Examination of the effects of correct alternative position of multiple-choice test items on differential performance of items using the Mantel-Hansel method

Ali Mohammad Khaled AL Saleh
Ministry of Education
alialsaleh333@gmail.com

Received : 26/07/2022 Accepted :11/12/2022

Abstract:

The study seeks to examine the impact of the position of the correct option in multiple-choice test items on differential item functioning (DIF) by employing the Mantel-Haenszel method. This method is one of the most commonly used approaches for identifying DIF and remains one of the most popular techniques for detecting bias in test items today. In order to achieve the objectives of the study, the study involved constructing an achievement test in educational measurement and evaluation, consisting of multiple-choice questions with four alternatives and a total of 34 items. The test included three forms with the same content, and the variable among these forms was the position of the correct alternative. The first form of test shows the location of the correct alternative on alternatives a, b, and the second form has the location of the correct alternative on alternatives c, d, and the third form is randomly distributing the correct alternative to the various four alternatives. The Mantel-Haenszel method was used through SPSS to detect the differential performance of the items for each of the three test models.

The results of the study showed that the percentage of items that showed a differential performance in favor of the first test model was 13%, in favor of the second model, 23%, and in favor of the third model, 28% of the total number of items in the test. The results also showed that there were no statistically significant differences between the arithmetic averages of the ability parameter estimates for individuals due to the location of the correct alternative.

The researcher recommended that when constructing a multiple-choice test, it is better to randomly distribute the correct alternative to the four different alternatives. It also recommended conducting more studies on the detection of differential performance in the light of other variables and in other ways, including the method that relies on the item response theory or the area between the curves.

Keywords: differential performance, Mantel-Hansel method, location of the right alternative, classical theory of measurement.

فحص تأثيرات موقع البديل الصحيح لفقرات اختبار اختيار من متعدد على الأداء التفاضلي للفقرات باستخدام طريقة مانتل–هانزل

علي محمد خالد الصالح
وزارة التربية والتعليم
alialsaleh333@gmail.com

الاستلام : 2022/07/26 القبول : 2022/12/11

الملخص:

هدفت الدراسة الكشف عن تأثيرات موقع البديل الصحيح لفقرات اختبار اختيار من متعدد، على الأداء التفاضلي للفقرات باستخدام طريقة مانتل–هانزل، التي تعدّ من الطرق الأكثر استخدامًا من أجل الكشف عن دالة الفقرة التفاضلية (DIF)، وهي من أشهر طرائق الكشف عن التحيز في الوقت الحاضر؛ وذلك لسهولة حساباتها وإجراءاتها. ولتحقيق أهداف الدراسة تم بناء اختبار تحصيلي في القياس والتقويم التربوي، من نوع الاختيار من متعدد بأربعة بدائل، مكون من (34) فقرة بصورته النهائية، اشتمل على ثلاثة نماذج لها المحتوى نفسه، والمتغير بين هذه النماذج الثلاثة هو موقع البديل الصحيح، الأنموذج الأول فيه موقع البديل الصحيح على البديلين (a)، (b)، والأنموذج الثاني فيه موقع البديل الصحيح على البديلين (c)، (d)، والأنموذج الثالث توزيع البديل الصحيح عشوائيًّا على مختلف البدائل الأربعة، وقد تم استخدام طريقة مانتل–هانزل من خلال برنامج (SPSS)؛ للكشف عن الأداء التفاضلي للفقرات لكل أنموذج من نماذج الاختبار الثلاثة. وأظهرت نتائج الدراسة أنّ نسبة الفقرات التي أظهرت أداءً تفاضليًّا لصالح أنموذج الاختبار الأول (13%)، ولصالح الأنموذج الثاني (23%)، ولصالح الأنموذج الثالث (28%)، من عدد الفقرات الكلي للاختبار. كما أظهرت عدم وجود فروق دالة إحصائيًّا بين المتوسطات الحسابية لتقديرات معلمة القدرة للأفراد، تُعزى لموقع البديل الصحيح. وفي نهاية الدراسة أوصى الباحث أنّه عند بناء اختبار اختيار من متعدد، من الأفضل أن يكون توزيع البديل الصحيح عشوائيًّا على مختلف البدائل الأربعة. كما أوصت بإجراء المزيد من الدراسات حول الكشف عن الأداء التفاضلي، في ضوء متغيرات أخرى، وبطرق أخرى، منها: الطريقة التي تعتمد على نظرية خصائص الفقرة (النظرية الحديثة في القياس) (Item Response Theory)، أو المساحات بين المنحنيات.

الكلمات المفتاحية: الأداء التفاضلي، طريقة مانتل–هانزل، موقع البديل الصحيح، النظرية الكلاسيكية في القياس.

المقدمة:

تستخدم الاختبارات على اختلاف أنواعها وأشكالها على مدى واسع، في الجامعات والمدارس لأغراض متعددة؛ من أجل تقييم المهارات والمعارف، والتصنيف والتقويم والتنبؤ بفرص النجاح في المستقبل، ولتحقيق هذه الأغراض، يجب أن يتسم الاختبار بخصائص عديدة، منها: معاملات صعوبة فقراته، والتمييز، والثبات، وفعالية المموهات، والصدق، والتحيز (الأداء التفاضلي للفقرة)، وفق معايير محددة في أدبيات القياس والتقويم التربوي؛ من أجل التوصل إلى استنتاجات مبنية على أسس علمية ذات دلالة وفائدة.

إنّ قواعد صياغة فقرات الاختبار التحصيلي، تمثل نتاج خبرة عملية، وأحكام الخبراء مبنية على المنطق في هذا المجال، وعلى الخلل أو العيوب التي قد يقع بها باني الاختبار (Weiten,1984)، فقد أشار المالكي (Al-Maliki,2015) إلى أنّ معلمي العلوم في المرحلة المتوسطة يلتزمون بنسبة (60%) من قواعد بناء أسئلة الاختيار من متعددِ أربعة بدائل، وقد يعزى ذلك إلى أنّ معظم المدرسين مضغوطين بالوقت ومتطلبات وظيفتهم، إضافة إلى عدم معرفة بعضهم بكيفية صياغة الأسئلة، وبالنتيجة فإنّ اختباراتهم قد تكون عرضة لأن تتضمّن فقرات ضعيفة الصياغة، وأنّ وجود هذه العيوب تكون مؤشرات لما يعرف بالحكمة في الإجابة لدى بعض الطلبة؛ مما قد يؤثر سلبًا على خصائص الفقرة والاختبار، بالإضافة إلى تضخيم درجات الاختبار بشكل مضلل؛ لأنّ فقرات الاختبار بدلًا 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).

تحظى قضية بناء مقياس تربوي عادل، باهتمام المختصين في القياس النفسي والتربوي، وذلك نظرًا لأهميّة القرارات التي يمكن أن يتوصلوا لها في ضوء نتائج تطبيق المقياس، إذ لا بدّ من توافر درجة مقبولة من الصدق والثبات لذلك المقياس. والمقياس يكون صادقًا بالنسبة لجميع الأفراد والمفحوصين إذا أعطى فرصًا متساوية للجميع في إبداء آرائهم وميولهم واتجاهاتهم وفق ما يتم قياسه، فإذا كان المقياس المصمم لقياس سمة ما، يميز مجموعة من المفحوصين عن المجموعة

أو المجموعات الأخرى من المستوى نفسه، فيمكن اعتباره متحيّزًا، وبالتالي ينتهك مبدأ العدالة بين المفحوصين (Rover, 2005)، ويعدّ التحيز أحد مهددات صدق المقياس، فعندما تميّز إحدى فقراته أيًّا من مجموعات المفحوصين عن المجموعات الأخرى، فإنّ ذلك يعني أنّ هذه المجموعة لها أفضلية دون المجموعات الأخرى في الظاهرة التي يتم قياسها، لذلك فإنّ التحقق من خلو المقياس وفقراته من التحيز، يعدّ خطوة مهمة من خطوات بنائه (Schumacher, 2010).

وينظر لمفهوم التحيز على أنّه مفهوم يخفي وراءه معنى سلبيًّا، يرتبط بإطاره العام بفكرة عدم العدالة، والاهتمام بالمفاهيم المشحونة بالجانب السياسي والاجتماعي، أكثر من ربطه بالمفهوم السيكومتري، لذلك عمل المتخصصون في مجال القياس والتقويم على استخدام مفهوم الأداء التفاضلي للفقرة؛ للتعبير عن الطرق والمعاني الإحصائية التي تستخدم للكشف عن التحيز، من منطلق أنّه من الأفضل استخدام المعاني الفنية السيكومترية، بدلًا من المعاني المشحونة اجتماعيًّا وسياسيًّا (Ellis, 2003).

إنّ طريقة مانتل–هانزل تعد من أكثر طرق نظرية القياس الكلاسيكية انتشارًا في الكشف عن تحيز الفقرة؛ وذلك لسهولة حساباتها وإجراءاتها (Nanda, 1993)، كما تعدّ إحدى طرق الكشف عن الأداء التفاضلي في الوقت الحاضر، وقد استخدمها التربويون والعاملون في القياس بشقيه النفسي والتربوي؛ بسبب سهولة تطبيقها وحساباتها وإجراءاتها، فهي تفحص الأداء التفاضلي بين المجموعتين المرجعية (R) والمستهدفة (G)، وهي المجموعة التي يعتقد بوجود فقرات تتحيز ضدها، وذلك وفق جدول توافق ثنائي البعد، وقد طورت لتصبح أكثر فاعلية مع الفقرات متعددة التدريج، حيث سميت بطريقة مانتل–هانزل العامة، والتي تركز على العلاقة بين مستوى القدرة واستجابات فئات المتغير، حيث تنظّم البيانات من خلالها وفق جدول توافق متعدد الأبعاد.

وتقوم على استقصاء التحيز بين مجموعتين فرعيتين من مجتمع الدراسة، مجموعة تسمى المجموعة المرجعية (Group Reference)، والمجموعة الأخرى تسمى المجموعة المستهدفة (Group Focal)، وهي المجموعة التي يعتقد بوجود فقرات تتحيز ضدها، ويتم بناء الفرضية الصفرية من خلالها، بحيث تنصّ على عدم وجود فروق دالة إحصائيًّا في نسب الإجابة الصحيحة بين المجموعتين المرجعية والمستهدفة، في مستوى القدرة الواحدة (Kamata, 2004).

طريقة مانتل–هانزل (HM):

تعدّ طريقة مانتل–هانزل من الطرق الأكثر استخدامًا من أجل الكشف عن دالة الفقرة التفاضلية (DIF)، وقدمت هذه الطريقة للكشف عن تحيز الفقرة في أواخر الثمانينات من القرن الماضي، من قبل هولند وثاير، وهي من أشهر طرائق الكشف عن التحيز في الوقت الحاضر، وذلك لسهولة حساباتها وإجراءاتها، وتقوم هذه الطريقة باستقصاء التحيز بين مجموعتين، إحداهما تسمى المجموعة المرجعية، والأخرى تسمى المجموعة المستهدفة، وهي المجموعة التي تتأثر بتحيز الفقرات، بمعنى أنّ الفقرات تتحيز ضدها كما يعتقد، ولتطبيق هذه الطريقة فإنّ البيانات تنظم في مصفوفة ثنائية (2x2) كما في الجدول (1).

الجدول (1): جدول البيانات الأساسية للكشف عن الأداء التفاضلي للفقرة

المجموعة	إجابة صحيحة (1)	إجابة صحيحة (0)	المجموع
المرجعية	At	Bt	Nrt
المستهدفة	Ct	Dt	Nft
المجموع	N1t	N0t	Nt

حيث إنّ:

At: عدد أفراد المجموعة المرجعية الذين أجابوا عن الفقرة إجابة صحيحة في مستوى القدرة (t).

Bt: عدد أفراد المجموعة المرجعية الذين لم يستطيعوا الإجابة عن الفقرة بشكل صحيح في مستوى القدرة (t).

Ct: عدد أفراد المجموعة المستهدفة الذين أجابوا عن الفقرة إجابة صحيحة في مستوى القدرة (t).

Dt: عدد أفراد المجموعة المستهدفة الذين لم يستطيعوا الإجابة عن الفقرة بشكل صحيح في مستوى القدرة (t).

N1t: عدد الأفراد الذين أجابوا عن الفقرة بشكل صحيح من كلا المجموعتين في مستوى القدرة (t).

N0t: عدد الأفراد الذين لم يستطيعوا الإجابة عن الفقرة بشكل صحيح من كلا المجموعتين في مستوى القدرة (t).

Nrt: عدد الأفراد الذين أجابوا عن الفقرة من المجموعة المرجعية عند مستوى القدرة (t).

Nft: عدد الأفراد الذين أجابوا عن الفقرة من المجموعة المستهدفة عند مستوى القدرة (t).

Nt: عدد الأفراد من المجموعتين المرجعية والمستهدفة عند مستوى القدرة (t).

واعتمادًا على الجدول السابق يتم تطبيق معادلة مانتل–هانزل كاي تربيع الآتية:

$$MH\chi^2 = \frac{\left(\left|\sum A_t - \sum E(A_t)\right| - 0.5\right)^2}{\sum Var(A_t)}$$

حيث إنّ E (At): هي القيمة المتوقعة لعدد أفراد المجموعة المرجعية، الذين أجابوا عن الفقرة إجابة صحيحة و Var (At) هو تباين (At). ويتبع الإحصائي $2MH\chi$ توزيع 2χ بدرجات حرية تساوي 1 كما

يمكن استخدام نسبة الأرجحية المشتركة للكشف عن التحيز، والتي تحسب وفق المعادلة الآتية:

$$\alpha MH = \frac{\sum \frac{A_t D_t}{N_t}}{\sum \frac{B_t C_t}{N_t}}$$

ويتم تحديد الأداء التفاضلي للفقرة من خلال قيمة دلالة إحصائي مانتل–هانزل، فإذا كانت قيمة الدلالة أقل من (0.05)، فإنّ ذلك يعني أنّ الفقرة تظهر أداءً تفاضليًّا، أمّا إذا كانت أكبر، فهذا يعني أنّ الفقرة لا تظهر أداءً تفاضليًّا، والفقرة التي تظهر أداءً تفاضليًّا، تظهر أداءً تفاضليًّا لصالح المجموعة المستهدفة، وإنْ كانت أكبر من (1)، فالفقرة تظهر أداءً تفاضليًّا لصالح المجموعة المرجعية Mubarak, (2010). وفي هذه الدراسة تم استخدام طريقة مانتل–هانزل للكشف عن الأداء التفاضلي لفقرات اختبار اختيار من متعدد، كون هذه الطريقة من أكثر الطرق الشائعة في الكشف عن الأداء التفاضلي، ولتوفر البرامج الإحصائية المستخدمة للكشف عن الأداء التفاضلي حسب هذه الطريقة.

المشكلة البحثية وأسئلة الدراسة:

تعتمد الفقرة الجيدة من نوع الاختيار من متعدد على خصائص متعددة، ومن معايير الفقرة الجيدة أن تكون ذات فاعلية وعادلة. ويعدّ الأداء التفاضلي للفقرة من المؤشرات والأساليب المستخدمة في بحث عدالة الفقرة، التي تعدّ جزءًا مهمًّا من صدقها. كما أنّه، لم يسبق تحديد العلاقة بين موقع البديل الصحيح في فقرات الاختيار من متعدد، على نطاق واسع- حسب علم الباحث– في أيّ دراسة سابقة؛ لذلك سعت الدراسة الحالية إلى كشف العلاقة بين الأداء التفاضلي للفقرة، وموقع البديل الصحيح في فقرات اختبارية من نوع الاختيار من متعدد. وتتميز الدراسة الحالية بأنّها ستفيد في تحديد موقع البديل الصحيح، الذي ينتج أداء تفاضليًّا للفقرة، حتى تكون أكثر عدالة بين المجموعات المختلفة من المفحوصين، وبالتحديد تحاول الدراسة الإجابة عن السؤال الرئيس الآتي:

ما العلاقة بين موقع البديل الصحيح والأداء التفاضلي للفقرة في اختبار في القياس والتقويم التربوي من نوع الاختيار من متعدد؟

وللإجابة عن هذا السؤال الرئيس تحاول الدراسة الإجابة عن الأسئلة الفرعية الآتية:

السؤال الأول: ما هي فقرات الاختبار المعرفي في القياس والتقويم، والتي تظهر أداء تفاضليًّا لمتغير موقع البديل الصحيح (موقع البديل الصحيح في الخيارين (a)، (b) مقابل توزيع موقع البديل الصحيح عشوائيا)؟

السؤال الثاني: ما هي فقرات الاختبار المعرفي في القياس والتقويم، والتي تظهر أداء تفاضليًّا لمتغير موقع البديل الصحيح (موقع البديل الصحيح في الخيارين (c)، (d) مقابل توزيع موقع البديل الصحيح عشوائيا)؟

السؤال الثالث: ما أثر موقع البديل الصحيح في اختبار اختيار من متعدد على تقديرات القدرة، لدى الأفراد وفق النظرية الكلاسيكية في القياس؟

أهداف الدراسة:

تهدف الدراسة الحالية إلى الآتي:

– مقارنة تأثير تغيير موقع البديل الصحيح: (توزيع موقع البديل الصحيح في الخيارين (a)، (b) مقابل توزيع موقع البديل الصحيح عشوائيًّا)، في اختبار اختيار من متعدد على الأداء التفاضلي للفقرات.

– مقارنة تأثير تغيير موقع البديل الصحيح: (توزيع موقع البديل الصحيح في الخيارين (c)، (d) مقابل توزيع موقع البديل الصحيح عشوائيا)، في اختبار اختيار من متعدد على الأداء التفاضلي للفقرات.

– مقارنة تأثير تغيير موقع البديل الصحيح في اختبار اختيار من متعدد، على تقديرات القدرة لدى الأفراد وفق النظرية الكلاسيكية في القياس.

الأهمية النظرية والعملية للدراسة:

تتسم أهمية الدراسة الحالية من الناحية النظرية، في تسليط الضوء على العلاقة بين الأداء التفاضلي لفقرة الاختبار من نوع الاختيار من متعدد، وموقع البديل الصحيح لها. ومن الناحية العملية تفيد الدراسة في تحديد أنماط موقع البديل الصحيح، في فقرات من نوع الاختيار من متعدد، المؤثرة في الأداء التفاضلي للفقرات، ويفيد ذلك في تحديد بعض مسببات الأداء التفاضلي للفقرات. وبمعالجة تلك المسببات خلال عملية مراجعة الفقرات؛ تتحسن عدالة الاختبار وصدقه، وهذا بدوره يساعد في زيادة دقة القرارات، وتحقيق العدالة.

مصطلحات الدراسة:

- الأداء التفاضلي للفقرة (DIF): هو مؤشر إحصائي للتعبير عن الفروق في الاستجابة الصحيحة، على الفقرة بين مجموعتين من المفحوصين، ممّن لهم المستوى نفسه من (Camilli & Shepard,1994).
- طريقة مانتل–هانزل (HM): طريقة من الطرق الأكثر استخدامًا من أجل الكشف عن الأداء التفاضلي للفقرة (DIF)، وتقوم هذه الطريقة باستقصاء التحيز بين مجموعتين، إحداهما تسمى المجموعة المرجعية، والأخرى تسمى المجموعة المستهدفة، وهي المجموعة التي تتأثر بتحيز الفقرات، بمعنى أنّ الفقرات تتحيز ضدّها كما يُعتقد.
- الأداء التفاضلي والتحيز: الأداء التفاضلي يهتم بالخصائص السيكومترية للفقرة، من حيث الكيفية التي تعمل بها في المجموعات المختلفة، بينما يهتم تحيز الفقرة بالمعنى الاجتماعي بشكل كبير، وعدم العدالة أو المساواة.

الحدود الزمنية والمكانية للدراسة:

اقتصرت هذه الدراسة جميع طلاب وطالبات جامعة تبوك، تخصص الدبلوم العام في العام الدراسي (1435/1434هـ).

الحدود الموضوعية للدراسة:

اقتصرت هذه الدراسة على فحص تأثيرات موقع البديل الصحيح، على دالة الأداء التفاضلي للفقرات باستخدام طريقة مانتل–هانزل.

الدراسات السابقة:

بعد مراجعة الأدب التربوي والدراسات السابقة المتعلقة بموضوع الدراسة، تمّت الاستفادة من عدّة دراسات باللغتين العربية والإنجليزيّة ومن أهمها:

- دراسة (Smith ,1993): هدفت الكشفَ عن التحيز في فقرات اختبارات ولاية مين الأمريكية للتقييم التربوي (Main Educational Assessment, (MEA وهو اختبار يتم تطبيقه لتقييم مستوى طلبة الصفوف: الرابع، والثامن، والحادي عشر؛ لتحديد صدق الاختبار عبر الثقافات المختلفة، وقد اقتصر تحليل النتائج على البيانات التي تمثل أداء طلبة الصف الثامن على (100) فقرة مشتركة، بين اختبار الرياضيات، واختبار القراءة، وتم تطبيقها على (732) طالبًا نصفهم تمامًا يتكلمون اللغة الإنجليزية فقط، والنصف الآخر يتكلمون اللغة الإنجليزية والفرنسية بطلاقة للعام الدراسي (1989)، وقد تم استخدام طريقة فرق المساحة بين منحنيات خصائص الفقرة، حسب الأنموذج اللوجستي ثلاثي المعلم؛ من أجل الكشف عن التحيز في فقرات الاختبار حسب متغير اللغة: (طلبة يتكلمون اللغة الإنجليزية فقط مقابل طلبة يتكلمون اللغتين الإنجليزية واللغة الفرنسية)، وقد بينت نتائج الدراسة وجود فرق في الأداء على الفقرات، تبعًا لمتغير اللغة التي يتكلمها الطلبة، ولصالح الطلبة الذين يتكلمون اللغتين الإنجليزية والفرنسية، ولذا يجعلها غير صادقة عبر الثقافات المختلفة.
- دراسة (Pae, 2004): هدفت الكشفَ عن الأداء التفاضلي في فقرات اختبار الجزء الإنجليزية، من اختبار القبول الكوري للعام (1998)، تبعًا لمتغير المجال الأكاديمي: (علوم إنسانية، علوم)، وذلك باستخدام طريقة نسبة الأرجحية، وطريقة مانتل–هانزل، حيث يضمّ الجزء الإنجليزي في هذا الاختبار (55) فقرة، منها (17) فقرة تقيس مهارة الاستماع، و(38) فقرة تقيس مهارة استيعاب المقروء، وطبق الاختبار على (839837) مفحوصًا، اختار الباحث منهم عينة عشوائية مكونة من (14000) مفحوصٍ، تمّ توزيعهم بالتساوي على المجالين الأكاديميين. وأظهرت نتائج الدراسة وجود (18) فقرة ذات أداء تفاضلي، فيها (7) فقرات كانت أسهل للعلوم الإنسانية، وتم تحليل محتوى الفقرات التي أظهرت أداء تفاضليًّا لمتغير المجال الأكاديمي في محاولة لتفسير نتائج الدراسة.
- دراسة (Hamadna, 2007): هدفت مقارنة الأداء التفاضلي لفقرات اختبارات تحديد الكفاءة اللغوية في اللغة الإنجليزية، في الجامعات الأردنية: (اليرموك، آل البيت، العلوم والتكنولوجيا)، تبعًا لمتغير الجنس وفرع الثانوية العامة، حسب طريقتي مانتل–هانزل وطريقة فرق المساحة بين منحنيات خصائص الفقرة في الأنموذج اللوجستي ثلاثي المعلم، والتعرف إلى درجة التوافق بين الطريقتين. وقد تمّ اختيار ثلاثة نماذج من اختبارات تحديد الكفاءة اللغوية في اللغة الإنجليزية. وكانت جميع فقرات الاختبارات من نوع الاختيار من متعدد، وبلغ عدد الفقرات في الجامعات الثلاثة (100،50،50) فقرة، وتقيس الفقرات في مجملها مهارات مشتركة في اللغة الإنجليزية، هي: القواعد والمفردات، استيعاب المقروء، والكتابة، وضمّت عينة الدراسة (1935) طالبًا وطالبة، وقد أظهرت نتائج الدراسة أنّ أكثر المهارات التي أظهرت فقراتها أداءً تفاضليًّا لمتغير الجنس، هما: مهارتا القواعد والمفردات، وكانت في معظمها لصالح الإناث، وجاءت نسب الفقرات التي أظهرت أداءً تفاضليًّا لمتغير الجنس: (24% ,21 %, 3%)، لجامعات: اليرموك، جامعة العلوم والتكنولوجيا، و آل البيت) على التوالي.
- دراسة (Sulaiman,2008): هدفت استقصاءَ أثر مخالفة قواعد صياغة فقرات اختبار الاختيار من متعدد، على الخصائص السيكومترية للاختبار وفقراته، ولتحقيق هذا الهدف تمّ إعداد اختبار تحصيليّ في مادة علم النفس التربوي، وهي إحدى المواد التي تطرح في الجامعة، وتكوّن الاختبار من (40) فقرة مخالفة من نوع اختيار من متعدد أربعة بدائل، بواقع (10) فقرات لكلّ نوع من المخالفات الأربع الآتية: طول البديل الصحيح، موقع البديل الصحيح، البديل: جميع ما ذكر صحيح، البديل: كلاهما صحيحان، ويناظره اختبارٌ مكوّنٌ من (40) فقرة متحررة من المخالفات، ولها المحتوى نفسه، طبق الاختبار على عينة عشوائية مكونة من (482) طالبًا من جامعة تبوك. وقد دلّت النتائج أنّ الفقرات المخالفة لقواعد الصياغة، أسهل من نظيرتها الفقرات المتحررة من المخالفات، وبفارق دال إحصائيًّا، ولا توجد فروق دالة إحصائيًّا بين معاملات تمييز الفقرات المخالفة الصياغة، ونظيرتها المتحررة من المخالفات، كما توجد فروق دالة إحصائيًّا بين معاملات صدق الاختبار وثباته، الذي فقراته مخالفة لقواعد صياغة الفقرات، والاختبار الذي فقراته متحررة من المخالفات، والفروق لصالح الاختبار الذي فقراته متحررة من المخالفات.
- دراسة (Zaza ،2012): تهدف الكشفَ 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 مخالفة الفقرة لقواعد صياغتها، وأظهرت النتائج أنّ (416) فقرة أي ما نسبته (12.6%) تضمّنت أكثر من مخالفة، وبشكل إجمالي وجدت (782) فقرة أي ما نسبته (26.9%) فقرة مخالفة، وتبيّن أن أكثر من قاعدتين انتهاك كانتا

للفقرتين: ترتيب البدائل عموديا بنسبة (67.7%)، وترتيب البدائل منطقيًا أو عدديًا أو أبجديًا 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.

- دراسة (Al-Kinani and Faydi,2012): بعنوان: فاعلية استخدام موقع البديل الصحيح عند أبعاد البديل الخاطئ في صعوبة الفقرات 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 الباحثين إلى الكشف عن أثر تغير موقع البديل الصحيح عند استخدام تقنية إبعاد البدائل الخاطئة على تحصيل الطلبة في أسئلة الاختيار من متعدد، وقد اشتملت عينة الدراسة على (60) طالبا من طلبة المرحلة الثالثة لكلية التربية الرياضية في جامعة المثنى في العراق ، حيث تم تقسيم عينة الدراسة إلى مجموعتين متساويتين، حيث أجاب (130) طالبًا وطالبة عن أسئلة الأنموذج الأول كما أجاب (130) طالبًا عن أسئلة الأنموذج الثاني. ودلّت النتائج على عدم وجود فروق معنوية بين أنموذجي الاختبار، من حيث الصعوبة والتمييز.
- دراسة (Al-Batoush, and Al-Quraan,2018): هدفت لفحص الأداء التفاضلي للفقرات، في أدوات تقييم الطلبة لجودة التعليم العالي في الأردن حسب الكلية الأكاديمية، باستخدام طريقة مانتل–هانزل، وتوصلت النتائج لوجود (4) فقرات أبدت أداء تفاضليًا في المقياس وفقًا للكلية، ولصالح طلبة كلية العلوم، على حساب كليتي التربية والآداب، كما توصلت النتائج لوجود إثر دالّ إحصائيًا للفقرات ذات الأداء التفاضلي في مؤشرات صدق البناء الداخليّ لأداة الدراسة.

التعقيب على الدراسات السابقة:

جميع الدراسات السابقة التي تم تناولها في هذه الدراسة، تبحث تأثير متغيرات، مثل: اللغة، والمجال الأكاديمي: (العلوم الإنسانية والعلوم)، ومتغير الجنس، وتأثير مخالفه قواعد صياغه فقرات اختبار اختيار متعدد، ومدى انتهاك قواعد صياغة فقرات الاختبار المتعدد متغير الكلية الأكاديمية، حيث إنّ جميع هذه المتغيرات تختلف عن الدراسه الحالية، في أنّ الدراسة الحالية تناولت تأثير متغير موقع البديل الصحيح، في الأداء التفاضلي للفقرات، هذا هو الاختلاف بين تلك الدراسات والدراسة الحالية، أي أنّه لا توجد دراسة تناولت تأثير موقع البديل الصحيح في الأداء التفاضلي للفقرات.

الطريقة والإجراءات.

مجتمع الدراسة وعينتها:

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

أداة الدراسة:

لقد مر بناء أداة الدراسة بالمراحل الآتية:

أولًا: تحديد الهدف من الاختبار والمتمثل في قياس التحصيل الدراسي للطلاب في مقرر القياس والتقويم، ومدى إتقانهم لمهارات القياس.

ثانيًا: تحديد محتوى الاختبار وذلك بالرجوع إلى 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. ثمّ صياغة الأهداف السلوكية التي تناولتها الموضوعات، والمتوقع إتقانها من قبل الطلبة.

ثالثًا: صياغة (40) فقرة بصورة أولية من نوع الاختيار من متعدد، أربعة بدائل، أحدها يمثل الإجابة الصحيحة لقياس كلّ هدف من الأهداف السلوكية، وقد تطلب قياس بعضها بأكثر من فقرة، وقد روعيت المبادئ الأساسية التي تحكم صياغة هذا النوع من الفقرات، ومطابقتها للهدف السلوكي الذي تقيسه، من حيث المحتوى والمستوى المعرفي.

رابعًا: إعداد ثلاثة نماذج من الاختبار، جميعها تتكون من المحتوى والفقرات نفسها، الأنموذج الأول موقع البديل الصحيح في النصف الأول (البديل a، أو البديل b)، والأنموذج الثاني محتوى النموذج الأول نفسه، لكن موقع البديل الصحيح في النصف الثاني من البدائل (البديل الصحيح c، أو البديل d)، والأنموذج الثالث يتكون من الفقرات نفسها باستثناء موقع البديل الصحيح، موزع بشكل عشوائي على بدائل الفقرة، وتم تطبيق النماذج الثلاثة من الاختبار على عينة مكونة من (1545) طالبًا وطالبة تخصص الدبلوم العام في التربية، تم تقسيمهم عشوائيًا على ثلاث مجموعات بواقع (515) في كل مجموعة.

صدق الأداة:

للتأكد من صدق المحتوى، عرضت فقرات الاختبار بصورتها الأولية وعددها (40) فقرة على عشرة محكمين مختصين، وطلب من كل منهم أن يبدي رأيه في فقرات الاختبار من حيث: درجة الاتفاق بين الأهداف السلوكية والهدف الأساسي للاختبار وتمثيلها لها، ومدى قياس الفقرات للهدف السلوكي المتعلق بها، ومدى اتساق الاختبار ككل مع الهدف الأساسي منه، وبناء على ملاحظاتهم لم يتم حذف فقرات، وإنّما جرى تعديل على بعض الفقرات، وأجمع المحكمون أن هناك فقرات تؤدي غرضها، وأنّها عينة ممثلة للمجال السلوكي الذي يقيسه.

ثبات أداة الدراسة:

التأكد من ثبات أداة الدراسة (الاختبار)، من خلال استخراج معامل الثبات طبقًا لمعادلة كرونباخ ألفا (Cronbach Alpha) للاتساق الداخلي، إذ بلغ معامل الثبات الكلي للاختبار (0.89)، وتعدّ هذه القيمة مناسبة لدرجة عالية جدًّا لأغراض الدراسة.

جمع البيانات وتحليلها:

لتحقيق الهدف من الدراسة، تم تطبيق نماذج الاختبار عشوائيًّا على أفراد عينة الدراسة، والبالغ عددهم (1545) طالبًا وطالبة، وبعد تطبيق نماذج الاختبار على أفراد عينة الدراسة، تمّ فرز كل أنموذج على حدة، وتصحيحه، ثمّ أدخلت البيانات إلى ذاكرة الحاسوب، واستخدمت البرامج الإحصائية لإجراء التحليلات الإحصائية للإجابة عن أسئلة الدراسة.

المعالجات الإحصائية:

للإجابة عن أسئلة الدراسة تمّ استخدام أساليب الإحصاء الوصفيّ والتحليليّ، وذلك باستخدام الحزمة الإحصائية (SPSS) وعلى النحو الآتي: للإجابة عن السؤالين الأول والثاني تم استخرج المتوسطات والانحرافات المعيارية لاستجابات الأفراد في عينة الدراسة، على كلّ فقرات الاختبار لكل أنموذج من نماذج الاختبار الثلاثة، كذلك تم تقدير قيم معاملات مانتل–هانزل للفقرات في كل أنموذج لفحص الأداء التفاضلي للفقرات. بينما استخدم الباحث للإجابة عن السؤال الثالث تحليل التباين الأحادي (One Way Anova).

النتائج ومناقشتها:

للإجابة عن السؤال الأول: "ما هي فقرات الاختبار المعرفي في القياس والتقويم، والتي تظهر أداء تفاضليًّا لمتغير موقع البديل الصحيح: (موقع البديل الصحيح في الخيارين (a)، (b) مقابل توزيع موقع البديل الصحيح عشوائيًّا)"؟.

تمّ تقدير قيم معاملات مانتل–هانزل للفقرات باستخدام برمجية (SPSS) لكل أنموذج من نماذج الاختبار: (النموذج الأول: الذي فيه توزيع موقع البديل الصحيح في الخيارين (a)، (b)، والأنموذج الثالث: الذي فيه توزيع موقع البديل الصحيح عشوائيًّا)، وتمّ أيضًا حساب المتوسطات الحسابية، والانحرافات المعيارية لكلّ فقرة من فقرات الأنموذجين. ويوضح الجدول (2) يبين ذلك.

الجدول (2): متوسطات استجابات الطلاب والطالبات على كلّ فقرة من فقرات اختبار الأنموذجين الأول والثاني

رقم الفقرة	أنموذج الاختبار				رقم الفقرة	أنموذج الاختبار			
	الأول: توزيع البديل الصحيح على a،b		الثالث: توزيع البديل الصحيح عشوائيًّا			الأول: توزيع البديل الصحيح على a،b		الثالث: توزيع البديل الصحيح عشوائيًّا	
	المتوسط	الانحراف المعياري	المتوسط	الانحراف المعياري		المتوسط	الانحراف المعياري	المتوسط	الانحراف المعياري
1	0.57	0.49	0.52	0.50	21	0.55	0.50	0.33	0.47
2	0.14	0.35	0.23	0.42	22	0.68	0.47	0.54	0.50
3	0.44	0.50	0.57	0.50	23	0.32	0.47	0.36	0.48
4	0.78	0.42	0.76	0.43	24	0.55	0.50	0.57	0.50
5	0.66	0.47	0.64	0.48	25	0.61	0.49	0.55	0.50
6	0.42	0.49	0.50	0.50	26	0.49	0.50	0.45	0.50
7	0.43	0.49	0.44	0.50	27	0.50	0.50	0.47	0.50
8	0.82	0.39	0.79	0.41	28	0.12	0.33	0.16	0.37
9	0.21	0.41	0.44	0.50	29	0.32	0.47	0.38	0.49
10	0.49	0.50	0.45	0.50	30	0.37	0.48	0.46	0.50
11	0.64	0.48	0.67	0.47	31	0.54	0.50	0.52	0.50
12	0.78	0.42	0.77	0.42	32	0.19	0.39	0.24	0.43
13	0.63	0.48	0.75	0.43	33	0.42	0.49	0.33	0.47
14	0.83	0.38	0.50	0.50	34	0.43	0.50	0.43	0.50
15	0.47	0.50	0.56	0.50	35	0.14	0.34	0.15	0.36
16	0.81	0.39	0.80	0.40	36	0.18	0.38	0.30	0.46
17	0.46	0.50	0.46	0.50	37	0.16	0.36	0.24	0.43
18	0.10	0.31	0.15	0.36	38	0.32	0.47	0.35	0.48
19	0.55	0.50	0.47	0.50	39	0.13	0.34	0.23	0.42
20	0.67	0.47	0.62	0.49	40	0.42	0.49	0.48	0.50

وللكشف عن الأداء التفاضلي لفقرات الاختبار تم استخدام طريقة مانتل–هانزل، والجدول (3) يبين ذلك.

الجدول (3): اختبار مانتل–هانزل للكشف عن الأداء التفاضلي لفقرات اختبار اختيار من متعدد

رقم الفقرة	اختبار مانتل–هانزل لفحص الأداء التفاضلي لفقرات اختبار أنموذجي الاختبار: الأول: توزيع البديل الصحيح على **b،a** الثالث **:**توزيع البديل الصحيح عشوائيا			رقم الفقرة	اختبار مانتل –هانزل لفحص الأداء التفاضلي لفقرات اختبار أنموذجي الاختبار: الأول: توزيع البديل الصحيح على **b،a** الثالث **:**توزيع البديل الصحيح عشوائيا		
	Chi-Squared Mantel-Haenszel	**الدلالة**	**الأداء التفاضلي لصالح**		**Chi-Squared Mantel-Haenszel**	**الدلالة**	**الأداء التفاضلي لصالح**
1	2.343	.126	لا يوجد	21	49.393	.000	الأنموذج الأول
2	11.909	.001	الأنموذج الثالث	22	20.779	.000	الأنموذج الاول
3	15.992	.000	الأنموذج الثالث	23	2.173	.140	لا يوجد
4	.458	.499	لا يوجد	24	.160	.689	لا يوجد
5	.442	.506	لا يوجد	25	3.236	.072	لا يوجد
6	5.853	.016	الأنموذج الثالث	26	1.310	.252	لا يوجد
7	.262	.609	لا يوجد	27	.682	.409	لا يوجد
8	.918	.338	لا يوجد	28	3.183	.074	لا يوجد
9	61.711	.000	الأنموذج الثالث	29	3.224	.073	لا يوجد
10	1.459	.227	لا يوجد	30	7.668	.006	الأنموذج الثالث
11	.643	.423	لا يوجد	31	.582	.446	لا يوجد
12	.000	1.000	لا يوجد	32	3.720	.054	لا يوجد
13	16.904	.000	الأنموذج الثالث	33	9.082	.003	الأنموذج الاول
14	116.883	.000	الأنموذج الأول	34	.000	.986	لا يوجد
15	6.777	.009	الأنموذج الثالث	35	.398	.528	لا يوجد
16	.231	.631	لا يوجد	36	20.055	.000	الأنموذج الثالث
17	.004	.949	لا يوجد	37	11.162	.001	الأنموذج الثالث
18	4.740	.029	الأنموذج الثالث	38	.880	.348	لا يوجد
19	6.128	.013	الأنموذج الأول	39	13.850	.000	الأنموذج الثالث
20	2.853	.091	لا يوجد	40	2.543	.111	لا يوجد

نلاحظ من الجدول (3) أنّ الفقرات التي أظهرت أداءً تفاضليًا لصالح أنموذج الاختبار الأول: (توزيع البديل الصحيح على b،a)، هي الفقرات: (14)، (19) ، (21)، (22)، (33)، أي ما نسبته (13%)، من عدد الفقرات الكلي للاختبار، أمّا الفقرات التي أظهرت أداءً تفاضليًا لصالح أنموذج الاختبار الثالث: (توزيع البديل الصحيح عشوائيا)، هي الفقرات: (2)، (3)، (6)، (9)، (13)، (15)، (18)، (30)، (36)، (37)، (39)، أي ما نسبته (28%) من عدد الفقرات الكلي للاختبار.

للإجابة عن السؤال الثاني: "ماهي فقرات الاختبار المعرفي في القياس والتقويم، والتي تظهر أداء تفاضليًا لمتغير موقع البديل الصحيح: (موقع البديل الصحيح في الخيارين (c)، (d)، مقابل توزيع موقع البديل الصحيح عشوائيا)؟".

تمّ تقدير قيم معاملات مانتل–هانزل للفقرات باستخدام برمجية (SPSS)، لكلّ أنموذج من نماذج الاختبار: (الأنموذج الثاني: الذي فيه توزيع موقع البديل الصحيح في الخيارين (c)، (d)، والأنموذج

الثالث: الذي فيه توزيع موقع البديل الصحيح عشوائيًا)، وتمّ أيضًا حساب المتوسطات الحسابية، والانحرافات المعيارية لكلّ فقرة من فقرات الأنموذجين. ويوضح الجدول (4) يبين ذلك.

الجدول (4): متوسطات استجابات الطلاب والطالبات على كل فقرة من فقرات اختبار الأنموذجين الثاني والثالث.

رقم الفقرة	أنموذج الاختبار				رقم الفقرة	أنموذج الاختبار			
	الثاني: توزيع البديل الصحيح على d،c		الثالث: توزيع البديل الصحيح عشوائيا			الثاني: توزيع البديل الصحيح على d،c		الثالث: توزيع البديل الصحيح عشوائيا	
	المتوسط	الانحراف المعياري	المتوسط	الانحراف المعياري		المتوسط	الانحراف المعياري	المتوسط	الانحراف المعياري
1	.4657	.49933	0.52	0.50	21	.1230	.32875	0.33	0.47
2	.3085	.46233	0.23	0.42	22	.4153	.49328	0.54	0.50
3	.6895	.46316	0.57	0.50	23	.4032	.49104	0.36	0.48
4	.7419	.43801	0.76	0.43	24	.5808	.49394	0.57	0.50
5	.6129	.48758	0.64	0.48	25	.4919	.50044	0.55	0.50
6	.5706	.49550	0.50	0.50	26	.4173	.49362	0.45	0.50
7	.4617	.49903	0.44	0.50	27	.4496	.49796	0.47	0.50
8	.7681	.42244	0.79	0.41	28	.1815	.38578	0.16	0.37
9	.6593	.47443	0.44	0.50	29	.4274	.49520	0.38	0.49
10	.4113	.49256	0.45	0.50	30	.5363	.49918	0.46	0.50
11	.6895	.46316	0.67	0.47	31	.4919	.50044	0.52	0.50
12	.7722	.41985	0.77	0.42	32	.2964	.45712	0.24	0.43
13	.8649	.34215	0.75	0.43	33	.2278	.41985	0.33	0.47
14	.2016	.40161	0.50	0.50	34	.4274	.49520	0.43	0.50
15	.6250	.48461	0.56	0.50	35	.1714	.37721	0.15	0.36
16	.7823	.41313	0.80	0.40	36	.4093	.49220	0.30	0.46
17	.4657	.49933	0.46	0.50	37	.3246	.46870	0.24	0.43
18	.1956	.39704	0.15	0.36	38	.3851	.48711	0.35	0.48
19	.3770	.48513	0.47	0.50	39	.3125	.46398	0.23	0.42
20	.5676	.49591	0.62	0.49	40	.5222	.50001	0.48	0.50

وللكشف عن الأداء التفاضلي لفقرات الاختبار، تمّ استخدام طريقة مانتل–هانزل، والجدول (5) يبين ذلك.

الجدول (5): اختبار مانتل–هانزل للكشف عن الأداء التفاضلي لفقرات اختبار اختيار من متعدد

رقم الفقرة	اختبار مانتل–هانزل لفحص الأداء التفاضلي لفقرات اختبار اختبار لنموذجي الاختبار: الثاني: توزيع البديل الصحيح على d،c توزيع البديل الصحيح عشوائي :الثالث			رقم الفقرة	اختبار مانتل–هانزل لفحص الأداء التفاضلي لفقرات اختبار لنموذجي الاختبار: الثاني: توزيع البديل الصحيح على d،c توزيع البديل الصحيح عشوائيا :الثالث		
	Chi-Squared Mantel-Haenszel	الدلالة	الأداء التفاضلي لصالح		Chi-Squared Mantel-Haenszel	الدلالة	الأداء التفاضلي لصالح
1	3.158	.076	لا يوجد	21	59.744	.000	لا يوجد
2	8.230	.004	الأنموذج الثاني	22	15.018	.000	الأنموذج الثالث
3	14.053	.000	الأنموذج الثاني	23	1.538	.215	لا يوجد
4	.263	.608	لا يوجد	24	.093	.760	لا يوجد

5	.520	.471	لا يوجد	25	2.941	.086	لا يوجد
6	4.680	.031	الأنموذج الثاني	26	1.049	.306	لا يوجد
7	.260	.610	لا يوجد	27	.260	.610	لا يوجد
8	.585	.444	لا يوجد	28	.490	.484	لا يوجد
9	45.744	.000	الأنموذج الثاني	29	2.214	.137	لا يوجد
10	1.482	.223	لا يوجد	30	5.515	.019	الأنموذج الثاني
11	.375	.541	لا يوجد	31	.487	.485	لا يوجد
12	.000	1.000	لا يوجد	32	3.461	.063	لا يوجد
13	20.313	.000	الأنموذج الثاني	33	12.040	.001	الأنموذج الثالث
14	95.553	.000	الأنموذج الثالث	34	.002	.963	لا يوجد
15	4.536	.033	لا يوجد	35	.474	.491	لا يوجد
16	.298	.585	لا يوجد	36	12.844	.000	الأنموذج الثاني
17	.016	.899	لا يوجد	37	7.936	.005	الأنموذج الثاني
18	2.798	.094	لا يوجد	38	.847	.357	لا يوجد
19	7.991	.005	لا يوجد	39	9.031	.003	النموذج الثاني
20	2.358	.125	لا يوجد	40	1.950	.163	لا يوجد

نلاحظ من الجدول (5) أنّ الفقرات التي أظهرت أداءً تفاضليًا لصالح أنموذج الاختبار الثاني: (توزيع البديل الصحيح على d،c)، هي الفقرات: (2)، (3)، (6)، (9)، (13)، (30)، (36)، (37)، (39)، أي ما نسبته (23%)، من عدد الفقرات الكلي للاختبار، أمّا الفقرات التي أظهرت أداءً تفاضليًا لصالح أنموذج الاختبار الثالث: (توزيع البديل الصحيح عشوائيًا) هي الفقرات: (14)، (22)، (33)، أي ما نسبته (8%) من عدد الفقرات الكلي للاختبار.

للإجابة عن السؤال الثالث: "ما أثر موقع البديل الصحيح في اختبار اختيار من متعدد، على تقديرات القدرة لدى الأفراد وفق النظرية الكلاسيكية في القياس؟

تمّ استخراج الإحصاءات الوصفية الخاصة بدرجات الطلبة على النماذج الثلاثة للاختبار، فكانت كما في الجدول (6).

الجدول (6): الإحصاءات الوصفية الخاصة بدرجات الطلبة على النماذج الثلاثة للاختبار

أنموذج الاختبار	N	Mean	Std. Deviation
موقع البديل الصحيح (a) أو (b)	496	18.69	4.22
موقع البديل الصحيح في (c) أو (d)	496	18.81	5.41
موقع البديل الصحيح موزع عشوائيًا	496	18.77	4.89
الكلي	1488	18.76	4.86

وتم استخدام تحليل التباين الأحادي (One Way Anova)؛ للكشف عن تأثيرات موقع البديل الصحيح في اختبار اختيار من متعدد، على تقديرات القدرة لدى الأفراد، وفق النظرية الكلاسيكية في القياس باستخدام برمجية (SPSS).

لقد أشارت نتائج التحليل الإحصائي، كما هو مبين في الجدول (7) تحليل التباين الأحادي (ANOVA) الآتي:

الجدول (7): تحليل التباين الاحادي (ANOVA) للكشف عن أثر موقع البديل الصحيح على القدرة باستخدام النظرية الكلاسيكية في القياس

Source of variance	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	3.500	2	1.750	.074	.929
Within Groups	35150.946	1485	23.671		
Total	35154.446	1487			

نلاحظ من الجدول (7) عدم وجود فروق دالة إحصائيًا، بين المتوسطات الحسابية لتقديرات القدرة للأفراد، تُعزى لموقع البديل الصحيح.

التوصيات:

في ضوء نتائج هذه الدراسة يوصي الباحث بالآتي:

– عند بناء اختبار اختيار من متعدد، من الأفضل أن يكون توزيع البديل الصحيح عشوائيًا على مختلف البدائل الأربعة.

- إجراء المزيد من الدراسات حول الكشف عن الأداء التفاضلي في ضوء متغيرات أخرى، وبطرق أخرى، منها الطريقة التي تعتمد على نظرية خصائص الفقرة: (النظرية الحديثة في القياس) (Item Response Theory)، أو المساحات بين المنحنيات.

12. Pae, T. (2004). DIF for examinees with different academic background. Language Testing, 21 (1). 53-72.
13. Rover, C. (2005). That is not fair Fairness, bias, and differential item functioning in language testing. Retrieved on 23 Feb. 2017.
14. Sarnacki, R (1979) an examination of test domain, Review of Educational Research, 49(2)252-279.
15. Schumacker, R.E. (2010). Test Bias and Differential Item Functioning. Applied Measurement Associates LLC.
16. Weiten, W. (1984). Violation of selected item construction principles in Educational measurement. Journal of Experimental Education, 52 (3), PP. 78 – 174.

References:

1. Al-Batoush, A., & Al-Quraan, M. (2018). Examining the differential functioning of items in student's assessment tools for the quality of higher education in Jordan according to the academic college using the general Mantel-Hanzel method. Journal of Al-Quds Open University for Educational and Psychological Research and Studies- Palestine, Volume (8). Issue (23), pp. 158-173.
2. Hamadna, E. (2007). The differential functioning of the items of tests to determine the language proficiency in the English language in Jordanian universities: a comparative study. Unpublished PhD thesis, Yarmouk University, Irbid, Jordan.
3. Al-Kinani, A. & Faydi, F. (2012). The effectiveness of using the site of the effective alternative when removing the wrong alternative in the difficulty of the items of multiple-choice questions and distinguishing them. Al-Qadisiyah Journal of Physical Education Sciences, 12 (1), pp. 207-225.
4. Al-Maliki, Y. (2015). The degree of commitment of middle school teachers to the rules of building objective questions, an unpublished master's thesis, Umm Al-Qura University, Saudi Arabia.
5. Mubarak, W. (2010). Differential Functioning of science test items in the international study PISA 2006. Unpublished PhD thesis, Yarmouk University, Irbid, Jordan.
6. Sulaiman, S. (2008). The Effect of Violating Multiple-Choice Items Writing Rules in on psychometric properties of test and its items. Paper presented at the fifth annual Meeting for psychology department. May 4-6, 2008, Tanta university Egypt.
7. Zaza, H. (2012). Exposing the extent of violation of the rules for drafting multiple-choice items in the questions of the general secondary school certificate exam in Jordan, The Jordanian Journal of Educational Sciences, 8 (1), pp. 81-94.
8. Camilli, G., & Shepard, A. (1994). Methods for identifying biased test items. Thousand Oaks, CA: SAGE Publications.
9. Ellis, B. & Raju, S. (2003). Test and item bias: What are not and How to detect them. Measuring up. Assessment issues for teachers, counselors and administrator. (ERIC Document Reproduction Service No. ED480042.
10. Kamata, A. & Vaughn, B. (2004). An Introduction to item functioning analysis. A Contemporary Journal, 2(2), 49-62.
11. Nanda, R. (1993). Simultaneous DIF amplification and Cancellation: Shealy Stout's test for DIF. Journal of Educational Measurement, 1973-90.