.١٩	١.٣٢	١.٣٢	٠.٢٩
الانحراف المعياري	٠.٥٦	١.٢٦	٠.١٠	٠.٦٥	٠.٩٧	٠.١٠	٠.٧٥	١	٠.٠٩

ويتضح من النتائج تبايناً ظاهرياً في مدى صعوبة الفقرات، حيث كان المدى في صعوبة الفقرات هو الأعلى في النموذج الاول (توزيع البديل الصحيح على أ ب) الذي تراوحت قيمه من ـ ١.٣٤ الى ٢.٩٨ بمتوسط حسابي مقداره(١.٠٩) في حين كان الأقل في النموذج الثالث (توزيع البديل الصحيح عشوائيا) الذي تراوحت قيمه من ـ ٠.٥٨ الى(٣)بمتوسط حسابي مقداره(١.٣٢) .

اما مدى التخمين للفقرات فكان الأعلى في النموذج الأول (توزيع البديل الصحيح على أ ، ب) الذي تراوحت قيمه من ٠.٠١ الى ٠.٤٨ بمتوسط حسابي مقداره ٠.٢٤ في حين كان الأقل في النموذج الثاني (توزيع البديل الصحيح على ج ، د) الذي تراوحت قيمه من ٠.٠٧ الى ٠.٠٥ ، بمتوسط حسابي مقداره أيضا ٠.١٩

إن التباين في مدى قيم معالم الفقرات أظهر تبايناً ملحوظاً في الأوساط الحسابية لقيم هذه المعالم ، ولمعرفة الاختلافات (إحصائيا) بين معالم الفقرات حسب متغير موقع البديل الصحيح فقد تم استخدام تحليل التباين متعدد المتغيرات (MANOVA) وذلك لوجود ثلاثة متغيرات تابعة وهي تقديرات معلمة التمييز و تقديرات معلمة الصعوبة و تقديرات معلمة التخمين ، ومتغير مستقل واحد وهو موقع البديل الصحيح بثلاثة مستويات وذلك لمعرفة دلالة الفروق في معالم الفقرات لكل نموذج من نماذج الاختبار تبعا لمتغير موقع البديل الصحيح . ويوضح الجدول (٣) نتائج تحليلا لتباين متعدد المتغيرات (Multivariate Analysis of Variance) لمعرفة دلالة الفروق

جدول (٣) نتائج تحليل التباين متعدد المتغيرات (Multivariate Analysis of Variance) لمعرفة دلالة الفروق في المتوسطات الحسابية لمعالم الفقرات باختلاف موقع البديل الصحيح

المتغير المستقل	الاختبار الاحصائي	القيمة الاختبار الاحصائي	F	درجات الحرية	خطأ درجات الحرية	الدلالة الاحصائية
موقع البديل الصحيح	Pillai's Trace	٠.٣٨٩	٨.١٢٢	٦	٢٠٢	٠٠٠
	Wilks' Lambda	٠.٦٢٠	٨.٩٩٧	٦	٢٠٠	٠٠٠
	Hotelling's Trace	٠.٥٩٨	٩.٨٧٥	٦	١٩٨	٠٠٠

يتضح من الجدول (٣) أن قيمة p-value (الدلالة الاحصائية) أقل من ٠.٠٥ للمتغير المستقل موقع البديل الصحيح وفقا لنتائج الاختبارات الاحصائية: Pillai's Trace ، Wilks Lambda ، Hotelling's Trace ، وبالتالي هذا يعني أنه يوجد تأثيرا دالا احصائيا لموقع البديل الصحيح على معالم الفقرات ولتحديد أي من هذه المعالم(الصعوبة ، التمييز ، التخمين) تتأثر بموقع البديل الصحيح نعرض الجدول (٤) والذي هو من نتائج تحليل متعدد المتغيرات (Manova)

جدول (٤) نتائج تحليل التباين متعدد المتغيرات عند دراسة تأثير موقع البديل الصحيح على تقديرات معالم الصعوبة ، التمييز ، التخمين

مصدر التباين	المتغير التابع	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمةF	الدلالة .الاحصائية
المجموعة	التمييز	١٨.٨٤٦	٢	٩.٤٢٣	٢١.٨٤٢	٠٠٠
	الصعوبة	١.٣٤٧	٢	٠.٦٧٣	٠.٥٧٤	٠.٥٦٥
	التخمين	٠.٠٩٧	٢	٠.٠٤٨	٥.١٠٣	٠.٠٠٨
الخطأ	التمييز	٤٤.٠٠٥	١٠٢	٠.٤٣١		
	الصعوبة	١١٩.٦٠١	١٠٢	١.١٧٣		
	التخمين	٠.٩٦٦	١٠٢	٠.٠٠٩		
الكلي	التمييز	٦٢.٨٥٠	١٠٤			
	الصعوبة	١٢٠.٩٤٧	١٠٤			
	التخمين	١.٠٦٢	١٠٤			

يتضح من الجدول (٤) انه يوجد فروق دالة إحصائياً في متوسطات معاملات التمييز يعزى لموقع البديل الصحيح حيث بلغت قيمة p-value = صفرا وبذلك تكون أقل من (٠.٠٥) وجاء هذا الفرق لصالح النموذج الثاني (يتوزع موقع البديل الصحيح على ج ، د). ويتضح كذلك من هذا الجدول أنه يوجد فرق دال إحصائيا في متوسطات معاملات التخمين يعزى لموقع البديل الصحيح حيث بلغت قيمة p-value = ٠.٠٠٨ وبذلك تكون أقل من ٠.٠٥ . بينما فيما يخص متوسطات معاملات الصعوبة يبين الجدول أنه لا توجد فروق دالة احصائيا في متوسطات معاملات الصعوبة يعزى لموقع البديل الصحيح وذلك لأن قيمة p-value = ٠.٥٦٥ أكبر من مستوى الدلالة الاحصائية α = ٠.٠٥ .

للإجابة عن السؤال الثاني "هل تختلف تقديرات معلمة القدرة للأفراد (θ) باختلاف موقع البديل الصحيح ؟ "

فقد تم إيجاد الأوساط الحسابية والانحرافات المعيارية لتقديرات معلمة القدرة للأفراد حسب متغير موقع البديل الصحيح ، ويبين الجدول(٥) هذه القيم.

جدول (٥) الأوساط الحسابية والانحرافات المعيارية لتقديرات معلمة القدرة للأفراد (θ)حسب متغير موقع البديل الصحيح لكل نموذج من نماذج الاختبار بعد حذف الافراد غير المطابقين للنموذج

نماذج الاختبار	العدد	الوسط الحسابى	الانحراف المعيارى
الاول: توزيع البديل الصحيح على أ ، ب	٥١٩	-٠.٧٨	١.٢٨
الثاني: مواقع البديل الصحيح توزعت في ج ، د	٥٠٩	- ٠.٥٨	٠.٩٢
الثالث: مواقع البديل الصحيح توزعت عشوائيا	٥١٣	-٠.٠٦	١.٢٠

يتضح من الجدول (٦) وجود فروق ظاهرية بين المتوسطات الحسابية لتقديرات معلمة القدرة للأفراد حسب متغير موقع البديل الصحيح ولمعرفة الدلالة الإحصائية لتلك الفروق الظاهرية تبعا لمتغير موقع البديل الصحيح تم استخدام تحليل التباين الاحادي ، (One-Way ANOVA) ويبين الجدول (٦) نتائج تحليل التباين الاحادي .

جدول (٦) نتائج تحليل التباين الاحادي (One-Way ANOVA) للمتوسطات الحسابية لتقديرات معلمة القدرة للأفراد حسب متغير موقع البديل الصحيح

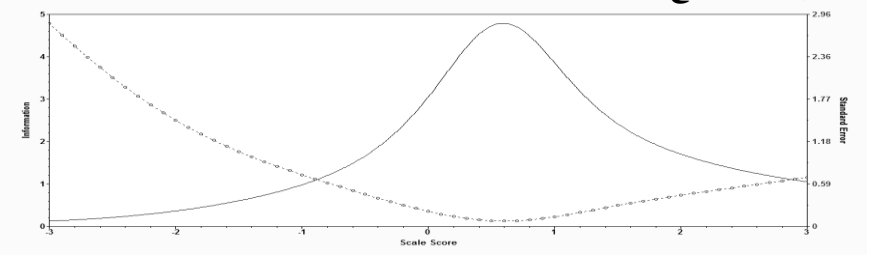
مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	الدلالة الإحصائية
موقع البديل الصحيح	٦.٩١٢	٣	٢.٣٠٤	١.٧٧١	٠.١٥١
الخطأ	٢٠٠٦.١٥٤	١٥٤٢	١.٣٠١		
الكلى	٢٠١٣.١٦٦	١٥٤٢			

يتضح من الجدول (٦) بأنه لا يوجد فروق ذات دلالة إحصائية عند مستوى الدلالة الإحصائية (α = ٠.٠٥) بين المتوسطات الحسابية لتقديرات معلمة القدرة للأفراد يعزى لمتغير البديل الصحيح .

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 البديل الصحيح ، وهذه التغيرات لم تؤثر على تقديرات معلمة القدرة للمفحوصين، فجميع الطلبة درسوا نفس محتوى مادة الاختبار، وبأساليب تدريس متنوعة، وأنشطة إثرائية كذلك، مما جعل تقديرات قدراتهم لا تختلف باختلاف موقع البديل الصحيح .

للإجابة عن السؤال الثالث " **هل تختلف دالة المعلومات للاختبار باختلاف موقع البديل الصحيح ؟** " 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 موقع البديل الصحيح من برنامج (BILOG- MG 3) فكانت على النحو التالي:

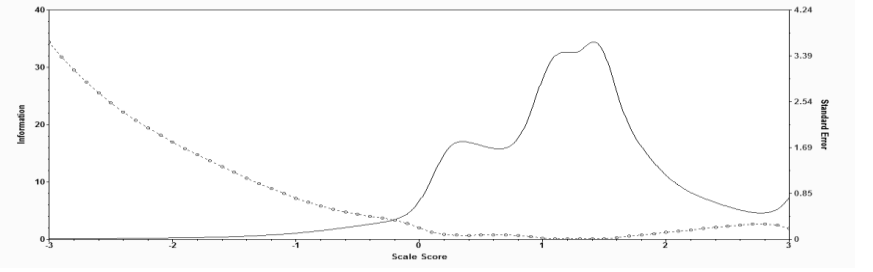
- أولا: دالة معلومات الاختبار للنموذج الاول والخطأ المعياري للقياس (توزيع البديل الصحيح على أ ، ب) كانت كما هو مبين في الشكل (١) :



شكل (١): دالة معلومات الاختبار للنموذج الاول للاختبار والخطأ المعياري للقياس

يتضح من الشكل (١) بأن أكبر كمية لدالة المعلومات للنموذج الاول كانت عند مستوى القدرة (θ) = ٠.٦ حيث بلغت كمية المعلومات ٤.٨ .

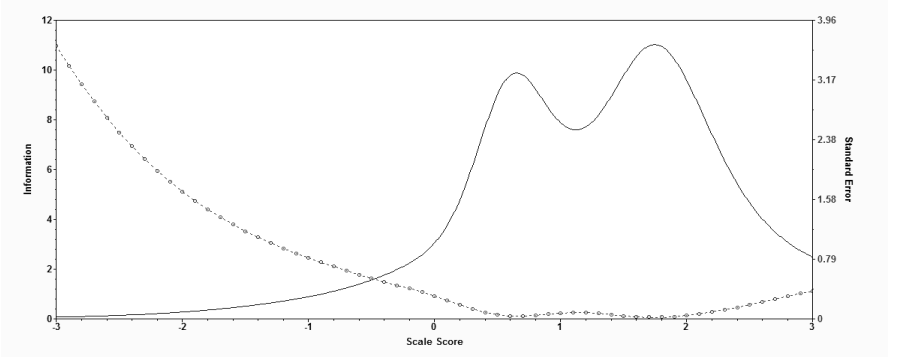
- ثانيا: دالة معلومات الاختبار للنموذج الثاني والخطأ المعياري للقياس (توزيع البديل الصحيح على ج ، د) كانت كما هو مبين في الشكل (٢) :



شكل (٢) : دالة معلومات الاختبار للنموذج الثاني للاختبار والخطأ المعياري للقياس

يتضح من الشكل (٢) بأن أكبر كمية لدالة المعلومات للنموذج الثاني كانت عندمستوى القدرة (θ) = ١.٦ حيث بلغت كمية المعلومات ٣٤ .

- ثالثا: دالة معلومات الاختبار للنموذج الثالث والخطأ المعياري للقياس (توزيع البديل الصحيح عشوائيا) كانت كما هو مبين في الشكل (٣) :



شكل (٣) : دالة معلومات الاختبار للنموذج الثالث للاختبار والخطأ المعياري للقياس

يتضح من الشكل (٣) بأن أكبر كمية لدالة المعلومات للنموذج الثالث كانت عند مستوى القدرة (θ) = ١.٨ حيث بلغت كمية المعلومات ١٠.٦ .

وعند مقارنة كميات المعلومات للنماذج الاختبارية الثلاثة وجد أنه أقل كمية لدالة المعلومات كانت للنموذج الأول (توزيع البديل الصحيح على أ ، ب) وأكبر كمية لدالة المعلومات كانت للنموذج الثاني (توزيع البديل الصحيح على ج ، د) ،وربما يعود ذلك إلى معلمة التمييز للفقرات، حيث كانت قيم التمييز للفقرات في النموذج الثاني هي الأعلى بينما كانت الأقل لفقرات النموذج الاول فكلما زاد معلم التمييز للفقرة زادت كمية المعلومات التي تسهم بها الفقرة حول معلمة الصعوبة .

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 (٧) هذه القيم.

جدول (٧) قيم معاملات الثبات الامبريقي لكل نموذج من نماذج الاختبار

نموذج الاختبار	الثبات الامبريقي
الأول : توزيع البديل الصحيح على أ ، ب	٠.٨٠
الثاني : توزيع البديل الصحيح على ج ، د	٠.٩٣
الثالث : توزيع البديل الصحيح عشوائيا	٠.٨٧

يتضح من الجدول (٧) تباين قيم معاملات الثبات الامبريقي للنماذج الاختبارية الثلاثة 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 ـ ١) ، والمعادلة الرياضية لهذا الاختبار هي :

$$M = \sum_{j=1}^{k} A_j - \frac{\sum_{j=1}^{k} A_j (1-\hat{\alpha}_j)^{-1/3}}{\sum_{j=1}^{k} A_j (1-\hat{\alpha}_j)^{-2/3}}$$

حيث إن :

$$A_j = \frac{(n_j - 1)(9N - 11)^2}{18 n_j (N_j - 1)}$$

NJ: عدد فقرات الاختبار j ، Nj: عدد المتقدمين للاختبار j ، $\hat{\alpha}_j$: معامل الثبات الامبريقي للاختبار j ، K (رقم نموذج الاختبار حسب موقع البديل الصحيح) = 1 ، 2 ، 3

وقد كانت قيمة الإحصائي M المحسوبة تساوي(١٧.٤٥٧) وهي أعلى من القيمة الحرجة لمربع كاي تربيع بدرجات حرية (٢) والذي يشير إلى وجود فروق ذات دلالة إحصائية بين قيم معاملات الثبات الامبريقي ، وقد جاءت تلك الفروق دالة لصالح النموذج الثاني (توزيع البديل الصحيح على ج ، د) . ويعزو الباحثان السبب في ذلك ربما إلى تباين قدرة المفحوصين واختلاف تمييز الفقرات التي جاءت لصالح النموذج الذي تتوزع البدائل على ج ، د وهي بذلك انسجمت مع نتائج الشكل (٢) والتحليل متعدد المتغيرات الوارد في الجدول (٨).

لقد تبين من النتائج التي توصلت إليها الدراسة الحالية بأنه لا يوجد أثر لموقع البديل الصحيح في اختبار اختيارٍ من متعدد على تقديرات معلمة الصعوبة للفقرة ، في حين كان هناك أثرا لموقع البديل الصحيح على معلمتي التمييز والتخمين للفقرة وعلى دالة المعلومات وذلك لصالح النموذج الذي يتوزع فيه البديل الصحيح على الخيارات ج ، د . وفي ضوء تلك النتائج فإن الباحثين يوصيان بإجراء دراسة مماثلة بحيث يتم فيها اقتراح توزيعات أخرى للبديل الصحيح وباستخدام نماذج أخرى من نماذج النظرية الحديثة في القياس في ضوء مقاييس أخرى .

- **المراجع** :

- أمبو سعيدي ، عبد الله. (٢٠٠٥) أثر تغير موقع المشتت القوي عند استخدام تقنية إبعاد البدائل الخطأ في صعوبة الفقرات لأسئلة الاختيار من متعدد في الفيزياء وتمييزها مجلة العلوم التربوية والنفسية ، جامعة البحرين، ٦ (٣) : ٨٥ ـ ٢٠٦ .
- بني عطا ، زايد والرباعي ، إبراهيم .(٢٠١٣) أثر عدد البدائل وتغير موقع المموه القوي في فقرات اختبار اختيار من متعدد على معالم الفقرة وقدرة الفرد ودالة المعلومات ، المجلة الأردنية في العلوم التربوية ، ٩ (٣) : ٣١٩ ـ ٣٣٣.

- الحارثي ، ابتسام. (٢٠٠٧) تقويم كفايت بناء الاختبارات التحصيلية لدى معلمات العلوم للمرحلة الثانوية بمحافظة جدة ، رسالة ماجستير غير منشورة ، جامعة الملك عبد العزيز جدة .
- سليمان ، شاهر. (٢٠٠٨) أثر مخالفة قواعد صياغة فقرات اختبار اختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته ، المؤتمر السنوي الخامس لقسم علم النفس تطبيقات العلوم النفسية ومشكلات المجتمع في الفترة من ٤ - ٦ مايو ، جامعة طنطا كلية الآداب ، جمهورية مصر العربية .
- سليمان ، شاهر. (٢٠١٤) حكمة الاختبار لدى طلاب جامعة تبوك في ضوء متغيرات التخصص والمعدل التراكمي والمستوى الدراسي ، مجلة العلوم التربوية والنفسية ، ١٥ (٢) : 245 - 273 .
- السويلم ، بندر .(٢٠٠٣) القواعد العامة في صياغة أسئلة الاختيار من متعدد ،ومدى التزام أساتذة كلية التربية بجامعة الملك سعود بهذه القواعد ، مركز بحوث كلية التربية ، جامعة الملك سعود ٣ (٢٠) : ٥٤ -٨٧
- علام ، صلاح الدين .(٢٠٠٨) نماذج لاستجابات المفردات الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي ، ط١ ، القاهرة : دار الفكر العربي
- الكناني ، عايد وفيضي ، فاضل (٢٠١٢) فاعلية استخدام موقع البديل الفعال عند ابعاد البديل الخاطئ في صعوبة فقرات الأسئلة اختيار من متعدد وتمييزها ، مجلة القادسية لعلوم التربية الرياضية ، ١٢ (١) : ٢٠٧ -٢٢٥ .
- المالكي ، يحي .(٢٠١٥) درجة التزام معلمي المرحلة المتوسطة بقواعد بناء الأسئلة الموضوعية ، رسالة ماجستير غير منشورة ، جامعة أم القرى ، المملكة العربية السعودية .
- يعقوب ، ابراهيم وأبو فودة ، باسل (٢٠١٢) أثر مخالفة قواعد صياغة فقرات الاختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته ، مجلة جامعة دمشق ، ٢٨ (١) ٤١٩ – ٤٤٢
- ظاظا ، حيدر .(٢٠١٢) الكشف عن مدى انتهاك قواعد صياغة فقرات الاختيار من متعدد في أسئلة امتحان شهادة الدراسة الثانوية العامة في الأردن ،المجلة الأردنية في العلوم التربوية ، ٨ (١) ، ٨١ -٩٤
- Adedoyin, O., Nenty, H. & Chilisa, B. (2008). Investigating the invariance of item difficulty parameter estimates based on CTT and IRT. Educational Research and Review. 3(2), 083-093
- Allen, M. & Yen, W. M., (1989). Introduction to Measurement theory California, Broks Cole publishing company.
- Brannick, M. (2003). Basic of IRT. [On-line] file: A:/Item Response theory. Htm
- Baker, F.B.(2001). The basics of item response theory (2nd Ed). College Park, MD: ERIC Clearing House on Assessment and Evaluation.
- Cantor, J. (1987). Developing Multiple-Choice Test Items. Training and Development Journal, 41(5), 85-88.

- Carter, K.(1986).Test wiseness for teachers and students. Educational Measurement: Issues and Practice, 5(4), 20 – 23.
- Crocker, L., & Algina, J.(1986). Introduction to classical and modern test theory. New York: Holt, Rinehart and Winston.
- Gleason, J., Alley, A., & Baker, S. (2010). The effect of item-writing rules on the reliability of an instrument to measure mathematical knowledge for teaching. Journal of Mathematical Sciences and Mathematics Education, 5(2), 21-27.
- Gronlund, N. E. (1993). How to make achievement tests and assessments. (5th Ed.). Boston: Allyn & Bacon.
- 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. (1995). Item response theory: Principles and applications. Boston: Kluwer-Nijhoff.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). Fundamentals of item response theory. London: Sage Publications, Inc.
- Hatti, J. (1984). An empirical study of various indices for determining unidimensionality. Multivariate Behavioral Research, 19, 49-78.
- Hoepfl, M. C. (1994). Developing and evaluation multiple choice tests. The Technology Teacher, 53(7), 25-26.
- Klimko, I. P. (1982). Item arrangement, cognitive entry characteristics, and sex as predictors of achievement test performance in college students. (Ph.D. dissertation), The University of Connecticut, United States -- Connecticut. Retrieved May 11, 2011, from Dissertations & Theses: Full Text. (Publication No. AAT 8216409).
- Lord, F. M.(1980). Application of item response theory to practical testing problems. New Jersey: Lawrence Erlbaum Associates, Publishers.
- -Nunnally, J.C, (1981); Psychometric Theory. 2nd ed. McGraq-Hill, New York.
- Rodriguez, M. C. (1997, April). The art & science of item –writing: a meta-analysis of multiple-choice item format effects. Paper

presented at the Annual Meeting of The American Educational Research Association, Chicago, IL. Retrieved from

- Rup, A.A. & Zumbo, B.D. (2006). Understanding parameter invariance in unidimensional IRT models. Educational and Psychological Measurement, 66, 63-84.
- Sarnacki, R (1979) an examination of test domain, Review of Educational Research, 49(2)252-279.
- -Tarrant, M., Knierim, A., Hayes, S. K., & Ware, J. (2006). The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments. Nurse Education in Practice, 6(6), 662-671. Retrieved from
- Warm, T.A.(1978). A Primer of Item Response Theory. Technical Report No, 940279, Oklahoma City: U.S, Coast Guard Institute.
- Weiten, W. (1984). Violation of selected item construction principles in educational measurement. Journal of Experimental Education, 52 (3), PP. 78 – 174.
- Wevric,L(2001) .Response set in a multiple-choice test , Educational and psychological Measurement, vol(33) pp58-65
- Withers, G., (2005). International Institute for Educational Planning Item writing for tests and examinations (Module 5).

❋❋❋❋❋❋❋