كلية التربية

كلية معتمدة من الهيئة القومية لضمان جودة التعليم

إدارة: البحوث والنشر العلمي (المجلة العلمية)

=======

أثر بنية اختبار اختيار من متعدد في الحاسب الآلي على معالم الفقرة وفق نظرية استجابة الفقرة

إعـداد

أ / محمد علي جابر المساوى

جامعة تبوك

﴿ المجلد الرابع والثلاثون– العدد الثاني – فبراير ٢٠١٨م ﴾

http://www.aun.edu.eg/faculty_education/arabic

الملخص:

هدف البحث الحالي إلى الكشف عن أثر بنية اختبار اختيار من متعدد على معالم الفقرة ودقة تقديرها وفق نظرية الاستجابة للفقرة، ولتحقيق أهداف البحث تم تصميم اختبار تحصيلي اختيار من متعدد أربعة بدائل في مقرر الحاسب الالي ،حيث تكونت الأداة بصورتها الأولية من (٤٠) فقرة، وتم إدخال أربع مخالفات على فقرات اختبار الاختيار من متعدد ، هي : طول البديل الصحيح ، موقع البديل الصحيح ، جميع ما ذكر صحيح ، كلاهما صحيحان ، بحيث أصبح الاختبار يتكون من نموذجين الأول يتكون من فقرات سليمة البنية ، والثاني يتكون من نفس محتوى فقرات النموذج الأول لكن أدخل عليها المخالفات الأربع السابقة ، ، ثم طبق النموذجين على عينة عشوائية تكونت من (٦٢٤) طالبا.

أظهرت نتائج البحث ما يأتي الفقرات المتحررة من المخالفات التالية: "موقع البديل الصحيح ، جميع ما ذكر صحيح كلاهما صحيحان " أكثر دقة في تقدير معلمة الصعوبة من الفقرات المخالفة . كما دلت على وجود فروق دالة إحصائياً بين متوسط معلمة التمييز للفقرات التي تتضمن المخالفة ،وأن دقة تقدير الفقرات المخالفة لمعلمة التمييز لا تختلف عن دقة تقدير الفقرات المتحررة من المخالفات. كما دلت النتائج على وجود فروق دالة إحصائياً بين المتوسط الحسابي لمعلمة تخمين الفقرات التي تتضمن المخالفات ،والمتوسط الحسابي لمعلمة تخمين للفقرات المتحررة من المخالفات وكانت الفرق لصالح الفقرات المتحررة من المخالفات.

الكلمات المفتاحية : بنية الاختبار ، اختبار اختيار من متعدد ، معالم الفقرات ، نظرية استجابة الفقرة .

Abstract

The main objective of the research is to explore the effect of a multiple choice test structure on the parameters of item and its accuracy and on the coefficient of Reliability according to Item Response theory, To achieve the objectives of the research, a four-choice multiple choice test was designed in the computer science course for the third secondary class (science and art). The questionnaire containing 40 item, was Entering four irregularities on the multiple choice test items: length of the correct alternative, position of the correct variant, all that is right , both right , so that the test consists of two forms. The first two items of the first model were intact. The two samples were based on a random sample consisting of (624) students.

The results of the research showed the following The items that are free of the following irregularities: "The location of the correct alternative, all that is true is both true" is more accurate in estimating the difficulty parameter of the violating paragraphs.There are statistically significant differences between the mean of the discrimination parameter for the items that contain the phrase "both are correct" and the average parameter of discrimination for the items that are free of the violation in favor of the liberated items. There are statistically significant differences between the means parameters of the guessing parameter for the items that include the following irregularities: the length of the correct variant, the correct position of the alternative, both true, and the mean of the guessing parameter for the items that were free of irregularities. There are statistically significant differences between the averages of the standard errors means of estimators the estimation parameter for the items that contain the following two violations: the correct alternative location, both correct, and the standard errors mean to estimate the guess parameter for vertebrae free items in favor of vertebrae.

المقدمة :

تعد فقرات الاختيار من متعدد من أكثر أنواع الاسئلة الاختيارية انتشاراً في الجامعات والمدارس ومن أكثرها مرونة ،ومما زاد من انتشارها وشيوعها كفاءتها وتعدد استعمالاتها حيث يمكن بواسطتها قياس أهداف عقلية متنوعة ، والتي يمكن تقويمها بالاختبارات المقالية باستثناء تلك الأهداف التي تتطلب مهارة في التعبير الكتابي وأصالة التفكير فهي لا تقتصر على استدعاء المعلومة بل يمكن بواسطتها قياس الفهم والاستنتاج والتطبيق والتحليل والتركيب وغيرها من النشاطات العقلية ، كما أنها تعتبر أقل تأثراً من غيرها من الفقرات الاختيارية بالتخمين والتهيؤ للاستجابة (Response-set) وتتكون فقرة الاختيار من متعدد من جزأين أو مكونين هما : المتن (stem) وهي التي تطرح المشكلة في الفقرة . أما الجزء أو المكون الآخر وهي قائمة البدائل التي يطلب فيها قراءة المتن والبدائل ثم يختار البديل الصحيح، وتسمى البدائل الخاطئة بالمموهات أو المشتتات ويكون المتن في الفقرة على شكل سؤال أو جملة غير تامة .(Aiken,2002)

كما أنه يمكن ذكر بعض الاستراتيجيات المتعلقة بحكمة الاختبار، هناك استراتيجيات تتعلق بإرشادات قواعد وبناء الاختبار مثل : الاستفادة من خبرة الطالب في تمييز البديل الصحيح كميل مصمم الاختبار لجعل البديل الصحيح هو الأطول من بين البدائل ، بالإضافة الى ميله الى وضع الفقرات الخاطئة أكثر من الصحيحة ، وميله في أغلب الأحيان الى وضع البديل الصحيح في موقع متميز بين البدائل ، بالاضافة الى ميله لوضع بعض الكلمات المحددة للإجابة مثل دائما– فقط – كل أو جميع ما ذكر) ، بالإضافة الى وجود بعض الكلمات الأخرى التي تتكرر في المتن أو البدائل(عوده،١٩٨٨).

والسؤال الذي يطرح نفسه هل أن هذه الاختبارات التي يتعرض الطالب لها يتم تصميمها بدقة وموضوعية تامة لتكون النتائج دقيقة في تحديد مستويات القدرة؟ وبالتالي تكون درجة القرار المعتمد على هذه النتائج أكبر دقة ، مع أن الكثير من كتب القياس والتقويم التربوي تحتوي على مجموعة كبيرة من النصائح الجيدة ، بالإضافة إلى تقديمها الشرح الوافي والكامل وذلك في مساعدة التربويين والمعلمين والأكاديميين في إعداد وتكوين وتحضير مجموعة من الاختبارات الجيدة والفعالة لطلابهم، لكن في الكثير من الأحيان يكون هناك قصوراً في هذه الاختبارات التي يعدها التربويون والمدرسون ، على الرغم من أن المبادئ العامة والقواعد الأساسية تكون متوفرة وموثقة في الكثير من كتب القياس بشكل جيد ، على الرغم أن كتابة الفقرات تعتبر فنا وعلما في آن واحد (Downing,2005).

أجرى داوننغ (Downing , 2005) دراسة بعنوان : أثر مخالفة قواعد كتابة الفقرات على الاختبار والطلاب، بهدف دراسة أثر مخالفة قواعد كتابة الفقرات على خصائص الاختبار و علامات الطلاب ، من خلال تطبيق أربعة اختبارات في العلوم الأساسية على طلاب السنة الأولى والثانية في كلية الطب ، فدلت النتائج على أن الفقرات المخالفة للقواعد كانت أصعب من الفقرات المعيارية غير المخالفة في الاختبارات الأربعة ، بينما كان تأثير الفقرات المخالفة للقواعد على صدق وثبات الاختبار ضئيلاً جداً .

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 اختباراً تحصيلياً 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 اختباراً مكوناً من (٤٠) فقرة متحررة من المخالفات ولها نفس المحتوى ، طبق الاختبار على عينة عشوائية مكونة من (٤٨٠) طالبا من طلاب جامعة تبوك وقد دلت النتائج أن الفقرات المخالفة لقواعد الصياغة ، أسهل من نظيرتها الفقرات المتحررة من المخالفات وبفارق دال إحصائياً ، 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 فقراته متحررة من المخالفات والفروق لصالح الاختبار الذي فقراته متحررة من المخالفات.

وفي دارسة الصبح (٢٠٠٨) ، قامت الباحثة بفحص أثر مخالفة قواعد بناء الفقرات على دقة تقدير معالم الفقرات والأفراد ودالة المعلومات من خلال بناء اختبار مكون من صورتين (معيارية ، مخالفة) في كل صورة منها (٣٥) فقرة ، حيث احتوت الصورة المخالفة فقرات تتضمن المخالفات التالية :

١. تضمين البديل "لا شيء مما ذكر" كإجابة صحيحة .
٢. صياغة المتن بصورة النفي .
٣. موقع البديل المميز (البديل الأول).
٤. البديل غير مقبول ظاهريا (غير المنطقي).
٥. وجود معلومات إضافية في المتن لا ضرورة لها .

دلت النتائج على أن دقة تقدير معالم الفقرات والأفراد كانت أعلى للفقرات المعيارية بشكل عام منها للفقرات المخالفة لقواعد بناء الفقرات ، أما بالنسبة لدالة المعلومات الاختبار فقد أظهرت النتائج أن الفقرات المعيارية قدمت معلومات أكثر فعالية للمفحوصين ذوا القدرات المتدنية والمتوسطة ، بينما قدم النموذج المخالف معلومات اكثر فاعلية للمفحوصين من ذوي القدرات العليا . أما بالنسبة لقيمة دالة المعلومات فقد كانت أعلى في النموذج المخالف عند جميع مستويات القدرة .

أما دراسة تيرنت و وير (Tarrant & Ware , 2008) فهدفت الى فحص أثر الفقرات المخالفة لقواعد البناء على الاختبار من نوع الاختيار من متعدد على خصائص الأفراد ، من خلال تجميع عشرة اختبارات في التمريض وتصنيف فقراتها الى معيبة أو معيارية ، حيث تعتبر الفقرة معيبة إذا خالفت إحدى قواعد بناء الاختبار أو أكثر ، وتعتبر الفقرة معيارية إذا التزمت بجميع قواعد في كتابة الفقرات . فكانت نسبة الفقرات المعيبة في الاختبارات جميعها ما نسبته (٤٧%) . ودلت النتائج على أن الفقرات المعيبة أثرت سلبا على الطلبة ذوي العلامات العليا ، ولم تؤثر على الطلبة ذوي العلامات المتدنية من الناجحين.

وبينت نتائج دراسة الشريفين و طعامنة (٢٠٠٩) والتي هدفت في الكشف عن أثر عدد البدائل في اختبار الاختيار من متعدد في دقة التقديرات لقدرة الأفراد ومعالم الصعوبة للفقرات بأنه لا يوجد فروق دالة إحصائيا بين متوسط الأخطاء المعيارية لصعوبة الفقرات تعزى لعدد البدائل للفقرة ، وكانت هناك فروق ذات دلالة احصائيا بين متوسطات الأخطاء المعيارية في تقديرات معلمة القدرة للأفراد تعزى لعدد البدائل حيث جاءت دقة النتائج لصالح الاختبار ذو الثلاثة بدائل .

ففي دراسة قام بها (المصري، ٢٠٠٩) هدفت للمقارنة بين أثر شكلين من أشكال الفقرات (الاختيار من متعدد والتكميل) على كل من تقدير قدرات المفحوصين والخطأ المعياري في تقديرها والخصائص السيكومترية للفقرات، وقد تم بناء اختبار في الثقافة الحاسوبية لطلاب الصف الأول الثانوي في مديرية التربية والتعليم لمنطقة اربد الثانية والبالغ عددهم (٢١٥٩) طالباً وطالبة، وقد أظهرت النتائج عدم وجود فروق ذات دلالة إحصائية عند مستوى الدلالة(٠.٠٥=α) في تقدير قدرات الأفراد تعزى إلى شكل فقرات الاختبار، وكذلك وجود فروق دالة إحصائيا عند نفس مستوى الدلالة في الآتي: (في الخطأ المعياري لتقدير قدرات الأفراد تعزى

لشكل فقرات الاختبار ولصالح فقرات الاختيار من متعدد ، وفي الوسط الحسابي لدالة معلومات الاختبار تعزى لشكل فقرات الاختبار ولصالح فقرات التكميل، و في تقدير الصعوبة والتمييز للفقرات تعزى إلى شكل الفقرات ولصالح فقرات التكميل، وفي الخطأ المعياري لتقدير صعوبة وتمييز الفقرات تعزى لشكل الفقرات ولصالح فقرات التكميل.

وأجرى الخصاونة (٢٠١٠) دراسة هدفت إلى فحص أثر الوزن النسبي للفقرات المخالفة لقواعد بناء الفقرات على الخصائص السيكومترية، حيث قامت الباحثة بتحكيم المخالفات على مقياس ثير ستون لإعطاء أوزان تعبر عن الأهمية النسبية لهذه المخالفات ، ثم قامت ببناء اختبار تحصيلي مكون من خمسة نماذج ، أربعة منها تحتوي على فقرات مخالفة بنسبة (٥٠%) ونموذج واحد سليم واستخدمت الباحثة المخالفات التالية في بحثها :

١. استخدام منبهات تدل على الإجابة الصحيحة .
٢. استخدام البديلين "جميع ما ذكر" و "لا شيء مما ذكر".
٣. بناء متن السؤال على صيغة أكمل الفراغ .
٤. أن يكون البديل الصحيح هو الأطول .

أظهرت النتائج أثراً سلبياً لمخالفة قواعد بناء الاختبارات يزداد بزيادة الأهمية النسبية للمخالفات ، على كل من الخصائص السيكومترية للاختبار وفقراته، من تمييز وصعوبة ومعاملات الثبات ومستوى الاداء على الاختبار .

وأجرى ظاظا(٢٠١٠) دراسة هدفت إلى الكشف عن مدى مخالفة قواعد صياغة فقرات الاختيار من متعدد في أسئلة امتحان شهادة الدراسة الثانوية العامة في الأردن في ضوء مراجعة (٢٩١٠) فقرات للأعوام (٢٠٠٧-٢٠١٠) باستخدام (١٨) قاعدة للكشف عن مخالفة الفقرة لقواعد صياغتها وأظهرت النتائج أن(٤١٦) فقرة أي ما نسبته ١٢.٦ % تضمنت أكثر من مخالفة ، وبشكل إجمالي وجدت (٧٨٢) فقرة أي ما نسبته ٢٦.٩% فقرة مخالفة ، وتبين أن أكثر قاعدتين انتهاكا كانتا: ترتيب البدائل عموديا بنسبة ٦٧.٧% وترتيب البدائل منطقياً أو عددياً أو أبجدياً 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 الكناني وفيضي (٢٠١٢) د راسة بعنوان فاعلية استخدام موقع البديل الصحيح عند ابعاد البديل الخاطئ في صعوبة فقرات لأسئلة الاختيار من متعدد وتمييزها. وهدفت 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 ، واشتملت عينة الدراسة على (٦٠) طالباً من طلبة المرحلة الثالثة لكلية التربية الرياضية في جامعة المثنى في العراق ، حيث تم تقسيم عينة الدراسة إلى مجموعتين متساويتين حيث أجاب (١٣٠) طالباً وطالبة عن أسئلة النموذج الأول، كما أجاب(١٣٠) طالباً عن أسئلة النموذج الثاني . دلت النتائج على عدم وجود فروق معنوية بين نموذجي الاختبار من حيث الصعوبة والتمييز .

كما أجرى يعقوب وأبو فودة (٢٠١٢) دراسة هدفت إلى معرفة أثر مخالفة قواعد صياغة فقرات الاختيار من متعدد في الخصائص السيكومترية لمفردات(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 حيث 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 على المخالفات في قواعد الصياغة أو التي احتوتها، وبينت النتائج أيضا أن الفرق بين قيم معاملي ثبات الاتساق الداخلي في نموذجي الاختبار كان دالاً إحصائياً ، وكان الفرق بين قيم معاملي صدقها دالاً أيضاً.

وفي دراسة ليامان الوارد في بني عطا والرباعي (٢٠١٣) ، والتي هدفت الى تحديد العدد الأمثل في البدائل في اختبار الاختيار من متعدد، وذلك بهدف مقارنة الخصائص السيكومترية لاختبارات الاختيار من متعدد ، حيث أشارت النتائج إلى أن مستوى ثبات فقرات الاختبار المكون من ثلاثة وخمسة بدائل أعلى منها في الفقرات المكونة من أربعة بدائل ، كما أشارت النتائج أيضا إلى عدم وجود فروق ذات دلالة إحصائية بين الاختبارات الثلاثة في الخصائص السيكومترية للفقرات الخاصة بمعلمة الصعوبة والتمييز لكل منها ، وقد اوصت الدراسة بأن فقرات الاختبار المكون من ثلاثة بدائل هي الأفضل ، ويعود السبب في ذلك إلى أن هذا النوع من الاختبارات يتم تطويره وتحليله بطريقه أسرع وأسهل .

يظهر مما سبق أهمية قواعد بناء فقرات الاختبار من الاختيار من متعدد والتي ظهر عدم تطبيقها أو مخالفتها في معظم الدراسات يؤثر على الخصائص السيكومترية الفقرات والاختبار حيث دلت بعض الدراسات على أن فقرات اختبار الاختيار من متعدد في النموذج صحيح البنية تتميز بمعامل ثبات عال بمقارنتها مع فقرات اختبار الاختيار المتعدد في النموذج لآخر النموذج المخالف للبنية.

كما أشارت هذه الدراسات الى وجود فروق ذات دلالة احصائية في تقدير الخطأ المعياري والوسط الحسابي وفي تقدير معاملات الصعوبة والتمييز والتخمين .

مشكلة البحث وأسئلته :

يهتم الكثير من علماء القياس والتقويم بقواعد كتابة فقرات الاختبارات بشكل عام ، وكتابة فقرات اختبار الاختيار من متعدد بشكل خاص ، لكنها تفتقر الى الكشف أو إيجاد البديل العلمي على صحة هذه القواعد (Haladyna,&Downing,1989).

كما وأن مخالفة هذه القواعد أصبح شائعا جدا (Tarrant & Ware , 2008) ومن هنا تبرز الحاجة لإجراء المزيد من الابحاث والدراسات لتسلط الضوء على هذه القواعد وبيان أهميتها ، وكذلك جاء البحث كمحاولة علمية لتقديم دليلاً علمياً تجريبيا ً على صحتها .

كما أن هذا البحث يحاول تسليط الضوء على بيان أهمية قواعد بناء فقرات اختبار الاختيار من متعدد وأثر مخالفتها على تقدير قدرات المفحوصين (Ability) وتقدير معالم الفقرات (Item Parameters) وخصائص الاختبار المتمثلة بدالة معلومات الاختبار (Test Information) باستخدام النموذج الثلاثي المعلم في نظرية استجابة الفقرة ويحاول البحث الاجابة عن الأسئلة التالية:

السؤال الأول : هل توجد فروق دالة إحصائياً بين المتوسطات الحسابية لمعلمة صعوبة الفقرة تعزى لاختلاف بنية الاختبار (سليم ، مخالف)؟

السؤال الثاني : هل توجد فروق دالة إحصائياً بين المتوسطات الحسابية للأخطاء المعيارية في دقة تقدير معلم صعوبة الفقرة تعزى لاختلاف بنية الاختبار (سليم ، مخالف)؟

السؤال الثالث : هل توجد فروق دالة إحصائياً بين المتوسطات الحسابية لمعلمة تمييز الفقرة تعزى لاختلاف بنية الاختبار (سليم ، مخالف)؟

السؤال الرابع : هل توجد فروق دالة إحصائياً بين المتوسطات الحسابية للأخطاء المعيارية في دقة تقدير معلم تمييز الفقرة تعزى لاختلاف بنية الاختبار (سليم ، مخالف)؟

السؤال الخامس : هل توجد فروق دالة إحصائياً بين المتوسطات الحسابية لمعلم تخمين الفقرة تعزى لاختلاف بنية الاختبار (سليم ، مخالف)؟

السؤال السادس : هل توجد فروق دالة إحصائياً بين المتوسطات الحسابية للأخطاء المعيارية في دقة تقدير معلم تخمين الفقرة تعزى لاختلاف بنية الاختبار (سليم ، مخالف)؟

أهمية البحث :

تكمن أهمية البحث الحالي في محاولتها لإثبات أهمية قواعد بناء اختبار الاختيار من متعدد ، حيث يقوم البحث الحالي بتوضيح الأضرار المتعلقة بمخالفة هذه القواعد على تقدير معالم الفقرات والاختبار ، وكذلك فيما إذا كان هناك أهمية لبعض القواعد عن غيرها .

كما و يعتبر البحث الحالي إضافة علمية لما سبق من أبحاث تربوية في هذا المجال حيث إن البعض من العلماء قام بدراسة مخالفة واحدة أو اثنتين ، أما البحث الحالي فقد استخدم أربع مخالفات في بناء وصياغة اختبار الاختيار من متعدد وهي :

١. طول البديل الصحيح.
٢. موقع البديل الصحيح.
٣. بديل مركب :جميع ما ذكر .
٤. بديل مركب كلاهما بديلان صحيحان.

كما أن ما يميز هذا البحث عن غيرها هو استخدامها للنظرية الحديثة في القياس وهو النموذج اللوجستي الثلاثي المعلم ، وهناك ندرة في الأبحاث التي استخدمت نماذج النظرية الحديثة في استبعاد أو استقصاء أهمية هذه القواعد حيث إن معظمها استخدمت النظرية التقليدية في هذا المجال.

أهداف البحث :

يهدف البحث الحالي إلى :

١. معرفة أثر مخالفة القواعد التالية لصياغة اسئلة الاختيار من متعدد وهي (طول البديل – موقع البديل – جميع ما ذكر – كلاهما بديلان صحيحان).
٢. معرفة أثر مخالفة القواعد السابقة على ثبات الاختبار وعلى الخصائص السيكومترية للفقرة (الصعوبة ، التمييز، التخمين)وفق النموذج اللوجستي ثلاثي المعلم لنظرية استجابة الفقرة.

مصطلحات البحث :

- **النموذج اللوجستي ثلاثي المعلم:** هو أحد نماذج نظرية الاستجابة للفقرة ثنائية التدريج ، حيث يُمكن لهذا النموذج تقدير ثلاثة معالم للفقرة وهي : الصعوبة، التمييز، التخمين، وهو النموذج الذي يتحدد به الأداء على الفقرة بناء على جميع معالم الفقرة.(Baker,2001).
- **معالم الفقرة:** وهي معالم الصعوبة والتمييز والتخمين المنبثقة عن النموذج اللوجستي ثلاثي المعلمة (Baker,2001).
- **معلمة الصعوبة:** هي نقطة على متصل القدرة تقابل احتمال (1+c/٢) للإجابة عن الفقرة (أ) إجابة صحيحة حيث يشير الرمز (c)الى معلمة التخمين (Baker,2001).
- **معلمة التمييز:** هي ميل منحنى خصائص الفقرة الذي يحدث عنده تغيير في اتجاه المنحنى (inflex point) الذي يقابل الصعوبة على متصل القدرة(Baker,2001) .
- **معلمة التخمين:** هي احتمال إجابة فرد ذو قدرة متدنية على فقرة ما إجابة صحيحة(Baker,2001).
- **معلمة القدرة (φ):**وهي وحدة قياس السمات الكامنة عند الأفراد وتقع على مقياس متصل السمة(Baker,2001).
- **دقة التقدير:** هو تعبير يشير الى جودة الفقرات التي يميزها الاحتمالية الكبيرة في أن التقدير قريب من القيمة الحقيقة باستخدام الأخطاء المعيارية في التقدير (Baker,2001).
- **اختبار الاختيار من متعدد :** هي مجموعة من الفقرات التي تتكون من جزأين هما المتن : والذي يتضمن معلومة وأربعة بدائل واحد منها صحيح.

- **فقرات اختيار من متعدد سليمة البنية أو الصياغة** : وهي الفقرات التي أعدها الباحث مراعيا في اعدادها قواعد وارشادات بناء فقرات الاختيار من متعدد.
- **فقرات اختيار من متعدد مخالفة البنية أو الصياغة** : وهي الفقرات التي أعدها الباحث بهدف إظهار المخالفات في قواعد البناء أو الصياغة ، والمخالفات هي :طول البديل الصحيح ، موقع البديل الصحيح ، جميع ما ذكر ، كلاهما صحيحان.

حدود البحث :

- يقتصر البحث على الطلبة الذكور في مدينة تبوك الذين يدرسون في النظام الفصلي للمستوى الثالث الفصل الأول للعام الدراسي ١٤٣٧ /١٤٣٨ هـ .
- يقتصر البحث على الوحدتين الأولى(الشبكات الالكترونية) والثانية (أمن البيانات والمعلومات) من منهج الحاسب الالي للمستوى الثالث للنظام الفصلي للمستوى الثالث للصف الثاني الثانوي.
- يقتصر البحث على إعداد نموذجين من اختبار اختيار من متعدد أحدهما صحيح البنية والأخر مخالف للقواعد التالية في صياغة فقرات الاختبار :
  ١. طول البديل الصحيح.
  ٢. موقع البديل الصحيح(النصف الأول إما(أ) او (ب) فقط)
  ٣. بديل مركب :جميع ما ذكر .
  ٤. بديل مركب كلاهما بديلان صحيحان.

ومن هنا تكمن أهمية البحث الحالي لمحاولته ليضيف شيئا جديدا في هذا المجال، كما بينت هذه الدراسات أن هناك قلة في الدراسات التي تناولت أثر قواعد وارشادات بناء فقرات الاختيار من متعدد في ضوء النظرية الحديثة بالإضافة إلى كونها تنفرد بمقارنة أهمية بعض القواعد من غيرها من حيث خطورة مخالفتها على معالم الفقرات وفق نظرية استجابة الفقرة

الطريقة والاجراءات

مجتمع البحث:

تكون مجتمع البحث من الطلاب الذكور الذين يدرسون في النظام الفصلي للمستوى الثالث للعام ١٤٣٧-١٤٣٨هـ ، وذلك من خلال الطلاب المنتظمين الذين يدرسون في المدارس الحكومية التابعة للإدارة تعليم مدينة تبوك ويقدر عددهم (١٧٥٦) بواقع (١٧) مدرسة .

عينة البحث :

تم اختيار عينة البحث بالطريقة العشوائية العنقودية حيث تكونت من (٦٢٤) طالبا موزعين على (١٧) مدرسة كما في

جدول(١) توزيع افراد عينة الدراسة

م	اسم المدرسة	العدد	النسبة المئوية من العينة الكلية
١	ثانوية الابناء الثانية	٥٥	%٨.٨
٢	ثانوية الملك عبدالله	٧٠	%١١.٢
٣	ثانوية الملك فيصل	٦٥	%١٠.٤
٤	ثانوية عبدالرحمن بن عوف	٣٩	%٦.٣
٥	ثانوية الامام تركي بن عبدالله	٤٣	%٦.٩
٦	ثانوية اسد بن الفرات	١٠٠	%١٦
٧	ثانوية الامام الشاطبي	٢٠	%٣.٢
٨	ثانوية الجزيرة	١٣٠	%٢٠.٩
٩	ثانوية اسامة بن زيد	١٠٢	%١٦.٣
	المجموع	٦٢٤	% ١٠٠

بناء أداة البحث :

تم بناء اختبار تحصيلي من نوع الاختيار من متعدد أربعة بدائل في مقرر الحاسب وتقنية المعلومات للصف الثاني ثانوي (مسار علمي ، مسار أدبي) للمستوى الثالث للنظام الفصلي في الوحدتين التالية : الوحدة الأولى (الشبكات الالكترونية) والوحدة الثانية (أمن المعلومات)، وقد استرشد في بناء الاختبار بالأسس العامة المتبعة في بناء الاختبارات التحصيلية التي أوردها (Gronlund&Linn,1995) كما يأتي :

١. تحديد الغرض من الاختبار وهو قياس تحصيل طلبة الصف الثاني الثانوي في مقرر الحاسب وتقنية المعلومات في وحدتي الشبكات الالكترونية وأمن المعلومات .
٢. تحليل المحتوى لوحدات مقرر الحاسب وتقنية المعلومات إلى مفاهيم ومصطلحات ورموز ومهارات وتطبيقات ومبادئ ، ثم صياغة الأهداف موضوع الدراسة ، ثم قام الباحث بإعداد لائحة مواصفات الاختبار تم فيه ربط مستويات الاهداف بمحتوى المادة الدراسية موضوع الاختبار كما في الملحق (١).

٣. قام الباحث بصياغة فقرات الاختبار من نوع الاختيار من متعدد لكل فقرة منها أربعة بدائل وقد بلغ عدد الفقرات بصورتها الأولية (٤٠) فقرة لقياس كل هدف من الأهداف . وقد تطلب قياس بعضها أكثر من فقرة واحدة ، وقد روعيت فيها الأسس الفنية في كتابة هذا النوع من الفقرات ومطابقتها للهدف الذي يقيسه من حيث المحتوى ومستوى الهدف ، كما روعي في كتابتها ألا تعتمد الإجابة على إحدى الفقرات على الإجابة على فقرات أخرى ، وهذا ما تتطلبه نظرية استجابة الفقرة (افتراض الاستقلال الموضعي).

صدق اداة البحث :

تم التأكد من الصدق الظاهري للأداة من خلال عرضه على مجموعة من المحكمين بلغ عددهم (١٥) محكماً منهم ثلاثة أعضاء من هيئة التدريس بالجامعة و مشرفين تربويين وعشرة معلمين حيث طلب من الكل أن يبدي رأيه في درجة الاتفاق بين الأهداف السلوكية والهدف الأساسي للاختبار ومدى تمثيلها له ، ومدى قياس الفقرات للهدف السلوكي المتعلق بها ، ومدى اتساق الاختبار ككل مع الهدف الأساسي، وبناء على ملاحظاتهم لم يتم حذف أي فقرة ، وإنما جرى تعديل بعض الفقرات ، وفي النهاية أجمع المحكمون على أن الفقرات تؤدي الغرض من الاختبار، و تمثل المجال السلوكي الذي يتم قياسه.

التجريب الأولي للأداة البحث :

طُبقت أداة البحث بصورتها الأولية على عينة استطلاعية مكونة من (٤٢) طالبا من طلاب مدرسة(ثانوية محمد الفاتح) موزعين على ثلاث شعب تم اختيارها بشكل عشوائي من شعب الصف الثاني الثانوي في المدرسة وذلك بتحديد الزمن الكافي للأداء على الاختبار، والتأكد من وضوح الصياغة اللغوية للفقرات والكشف عن الفقرات الغامضة ، وتم إدخال البيانات على برنامج التحليل الاحصائي(SPSS 22) ثم حساب معامل الصعوبة لكل فقرة من خلال إيجاد نسبة الطلبة الذين أجابوا عن الفقرة إجابة صحيحة الى الطلبة الذين حاولوا الإجابة عن الفقرة ، كما تم حساب معامل التمييز للفقرات والتي تمثل معاملات ارتباط بوينت بايسيريل Biserial Correlation بين درجة الفقرة والدرجة الكلية على الاختبار كما في الجدول (٢) .

جدول (٢)

معاملات الصعوبة والتمييز لفقرات الاختبار للعينة الاستطلاعية وفق النظرية الكلاسيكية

الفقرة	معامل الصعوبة	معامل التمييز	الفقرة	معامل الصعوبة	معامل التمييز	الفقرة	معامل الصعوبة	معامل التمييز	الفقرة	معامل الصعوبة	معامل التمييز
١	٠.٧٤	٠.٤٨	٩	٠.٣٦	٠.٣٣	١٧	٠.٧١	٠.٦٠	٢٥	٠.٧٤	٠.٣٣
٢	٠.٥٧	٠.٣٨	١٠	٠.٩٠	٠.٣١	١٨	٠.٧٤	٠.٨١	٢٦	٠.٧٦	٠.٣٠
٣	٠.٦٧	٠.٦٠	١١	٠.٦٤	٠.٣٠	١٩	٠.٨٦	٠.٤٢	٢٧	٠.٥٧	٠.٣٣
٤	٠.٨٨	٠.٢٩	١٢	٠.٦٠	٥	٢٠	٠.٧٩	٠.٥٧	٢٨	٠.٧١	٠.٦٧
٥	٠.٧٦	٠.٣٣	١٣	٠.٧٤	٠.٠٦	٢١	٠.٦٤	٠.٣٥	٢٩	٠.٩٥	٠.٣٥
٦	٠.٧١	٠.٤٤	١٤	٠.٨١	٠.٣٩	٢٢	٠.٤٥	٠.٢٩	٣٠	٠.٧٤	٠.٦١
٧	٠.٣٨	-٠.٠٣	١٥	٠.٧٦	٠.٥٤	٢٣	٠.٧٤	٠.٥٥	٣١	٠.٦٨	٠.٤٧
٨	٠.٧٤	٠.٤٨	١٦	٠.٨١	٠.٤٢	٢٤	٠.٧٤	٠.٢٨	٣٢	٠.٧١	٠.٢٩

معامل الصعوبة		معامل التمييز	
أعلى قيمة	٠.٩٥	أعلى قيمة	٠.٨٣
أقل قيمة	٠.٢٦	أقل قيمة	-٠.٠٣
المتوسط	٠.٧١	المتوسط	٠.٤٢

يتضح من الجدول (٢) أن معاملات الصعوبة للفقرات تراوحت قيمتها بين (٠.٩٥) و (٠.٢٦) بمتوسط حسابي (٠.٧١)، أما معاملات التمييز فقد تراوحت قيمتها بين (٠.٨٣) و (-٠.٠٣) بمتوسط حسابي مقداره (٠.٤٢) وتم حذف الفقرتين (٣ ، ٧) ولذلك لتدني معاملات تمييزهما فقد

كانت معاملات التمييز لهاتين الفقرتين على النحو التالي : ٠.٠٦ ، ٠.٠٢٩ ، أما معاملات الصعوبة للفقرات فكانت جميعها مقبولة باستثناء الفقرات التالية (٤، ١٠ ، ١٤ ، ١٦ ، ١٩ ، ٢٩ ٣١) وبعد حذف الفقرات أصبح عدد الفقرات (٣١) فقرة ، هذا وقد تم الاحتفاظ بالفقرات على المعايير الإحصائية التالية التي أوردها ايبل (Ebel,1972) :

- الفقرات التي معامل صعوبتها يتراوح بين (٠.٢٠- ٠.٨٠) يمكن أن تكون مقبولة وينصح بالاحتفاظ بها .
- الفقرة التي معامل تمييزها يتراوح بين ٠.٢٠ – ٠.٣٩ ذات تمييز مقبول وينصح بتحسينها.
- الفقرة التي معامل تمييزها أكبر من ٠.٣٩ ذات تمييز جيدة يمكن الاحتفاظ بها.

كما حُسب معامل الثبات للاختبار بطريقة التجزئة النصفية وطريقة ألفا لكرونباخ والنتائج كما في جدول (٣).

جدول (٣) معامل الثبات للأداة البحث

معاملات الثبات	القيمة
معامل الثبات الفا كرونباخ	٠.٧٨
معامل الثبات التجزئة النصفية	٠.٨٥

يتضح من الجدول (٣) أن معاملات الثبات لأداة البحث باستخدام معامل الثبات الفا كرونباخ بلغت قيمته (٠.٨٧) وباستخدام طريقة التجزئة النصفية بلغت قيمته (٠.٨٥) وهي قيم جيدة أو مقبولة لغايات البحث العلمي.

أداة البحث :

تم إدخال أربع مخالفات أو إرشادات لبناء فقرا ت الاختيار من متعدد على الاختبار الذي أعده الباحث ، والمخالفات هي : طول البديل الصحيح أي البديل الأطول هو البديل الصحيح ، موقع البديل الصحيح أي أن موقع البديل الصحيح هو البديل الأول أو الثاني ، البديل المركب جميع ما ذكر، البديل المركب كلاهما صحيحان، بحيث أصبح الاختبار يتكون من نموذجين : الأول يتكون من ٣١ فقرة سليمة البنية و النموذج الثاني يتكون من ٣١ فقرة لها نفس محتوى فقرات النموذج الأول لكن أدخل عليها المخالفات الأربع لقواعد بناء فقرات الاختيار من متعدد وبذلك تكون فقرات النموذج الثاني التي أدخلت عليها المخالفات اختبار غير سليمة البنية والفقرات التي لم يدخل عليه المخالفات اختبار سليم البنية .

متغيرات البحث :

١. **المتغير المستقل بنية الاختبار وتتكون من مستويين** : سليمة البنية أو الصياغة ، مخالفة الصياغة.
٢. **المتغيرات التابعة** : معالم الفقرة (الصعوبة ، التمييز ، التخمين) ، معامل ثبات الاختبار.

جمع البيانات (التطبيق النهائي) :

لتحقيق الهدف من البحث قام الباحث بتطبيق نموذجي الاختبار السليم البنية والمخالف البنية على عينة الدراسة التي اشتملت على (٦٢٤) طالبا ، وبعد إجابة أفراد عينة الدراسة على النموذجين، أُدخلت البيانات إلى جهاز الحاسب الالي ، واستخدمت البرامج الاحصائية الآتية : برنامج (Spss22) وبرنامج (Bilog–Mg3) لإجراء التحليلات الإحصائية.

المعالجات الاحصائية :

من أجل الإجابة على أسئلة البحث ، فقد تم استخدام المعالجات الاحصائية التالية :

- معامل ارتباط بوينت بايسيريل.
- معامل الصعوبة والتمييز والتخمين لكل فقرة في الاختبار السليم والمخالف للصياغة .
- اختبار (T- Test) لعينتين مستقلتين لمعرفة دلالة الفروق بين المتوسطات الحسابية لمعالم الفقرات : الصعوبة والتمييز والتخمين باختلاف بنية الاختبار .
- مربع كاي (x^2).

النتائج وتفسيرها

أولا :النتائج المتعلقة بالتحقق من افتراض النموذج اللوجستي ثلاثي المعلم

١. أحادية البعد:

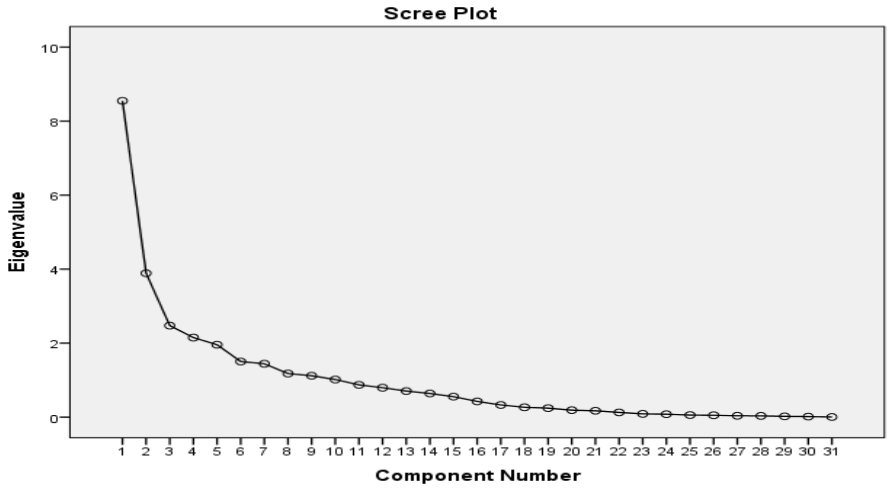
تم التحقق من افتراضات النموذج اللوجستي ثلاثي المعلمة وهي افتراض أحادية البعد ، لما له أثر على دقة التقديرات وكان ذلك باستخدام نموذج التحليل العاملي الاستكشافي وذلك باستخدام طريقة المكونات الرئيسية (Principle Component Analysis)، و يظهر الجدول (٤) قيمة الجذر الكامن (Eigen Values) ونسبة التباين المفسر (Explained Variance) للعامل الأول والثاني وناتج قسمة الجذر الأول على العامل الثاني في النموذج الصحيح، كما في جدول (٤) .

جدول (٤)

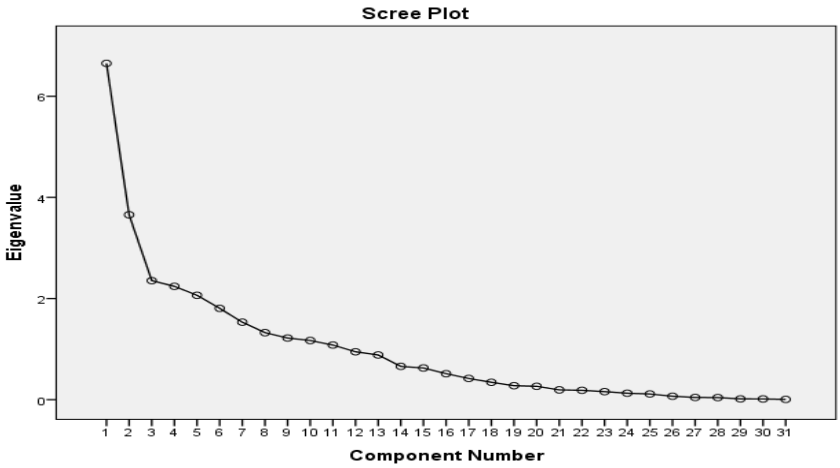
قيم الجذر الكامن ونسب التباين المفسر للعامل الأول والثاني وناتج قسمة الجذر الأول على العامل الثاني

نموذج الاختبار		العامل الأول	العامل الثاني	الجذر الكامن للعامل الأول / الجذر الكامن للعامل الثاني
النموذج الأول (النموذج الصحيح)	الجذر الكامن	٨.٥٥	٣.٨٨	٢.٢٠
	التباين المفسر	%٢٧.٥٧	%١٢.٥٤	
النموذج الثاني (النموذج المخالف)	الجذر الكامن	٦.٤٢	٢.٩٥	٢.١٧
	التباين المفسر	%٢١.٤٥	%١١.٧٩	

يتضح من الجدول (٤) ان في كل من نموذجي الاختبار السليم البنية والمخالف كانت نسبة التباين المفسر من العامل الأول أعلى من ٢٠% ، وكما أن ناتج قسمة العامل الأول على العامل الثاني كانت أكبر من (٢) مما يدل على وجود سمة سائدة للاختبار ، وهذا يشير الى تحقق افتراض أحادية البعد للاختبار(Hatti,1984) ، كما يمكن الإشارة إلى ذلك من خلال مخطط سكري Scree Plot لكل من الاختبارين كما في الشكلين (١ ، ٢) .



شكل (١) مخطط سكري لنموذج الاختبار السليم البنية



شكل (٢) مخطط سكري لنموذج الاختبار المخالف البنية

٢. الاستقلال الموضعي Local Independence:

ويعتبر هذا الافتراض مكافئ لافتراض أحادية البعد حيث أشار هامبلتون وسوا فثبان (Hambleton and Swaminathan ,1985) أنه إذا تحقق افتراض أحادي البعد يعني تحقق افتراض الاستقلال الموضعي ، لأن أحادية البعد والاستقلال الموضعي مفهومان متكافئان .

٣. التحرر من السرعة في الأداء Speediness :

تأكد الباحث أن إخفاق الأفراد في الإجابة عن فقرات الاختبار يعود إلى انخفاض قدراتهم، وليس إلى تأثير عامل السرعة في الإجابة عن طريق إعطاء الوقت الكافي لهم للإجابة عن فقرات الاختبار حيث كان زمن تطبيق الاختبار (٤٥) دقيقة كاملة ، بالإضافة إلى أنه لم يشتكي أي طالب في أثناء التطبيق من ضيق الوقت، وعدم كفايته (Allen & Yen ,1989, p97).

ثانيا: مطابقة الأفراد والفقرات للنموذج اللوجستي ثلاثي المعلم

١.مطابقة الأفراد للنموذج اللوجستي ثلاثي المعلم :

وذلك بهدف التحقق من مطابقة الأفراد للنموذج الثلاثي المعلمي من خلال برنامج (Bilog MG3) تم إجراء اختبار التوزيع الاحتمالي الهامشي حيث كانت قيمة الدلالة الإحصائية لاختبار التوزيع الهامشي أقل من (٠.٠١) لجميع الطلاب في نموذج الاختبار السليم البنية والمخالف للبنية التي تناولها البحث الحالي مما يعني مطابقة جميع الأفراد للنموذج وعدم استبعاد أي من المفحوصين في نموذجي الاختبار .

٢.مطابقة الفقرات للنموذج اللوجستي ثلاثي المعلم :

بهدف التحقق من افتراضات نظرية استجابة الفقرة ، تم مطابقة فقرات الاختبار السليم البنية للنموذج ثلاثي المعلم باستخدام برنامج (Bilog Mg3) باستخدام مربع كأي على أن يكون مستوى الدلالة (٠.٠١) والنتائج كما في الجدول (٥) .

جدول (٥)

نتائج مطابقة فقرات الاختبار السليم البنية للنموذج اللوجستي ثلاثي المعلمة

رقم الفقرة	قيمة مربع كأي	مستوى الدلالة	رقم الفقرة	قيمة مربع كأي	مستوى الدلالة	رقم الفقرة	قيمة مربع كأي	مستوى الدلالة
١	٧,١	٠.٤١	١٢	٥.١	٠.٥١	٢٣	٤,٦	٠.٧٠
٢	٦,٨	٠.٤٠	١٣	٦,٥	٠.٤٨	٢٤	٢,٣	٠.٩٤
٣	٢,٥	٠.٩٢	١٤	٨,٧	٠.٢٧	٢٥	٣,٦	٠.٨٢
٤	١٠,٤	٠.١٦	١٥	٤,٣	٠.٧٣	٢٦	٤,٣	٠.٧٣
٥	٣,٧	٠.٨١	١٦	٤,٣	٠.٧٤	٢٧	٣,٩	٠.٧٩
٦	٣,٢	٠.٨٦	١٧	٢,٦	٠.٩١	٢٨	٧,١	٠.٤٢
٧	٦,٥	٠.٣٦	١٨	٩,١	٠.١٦	٢٩	٣,٥	٠.٨٣
٨	٦,٦	٠.٤٧	١٩	١٢,٢	٠.٠٥	٣٠	٦,٤	٠.٤٨
٩	٨,٩	٠.٢٥	٢٠	٠,٥٠	٠.٩٩	٣١	١,٤	٠.٩٦
١٠	٣,٣	٠.٨٥	٢١	٣,٩	٠.٧٨			
١١	٧,٤	٠.٣٨	٢٢	١,٣	٠.٩٨			

يتضح من الجدول (٥) مطابقة فقرات الاختبار السليم البنية للنموذج ثلاثي المعلمة (3PLM) ، حيث كانت قيمة دلالة كأي تربيع لجميع الفقرات أكثر من (٠.٠١)، لذا تم الابقاء على جميع الفقرات .

كما تم مطابقة الفقرات في الاختبار المخالف البنية للنموذج اللوجستي ثلاثي المعلم باستخدام مربع كأي على أن يكون مستوى الدلالة (٠.٠١) والنتائج كما في الجدول (٦)

جدول (٦)

نتائج مطابقة الفقرات المخالفة للبنية للنموذج اللوجستي ثلاثي المعلمة

رقم الفقرة	قيمة مربع كأي	مستوى الدلالة	رقم الفقرة	قيمة مربع كأي	مستوى الدلالة	رقم الفقرة	قيمة مربع كأي	مستوى الدلالة
١	٨.٤	٠.١٣	١٢	٦.٣	٠.٣٧	٢٣	١١.٢	٠.٠٤
٢	٣.١	٠.٧٩	١٣	٨.٣	٠.١٤	٢٤	٣.١	٠.٧٩
٣	٣.٩	٠.٦٧	١٤	١١.١	٠.٠٥	٢٥	٣.٣	٠.٧٧
٤	٩.٦	٠.٠٧	١٥	١٠.٦	٠.٠٦	٢٦	٣.٩	٠.٦٦
٥	٨.١	٠.١٥	١٦	٤.٧	٠.٥٩	٢٧	٥.٦	٠.٢٨
٦	٩.١	٠.١٠	١٧	٨.٢	٠.١٢	٢٨	٩.٥	٠.٠٨
٧	٦.٦	٠.٢٦	١٨	٩.١	٠.١٠	٢٩	٥.٢	٠.٣٧
٨	٦.٧	٠.٢٤	١٩	٨.٣	٠.١٢	٣٠	٥.٠	٠.٣٩
٩	٨.٢	٠.١٦	٢٠	٣.٨	٠.٦٩	٣١	٥.٩	٠.٣١
١٠	٨.٩	٠.١٤	٢١	٦.٤	٠.٣٦			
١١	٩.١	٠.١٠	٢٢	٦.٥	٠.٣٦			

يتضح من الجدول (٦) مطابقة فقرات الاختبار المخالف البنية للنموذج ثلاثي المعلمة 3PLMحيث كانت قيمة الدلالة الاحصائية إلى قيمة مربع كأي لجميع الفقرات أكثر من (٠.٠١)، لذا تم الابقاء على جميع الفقرات .

ثالثا : النتائج المتعلقة بالإجابة عن أسئلة البحث:

السؤال الأول : هل توجد فروق دالة إحصائياً بين المتوسطات الحسابية لمعلم صعوبة الفقرة تعزى لاختلاف بنية الاختبار (سليم ، مخالف)؟

للإجابة عن هذا السؤال حسبت المتوسطات الحسابية والانحرافات المعيارية لمعلمة الصعوبة لأنماط بنى الفقرات ولنموذجي الاختبار السليم والمخالف البنية واستخدم اختبار (T-test) لعينتين مستقلتين للكشف عن الفروق بين المتوسطات الحسابية لمعلمة صعوبة الفقرات لكل نمط من أنماط بنى الفقرات ولنموذجي الاختبار والنتائج كما في الجدول (٧) .

جدول (٧)

اختبار (T) للكشف عن الدلالة الاحصائية بين متوسطات معلمة الصعوبة لنموذجي الاختبار السليم البنية والمخالف

نوع المخالفة	عدد الفقرات		المتوسط الحسابي لمعلمة الصعوبة		قيمةT	الدلالة الاحصائية
	السليمة	المخالفة	السليمة	المخالفة		
طول البديل الصحيح	٧	٧	١.١٧	١.٥٦	٠.٢٧٠	٠.٧٩٢
موقع البديل الصحيح	٦	٦	١.٢٨	١.٦٣	-٠.٨٦٧	٠.٤٠٦
جميع ما ذكر صحيح	٩	٩	١.٢٩	١.٦٣	-٠.٥١٤	٠.٦١٤
كلاهما صحيحان	٩	٩	١.٦٠	٠.٤٤	*٢.٦٩٩	٠.٠١٨

* القيمة دالة إحصائيا على مستوى ٠.٠٥ ، ** القيمة دالة إحصائيا على مستوى ٠.٠١

يتضح من الجدول (٧) عدم وجود فروق دالة إحصائية عند مستوى ($\alpha \leq$ ٠.٠٥) بين المتوسطات الحسابية لمعاملات الصعوبة لأنماط المخالفات التالية :طول البديل الصحيح ، موقع البديل الصحيح ، جميع ما ذكر صحيح ، والمتوسطات الحسابية لمعاملات صعوبة الفقرات المتحررة من المخالفات . ووجود فروق ذات دلالة احصائية عند مستوى الدلالة ($\alpha \leq$ ٠.٠٥) بين المتوسط الحسابي لمعلمة الصعوبة للفقرات المخالفة البنية " كلاهما صحيحان " و المتوسط الحسابي لمعلمة الصعوبة للفقرات المتحررة من المخالفة وكان الفرق لصالح الفقرات السليم البنية أي المتحررة من المخالفة ، كما اتفقت هذه النتائج مع دراسة سليمان(٢٠٠٨) ودراسة الكيلاني وفيضي (٢٠١٢). كما تعارضت هذه النتائج مع دراسة داوننغ (Downing,2005) و الصبح (٢٠٠٨).

السؤال الثاني : هل توجد فروق دالة إحصائياً بين المتوسطات الحسابية للأخطاء المعيارية لدقة تقدير معلم صعوبة الفقرة تعزى لاختلاف بنية الاختبار (سليم ، مخالف)؟

للإجابة عن هذا السؤال حسبت المتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة الصعوبة لكل نمط من أنماط البنى (المخالفات) وللكشف عن الدلالة الاحصائية للفروق بين المتوسطات الحسابية للأخطاء المعيارية بين الفقرات السليمة والفقرات المخالفة . استخدم اختبار (T- test) لعينتين مستقلتين والنتائج كما في الجدول (٨) .

جدول (٨)

اختبار (T) للكشف عن دلالة الفروق بين الأخطاء المعيارية في دقة التقدير لمعلمة صعوبة الفقرات السليمة البنية والمخالفة

نوع المخالفة	عدد الفقرات		المتوسط الحسابي للأخطاء المعيارية لتقدير معلمة الصعوبة		قيمةT	الدلالة الاحصائية
	السليمة	المخالفة	السليمة	المخالفة		
طول البديل الصحيح	٧	٧	٠.٥٩٣	٠.٧٣٠	-٠.٥٩٠	٠.٥٦٦
موقع البديل الصحيح	٦	٦	٠.٢٥٨	٠.٦٦٢	-٠.٠٢٤**	٠.٠٠١
جميع ما ذكر	٩	٩	٠.٠٩٧	٠.٦٢٨	-٢.٧٧٣*	٠.٠٢١
كلاهما صحيحان	٩	٩	٠.٣٣٣	٠.٥٥٦	-٢.١٢٣*	٠.٠٥٠

* القيمة دالة إحصائيا على مستوى ٠.٠٥ ، ** القيمة دالة إحصائيا على مستوى ٠.٠١

يتضح من الجدول (٨) عدم وجود فروق ذات دلالة احصائية عند مستوى ($\alpha \leq$ ٠.٠٠٥) بين متوسط الخطأ المعياري لدقة تقدير معلمة الصعوبة للفقرة المخالفة (طول البديل الصحيح) ومتوسط الخطأ المعياري لدقة تقدير معلمة الصعوبة للفقرة المتحررة من المخالفة، ووجود فروق ذات دلالة احصائية عند مستوى($\alpha \leq$ ٠.٠٠٥) بين المتوسطات الحسابية للأخطاء المعيارية لدقة تقدير معلمة الصعوبة للفقرات التي تتضمن المخالفات التالية : موقع البديل الصحيح ، جميع ما ذكر صحيح ، كلاهما صحيحان والمتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة الصعوبة للفقرات المتحررة من المخالفات ، أي أن الفقرات المتحررة من المخالفات التالية: موقع البديل الصحيح ، جميع ما ذكر صحيح كلاهما صحيحان أكثر دقة في تقدير معلمة الصعوبة من الفقرات المخالفة ، وهذا يتفق مع دراسة (المصري ، ٢٠٠٩) والتي تشير الى وجود فروق ذات دلالة احصائية لدقة تقدير معلمة الصعوبة (موقع البديل ، جميع ما ذكر ، كلاهما).

السؤال الثالث :هل توجد فروق دالة إحصائياً بين المتوسطات الحسابية لمعلمة تمييز الفقرة تعزى لاختلاف بنية الاختبار (سليم ، مخالف)؟

للإجابة عن هذا السؤال حُسبت المتوسطات الحسابية لمعلمة التمييز لكل نمط من أنماط المخالفات وللكشف عن الدلالة الاحصائية للفروق بين المتوسطات الحسابية لمعاملات التمييز بين الفقرات السليمة والفقرات المخالفة استخدم اختبار (T–test) لعينتين مستقلتين والنتائج كما في الجدول(٩).

جدول (٩)

اختبار (T) للكشف عن الدلالة الاحصائية بين متوسطات معلمة التمييز لنموذجي الاختبار السليم البنية والمخالف

نوع المخالفة	عدد الفقرات		المتوسط الحسابي لمعلمة التمييز		قيمة(T)	الدلالة الاحصائية
	السليمة	المخالفة	السليمة	المخالفة		
طول البديل الصحيح	٧	٧	٠.٧٦٩	١.٢٩٠	١.٠٤٤	٠.٣١٧
موقع البديل الصحيح	٦	٦	١.٧٩٠	١.٢٦٦	١.٩٢٦	٠.٠٨٢
جميع ما ذكر	٩	٩	١.٨١٦	١.٤٠٣	١.٢٥٣	٠.٢٢٨
كلاهما صحيحان	٩	٩	٢.٢١٥	١.٤١٤	*٢.١٩٢	٠.٠٤٤

* القيمة دالة إحصائيا على مستوى ٠.٠٥ ، ** القيمة دالة إحصائيا على مستوى ٠.٠١

يتضح من الجدول (٩) وجود فروق دالة إحصائيا بين متوسط معلمة التمييز للفقرات التي تتضمن المخالفة "كلاهما صحيحان " ومتوسط معلمة التمييز للفقرات المتحررة من المخالفة لصالح الفقرات المتحررة .وعدم وجود فروق دالة إحصائيا بين متوسطات معلمة التمييز للفقرات التي تتضمن المخالفات التالية : طول البديل الصحيح ، وموقع البديل الصحيح ، وجميع ما ذكر صحيح ،وبين متوسطات معلمة التمييز للفقرات المتحررة من المخالفات .

كما تعارضت هذه النتائج مع دراسة كلا من الشريفين وطعامنه(٢٠٠٩ و دراسة سليمان(٢٠٠٨) ودراسة الكيلاني وفيضي (٢٠١٢) حيث أنها لم توجد فروقا تعود لمعلمة تمييز الفقرة.

السؤال الرابع : هل توجد فروق دالة إحصائياً بين المتوسطات الحسابية للأخطاء المعيارية في دقة تقدير معلم تمييز الفقرة تعزى لاختلاف بنية الاختبار (سليم ، مخالف)؟

للإجابة عن هذا السؤال حسبت المتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة التمييز لكل نمط من أنماط البنى (المخالفات) والمتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة التمييز للفقرات المتحررة من المخالفات ، وللكشف عن الدلالة الاحصائية للفروق بين المتوسطات الحسابية للأخطاء المعيارية بين الفقرات السليمة والفقرات المخالفة . استخدم اختبار لعينتين مستقلتين (T– test) والنتائج كما في الجدول(١٠).

جدول (١٠)

اختبار (T) للكشف عن دلالة الفروق بين متوسطات الأخطاء المعيارية في دقة تقدير معلمة التمييز للفقرات السليمة البنية والمخالفة

نوع المخالفة	عدد الفقرات		المتوسط الحسابي للأخطاء المعيارية في تقدير معلمة التمييز		قيمةT	الدلالة الاحصائية
	السليمة	المخالفة	السليمة	المخالفة		
طول البديل الصحيح	٧	٧	٠.٦٧٨	٠.٦٨١	-٠.٠٠٨**	٠.٩٩٣
موقع البديل الصحيح	٦	٦	٠.٧٠٧	٠.٥٧٧	٠.٦٩٣	٠.٥٠٤
جميع ما ذكر	٩	٩	٠.٥٨٧	٠.٥٨٥	٠.٠١٠**	٠.٩٩٣
كلاهما صحيحان	٩	٩	٠.٨٨٦	٠.٨٢٧	٠.٢٤٥	٠.٨١٢

* القيمة دالة إحصائيا على مستوى ٠.٠٥ ، ** القيمة دالة إحصائيا على مستوى ٠.٠١

يتضح من الجدول(١٠) عدم وجود فروق دالة إحصائيا بين متوسطات معلمة تمييز الفقرات التي تتضمن المخالفات التالية : طول البديل الصحيح ، موقع البديل الصحيح ، جميع ما ذكر صحيح ، كلاهما صحيحان ، ومتوسطات معلمة تمييز الفقرات المتحررة من المخالفات .أي أن دقة تقدير الفقرات المخالفة لمعلمة التمييز لا تختلف عن دقة تقدير الفقرات المتحررة من المخالفات، وهذا يتعارض مع دراسة (المصري،٢٠٠٩) التي أشارت الى وجود فروق دالة إحصائيا في دقة تقدير معلم التمييز للفقرات .

السؤال الخامس : هل توجد فروق دالة إحصائياً بين المتوسطات الحسابية لمعلمة تخمين الفقرة تعزى لاختلاف بنية الاختبار (سليم ، مخالف)؟

للإجابة عن هذا السؤال حُسبت المتوسطات الحسابية لمعلمة التخمين لكل نمط من أنماط المخالفات والمتوسطات الحسابية لمعلمة التخمين للفقرات المتحررة من المخالفات ،وللكشف عن الدلالة الاحصائية للفروق بين المتوسطات الحسابية لمعاملات تخمين الفقرات السليمة و ومتوسطات معاملات تخمين الفقرات المخالفة استخدم اختبار (T-test) لعينتين مستقلتين والنتائج كما في الجدول (١١).

جدول(١١)

اختبار (T) للكشف عن الدلالة الاحصائية بين متوسطات معلمة التخمين لنموذجي الاختبار السليم البنية والمخالف

نوع المخالفة	عدد الفقرات		المتوسط الحسابي لمعلمة التخمين		قيمةT	الدلالة الاحصائية
	السليمة	المخالفة	السليمة	المخالفة		
طول البديل الصحيح	٧	٧	٠.٢٢٣	٠.٤٢٥	-٣.٠٦٦ **	٠.٠١٠
موقع البديل الصحيح	٦	٦	٠.٢٢٤	٠.٤٥٨	-٧.٢٣٦**	٠.٠٠٠
جميع ما ذكر	٩	٩	٠.٣٠٨	٠.٣٩٦	-١.٩٧٧	٠.٠٦٦
كلاهما صحيحان	٩	٩	٠.٢٣١	٠.٣٨٨	-٤.٧٥٥**	٠.٠٠٠

* القيمة دالة إحصائيا على مستوى ٠.٠٥ ، ** القيمة دالة إحصائيا على مستوى ٠.٠١

يتضح من الجدول (١١) وجود فروق دالة إحصائيا بين المتوسطات الحسابية لمعلمة التخمين للفقرات التي تتضمن المخالفات التالية : طول البديل الصحيح ، موقع البديل الصحيح ، كلاهما صحيحان ، والمتوسطات الحسابية لمعلمة التخمين للفقرات المتحررة من المخالفات لصالح الفقرات المتحررة من المخالفات ،أي أن متوسطات معلمة التخمين للفقرات السليمة أقل من متوسطات معلمة التخمين للفقرات المخالفة .

ولا توجد فروق دالة إحصائيا بين المتوسطات الحسابية لمعلمة التخمين التي تتضمن مخالفة " جميع ما ذكر صحيح " والمتوسطات الحسابية لمعلمة التخمين للفقرات المتحررة من المخالفات .

ربما يعود ذلك إلى قدرة الطالب على استخدام ما يسمى بحكمة الاختبار في الاجابة عن الفقرات التي تتضمن المخالفات التالية: طول البديل الصحيح ، وموقع البديل الصحيح حيث من الممكن أن يكون ذلك مؤشرا جيدا للوصول الى البديل بدون أي تفكير وجهد، والاستفادة من الخلل أو العيوب الموجودة في صياغة الفقرة واستغلال هذا الخلل أو العيب لمعرفة الإجابة الصحيحة رغم عدم امتلاك الطالب المعرفة اللازمة للإجابة عن الفقرة والتي يطلق عليها المنبهات .

واختلفت هذه النتائج مع دراسة كلا من الشريفين وطعامنه(٢٠٠٩)، ودراسة الكيلاني وفيضي (٢٠١٢).

السؤال السادس : هل توجد فروق دالة إحصائياً بين المتوسطات الحسابية للأخطاء المعيارية في دقة تقدير معلم تخمين الفقرة تعزى لاختلاف بنية الاختبار (سليم ، مخالف)؟

للإجابة عن هذا السؤال حسبت المتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة التخمين لكل نمط من أنماط البنى (المخالفات) والمتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة التخمين للفقرات المتحررة من المخالفات ، وللكشف عن الدلالة الاحصائية للفروق بين المتوسطات الحسابية للأخطاء المعيارية بين الفقرات السليمة والفقرات المخالفة . استخدم اختبار (T– test) لعينتين مستقلتين والنتائج كما في الجدول (١٢)

جدول(١٢)

اختبار (T) للكشف عن دلالة الفروق بين متوسطات الأخطاء المعيارية لدقة تقدير معلمة التخمين للفقرات السليمة البنية والمخالفة

نوع المخالفة	عدد الفقرات		المتوسط الحسابي لمعامل التخمين		قيمةT	الدلالة الاحصائية
	السليمة	المخالفة	السليمة	المخالفة		
طول البديل الصحيح	٧	٧	٠.٠٤٧	٠.٠٦٥	١.٤٦٥-	٠.١٦٩
موقع البديل الصحيح	٦	٦	٠.٠٤٦	٠.٠٦٣	٢.٣٠٣- *	٠.٠٤٤
جميع ما ذكر	٩	٩	٠.٠٥٠	٠.٠٦٧	١.٨٨٦-	٠.٠٨٥
كلاهما صحيحان	٩	٩	٠.٠٣٨	٠.٠٨٢	٦.٧٧٤- **	٠.٠٠٠

* القيمة دالة إحصائيا على مستوى ٠.٠٥ ، ** القيمة دالة إحصائيا على مستوى ٠.٠١

يتضح من الجدول (١٢) وجود فروق دالة إحصائياً بين المتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة التخمين للفقرات التي تتضمن المخالفتين التاليتين : موقع البديل الصحيح ، وكلاهما صحيحان ، والمتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة التخمين للفقرات المتحررة من المخالفات لصالح الفقرات المتحررة من المخالفات ،أي أن الفقرات المتحررة من المخالفتين السابقتين أكثر دقة في تقدير معلمة التخمين من الفقرات التي تتضمن المخالفتين وعدم وجود فروق دالة إحصائيا بين المتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة التخمين للفقرات التي تتضمن المخالفتين التاليتين : طول البديل الصحيح ، جميع ما ذكر صحيح ، والمتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة التخمين للفقرات المتحررة من المخالفتين .

ملخص النتائج :

- عدم وجود فروق دالة إحصائية عند مستوى ($\alpha \leq$ ٠.٠٥) بين المتوسطات الحسابية لمعاملات الصعوبة لأنماط المخالفات التالية :طول البديل الصحيح ، موقع البديل الصحيح ، جميع ما ذكر صحيح ، والمتوسطات الحسابية لمعاملات صعوبة الفقرات المتحررة من المخالفات . ووجود فروق ذات دلالة احصائية عند مستوى الدلالة ($\alpha \leq$ ٠.٠٥) بين المتوسط الحسابي لمعلمة الصعوبة للفقرات المخالفة البنية " كلاهما صحيحان " و المتوسط الحسابي لمعلمة الصعوبة للفقرات المتحررة من المخالفة وكان الفرق لصالح الفقرات السليم البنية أي المتحررة من المخالفة

- عدم وجود فروق ذات دلالة احصائية عند مستوى ($\alpha \leq$ ٠.٠٥) بين متوسط الخطأ المعياري لدقة تقدير معلمة الصعوبة للفقرة المخالفة (طول البديل الصحيح) ومتوسط الخطأ المعياري لدقة تقدير معلمة الصعوبة للفقرة المتحررة من المخالفة، ووجود فروق ذات دلالة احصائية عند مستوى ($\alpha \leq$ ٠.٠٥) بين المتوسطات الحسابية للأخطاء المعيارية لدقة تقدير معلمة الصعوبة للفقرات التي تتضمن المخالفات التالية : موقع البديل الصحيح ، جميع ما ذكر صحيح ، كلاهما صحيحان والمتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة الصعوبة للفقرات المتحررة من المخالفات ، أي أن الفقرات المتحررة من المخالفات التالية: موقع البديل الصحيح ، جميع ما ذكر .

- وجود فروق دالة إحصائيا بين متوسط معلمة التمييز للفقرات التي تتضمن المخالفة "كلاهما صحيحان " ومتوسط معلمة التمييز للفقرات المتحررة من المخالفة لصالح الفقرات المتحررة .وعدم وجود فروق دالة إحصائيا بين متوسطات معلمة التمييز للفقرات التي تتضمن المخالفات التالية : طول البديل الصحيح ، وموقع البديل الصحيح ، وجميع ما ذكر صحيح ،وبين متوسطات معلمة التمييز للفقرات المتحررة من المخالفات .

- دقة تقدير الفقرات المخالفة لمعلمة التمييز لا تختلف عن دقة تقدير الفقرات المتحررة من المخالفات .

– وجود فروق دالة إحصائيا بين المتوسطات الحسابية لمعلمة التخمين للفقرات التي تتضمن المخالفات التالية : طول البديل الصحيح ، موقع البديل الصحيح ، كلاهما صحيحان ، والمتوسطات الحسابية لمعلمة التخمين للفقرات المتحررة من المخالفات لصالح الفقرات المتحررة من المخالفات ،أي أن متوسطات معلمة التخمين للفقرات السليمة أقل من متوسطات معلمة التخمين للفقرات المخالفة .

ولا توجد فروق دالة إحصائيا بين المتوسطات الحسابية لمعلمة التخمين التي تتضمن مخالفة " جميع ما ذكر صحيح " والمتوسطات الحسابية لمعلمة التخمين للفقرات المتحررة من المخالفات .

– وجود فروق دالة إحصائياً بين المتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة التخمين للفقرات التي تتضمن المخالفتين التاليتين : موقع البديل الصحيح ، وكلاهما صحيحان ، والمتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة التخمين للفقرات المتحررة من المخالفات لصالح الفقرات المتحررة من المخالفات ،أي أن الفقرات المتحررة من المخالفتين السابقتين أكثر دقة في تقدير معلمة التخمين من الفقرات التي تتضمن المخالفتين .

وعدم وجود فروق دالة إحصائيا بين المتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة التخمين للفقرات التي تتضمن المخالفتين التاليتين : طول البديل الصحيح ، جميع ما ذكر صحيح، والمتوسطات الحسابية للأخطاء المعيارية لتقدير معلمة التخمين للفقرات المتحررة من المخالفتين. – وجود فروق دالة احصائيا عند ($\alpha \leq$ ٠.٠٥)بين معامل ثبات الاختبار السليم البنية ومعامل ثبات الاختبار المخالف البنية لصالح الاختبار السليم البنية.

التوصيات :

١. يفضل التزام واضعي الاختبارات بالإرشادات والأسس الخاصة بكتابة فقرات الاختيار من متعدد وتحديدا (موقع البديل الصحيح و كلاهما صحيحان) حيث كانت هذه المخالفات تدل على وجود فروق دالة إحصائيا .

٢. ضرورة الوعي بأن الاختبارات السليمة البنية (غير المخالفة) لإرشادات كتابة الفقرات تتمتع بخصائص سيكومترية أكثر دقة من الاختبارات التي تنتهك لمثل هذه القواعد.

المقترحات:

١. تناول أنماطاً أخرى من المخالفات في كتابة فقرات اختبار الاختيار من متعدد في مقررات دراسية أخرى.

٢.تناول أنماط المخالفات في كتابة الفقرات في اختبارات أخرى مثل اختبار الصواب والخطأ واختبار التكميل واختبار الخطأ المتعدد.

المراجع العربية :

بني عطا ، زايد و الرباعي ، ابراهيم (٢٠١٣). **أثر عدد البدائل وتغيير موقع المموه القوي في فقرات اختبار الاختيار من متعدد على معالم الفقرات ، وقدرة الفرد ودالة المعلومات** .المجلة الاردنية في العلوم التربوية . مجلد ٩ ، عدد ٢ ، ١٥١-١٦٦ &٣١٩-٣٣٣.

الخصاونة ، ابتسام (٢٠١٠) .**أثر اختلاف الأوزان النسبية لقواعد انتهاك صياغة فقرات الاختيار من متعدد في الاختبارات على خصائصها السيكومترية**. رسالة دكتوراه غير منشورة . جامعة اليرموك.

الكناني ، عايد و فضي ، فاضل (٢٠١٢). **فاعلية استخدام موقع البديل الفعال عند أبعاد البديل الخاطئ في صعوبة الفقرات لأسئلة اختبار الاختيار من متعدد وتمييزها** ، مجلة القادسية لعلوم التربية الرياضية، ١٢ (١)، ص ٢٠٧-٢٢٥ .

سيلمان ، شاهر (٢٠٠٨) **أثر مخالفة قواعد صياغة فقرات اختبار اختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته** ، المؤتمر السنوي لقسم علم النفس تطبيقات العلوم النفسية ومشكلات المجتمع في الفترة ٤-٦ مايو ، جامعة طنطا، كلية الآداب ، جمهورية مصر العربية .

الشريفين ، نضال و طعامنه، إيمان . (٢٠٠٩). **أثر عدد البدائل في اختبار الاختيار من متعدد على تقديرات القدرة للأفراد والخصائص السيكومترية للفقرات والاختبار وفق نموذج راش في نظرية الاستجابة للفقرة** . المجلة الأردنية في العلوم التربوية ، مجلد٥ ،عدد (٢):٣٠٩ -٣٥٥ .

الصبح رانيا(٢٠٠٨). **أثر بنية فقرات الاختيار من متعدد ومستوى القدرة لدى الأفراد على دقة التقديرات لمعالم الفقرات والأفراد وفق نظرية الاستجابة للفقرة** . رسالة ماجستير غير منشورة . جامعة اليرموك .

عوده، أحمد (١٩٨٨) **أثر تعليم استراتيجيات التخمين واكتساب الطلبة بالمستوى الجامعي لمهارة حكمة الاختبار** ، المجلة التربوية ، ٦(٢) : ١٢٣-١٣٧.

ظاظا، حيدر (٢٠١٢) **الكشف عن مدى انتهاك قواعد صياغة فقرات الاختيار من متعدد في اسئلة امتحان شهادة الثانوية العامة في الأردن** ، المجلة الأردنية في العلوم التربوية ، ٨(١)، ص ٨١-٨٤ .

المصري، أحمد محمود (٢٠٠٩).**أثر شكل فقرات الاختبار في دقة تقدير قدرات المفحوصين ومعالم الفقرات وفق نظرية استجابة الفقرة.** رسالة دكتوراه غير منشورة، جامعة اليرموك، اربد، الأردن.

يعقوب ، ابراهيم و أبو فودة ، باسل (٢٠١٢) .**أثر مخالفة قواعد صياغة فقرات الاختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته** ، مجلة دمشق ٢٨ (١) ص ٤١٩-٤٤٢.

المراجع الأجنبية :

Aiken ,i.(2002) Testing with multi – choice items**, journal of evelopment in Education,** vol30 (8),pp144-157.

Allen, M. & Yen, W. M., (1989): **Introduction to Measurement theory** California, Broks Cole publishing company.

Baker , F.(2011). **The Basic of item Response Theory** . Eric learing house on Assessment and Evaluation .USA.

Downing, S.(2005). **The Effect of Violating Standard Item Writing Principles on Tests and Students :The Consequences of Using Flawed Test Item on Achievement Examination in Medical Education Advances in health Sciences Education** .10: 133- 143.

Ebel ,R(1972) **Essential of Educational Measurement. New Jersey;** prentice-Hall,Inc.

Gronlund, N . & Line , R (1990) . **Measurement and Evaluation in Teaching**(6 th ed.)New york : Macmiuan publishing company , pp 166-189.

Haladyna, T . M , & Downing , S . M . (1989). A taxonomy of multiple-choice item writing rules . **Applied Measurement in Education** , 2 (1), 37-50.

Hatti, J. (1984). An empirical study of various indices for determining unidimensionality. **Multivariate Behavioral Research**, 19,49-78.

Tarrant , M.& Ware, j . , (2008) .Impact of item-writing flaws in multiple-choice question of student achievement in high-stakes nursing assessments. **Medical Education** . 2008 ;42:198-206 .