

# مقارنة تأثيرات تغيير موقع البديل الصحيح في اختبار اختيار من متعدد على تقديرات الصعوبة للفقرات وتقديرات القدرة لدى الأفراد باستخدام نموذج راش

علي محمد خالد الصالح\*

الملخص\_ هدفت الدراسة الحالية إلى الكشف عن 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 الاختيار من متعدد بأربعة بدائل في القياس والتقويم التربوي المقرر لطلاب دبلوم التربية العام مكون من (34) فقرة بصورته النهائية، وقد أشتمل على ثلاثة نماذج لها نفس المحتوى والمتغير هو موقع البديل الصحيح: النموذج الأول 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 عشوائياً على مختلف البدائل الأربعة، وطبقت النماذج الثلاثة على عينة عشوائية مكونة (1541) من طلاب وطالبات جامعة تبوك. استخدم الباحث برنامجي SPSS15، Bilog-Mg3 لتحليل البيانات لكل نموذج من نماذج الاختبارات الثلاثة وفقاً للنموذج أحادي المعلمة (نموذج راش)، ودلت نتائج الدراسة على عدم وجود فروق دالة إحصائية بين المتوسطات الحسابية لمعاملات صعوبة الفقرات، تعزى لموقع البديل الصحيح، وعدم وجود فروق دالة إحصائية بين المتوسطات الحسابية لتقديرات معلمة القدرة للأفراد تعزى لموقع البديل الصحيح.

الكلمات المفتاحية: معلمة صعوبة الفقرة، معلمة القدرة، موقع البديل الصحيح، نموذج راش، نظرية الاستجابة للفقرة.

## مقارنة تأثيرات تغيير موقع البديل الصحيح في اختبار اختيار من متعدد على

### تقديرات الصعوبة للفقرات وتقديرات القدرة لدى الأفراد باستخدام نموذج راش

#### 1. المقدمة

لقد أسفرت جهود العلماء عن ظهور بعض الاتجاهات الحديثة في مجال القياس والتقويم، ومن بين هذه الاتجاهات، نظرية الاستجابة للفقرة (IRT) أو نظرية السمات الكامنة Latent Traits Theory، وحظي هذا المدخل الجديد باهتمام الباحثين حيث يتغلب على كثير من مشكلات القياس التقليدي.

ويخلص عالم [8] الفكرة الأساسية لنماذج الاستجابة للفقرة في أنها تحاول اشتقاق قيم تقديرية للسمات التي تنطوي عليها مجموعة من الاستجابات لمجموعة من الفقرات، وعادة يفترض أن السمة المقاسة هي قدرة معينة أو خاصية من خصائص الفرد الذي يختبرها، بحيث لا توجد علاقة منتظمة بين مستويات السمة المقاسة لدى أفراد مختلفين، واحتمالات الاستجابة الصحيحة لفقرات مختلفة بمعنى آخر.

تقوم نظرية الاستجابة للفقرة على افتراضات محددة لا بد من التحقق من وجودها؛ إذ لا تؤدي هذه النظرية بدونها إلى نتائج صحيحة. ويندرج تحت إطار نظرية الاستجابة للفقرة نماذج متعددة، فقد تكون هذه النماذج: أحادية البعد أو متعددة الأبعاد، وتستخدم جميعها في التنبؤ بأداء المفحوصين على فقرة باستخدام خواص الفقرة والسمات التي يقيسها الاختبار، ومستوى الإجابة. ومن بين النماذج نماذج أحادية البعد.

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 [9]، وتعد استقلالية القياس بمثابة الفرق الجوهرية بين النظرية التقليدية ونظرية استجابة الفقرة، وهو ما عبر عنه لورد [10] بخاصية اللاتغير [11] Invariance، هذا وقد بنيت هذه النظرية على افتراضات ينبغي تحقيقها في البيانات لكي تؤدي إلى نتائج يمكن الوثوق بها، وهذه الافتراضات هي:

- أحادية البعد (Unidimensionality) يعني أن هناك قدرة واحدة تفسر أداء الفرد في الاختبار لذا تسمى بالنماذج أحادية البعد، أما النماذج التي تفترض وجود أكثر من قدرة واحدة تكمن وراء هذا الأداء تسمى متعددة الأبعاد.

- الاستقلال الموضعي (Local Independence) أي أن تكون استجابات الفرد على الفقرات المختلفة في الاختبار مستقلة استقلالاً إحصائياً عند مستوى معين من القدرة، أي أن استجابة الفرد عن فقرة ما يجب ألا تؤثر سلباً أو إيجاباً على استجابته لفقرة أخرى [12].

- منحى خصائص الفقرة (ICC) Item Characteristic Curve وهو اقتران رياضي يربط بين احتمال استجابة الفرد استجابة صحيحة على فقرة الاختبار، وبين القدرة التي تقيسها مجموعة الفقرات التي تحتوي

تفترض الاختبارات النفسية والتربوية أن هناك سمات أو خصائص معينة يشترك فيها جميع الأفراد، ولكنهم يختلفون في مقدارها وبالرغم من أن هذه السمات غير منظورة، إلا أنه يمكن الاستدلال على مقدارها من السلوك الملاحظ للفرد المتمثل في استجاباته على مفردات الاختبار، وهذا ما يبرر تسميتها بالسمات الكامنة. فالسمة التي تكمن وراء استجابة الفرد على مفردات الاختبار اللفظي تختلف عن السمة التي تكمن وراء استجاباته على مفردات اختبار عددي أو مكاني، ولكن يمكن أن تكمن سمة واحدة وراء استجاباته على مفردات اختبارين مختلفين متعلقين بنفس المحتوى [1].

ويرى هاملتون وسوامنيانان [2] 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 المموهات. أما بالنسبة لصعوبة الفقرة، ويرى عالم [3] أن من العوامل المهمة المؤثرة على صعوبة فقرات الاختبار من متعدد، إضافة إلى قدرة الطالب على الإجابة، عوامل فنية تتعلق بقواعد صياغة فقرات الاختبار من متعدد مثل: درجة تجانس البدائل، وسلامة الصياغة اللغوية، والتوزيع العشوائي للبدائل الصحيحة.

إن 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 [4].

أشار المالكي [5] إلى أن معلمي العلوم في المرحلة المتوسطة يلتزمون ب (60%) من قواعد بناء أسئلة الاختبار 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 [6].

إن الشخص الذي يعرف الإجابة فإن بديلاً واحداً من بدائل الفقرة يشده، وهو المقبول لديه بأنه 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 يجولون الإجابة عن الفقرة، مما يغير نسبة احتمال الإجابة عن ذلك البديل [7].

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 (97) عضواً من أعضاء هيئة التدريس بكلية التربية، كما اشتملت عينة تحليل محتوى الأسئلة على 489 سؤالاً، وقد دلت نتائج الدراسة أن أكثر أنواع الأسئلة استخداماً في الاختبارات الأكاديمية هو أسئلة الاختيار من متعدد، إلا أنها كشفت عن قصور وضعف صدقها المحكي، حيث أن (3.65%) 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 نظراً لضخامة القرارات المرتبطة بنتائج تلك الاختبارات.

أجرى ويفرك [19] دراسة بعنوان الاستعداد الخاص للاستجابة في اختبار اختيار من متعدد إذا كان البديل الصحيح يحتل موقعاً متحيزاً (البديل الأوسط)، لتحقيق هدف الدراسة صمم الباحث اختبار اختيار من متعدد ذو ثلاثة بدائل يتكون من (100) فقرة، يشمل ثلاثة نماذج: الأول يكون موقع البديل الصحيح في الوسط، والثاني البديل الوسط يتضمن مموه للإجابة، والنموذج الثالث موقع البديل الصحيح موزع توزيعاً عشوائياً، طبقت نماذج الاختبار على عينة عشوائية تكونت من (467) طالباً جامعياً، وقد 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 موزعاً عشوائياً فإنه لا تأثير يذكر لما يسمى باستعداد الطالب الخاص للإجابة عن الفقرات بطريقة معينة.

كما أجرى سليمان [20] دراسة هدفت إلى استقصاء أثر مخالفة قواعد صياغة فقرات اختبار الاختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته، ولتحقيق هذا الهدف تم إعداد اختباراً تحصيلياً في مادة علم النفس التربوي، وهو أحد المواد التي تطرح في الجامعة، وتكون الاختبار من (40) فقرة مخالفة من نوع اختيار من متعدد أربعة بدائل بواقع (10) فقرات لكل نوع من المخالفات الأربع الآتية: طول البديل الصحيح، موقع البديل الصحيح، البديل جميع ما ذكر صحيح، البديل كلاهما صحيحان. وبنظره اختباراً مكوناً من (40) فقرة متحررة من المخالفات ولها نفس المحتوى، طبق الاختبار على عينة عشوائية مكونة من (482) طالباً جامعياً، دلت النتائج أن الفقرات المخالفة لقواعد الصياغة، أسهل من نظيرتها الفقرات المتحررة من المخالفات ويفارق دال إحصائياً، ولا توجد فروق دالة إحصائية بين معاملات تمييز الفقرات المخالفة للصياغة ونظيرتها المتحررة من المخالفات، كما يوجد فروق دالة إحصائية 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.

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 [13].

ويوجد مجموعة من الافتراضات خاصة بكل نموذج من نماذج استجابة الفقرة البارامترية، حيث أن النموذج الأحادي المعلمة (نموذج راش) النموذج المستخدم في هذه الدراسة، يعتبر أحد نماذج الاستجابة للفقرة. ويعتبر نموذج راش أبسط نماذج السمات الكامنة وأكثرها استخداماً في بناء الاختبارات والمقاييس النفسية والتربوية وتطويرها، والفكرة الأساسية التي يقوم عليها هذا النموذج هي أن كل فقرة تحمل شحنة انفعالية تساهم مع الفقرات الأخرى في القائمة في تكوين شحنة انفعالية إجمالية تعبر عن اتجاه الفرد بما يتفق مع تقديره لتلك الفقرة حسب عدد فئات التدرج المستخدم في القائمة، ويقوم النموذج بتقدير هذه الشحنة لكل فقرة حسب الدالة الرياضية الاحتمالية التي يعتمدها النموذج، ثم التحقق من شروط ملائمة الفقرات [14]. ويستخدم نموذج راش أسلوب يسمى طريقة الأرجحية العظمى (ML) والتي بدورها تعتمد على ثلاث طرق رئيسية مختلفة في تقدير معلمة صعوبة الفقرات في النماذج أحادية البعد [15,16] المشار إليهم في الشريطين، [17] ، وهذه الطرق هي تقديرات طريقة الأرجحية العظمى المشتركة (Joint Maximum Likelihood Estimation)، وطريقة تقديرات الأرجحية العظمى الشرطية (Conditional Maximum Likelihood Estimation)، وطريقة (Marginal Maximum Likelihood Estimation)، أما طرق تقدير معالم القدرة للأفراد، فيتم تقديرها بطريقتين: الطريقة الأولى تعتمد على تقديرات الأرجحية العظمى (ML)، بينما تعتمد الطريقة الثانية على تقديرات بايز (Bayesian Estimation)

يعد نموذج راش (Rash Model) أكثر هذه النماذج شهرة، ويسمح هذا النموذج باختلاف الفقرات في معاملات صعوبتها. ويُعبر عن نموذج راش رياضياً بالصورة:

حيث أن: معلمة قدرة المفحوص.

احتمال إجابة المفحوص الذي قدرته (i) على الفقرة (i) إجابة صحيحة.

D: معلم التدرج وقيمتها (1,702)، bi معلمة صعوبة الفقرة (i)

e: اللوغاريتم النيبييري (Naperian log) والذي قيمته (2,71828) تقريباً.

### 3. الدراسات السابقة

في دراسة قام بها السويلم [18] 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

## مقارنة تأثيرات تغيير موقع البديل الصحيح في اختبار اختيار من متعدد على تقديرات الصعوبة للفقرات علي الصالح

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، تكوّن من (40) فقرة ذات أربعة بدائل، ثم جرى صياغة فقرات الاختبار في نموذج آخر مخالف لقواعد الصياغة، وقد جرى تطبيق نموذجي الاختبار على عيّنة مكوّنة من 300 طالب، بواقع 150 طالباً لكل نموذج، وفي ظروف صافية 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.

وتأتي هذه الدراسة للكشف عن تأثير تغيير موقع البديل الصحيح في اختبار اختيار من متعدد في مطابقة الفقرات لنموذج راش، وتأثيرات تغيير موقع البديل الصحيح في اختبار اختيار من متعدد أربعة بدائل على معلمة صعوبة الفقرات وتقديرات القدرة للأفراد باستخدام نموذج أحادي المعلمة (نموذج راش).

### أ. أهداف الدراسة

تهدف الدراسة الحالية إلى الآتي:

- مقارنة تأثير تغيير موقع البديل الصحيح في اختبار اختيار من متعدد على معلمة صعوبة الفقرة وفق نموذج راش.
- مقارنة تأثير تغيير موقع البديل الصحيح في اختبار اختيار من متعدد على تقديرات القدرة لدى الأفراد وفق نموذج راش.

### ب. أسئلة الدراسة

السؤال الأول: هل تختلف تقديرات معلمة صعوبة الفقرة المقدرّة باستخدام نموذج راش باختلاف موقع البديل الصحيح في اختبار اختيار من متعدد؟

السؤال الثاني: هل تختلف تقديرات معلمة قدرة الفرد المقدرّة (θ) باستخدام نموذج راش باختلاف موقع البديل الصحيح في اختبار اختيار من متعدد؟

### ج. حدود الدراسة

- الحدود الموضوعية:

اقتصرت الدراسة على دراسة تأثيرات تغيير موقع البديل الصحيح في اختبار اختيار من متعدد على صعوبة الفقرات وتقديرات القدرة للأفراد في ثلاثة نماذج للاختبار هي كالتالي:

النموذج الأول يكون فيه موقع البديل الصحيح في النصف الأول (البديلين أ، ب)، والنموذج الثاني يكون فيه موقع البديل الصحيح في النصف الثاني (البديلين ج، د)، والنموذج الثالث يكون فيه توزيع موقع البديل الصحيح عشوائياً على البدائل، باستخدام نموذج راش كأحد نماذج نظرية استجابة الفقرة.

- الحدود البشرية: طلاب وطالبات جامعة تبوك تخصص دبلوم عام في التربية للعام الدراسي 1434 هـ / 1435 هـ.

- الحدود المكانية: جامعة تبوك - كلية التربية والآداب.

وأجرت الحارثي [21] دراسة هدفت إلى تقويم كفايات بناء الاختبارات التحصيلية لدى معلمات المرحلة الثانوية تبعاً للمتغيرات (التخصص، والخبرة، التدريب)، وقد استخدمت الباحثة المنهج الوصفي في دراستها من خلال بطاقة تحليل الأسئلة المعدّة من قبل عينة من معلمات المرحلة الثانوية التي بلغت (251) معلمة بمحافظة جدة، دلّت النتائج على تدني درجة معرفة وممارسة معلمات العلوم لكفايات بناء الاختبار التحصيلية، وعلى وجود فروق ذات دلالة إحصائية في متغير التدريب لصالح المتدربات.

وأجرى الكنانة وفيضي [22] دراسة بعنوان فاعلية استخدام موقع البديل الصحيح عند إبعاد 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 تغيير موقع البديل الصحيح عند استخدام تقنية إبعاد البدائل الخاطئة على تحصيل الطلبة في أسئلة الاختيار من متعدد، اشتملت عينة الدراسة على (60) طالباً من طلبة المرحلة الثالثة لكلية التربية الرياضية في جامعة المثنى في العراق، حيث تم تقسيم عينة الدراسة إلى مجموعتين متساويتين، حيث أجاب (130) طالباً وطالبة عن أسئلة النموذج الأول، كما أجاب (130) طالباً عن أسئلة النموذج الثاني، دلّت النتائج على عدم وجود فروق معنوية بين نموذجي الاختبار من حيث الصعوبة والتمييز.

وأجرى ظاظا [23] 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 (2910) فقرات، للأعوام 2007 - 2010 باستخدام (18) قاعدة للكشف عن مخافة الفقرة لقواعد صياغتها، أظهرت النتائج أن (416) فقرة أي ما نسبته 12.6% تضمنت أكثر من مخالفة، وبشكل إجمالي وجدت 782 فقرة أي ما نسبته 26.9% فقرة مخالفة، وتبين أن أكثر قاعدتين انتهاكاً كانتا: ترتيب البدائل عمودياً بنسبة 67.7%، وترتيب البدائل منطقياً أو عددياً أو أبجدياً بنسبة 9.8%، في حين حافظت الفقرات على عدم انتهاك ست قواعد، كما توصلت الدراسة إلى أنّ امتحان الثانوية العامة في الأردن بشكل عام يراعي قواعد صياغة فقرات الاختيار من متعدد.

وأجرى بني عطا، والرباعي [24] 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 مكوّن من (41) فقرة بصورته النهائية، وقد اشتمل على أربعة نماذج حسب عدد البدائل، وكان عدد أفراد عينة الدراسة (2111) طالباً وطالبة لجميع نماذج الاختبارات الأربعة، استخدم الباحثان النموذج اللوجستي ثلاثي المعلمة، وأشارت النتائج أنه لا توجد فروق ذات دلالة إحصائية بين متوسطات معالم الصعوبة والتمييز والتخمين تعزى لعدد البدائل أو موقع المموه.

كما أجرى يعقوب، وأبو فودة [25] دراسة هدفت إلى معرفة أثر

السلوكي الذي تقيسه من حيث المحتوى والمستوى المعرفي، كما تم إعداد 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 ج، أو البديل د)، والنموذج الثالث يتكون من نفس الفقرات باستثناء موقع البديل الصحيح موزع بشكل عشوائي وتم تطبيق النماذج على عينة مكونة من (1545) طالبا وطالبة تخصص الدبلوم العام في التربية. صدق الأداة

للتأكد من صدق المحتوى، عرضت فقرات الاختبار بصورتها الأولية وعددها (40) 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 لها، ومدى قياس الفقرات للهدف السلوكي المتعلق بها، ومدى إتساق الاختبار ككل مع الهدف الأساسي منه، وبناء على ملاحظاتهم لم يتم حذف فقرات، وإنما جرى تعديل على بعض الفقرات، وأجمع المحكمون أن هناك فقرات تؤدي غرضها، وأنها عينة ممثلة للمجال السلوكي الذي يقيسه.

جمع وتحليل البيانات

لتحقيق الهدف من الدراسة، تم تطبيق نماذج الاختبار عشوائيا على أفراد عينة الدراسة والبالغ عددهم 1545 طالبا وطالبة، 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 برنامج Bilog-Mg3، وبرنامج SPSS15 للإجابة عن أسئلة الدراسة، وقد تمت عملية التحليل وفق المراحل التالية:

أولاً: التحقق من افتراضات النموذج:

تم التحقق من افتراضات نموذج راش وهي أحادية البعد لما لها من أثر على دقة التقديرات، وذلك باستخدام التحليل العاملي لكل نموذج من نماذج الاختبار الثلاثة باستخدام طريقة المكونات الرئيسية (Principle Component Analysis)، ويظهر الجدول (1) قيما للجذر الكامن (Eigen Values)، ونسب التباين المفسر (Explained Variance) للعامل الأول والثاني وناتج قسمة الجذر الأول على العامل الثاني في نماذج الاختبار الثلاثة.

جدول 1. 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 الأول	الجذر الكامن للعامل الثاني
الأول	9.34	3.79	2.46	
	23.5	9.4		
الثاني	11.43	5.01	2.28	
	24.11	10.42		
الثالث	10.2	4.02	2.54	
	23.79	9.46		

نموذج راش

يطلق عليه نموذج أحادي المعلمة One Parameter Model ، وقد اقترحه جورج راش، ويهتم بتحديد موقع الفقرة الاختبارية على سلم صعوبة لجميع الفقرات التي تشكل الاختبار، كما يهتم بتدرج مستويات قدرة الفرد باختبار معين على نفس متصل الفقرات [8].

معلمة الصعوبة

القيمة التي تندرج على متصل السمة ويكون احتمال إجابة الفقرة عندها إجابة صحيحة مساويا 50%.

قدرة الفرد

قيمة يتم تقديرها تعظم أرجحية استجابات الفرد عن فقرات الاختبار.

اختبار الاختيار من متعدد

اختبار تتكون كل فقرة من فقراته أربعة بدائل لكل فقرة بديل واحد فقط صحيح، يعطى الفرد العلامة (1) عند اختيار البديل الصحيح والعلامة (0) عند اختيار البديل الخاطئ. وله ثلاثة صور تختلف فقط في موقع البديل الصحيح.

#### 4. الطريقة والإجراءات

أ. عينة الدراسة

تكونت من جميع أفراد مجتمع الدراسة وهم طلاب وطالبات جامعة تبوك تخصص الدبلوم العام في العام الدراسي 1434 / 1435 هـ، وبلغ عددهم (1545) طالبا وطالبة.

ب. أداة الدراسة

لقد مر بنا أداة 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 الأهداف السلوكية التي تناولتها الموضوعات والمتوقع إتقانها من قبل الطلبة.

- صياغة (40) 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 روعي المبادئ الأساسية التي تحكم صياغة هذا النوع من الفقرات ومطابقتها للهدف

## مقارنة تأثيرات تغيير موقع البديل الصحيح في اختبار اختيار من متعدد على تقديرات الصعوبة للفقرات علي الصالح

الاختبارات الثلاثة والبالغ عددها (6) فقرات أصبح الاختبار بنماذجه الثلاثة مكون من (34) فقرة.  
المعالجات الإحصائية  
من أجل الإجابة عن أسئلة الدراسة، تم استخدام المعالجات الإحصائية التالية:

- إيجاد مؤشر مربع كاي عند مستوى الدلالة  $\alpha = 0.05$  للكشف عن مطابقة الافراد والفقرات لنموذج راش المستخدم في هذه الدراسة (باستخدام البرنامج Bilog-Mg).
- إيجاد معاملات الصعوبة في كل نموذج من نماذج الاختبار الثلاثة حسب متغير موقع البديل الصحيح باستخدام نموذج راش (باستخدام البرنامج Bilog-Mg).
- استخدام تحليل التباين الأحادي لاختبار أثر موقع البديل الصحيح على تقديرات معلمة الصعوبة للفقرات ومعلمة القدرة لدى الأفراد (باستخدام البرنامج SPSS).

### 5. النتائج ومناقشتها

للإجابة عن السؤال الأول " هل تختلف تقديرات معلمة صعوبة الفقرة المقدره باستخدام نموذج راش باختلاف موقع البديل الصحيح في اختبار اختيار من متعدد؟"

تم تقدير قيم معلمة الصعوبة للفقرات باستخدام نموذج راش لكل نموذج من نماذج الاختبار الثلاثة: النموذج الاول: توزيع البديل الصحيح على البدائل أ، ب، النموذج الثاني: توزيع البديل الصحيح على البدائل ج، د، النموذج الثالث: توزيع البديل الصحيح عشوائيا على البدائل، وتم أيضا حساب المتوسطات الحسابية والانحرافات المعيارية لقيم هذه المعالم. ويوضح الجدول (2) قيم معالم الفقرات والمتوسطات الحسابية لكل نموذج من نماذج الاختبار.

### جدول 2

قيم معالم الصعوبة للفقرات والمتوسطات الحسابية والانحرافات المعيارية لها لكل نموذج من نماذج الاختبار حسب متغير موقع البديل الصحيح

رقم الفقرة	الأول: توزيع البديل الصحيح على أ، ب	الثاني: توزيع البديل الصحيح على ج، د	الثالث: توزيع البديل الصحيح عشوائيا
	معلمة الصعوبة	معلمة الصعوبة	معلمة الصعوبة
2	2.69	1.53	2.86
3	0.74	-1.52	-0.69
4	-3.94	-2.00	-2.65
5	-2.08	-0.88	-1.32
8	-2.66	-2.26	-3.07
10	0.12	0.67	0.45
11	-1.86	-1.52	-1.65
13	-1.64	-3.46	-2.55
14	-2.87	2.59	0.00
15	0.37	-0.98	-0.53
16	-2.58	-2.41	-3.18
17	0.58	0.25	0.38
18	2.62	2.66	3.94
19	-0.61	0.95	0.30

1.65	3.67	-0.68	21
-0.53	0.51	-2.67	22
1.23	0.74	2.20	23
-0.61	-0.57	-0.68	24
-0.38	-0.06	-0.92	25
0.60	0.86	0.14	26
0.15	0.37	-0.37	27
3.68	2.48	2.68	28
1.07	0.66	2.11	29
0.68	-0.17	2.17	30
-0.16	0.05	-0.53	31
2.27	1.37	3.94	32
1.82	2.64	1.00	33
0.62	0.55	0.87	34
3.94	2.96	2.74	35
1.90	0.56	2.87	36
2.66	1.53	2.31	37
1.46	0.77	2.61	38
2.89	1.62	2.40	39
0.07	-0.37	1.24	40
0.51	0.41	0.36	المتوسط الحسابي
1.91	1.66	2.07	الانحراف المعياري

إن التباين في مدى قيم معالم صعوبة الفقرات أظهر تبايناً ملحوظاً في الأوساط الحسابية لقيم هذه المعالم، ولمعرفة دلالة الفروق (إحصائياً) بين معالم الفقرات حسب متغير موقع البديل الصحيح، فقد تم استخدام تحليل التباين الأحادي (One Way Anova)، وذلك لوجود متغير مستقل واحد، وهو موقع البديل الصحيح بثلاثة مستويات ومتغير تابع واحد وهو تقديرات معلمة الصعوبة. ويوضح الجدول (3) نتائج تحليل التباين الأحادي.

ويتضح من النتائج في الجدول رقم (2) تبايناً ظاهرياً في مدى صعوبة الفقرات، حيث كان المدى في صعوبة الفقرات هو الأعلى في النموذج الأول (توزيع البديل الصحيح على أ، ب)، الذي تراوحت قيمه من 3.94 - إلى 3.94 بمتوسط حسابي مقداره (0.36)، في حين كان الأقل في النموذج الثاني (توزيع البديل الصحيح على ج، د)، الذي تراوحت قيمه من 3.18 - إلى 3.94 بمتوسط حسابي مقداره (0.51).

### جدول 3

نتائج تحليل التباين الأحادي (One Way Anova) لمعرفة دلالة الفروق في المتوسطات الحسابية لتقديرات معلمة الصعوبة للفقرات باختلاف موقع البديل الصحيح

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	الدلالة الإحصائية
موقع البديل الصحيح	.386	2	.193	.054	.947
الخطأ	352.977	99	3.565		
الكل	353.364	101			

المقدرة ( $\theta$ ) باستخدام نموذج راش باختلاف موقع البديل الصحيح في اختبار اختيار من متعدد؟ " فقد تم إيجاد الأوساط الحسابية والانحرافات المعيارية لتقديرات معلمة القدرة للأفراد حسب نموذج الاختبار (متغير موقع البديل الصحيح). ويبين الجدول (4) هذه القيم.

يتضح من الجدول (3) أن قيمة p-value الدلالة الإحصائية أكبر من 0.05 للمتغير المستقل موقع البديل الصحيح وفقاً لنتائج الاختبارات الإحصائية، وبالتالي هذا يعني أنه لا يوجد تأثيراً دالاً إحصائياً لموقع البديل الصحيح على تقديرات معالم الصعوبة للفقرات.

للإجابة عن السؤال الثاني " هل تختلف تقديرات معلمة قدرة الفرد

جدول 4. الأوساط الحسابية والانحرافات المعيارية لتقديرات معلمة القدرة للأفراد ( $\theta$ ) حسب متغير موقع البديل الصحيح لكل نماذج الاختبار

نماذج الاختبار	العدد	الوسط الحسابي	الانحراف المعياري
الأول: توزيع البديل الصحيح على أ، ب	519	- 0.007	1.52
الثاني: مواقع البديل الصحيح توزعت في ج، د	509	- 0.004	1.21
الثالث: مواقع البديل الصحيح توزعت عشوائياً	513	- 0.0047	1.29

## مقارنة تأثيرات تغيير موقع البديل الصحيح في اختبار اختيار من متعدد على تقديرات الصعوبة للفقرات علي الصالح

يتضح من الجدول (4) وجود فروق ظاهرية بين المتوسطات الحسابية لتقديرات معلمة القدرة للأفراد حسب متغير موقع البديل الصحيح، ولمعرفة الدلالة الإحصائية لتلك الفروق الظاهرية تبعاً لمتغير

جدول 5

نتائج تحليل التباين الأحادي (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	.002	2	.001	001	.999
الخطأ	2710.597	1485	1.825		
الكل	2710.599	1487			

[5] المالكي، يحيى (2015). درجة التزام معلمي المرحلة المتوسطة بقواعد بناء الأسئلة الموضوعية، رسالة ماجستير غير منشورة، جامعة أم القرى، المملكة العربية السعودية.

[8] علام، صلاح الدين (2000). القياس والتقويم التربوي والنفسي - أساسياته وتطبيقاته وتوجهاته المعاصرة - القاهرة، دار الفكر العربي.

[14] عودة، أحمد (1992). مدى التوافق بين نموذج راش والمؤشرات التقليدية في اختبار فقرات قائمة اتجاه سباعي التدريج، مجلة كلية التربية بجامعة الإمارات، العدد 8 يونيو، 153 - 179.

[17] الشريفي، نضال (2012). أثر طريقة تقدير معالم الفقرة وقدرات الأفراد على قيم معالم الفقرة والخصائص السيكمترية للاختبار، في ضوء تغير حجم العينة. المجلة الأردنية في العلوم التربوية، 12، (104): 22-67.

[18] السويلم، بندر (2003). القواعد العامة في صياغة أسئلة الاختبار من متعدد، ومدى التزام أساتذة كلية التربية بجامعة الملك سعود بهذه القواعد، مركز بحوث كلية التربية، جامعة الملك سعود 3 (20)، ص 54.

[20] سليمان، شاهر (2008). أثر مخالفة قواعد صياغة فقرات اختبار اختيار من متعدد على الخصائص السيكمترية للاختبار وفقراته، المؤتمر السنوي الخامس لقسم علم النفس تطبيقات العلوم النفسية ومشكلات المجتمع في الفترة من 4-6 مايو، جامعة طنطا، كلية الآداب، جمهورية مصر العربية.

[21] الحارثي، ابتسام (2007). تقويم كفايت بناء الاختبارات التحصيلية لدى معلمات العلوم للمرحلة الثانوية بمحافظة جدة، رسالة ماجستير غير منشورة، جامعة الملك عبد العزيز، جدة.

[22] الكنانى، عايد وفيضي، فاضل (2012). فاعلية استخدام موقع البديل الفعال عند ابعاد البديل الخاطئ في صعوبة فقرات الأسئلة اختيار من متعدد وتمييزها، مجلة القادسية لعلوم التربية الرياضية، 12 (1)، ص 207-225.

[23] ظاظا، حيدر (2012). الكشف عن مدى انتهاك قواعد صياغة فقرات الاختبار من متعدد في أسئلة امتحان شهادة الدراسة الثانوية العامة في الأردن، المجلة الأردنية في العلوم التربوية، 8 (1)، ص 81-94.

يتضح من الجدول (5) بأنه لا يوجد فروق ذات دلالة إحصائية عند مستوى الدلالة الإحصائية ( $\alpha = 0.05$ ) 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. لقد اختلفت نتائج هذه الدراسة مع النتائج التي توصلت إليها دراسة ويفرك [19] والتي أشارت الى وجود تأثيرات لموقع البديل الصحيح في اختبار اختيار من متعدد على معاملات الصعوبة للفقرات، وقد اتفقت نتائج هذه الدراسة مع النتائج التي توصلت إليها دراسة الكنانى وفيضي [22]، والتي اشارت الى عدم وجود تأثيرات لتغيير موقع البديل الصحيح على تحصيل الطلبة في أسئلة الاختبار من متعدد.

### 6. التوصيات

يوصي الباحث القيام بالدراسات التالية:

- دراسة تأثيرات تغيير موقع البديل الصحيح في اختبار اختيار من متعدد بخمسة بدائل وستة بدائل على تقديرات معلمة صعوبة الفقرات وتقديرات معلمة القدرة لدى الافراد.
- دراسة تأثيرات تغيير موقع البديل الصحيح في اختبار اختيار من متعدد باستخدام نماذج أخرى من نظريات الاستجابة للفقرة.
- مقارنة نتائج هذه الدراسة مع النتائج فيما لو أعيد التحليل في ضوء النظرية الكلاسيكية في القياس.

### المراجع

#### أ. المراجع العربية

- [1] علام، صلاح الدين (1987). دراسة موازنة ناقدة لنماذج السمات الكامنة، والنماذج الكلاسيكية في القياس النفسي والتربوي - جامعة الكويت، المجلة العربية للعلوم الإنسانية، العدد (27)، 18، - 44.
- [3] علام، صلاح الدين (2008). نماذج لاستجابات الفقرات الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي، ط1، القاهرة: دار الفكر العربي.



- [11] Rup, A.A., Zumbo, B.D. (2006). *Understanding parameter invariance in unidimensional IRT models*. Educational and Psychological Measurement, 66, 63-84.
- [12] Crocker, L., Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston.
- [13] Allen, M. & Yen, W. M., (1989): *Introduction to Measurement theory* California, Brooks Cole publishing company
- [15] Degrujter, N., & Van der Kamp, J. (2005). *Statistical test theory for education and Psychology*. Retrieved in 2013 from: [www.leidenuniv.nl/griuijterdne](http://www.leidenuniv.nl/griuijterdne).
- [16] Hambleton, R., Rogers, H. (1989). *Detecting potentially biased test items*. Comparison of IRT area and Mantel-Hanszel method. *Applied Measurement in Education*, 2(4), 313-334.
- [19] Wevric, L (2001). Response set in a multiple-choice test, *Educational and psychological Measurement*, vol (33) pp58-65.
- [26] Hattie, J. (1984). *An empirical study of various indices for determining unidimensionality*. *Multivariate Behavioral Research*, 19, 49-78.
- [24] بني عطا، زايد والرياعي، إبراهيم (2013). 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*، 9 (3)، ص319-333.
- [25] يعقوب، ابراهيم وأبو فودة، باسل(2012). أثر مخالفة قواعد صياغة فقرات الاختبار من متعدد على الخصائص السيكومترية للاختبار وفقراته، *مجلة جامعة دمشق*، 28 (1) ص419 – 442.
- ب. المراجع الاجنبية
- [2] Hambleton, R. K., & Swaminathan, H. (1995). *Item response theory: Principles and applications*. Boston: Kluwer-Nijhoff.
- [4] Weiten, W. (1984). Violation of selected item construction principles in - Educational measurement. *Journal of Experimental Education*, 52 (3), PP. 78 – 174.v
- [6] Sarnacki, R (1979). *An examination of test domain*, *Review of Educational Research*, 49(2)252-279.
- [7] Nunnally, J.C, (1981); *Psychometric Theory*. 2nd ed. McGraw-Hill, New York.
- [9] Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1999). *Fundamentals of item response theory*. London: Sage Publications, Inc.
- [10] Lord, F. M. (1980). *Application of item response theory to practical testing problems*. New Jersey: Lawrence Erlbaum Associates, Publishers.

# COMPARISON OF THE EFFECTS OF CHANGING THE CORRECT ALTERNATIVE LOCATION IN A MULTIPLE CHOICE TEST ON THE ITEMS DIFFICULTY ESTIMATES AND INDIVIDUAL ABILITY ESTIMATES USING RASH MODEL

ALI MOHAMMAD KHALID AL-SALEH  
Ministry of Education – Jordan

**ABSTRACT\_** *The current study aimed at detecting the effects of changing the location of the correct alternative in a multiple choice test on the difficulty estimates of items and estimates of the ability of individuals using the Rush model. To achieve the aims of the study were built achievement test of multiple choice with four alternatives in educational measurement and evaluation scheduled for students of General Education Diploma consisting of (34) items in the final form of test. It contains three models that had the same content and the variable was the correct alternative site. The first model was the correct alternative site on alternatives A and B. The second model was the correct alternative position on alternatives C, D, and the third model randomly distributed the correct alternative on the four different alternatives. The three models were applied to a random sample of (1541) students from the University of Tabuk. The researcher used the SPSS15, Bilog-Mg3 software to analyze the data for each of the three test models according to Rush model. The results of the study showed that there were no statistically significant differences between the arithmetic means of items difficulty coefficients due to the location of the correct alternative, and no statistically significant differences between the arithmetic means of the individual's ability parameter estimates due to the location of the correct alternative.*

**KEY WORDS:** *Item difficulty parameter, Ability parameter, correct alternative position, Rash Model, Item of Response Theory.*