مجلة جامعة الحسين بن طلال للبحوث ، مجلة علمية محكّمة تصدر عن عمادة البحث العلميّ والدّراسات العليا، ISSN 2519-7436

المجلد (9) العدد (3) لعام 2023 م

أثر استخدام البدائل المركبة في تقدير معالم فقرات ودالة معلومات الاختبار بالاعتماد على النموذج ثلاثي المعلمة في النظرية الحديثة في القياس

د. حابس سعد الزبون

جامعة الحسين بن طلال

الملخص

هدفت الدراسة التعرف إلى أثر استخدام البدائل المركبة في تقدير معالم فقرات ودالة معلومات الاختبار بالاعتماد على النموذج ثلاثي المعلمة، ولتحقيق أهداف الدراسة تم بناء اختبار تحصيليّ تكون بصورته النهائية من (36) فقرة، وبالاعتماد على هذا الاختبار تم اشتقاق صورة أخرى من الاختبار تضمنت بدائل مركبة ضمن بدائل فقرات الاختبار، طبقت صورتي الاختبار على عينة تكونت من (450(معلماً ومعلمة من معلمي محافظة معان ولواء وادي السير اختيروا عشوائياً بالاعتماد على طريقة العينة العشوائية البسيطة.

وبعد التأكد من الخصائص السيكومترية للاختبار والفقرات تم الاعتماد على برنامج (Bilog–MG3) لمطابقة الأفراد والفقرات للنموذج الثلاثي المعلمة وإيجاد دالة المعلومات لفقرات صورتي الاختبار.

توصلت نتائج الدراسة إلى وجود فروق بين المتوسطات الحسابية لمعالم الفقرات تبعاً لنوع البديل (البدائل التقليدية، البدائل المركبة) ودالة معلومات الاختبار ولصالح صورة الاختبار ذي البدائل التقليدية.

الكلمات المفتاحية: *البدائل المركبة، دالة المعلومات، النظرية الحديثة في القياس.*

The effect of the using the composite alternatives in estimating the items parameters and the test information function based on the three-parameter model in item Response Theory

Abstract

The study aimed to identify the impact of the use of composite alternatives in estimating the items parameters of the test items and the test information function based on the three-parameter model under item response theory. To achieve the objectives of the study the researcher built an achievement test. The test is formed of (36) items. Another form of the test was derived included composite alternatives within the alternatives to the test items. The two test forms were applied to a sample of (450) teachers from Ma'an governorate selected randomly. After finding the psychometrics properties of the test، Bilog-MG3 software was used to analyze the data and to match individuals and items to the three-parameter logistics model and to find the information function for the test items for the two forms.

The results of the study showed that there are differences between the mathematical means of items parameters depending on the type of alternative (traditional alternatives، composite alternatives) and the test information function in favor of the test forms with traditional alternatives.

Key words: *Composite Alternatives، Test Information Function، Item Response Theory (IRT).*

المقدمة والإطار النظري:

يُعد الاختبار التحصيليّ أداة من أهم الأدوات التي يُعتمد عليها في تحديد مستويات الطلبة؛ وذلك من أجل معرفة مدى مستواهم التحصيليّ، ومعرفة درجة تحقق الأهداف والنواتج التعليمية، وما يقدمه المعلم من معلومات تساعد في تنمية تحصيل الطلبة؛ لذلك حرص المشرفون التربويّون على أنْ تتمتع هذه الاختبارات بخصائص جيده، وتُمثل الاختبارات التربويّة عنصراً حيوياً في العمليّة التعليمية؛ لما لها من أثر فاعل على العمليّة التربويّة كلها، فعن طريقها تصدر الأحكام القيمية على قدرة العمليّة التربويّة أو فشلها في تحقيق الأهداف المرجوة. ولهذا فقد اهتمت المؤسسات التربويّة اهتماماً خاصاً بالاختبارات التحصيليّة وأعطتها اهتماماً بالغاً، لتصبح الاختبارات التحصيليّة تحتل المكانة الأولى في المدارس والمعاهد والجامعات (الدوسري، 2004).

وتتنوع الاختبارات تبعاً للأهداف التي وضعت من أجلها، فما يصلح لمادة دراسية قد لا يصلح لأخرى، وما يصلح لهدف قد لا يصلح لآخر، وما يصلح لمستوى من الطلاب قد لا يصلح لمستوى آخر، فعلى المدرس أنْ يختار النوع المناسب، ومن أهم هذه الأنواع اختبارات التحصيل (Achievement Tests) التي تهدف لقياس تحصيل الطلبة (الكيلاني وعدس، 1986).

وتعدد أشكال الاختبارات فمنها الموضوعية والإنشائية، وتعدّ اختبارات الاختيار من متعدد أكثر أنواع الأسئلة الموضوعية استخداماً وانتشاراً عند بناء الاختبارات التحصيليّة وبنوك الأسئلة، وذلك لما تتمتع به من مرونة وسهولة تطبيقه على عدد كبير من المفحوصين، وسهولة التصحيح والموضوعية وثبات درجات المفحوصين، وتوفر تغطية جيدة للمادة الدراسية، ويتكون هذا النوع من الاختبارات من مجموعة من البدائل بالإضافة إلى الإجابة الصحيحة، ويعد اختيار نوع البديل وعدد البدائل المناسب من القضايا الهامة التي يجب الاهتمام بها عند بناء مثل هذا النوع من الاختبارات (Sahin&Anil، 2017).

وتتكون أسئلة الاختيار من متعدد من سؤال أو مشكلة رئيسة تشمل على الجذر وعدد من البدائل المطروحة لاستثارة تفكير المفحوص بحيث تشتمل هذه البدائل على الإجابة الصحيحة وثلاثة أو أربعة مشتتات أو بدائل أو أكثر لذا فإنّ تصميم هذه البدائل وصياغتها وعددها يؤثر تأثيراً مباشراً على اختيار الإجابة الصحيحة ومن ثم على معامل تمييز الفقرة وبالتالي على ثبات وصدق الاختبار ككل (الزيات، 1989).

وتعد اختبارات الاختيار من متعدد من أهم أدوات تقويم الطالب وأكثرها شيوعاً واستخداماً، وربما يعود ذلك لما تمتاز به من قدرة على تمثيل المجال السلوكي للسمة موضوع القياس، وسهولة التصحيح، والموضوعية .... غير أن هذه الاختبارات، لا تخلو من أخطاء القياس، التي تؤثر بشكل كبير على عمليّة التقييم والقرارات المترتبة عليها التي تكون أحياناً غير صحيحة نتيجة هذه الأخطاء Ban–Simon،(1997).

وقد اهتمت العديد من الدراسات بتقديم إرشادات لصياغة أسئلة الاختيار من متعدد وتؤكد هذه الإرشادات على أن صياغة بدائل فعالة يعد من الموضوعات الهامة والأساسية لبناء أسئلة ذات خصائص سيكومترية جيده الجودة، أمّا على أرض الواقع وبرغم من وجود هذه الإرشادات، إلّا أن صياغة بدائل حقيقية جيدة يعد مهمة صعبة، لوجود عدد من الاشتراطات التي يجب أنْ يحافظ عليها واضع الاختبار، ومع زيادة عدد الأسئلة تصبح المهمة أكثر صعوبة، فيلجأ البعض لاستخدام البدائل المركبة للحفاظ على عدد ثابت للبدائل في اختبار الاختيار من متعدد (Rich & Johanson،1990).

ويُعد استخدام النظرية الحديثة في القياس (Item Response Theory) في بناء وتحليل البيانات المتعلقة بالاختبارات النفسية والتربويّة مثابة حلّ لتلافي معظم جوانب الضعف والمشاكل الموجودة في النظرية الكلاسيكية في القياس (Classical Test Theory)، ونذكُر منها على وجه التحديد تأثر الخصائص السيكومترية للفقرات بخصائص قدرة عينة المفحوصين الذين أجابوا عليها، وتأثر درجات عينة المفحوصين بالخصائص السيكومترية لعينة الفقرات المكونة للاختبار، و كذلك قصورها في معالجة معادلة درجات الاختبار و تحيّز الفقرات وبنوك الأسئلة، ومن أهم ما جاءت به نظرية استجابة الفقرة خاصية اللاتغير في تقديرات معالم الفقرات المكونة للاختبار بتغير شكل توزيع قدرة المفحوصين (Person Free)، وكذلك اللاتغير في تقديرات معلمة القدرة للمفحوصين بتغير معالم الفقرات التي أجابوا عليها (Item Free)، وهذا النجاح الذي حققته نظرية استجابة الفقرة في معالجة المشاكل الموجودة في النظرية الكلاسيكية، جاء نتيجة بحوث سيكومترية وإحصائية مكثفة لتطوير نظرية استجابة الفقرة، وذلك من أجل توفير الطرق الإحصائية والاحتمالية المختلفة لحل مشكلات النظرية الكلاسيكية في القياس، و كما أنّ نتائج هذه البحوث جعلت من النظرية الحديثة في القياس منهجاً جديداً للقياس النفسي والتربويّ، وكأي نظرية إحصائية أو احتمالية نظرية فإنّ استجابة الفقرة قامت على مجموعة من الافتراضات الواجب تحقيقها في البيانات من أجل الوصول إلى نتائج دقيقة(Hambleton & Swaminathan،1985) وEmbretson & Reise)، 2000) ومن هذه الافتراضات افتراض أحادية البعد

(Unidimensionality) الذي يبين وجود سمة وحيدة تُفسر أداء المفحوصين على الاختبار، بمعنى أنّ درجة المفحوص على الاختبار تعكس السمة التي يقيسها الاختبار فقط، وهناك طرق إحصائية مختلفة تُستخدم لفحص تحقيق البيانات لهذا الافتراض ومن أهمها أسلوب التحليل العاملي، ثانيا افتراض الاستقلال الموضعي (Local Independence) و تحقيق هذا الافتراض يتطلب أن تكون استجابات المفحوصين على فقرات الاختبار مستقلة عن بعضها بعضاً عند مستوى قدرة معين، بمعنى آخر أنه لا تتأثر الإجابة على فقرة ما إيجاباً أو سلباً بإجابة المفحوص على الفقرات الأخرى(Crocker & Algina،1986). وهذا الافتراض يتحقق فقط في حالة أن الاختبار أحادي البعد الذي يعني أن الأفراد من مستوى قدرة معين لا يتأثر أداؤهم على فقرة ما بأدائهم على فقرة أخرىHambleton & Swaminathan،(1985)، أما الافتراض الثالث الذي يسمى بمنحى خصائص الفقرة (ICC) (Item Characteristics Curve) حيث إنّ مفهوم منحنى نظرية استجابة الفقرة عبارة عن اقتران رياضي يربط بين احتمالية نجاح المفحوص على الفقرة و القدرة التي تقيسها مجموعة من الفقرات المكونة للاختبار، و أخيراً افتراض السرعة في الأداء (Speediness)، الذي يفترض أن عامل السرعة لا يؤثر في الإجابة على فقرات الاختبار، أي أن الإجابة الخاطئة على الفقرة سببها القدرة و ليس الوقت المخصص للاختبار (علام،2005).

ونتج عن نظرية استجابة الفقرة عدد من النماذج، عُرفت باسم نماذج السمات الكامنة (Latent Trait Models)، والتي تهدف إلى دراسة العلاقة بين أداء المفحوص على الاختبار والسمة التي تُفسر هذا الأداء، وتعد النماذج اللوجستية ذات المعلمة الواحدة (نموذج راش) والمعلمتين والثلاثة معالم، من أهم النماذج واسعة الانتشار، وتتميز النماذج عن بعضها في عدد معالم الفقرة وتقديرها، ويمثل النموذج ثلاثي المعلمة الصيغة العامة للنماذج اللوجستية؛ لأنّه يحتوي على المعالم الثلاثة الممكنة للفقرة، وهي الصعوبة ، والتمييز ، التخمين (c_i،a_i،b_i) على الترتيب، وتعبر المعادلة الآتية عن النموذج الثلاثي، التي تقيس احتمال إجابة المفحوص ذي القدرة (θ) على الفقرة (i).

$$P_i(\theta) = C_i + \frac{1 - C_i}{1 + e^{-1.7a_i(\theta_i - b_i)}}$$

$P_i(\theta)$: احتمال إجابة مفحوص تم اختياره عشوائياً من مستوى القدرة (θ)إجابة صحيحة على الفقرة. وحيث إنّ:

b_i: معلمة الصعوبة للفقرة (i). عدد فقرات المكونة للاختبار. e: الأساس اللوغاريتمي الطبيعي ويساوي (2.718) θ: قيمة مستوى القدرة. D: عامل التدريج ويساوي (1.7). (c_i) :معلمة التخمين للفقرة (i)

وقد أوضح هامبلتون وسوامينثان (Hambleton & Swaminathan،1985) أن إحدى الطرق التي تمكننا من خلال نماذج السمات الكامنة بناء اختبارات أفضل، بحيث تعطي معلومات أكثر عن السمة الكامنة المقدرة هو استخدام اقتران المعلومات، ويمكن استخدامها أيضاً للتأكد من دقة تقدير معالم الفقرة ومعلمة القدرة، وتأتي أهمية اقتران معلومات الفقرة بأنّ مساهمة كل فقرة باقتران معلومات الاختبار مستقل عن باقي فقرات الاختبار، أمّا في النظرية الكلاسيكية فإنّ مساهمة الفقرات بثبات الاختبار وصدقه لا يتحدد بشكل مستقل عن باقي فقرات الاختبار، ويتناسب اقتران معلومات الاختبار عكسياً مع الخطأ المعياري في التقدير، وبذلك تُعدّ تقديراً لثبات الاختبار؛ لأنّه بزيادة الخطأ المعياري في التقدير يقل الثبات ويزيد الثبات عندما تقل قيمة الخطأ المعياري، واقتران معلومات الاختبار هو ناتج جمع اقترانات المعلومات لجميع الفقرات ويُعبر عنه بالمعادلة الآتية :

$$I(\theta) = \sum I_i(\theta)$$

$\boldsymbol{I_i(\theta)}$: كمية المعلومات التي يوفرها الاختبار عند مستوى القدرة $\boldsymbol{\theta}$.

$\boldsymbol{I(\theta)}$: كمية المعلومات التي توفرها الفقرة (I) عند مستوى القدرة $\boldsymbol{\theta}$

ويشير كروكر والجينا (Croker & Algina،1987) أنّ اقتران المعلومات له فوائد متعددة مثل مقارنة اختبارات مختلفة تقيس نفس السمة، أو مقارنة طريقتين مختلفتين لتصحيح نفس الاختبار، بالإضافة إلى مقارنة استراتيجيات مختلفة في تشكيل اختبارات مختلفة مثل المقارنة بين الاختبار التقليدي ولاختبار التكيفي وذلك بمقارنة كمية المعلومات الناتجة عن كل اختبار.

وتتأثر دالة المعلومات بمعالم الفقرة، ففي النماذج اللوجستية الأحادية والثنائية تكون أعلى كمية معلومات عند معلم الصعوبة للفقرة (عند تساوي القدرة مع الصعوبة)؛ لأنّ شكل دالة معلومات الفقرة بشكل عام قريب من الشكل الجرسي (bellshape)، بينما في حالة النموذج اللوجستي الثلاثي يمكن ملاحظة تأثير معالم الفقرة على دالة معلومات الاختبار من خلال تبسيط المعادلة الرياضية لدالة معلومات الاختبار على النحو الآتي (Hambleton& Swaminathan،1993).

$$I_i(\theta) = \frac{2.89a_i{}^2(1-c_i)}{(c_i + e^{(1.7a_i(\theta-b_i))})(1+e^{(1.7a_i(\theta-b_i)})^2}$$

وبالاعتماد على المعادلة الرياضية السابقة يمكن ملاحظة ما يأتي:

1. تكون كمية المعلومات أكبر بشكل عام عندما تكون معلمة التمييز كبيرة.
2. تزداد كمية المعلومات كلما اقتربت معلمة التخمين من الصفر.
3. تكون كمية المعلومات عندما تكون (θ) قريبة من (b_i) أكبر من كمية المعلومات عندما تكون عندما تكون (θ) بعيدة من (b_i) .

ويشير هامبيلتون (Hambleton،1989) أن دالة معلومات الفقرة تعتبر من المفاهيم الأساسية والهامة في بناء وتطوير الاختبارات، لذلك يجب اختيار الفقرات التي تقدم دقة قياس مرتفعة عند نقطة معينة على متصل السمة، وذلك بسبب أن دالة معلومات الفقرة تختلف عند المستويات المختلفة للسمة المقاسة، كما أن الفقرات ذات التمييز العالي تُعطي معلومات أكبر عن قدرة المفحوصين وبالتالي زيادة دقة القياس.

ويشير هامبلتون (Hambelton،1989) أنّ دالة معلوماتية الفقرة عند كل مستوى من مستويات القدرة تعطى بالعلاقة الرياضية الآتية:

$$I_i(\theta) = \frac{\left(P'_i(\theta)\right)^2}{P(\theta)_i Q_i(\theta)}$$

إذ تشير العلاقة الرياضية السابقة إلى أنّ دالة معلوماتية الفقرة هي القيمة العددية لحاصل النسبة بين مربع المشقة الأولى لاقتران منحنى خصائص الفقرة إلى تباين الأداء على الفقرة. وتبعاً للنماذج اللوجستية لنظرية استجابة الفقرة فان داله المعلوماتية تتخذ عنده صيغ رياضية تبعاً للنموذج اللوجستي المستخدم، فالصيغة الرياضية لدالة معلوماتية الفقرة تبعاً للنموذج الأحادي المعلمة وثنائي المعلمة وثلاثي المعلمة لنظرية استجابة الفقرة تعطى بالصيغ الرياضية الآتية وعلى التوالي:

$$I_i(\theta) = P_i(\theta)Q_i(\theta)$$

$$I_i(\theta) = a_i{}^2 P_i(\theta)Q_i(\theta)$$

$$I_i(\theta) = a^2 \left[\frac{Q_i(\theta)}{P_{ii}(\theta)} \right] \left[\frac{P_i(\theta) - c}{1 - c} \right]$$

وهناك العديد من الدراسات التي تناولت الاختبارات الموضوعية وخصائصها وعدد بدائل الفقرة بالاعتماد على النظرية الكلاسيكية والنظرية الحديثة في القياس إلا أنّ هذه الدراسات لم تركز على موضوع البدائل المركبة وفقا للنظرية الحديثة في القياس ومن هذه الدراسات دراسة كريهان وهالدينا وبريور(Crehan،Haladyna & Brewer، 1993) التي هدفت إلى معرفة العدد المناسب لبدائل أسئلة اختبار الاختيار من متعدد، ومعرفة أثر تضمين البديل "لا شيء مما ذكر" من ضمن البدائل، في صعوبة الفقرات وتمييزها، ولتحقيق أهداف الدراسة تم اختيار (48) فقرة من دليل المعلم لمادة علم النفس في جامعة "وسترن" من قبل مدرسي المادة وأعيد كتابة بدائل فقراتها بأربع صور، الصورة الأولى كانت لفقراتها أربعة بدائل دون تضمين البديل "لا شيء مما ذكر" والثانية كان لفقراتها أربعة بدائل أحدها البديل "لا شيء مما ذكر" أما الثالثة فكان لفقراتها ثلاثة بدائل دون تضمين البديل "لا شيء مما ذكر" والرابعة لفقراتها ثلاثة بدائل أحدهما البديل "لا شيء مما ذكر" وتم الحصول على الصورة التي لفقراتها ثلاثة بدائل بحذف أقل المموهات تمييزا من الصورة التي لفقراتها أربعة بدائل وطبقت الصور الثلاثة على عينة تكونت من (330) طالباً توزعوا على الصور الثلاثة بالتساوي، وتوصلت نتائج الدراسة إلى أن الفقرات ذات البدائل الثلاثة أصعب من الفقرات ذات البدائل الأربعة، بالإضافة إلى عدم وجود أثر ذي دلالة إحصائية لعدد البدائل على تمييز الفقرات.

وهدفت دراسة الشريفين وطعامنة (2009) إلى معرفة أثر عدد بدائل اختبار الاختيار من متعدد في تقدير قدرة المفحوص، وتقدير معلمة صعوبة ودالة معلومات الفقرة والاختبار.تكونت عينة الدراسة من (600) طالبٍ وطالبةٍ تم اختيارهم بشكل عشوائي من مجتمع الدراسة، وتم بناء اختبار اختيار من متعدد في مادة الرياضيات للصف العاشر الأساسي تكون من (40) فقرة، وبعد التأكد من خصائصه السيكومترية تم اشتقاق ثلاثة نماذج من الاختبار الأساسي اختلفت في عدد البدائل الخاصة بكل فقرة، حيث تكونت فقرات النموذج الأول من ثلاثة بدائل، وفقرات النموذج الثاني من أربعة بدائل، وخمسة بدائل لفقرات النموذج الثالث، وتوصلت الدراسة إلى عدد من النتائج من أهمها عدم وجود فروق دالة إحصائيّاً بين متوسطات الأخطاء المعيارية لتقدير معلمة صعوبة الفقرة، وعدم وجود فروق ذات دالة إحصائيّاً تعزى لعدد البدائل في معاملات الثبات .

وهدفت دراسة ملياني (2010) إلى المقارنة بين معاملات صعوبة الفقرة وقدرة الأفراد تبعاً لاختلاف عدد بدائل الفقرة بالاعتماد على النظرية الكلاسيكية والنموذج أحادي المعلمة (نموذج راش). وتم بناء اختبار تحصيليّ من نوع الاختيار من متعدد في الرياضيات للصف الثاني الثانوي تكون من (52) فقرة من نوع الاختيار من متعدد أُشتق منه ثلاثة نماذج (ثلاثة البدائل، أربعة بدائل، خمسة بدائل)، وتوصلت الدراسة إلى عدد من النتائج أهمها وجود فروق دالة إحصائيّاً في قدرات الافراد المقدرة بالاعتماد نموذج راشو لصالح النموذج ذي الثلاثة بدائل، عدم وجود أثر دال إحصائيّاً يعزى للنموذج في قدرات الافراد المقدرة وفق النموذج أحادي المعلمة، وجود علاقة ارتباطيَّة دالة إحصائيّاً بين قيم معاملات الصعوبة المقدرة وفق النَّظريَّة الكلاسيكية ونموذج راش في الثلاثة نماذج، وأظهرت النتائج عدم وجود فروق في قيم معاملات الصعوبة المقدرة وفق النظرية الكلاسيكية تعزى للنموذج، وتبين وجود أثر دال إحصائيّاً في قيم معاملات الصعوبة المقدرة وفق النموذج أحادي المعلمة ولصالح النموذج المكون من ثلاثة بدائل.

وقام نوادنجول ونابل (Nwadinigwel & Naibil, 2013) بدراسة هدفت لمعرفة أثر عدد البدائل في اختبار الاختيار من متعدد على الخصائص السيكومترية للاختبار، تم بناء اختبار في مادة الرياضيات للصف السادس الأساسي تكون من نموذجين من نوع الاختيار من متعدد تكونت فقرات النموذج الأول من خمسة بدائل، وفقرات النموذج الثاني تكونت من ثلاثة بدائل، طُبقت صورتي الاختبار على عينة تكونت من (159) مفحوصاً من طلبة الصف السادس الأساسي في نيجيريا، وتوصلت نتائج الدراسة إلى أن هناك أثراً دالاً إحصائيّاً لعدد البدائل على معاملات صعوبة وتمييز الفقرة ولصالح الاختبار ذي الثلاثة بدائل.

وهدفت دراسة أتلامس وكنكستون (Atalmis& Kingston, 2017) إلى اختبار عدد البدائل وأثر استخدام البديل "لا شيء مما سبق" على معاملات تمييز وصعوبة وثبات الفقرة، حيث تم بناء اختبار في مادة الرياضيات طُبق على عينة تكونت من (585) مفحوصاً من طلبة الصف السابع والثامن، ومن أهم نتائج الدراسة بأنه لا توجد فروق في صعوبة السؤال وتميزه ومعامل ثباته بين الأسئلة ذات الاربعة بدائل بنوعيها ذات البدائل الحقيقية والبدائل التي تتضمن البديل "لا شيء مما سبق"

وهدفت دراسة النصراوين (2019) التعرف إلى أثر عدد البدائل في اختبار الاختيار من متعدّد على دالة معلومات الفقرة والاختبار بالاعتماد على النظرية الحديثة في القياس باستخدام النموذج ثلاثي المعلمة؛ وقام الباحث ببناء

اختبار تحصيليّ في مادة الرياضيات للصف العاشر الأساسي تكون بصورته النهائية من (83) فقرة من نوع الاختيار من متعدد. أُشتق منه ثلاثة صور (ثلاثة بدائل، اربعة بدائل، خمسة بدائل) تم تطبيقها على عينة تكونت من (1530) طالبًا، وتوصلت الدراسة إلى عدد من النتائج منها وجود فروق دالة إحصائيّاً في معاملات الثبات لصالح صورة الاختبار ذي الخمسة والأربع بدائل، وعدم وجود فروق دالة إحصائيّاً في تقدير قيم دالة المعلومات تعزى لعدد بدائل الفقرة.

وبمراجعة الدراسات السابقة نلاحظ أن الدراسات السابقة تطرقت لدراسة خصائص فقرات اختبارات الاختيار من متعدد ودالة معلومات الاختبار باختلاف عدد بدائل الفقرة أو شكل البديل ولا يوجد دراسات في حدود علم الباحث تناولت موضوع أثر استخدام البدائل المركبة في تقدير معالم فقرات ودالة معلومات الاختبار بالاعتماد على النموذج الثلاثي في نظرية استجابة الفقرة لذلك جاءت هذه الدراسة والتي تحددت مشكلتها بما يلي:

مشكلة الدراسة وأسئلتها:

تُعد الاختبارات التحصيليّة الوسيلة الأكثر استخداماً لقياس تحصيل الطلبة في جميع مراحل التعليم، وأكثر أنماط هذه الاختبارات استخداماً في مرحلة التعليم الجامعي والمدرسي هي الاختبارات ذات الإجابة المنتقاة وخاصة اختبارات الاختيار من المتعدد، ويسعى مُعدّ هذا النوع من الاختبارات إلى تهيئة كافة الظروف التي تجعل القياس أكثر دقة وموضوعية، وخاصة ما يتعلق ببدائل الفقرة وقوة هذه البدائل وما هو العدد الأنسب للبدائل وذلك لما لذلك من أثر كبير على خصائص الاختبار وخصائص فقراته، لذلك لا يستطيع واضعي الاختبارات أحياناً وضع بديل فعال وقوي مما يدفعهم للاعتماد على بديل مركب من البدائل الموجودة للفقرة مما قد يؤثر في دقة القياس، وبالتالي على خصائص الاختبار لذلك جاءت هذه الدراسة بهدف التعرف أثر البدائل المركبة في تقدير الخصائص معالم فقرات ودالة معلومات الاختبار باستخدام النموذج الثلاثي لنظرية استجابة الفقرة وبالتحديد سيحاول البحث الإجابة عن الأسئلة الآتية :

1- هل توجد فروق ذات دلالة إحصائية عند مستوى ($\alpha \leq 0.05$) بين المتوسطات الحسابية لمعالم الفقرة (الصعوبة، التمييز، التخمين) تُعزى لنوع البديل (البدائل التقليدية، البدائل المركبة)؟
2- هل تختلف دالة معلومات الاختبار باختلاف نوع البديل (البدائل التقليدية، البدائل المركبة)؟

أهداف الدراسة:

تهدف الدراسة إلى معرفة تأثير البدائل المركبة في فقرات الاختبار التحصيليّ على معالم فقرات ودالة معلومات الاختبار باختلاف نوع البديل (البدائل التقليدية، البدائل المركبة).

أهمية الدراسة:

تأتي الأهمية العمليّة لهذه الدراسة؛ لاعتماد كثيرين على علامة الاختبار؛ وذلك من أجل ترتيب المفحوصين وإعطائهم درجات، وأيضاً لمفاضلتهم وتعيينهم، وأحد الطرق التي تستخدم للحصول على اختبارات جيده للمفاضلة بين المفحوصين الاعتماد على دالة معلومات الفقرة لاختيار الفقرات عند بناء الاختبارات، بالاعتماد على النظرية الحديثة في القياس، على افتراض أنّ معلومات الفقرة تتغير عبر مستويات السمة المختلفة؛ وعليه فإنّه من الممكن اختيار فقرات تقدم دقة قياس مرتفعة عند نقطة معينة على متصل السمة وذلك بانتقاء الفقرات التي تقدم أعلى كمية معلومات وأقل خطأ معياري في تقدير القدرة للاختبار، ومما يزيد أهمية هذه الدراسة أيضاً قلة وجود دراسات محلية تناولت هذا الموضوع، حيث تسعى هذه الدراسة إلى معرفة أثر البدائل المركبة في تقدير معالم فقرات الاختبار ودالة معلوماته بالاعتماد على النموذج ثلاثي المعلمة، وكذلك لتعطي ملاحظة إلى الأشخاص الذين يضعون الاختبارات بأنْ يهتموا بنوع البديل المناسب عند إعدادهم الاختبار.

مصطلحات الدراسة:

معالم الفقرة (item parameter): هيمعلمةالتمييز، معلمة الصعوبة، ومعلمة التخمين.

معلمة الصعوبة(Threshold): هيالنقطة على متصل القدرة التي تقابل احتمال 2/(1+ci) للإجابة الصحيحة على الفقرة (i)، ويشير الرمز (ci) إلى معلمة التخمين.

معلمة التمييز(Slope): نسبة ميل منحنى خصائص الفقرة الذي يقابل النقطة التيتتساوى عندهاصعوبة الفقرة مع علامة القدرة.

معلمة التخمين(Asymptote): هي خط المقاربة الأدنى Lower Asymptote من منحنى خصائص الفقرة ويمثل احتمال إجابة المفحوص ذي القدرة المتدنية على الفقرة إجابة صحيحة.

دالة معلومات **الاختبار(Test Information Function):** هو كمية المعلومات التي يوفرها الاختبار عند مستوى القدرة θ والذي يعبر عن مجموع دوال معلومات فقرات الاختبار عند مستوى معين من القدرة ويمكن التعبير عنه بالصيغة الآتية:

$$I(\theta) = \sum I_i(\theta)$$

اختبار الاختيار من متعدد: اختبار تتكون كل فقرة من فقراته من متن الفقرة وله بدائل عدة، يعطى الفرد العلامة (1) عند اختيار البديل الصحيح والعلامة (0) عند اختيار البديل الخاطئ. وفي البحث الحالي له صورتان (البدائل التقليدية، البدائل المركبة).

البدائل المركبة: هو بديل يتكون من بديلين من البدائل الخاصة بالفقرة، يتم بناءه بالاعتماد على الخيارات الموجودة الخاصة بالفقرة.

الطريقة والإجراءات:

مجتمع الدراسة:

تكون من جميع معلمي ومعلمات محافظة معان ولواء وادي السير بواقع (2853) معلماً ومعلمة.

عينة الدراسة

تكونت من (450) معلماً ومعلمة من معلمي محافظة معان ولواء وادي السير تم اختيارهم بالطريقة العشوائية البسيطة.

أداة الدراسة

تم بناء اختبار تحصيليّ تكون من (40) فقرة من نوع الاختيار من متعدد لكل فقرة منها أربعة بدائل أحدهما الإجابة الصحيحة، في مناهج البحث يقيس مهارات المعلمين بالبحث التربويّ ولا سيما انّ هذه المهارات أصبحت ضرورية بالنسبة للمعلمين في ظل حاجتهم إلى هذه المهارات لعمل الابحاث وذلك للاستفادة من نظام رتب المعلمين.

الخصائص السيكومترية للاختبار:

صدق المحكمين:

عرض الاختبار بصورته الأولية على عدد من المحكمين من أعضاء هيئة التدريس في كلية العلوم التربويّة تخصص المناهج العامة والقياس والتقويم، وعلم النفس التربويّ، في جامعة الحسين بن طلال والجامعات الأخرى وطُلب منهم الحكم على فقرات الأداة من حيث علاقة الفقرات بموضوع الدراسة، وضوح الصياغة اللغوية، ومدى تغطية الفقرات لموضوع البحث وتم حذف فقرة بناء على رأي المحكمين.

صدق الاتساق الداخلي:

للتحقق من دلالات صدق الاتساق الداخلي للاختبار بصورته الأولية والمُكَّون من (39) فقرة وبعد إجراء التعديلات بناء على ملاحظات المحكِّمين تم تطبيقه على عينة استطلاعية قوامها (30) معلماً ومعلمة من مجتمع الدراسة ومن خارج عينتها، واستخرجت معاملات الارتباط بين درجة كل فقرة والدرجة الكلية على الاختبار، وقد انحصرت قيم معاملات ارتباط درجات الفقرات مع الدرجة الكلية ما بين (0.33 – 0.82) وكانت جميع معاملات الارتباط كانت ذات درجات مقبولة ودالة إحصائيّاً.

ثبات الاختبار:

تم حساب معامل ثبات الاختبار التحصيليّ بعد تطبيقه على أفراد عينة الدراسة الاستطلاعية التي بلغت (30) معلماً ومعلمة وذلك بالاعتماد طريقة التجزئة النصفية، وللتخلص من أثر التجزئة النصفية تم الاعتماد على معادلة سبيرمان بروان، وكانت قيمة معامل الثبات بالاعتماد على طريقة التجزئة النصفية (0.88) وهذا يدل على أن الاختبار يتمتع بدرجة عالية من الثبات مما يطمئن الباحث لتطبيقه على عينة الدراسة.

الخصائص السيكومترية لفقرات الاختبار:

تم التحقق من الخصائص السيكومترية لفقرات الاختبار بإيجاد قيم معاملات الصعوبة والتمييز لفقرات الاختبار بعد تطبيقه على عينة استطلاعية تكونت من (30) معلماً ومعلمة وبالاعتماد على القيم الناتجة فقد تم حذف (3) فقرات من فقرات الاختبار ذات معاملات التمييز المنخفض (أقل من 0.19) والفقرات ذات معاملات الصعوبة المتدنية (أقل من 0.2) وذات معاملات الصعوبة المرتفعة (أعلى من 0.8) وبذلك أصبح الاختبار بصورته النهائية مكوناً من (36) فقرة. مع الأخذ بعين الاعتبار أن الفقرات المتبقية بعد الحذف تحقق شموليتها وتمثيلها لمهارات البحث العلمي التي يتوقع من المعلمين معرفتها والخاصة بالبحث التربويّ، وقد تراوحت قيم معاملات الصعوبة لفقرات الاختبار التحصيليّ بين (0.30–0.78) وتعتبر جميع معاملات الصعوبة جيده ومناسبة ويعد ذلك مؤشراً على جودة فقرات الاختبار ومناسبتها للهدف الذي أعد من أجله وتراوحت باقي قيم معاملات التمييز لفقرات الاختبار التحصيليّ بين (0.28–0.67) والتي تعتبر معاملات تمييز موجبة ومناسبة ، وتشير إلى أن فقرات الاختبار تتمتع بمعامل تمييز جيدة.

وبعد التأكد من الخصائص السيكومترية للاختبار وفقراته تم حذف أحد البدائل عشوائياً من بدائل كل فقرة من فقرات الاختبار واضافة بديل مركب لبدائل الفقرة لتتكون صورة جديدة للاختبار، ومن ثم تطبيق صورتي الاختبار بالصورة النهائية على عينة الدراسة وجمع استجابات عينة الدراسة وإدخالها وتحليلها بالاعتماد على البرنامج الإحصائي (spss).

التحقق من افتراضات نظرية الاستجابة للفقرة.

1-افتراض أحادية البعد Unidimensionality:

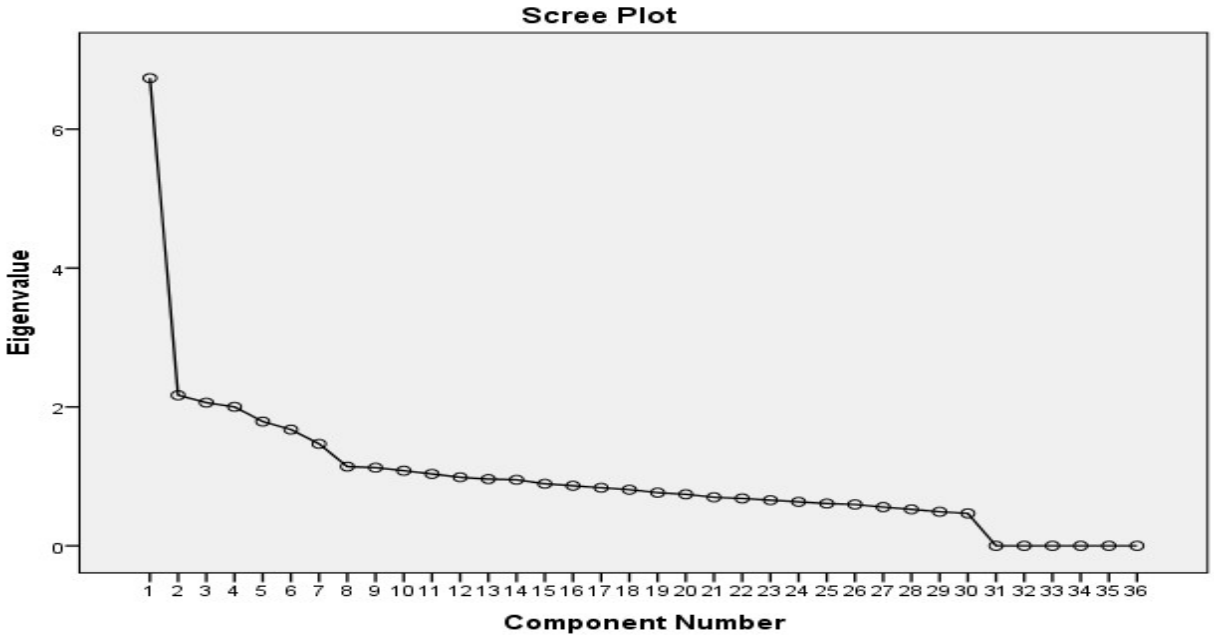
تم فحص هذا الافتراض بالاعتماد على التحليل العاملي الاستكشافي لاستجابات المفحوصين (Factor Analysis) بطريقة المكونات الرئيسية (Principal Component) والاعتماد على طريقة (varimax) لتدوير المحاور. ويظهر في الجدول (1) قيم الجذور الكامنة (Eigen Values) ونسبة التباين الذي يفسره كل عامل (Explained Variance) وكذلك التباين المفسر التراكمي لكل عامل من العوامل:

جدول (1): قيم الجذور الكامنة ونسبة التباين المفسر والتراكمي للاختبار.

العامل	الجذر الكامن	نسبة التباين المفسر	نسبة التباين المفسر التراكمي	الجذر الكامن الأول / الجذر الكامن الثاني
1	6.74	18.71	18.71	3.12
2	2.16	6.02	24.72	
3	2.06	5.73	30.46	
4	2.00	5.56	36.01	
5	1.79	4.97	40.98	
6	1.67	4.64	45.63	

يتضح من الجدول رقم (1) أن العامل الأول يفسر أعلى نسبة تباين بالمقارنة مع العوامل الأخرى. كما يتبين أن ناتج قسمة الجذر الكامن للعامل الأول على الجذر الكامن للعامل الثاني أكبر من (2). مما يدل على وجود سمة سائدة للاختبار، وهذا مؤشر على تحقق افتراضية أحادية البعد للاختبار (Hattie،1985).

كما تم الاعتماد على التمثيل البياني باستخدام (Scree Plot) للعوامل مع الجذور الكامنة والشكل التالي يوضح ذلك.



شكل رقم (1): الرسم البياني لقيم الجذور الكامنة للعوامل.

يتضح من التمثيل السابق أنّ العامل الأول هو العامل الطاغي حيث فسر أعلى نسبة من التباين مقارنة مع العوامل الأخرى، وهذا يعتبر أيضًا مؤشرًا على أحادية البعد.

2-افتراض الاستقلال الموضعي Local Independence:

يوضح هذا الافتراض أن إجابة المستجيب عن الفقرة لا تتأثر بإجابته عن فقرة أخرى والعامل الوحيد المؤثر على احتمال الإجابة عن الفقرة هو مقدار القدرة (θ) وخصائص الفقرة. وللتحقق من هذا الافتراض تم اعتماد أنْ تَحقُّق افتراضاً أحادية البُعد لفقرات الاختبار يُعَدّ دليلًا على تَحقُّق افتراض الاستقلال الموضعي، إذ إنّ هذا الافتراض مكافئ لافتراض أحادية البعد،(Hambleton & Swaminathan, 1985).

3-التحرر من السرعة في الأداء:

يمكن التأكد من هذا الافتراض من خلال معرفة نسبة المفحوصين الذين أكملوا الاختبار، ومعرفة الفقرات التي لم يجب عنها المفحوصين. ووفقًا لما يراه هامبلتون وآخرون (Hambleton et al., 1985). فإنه إذا أكمل (75%) من الطلبة الإجابة عن الاختبار، وتمت الإجابة عن (80%) من فقرات الاختبار، فإن السرعة لا تعد في هذه الحالة عاملًا

مهمًا في الأداء على الاختبار. علماً بأن جميع الطلبة أجابوا على هذا الاختبار وبنسبة (100%)، وأنه تمت الإجابة على نسبة من الفقرات تجاوزت (92%) وهذا يؤكد أن الاختبار يقيس القوة وليس السرعة.

التحقق من مطابقة الأفراد والفقرات للنموذج الثلاثي لنظرية استجابة الفقرة:

استخدم برنامج (Bilog-MG3)؛ للتأكد من مطابقة الأفراد (Person-Fit)، ومطابقة الفقرات (Item-Fit) للنموذج الثلاثي باستخدام اختبار مربع كاي (χ^2) عند مستوى دلالة ($\alpha = 0.01$)، وقد أظهرت نتائج التحليل عدم مطابقة استجابات ثمانية أفراد للنموذج ثلاثي المعلمة، حيث كانت القيم الاحتمالية لهم أقل من (0.01)؛ لذا تم حذف استجابتهم وأصبح حجم العينة في صورتها النهائية (442) مفحوصاً.

وفيما يتعلق بفحص مطابقة الفقرات للنموذج المستخدم أُعيد التحليل باستخدام برنامج (Bilog-MG3) بعد حذف استجابات الأفراد غير المطابقين للنموذج بالاعتماد على مؤشر (χ^2) عند مستوى دلالة ($\alpha = 0.01$)؛ وأظهرت نتائج التحليل مطابقة جميع فقرات الاختبار للنموذج ثلاثي المعلمة، حيث كانت القيمة الاحتمالية أكبر من (0.01) لجميع الفقرات.

نتائج الدراسة ومناقشتها:

السؤال الأول: هل توجد فروق ذات دلالة إحصائية عند مستوى ($\alpha \leq 0.05$) بين المتوسطات الحسابية لمعالم الفقرة (الصعوبة، التمييز، التخمين) تُعزى لنوع البديل (البدائل التقليدية، البدائل المركبة)؟

للإجابة على هذا السؤال تم حساب معالم الفقرات بالاعتماد باستخدام برنامج (Bilog-mg3) وفق النموذج ثلاثي المعلمة ثم تم حساب المتوسطات الحسابية والانحرافات المعيارية لمعالم الفقرة (الصعوبة، التمييز، التخمين) تبعاً لنوع البديل (البدائل التقليدية، البدائل المركبة) واستخدام اختبار(ت) للعينات المستقلة (Independent Samples Test) لمعرفة دلالة الفروق بالاعتماد على برنامج (SPSS)، والجدول (2) يوضح هذه النتائج.

جدول (2): اختبار (ت) للعينات المستقلة والمتوسطات الحسابية والانحرافات المعيارية لمعالم الفقرة (الصعوبة، التمييز، التخمين) تبعاً لنوع البديل (البدائل التقليدية، البدائل المركبة).

معلم الفقرة	نوع البديل	المتوسط الحسابي	الانحراف المعياري	قيمة "ت"	درجات الحرية	الدلالة الإحصائية
الصعوبة	البدائل التقليدية	1.808-	1.57	2.74	70	0.008
	البدائل المركبة	0.871-	1.31			
التمييز	البدائل التقليدية	0.318	0.07	2.37	70	0.02
	البدائل المركبة	0.167	0.10			
التخمين	البدائل التقليدية	0.001	0.001	2.34	70	0.02
	البدائل المركبة	0.081	0.202			

يتبين من الجدول (2) وجود فروق دالة إحصائيّاً بين المتوسطات الحسابية لمعلمة الصعوبة تبعاً لنوع البديل (البدائل التقليدية، البدائل المركبة) ولصالح صورة الاختبار ذي البدائل المركبة، أي أن المفحوصين يفضلون البدائل المركبة أفضل من البدائل التقليدية وبالنسبة لهم سيكون الاختبار أسهل منه في حال الاعتماد على البدائل التقليدية، وربما يعود السبب في ذلك أنه عند اعتماد الطالب على البدائل المركبة فإنّ معرفته لأحد البدائل المتضمنة في البديل المركب أنه غير صحيح سيدفعه إلى استبعاد البديل المركب، أي إنّ اختيار البديل المركب أو استبعاده مرتبط باختيار أو استبعاد البدائل الأخرى مما يزيد من نسبة التخمين للإجابة الصحيحة عند المفحوص فيصبح السؤال سهلاً بالنسبة له وبالتالي ستزيد قيمة معاملات الصعوبة.

كما يتبين من الجدول (2) وجود فروق دالة إحصائيّاًبين المتوسطات الحسابية لمعلمة التمييز تبعاً لنوع البديل (البدائل التقليدية، البدائل المركبة) ولصالح نوع البديل التقليدي، وهذا يؤكد أن معاملات التمييز للفقرات في حالة الاعتماد على البدائل التقليدية ستكون أعلى بالمقارنة باستخدام البدائل المركبة وربما يعود السبب في ذلك أنّ الاعتماد على البدائل المركبة يزيد من فرصة التخمين لدى المفحوصين وبالتالي سيؤثر ذلك على دقة القياس وعدم قدرة الفقرة على التصنيف الحقيقي للمفحوصين حسب ما لديهم من قدرة.

كما يتضح من الجدول (2) وجود فروق دالة إحصائيّاً بين المتوسطات الحسابية لمعلمة التخمين تبعاً لنوع البديل (البدائل التقليدية، البدائل المركبة) ولصالح نوع البديل المركب. أي تزداد قيمة التخمين في حال استخدام البدائل المركبة من قبل المفحوصين وذلك بسبب أنه معرفة الطالب لأحد البدائل المتضمنة في البديل المركب أنه غير صحيح سيدفعه إلى استبعاد البديل المركب مما يزيد من نسبة التخمين للإجابة الصحيحة.

السؤال الثاني: هل تختلف دالة معلومات الاختبار باختلاف نوع البديل (البدائل التقليدية، البدائل المركبة)؟

للإجابة على السؤال تم حساب دالة المعلومات لكل فقرة من فقرات صورتي الاختباربالاعتماد على برنامج (Bilog–mg3) وفق النموذج ثلاثي المعلمة ثم تم حساب المتوسطات الحسابية والانحرافات المعيارية لقيم دوال المعلومات لكل صورة من صور الاختبار (البدائل التقليدية، البدائل المركبة) واستخدام اختبار (ت) للعينات المستقلة (Independent Samples Test) لمعرفة دلالة الفروق بالاعتماد على برنامج (SPSS)، والجدول (3) يوضح هذه النتائج.

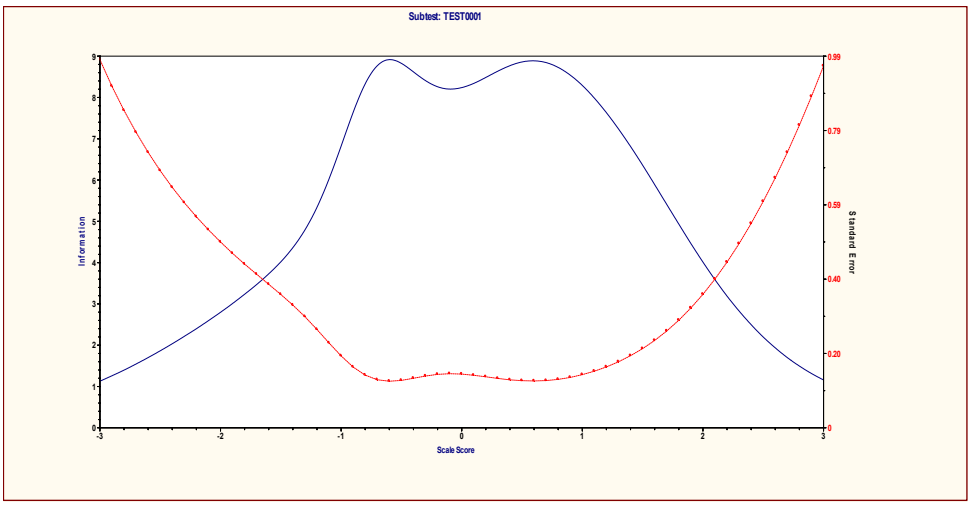
جدول (3):اختبار ت للعينات المستقلة والمتوسطات الحسابية والانحرافات المعيارية لدالة معلومات الفقرة لصورتي الاختبار (البدائل التقليدية، البدائل المركبة).

صورة الاختبار	المتوسط الحسابي	الانحراف المعياري	قيمة "ت"	درجات الحرية	الدلالة الإحصائية
الصورة الأولى (البدائل التقليدية)	0.37	0.12	4.71	70	0.00
الصورة الثانية (البدائل المركبة)	0.22	0.13			

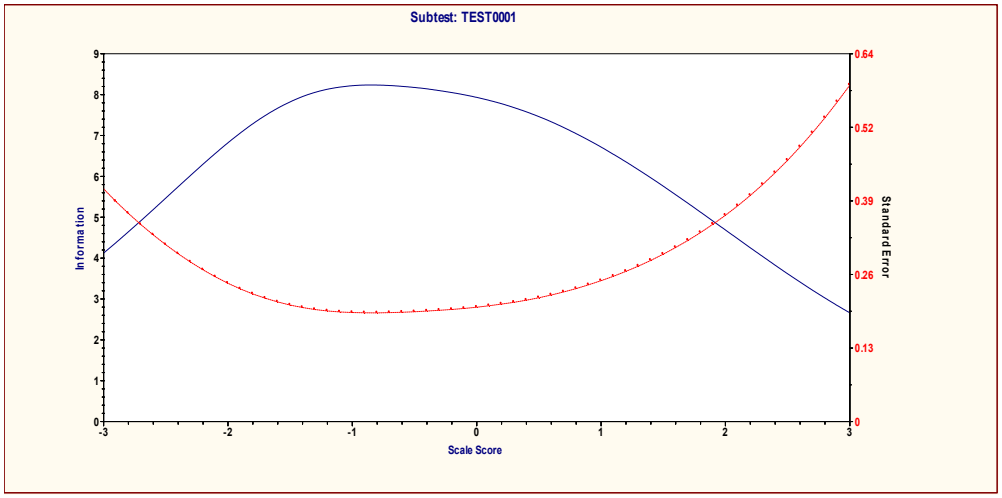
يتضح من الجدول (3) وجود فروق دالة إحصائيّاًبين المتوسطات الحسابية لدالة معلومات الفقرة لصورتي الاختبار (البدائل التقليدية، البدائل المركبة) ولصالح نوع البديل التقليدي. وقد يعزى السبب في ذلك إلى أن البدائل المركبة قد تؤثر سلباً على دالة معلومات الفقرات وبالتالي دالة معلومات الاختبار؛ إذْ تزداد فرصة التخمين في حالة وجود البدائل المركبة ضمن فقرات الاختبار إذْ إنّ اعتماد الطلبة على التخمين العشوائي سيؤثر سلباً على كل من صدق وثبات الاختبار وزيادة الخطأ المعياري في القياس وبالتالي سيؤثر ذلك أيضاً سلبا على دالة معلومات الاختبار إذْ إنّ اعتماد الطلبة على التخمين سيضيف تباينا لا يرتبط بالتباين في القدرة الحقيقية. حيث إنّه عند اعتماد الطالب على البدائل المركبة يؤدي إلى أنّه عند معرفة الطالب لأحد الخيارات المتضمنة في البديل المركب أنه غير صحيح سيدفعه إلى استبعاد البديل المركب مما يزيد من نسبة التخمين للإجابة الصحيحة عند المفحوص. وربما يعود السبب في ذلك أيضاً أن معامل التمييز سيكون منخفضاً في حال الاعتماد على البدائل المركبة بالمقارنة مع البدائل التقليدية أي أنه في حال الاعتماد على البدائل المركبة سوف تعطي دالة معلومات أقل للفقرات وبالتالي للاختبار ككل وهذا أمر منطقي حيث إنّ الفقرة ذات التمييز المنخفض لديها قدرة ضعيفة على التمييز بين المجموعتين العليا والدنيا من الطلبة حيث انه كل ما كان معامل التمييز للفقرة عالي كلما أشار ذلك إلى قدرة الفقرة على تصنيف المفحوصين وتدريجهم حسب ما لديهم من قدرة. وهذا يتضح من المعادلة الآتية:

$$I_i(\theta) = \frac{2.89a_i^{\,2}(1-c_i)}{(c_i + e^{(1.7a_i(\theta-b_i))})(1+e^{(1.7a_i(\theta-b_i)})^2}$$

حيث يتضح من المعادلة السابقة أنّه إذا زادت نسبة التخمين ستقل قيمة البسط للمعادلة وبالتالي يقل المقدار الكلي أي تقل قيمة دالة المعلومات عند استخدام البدائل المركبة في الاختبارات. وهذا يتضح أيضاً من الرسم البياني التالي لدالة المعلومات للاختبار إذْ يتضح أن قيمة دالة المعلومات في حالة الاعتماد على البدائل المركبة أقل منها في حالة الاعتماد على البدائل التقليدية.



منحى دالة المعلومات لصورة الاختبار الأولى (البدائل التقليدية)



منحى دالة المعلومات لصورة الاختبار الثانية (البدائل المركبة)

التوصيات:

اعتماداً على نتائج الدراسة يوصي الباحث:

1. الابتعاد عن استخدام البدائل المركبة ضمن فقرات اختبار الاختيار من متعدد من قبل واضعي الاختبارات.
2. إجراء دراسات للمقارنة بين أنواع اخرى من بدائل فقرات اختبار الاختيار من متعدد مثل " لا شيء مما ذكر " " جميع ما ذكر صحيح"

المصادر والمراجع

المراجع باللغة العربية:

- الشريفين، نضال وطعامنة، إيمان (2009). أثر عدد البدائل في اختبار الاختيار من متعدد في تقديرات القدرة للأفراد والخصائص السيكومترية للفقرات والاختبار وفق نموذج راش في نظرية الاستجابة للفقرة، المجلة الأردنيّة في العلوم التربويّة، 5(4)، 309-335.

- الزيات، فتحي (1989). أثر قوة تشتيت البدائل في أسئلة الاختيار من متعدد على صدق الاختبار وثباته: دراسة تحليلية، مجلة كلية التربية، جامعة المنصورة، مجلد 11، ص ص 86-108.

- النصراوين، معين (2019). أثر عدد البدائل في اختبار الاختيار من متعدّد على دالة معلومات الفقرة والاختبار وفق النموذج الثلاثي المعلمة في إطار نظرية استجابة للفقرة، المجلة الدولية للأبحاث التربويّة، 43(3)، 159-181.

- الدوسري، راشد حماد (2004). القياس والتقويم التربويّ الحديث مبادئ وتطبيقات وقضايا معاصرة. الأردن، عمان: دار الفكر.

- الكيلاني، عبدالله وعدس، عبد الرحمن (1986). القياس والتقويم في التربية وعلم النفس. مركز الكتاب الأردنيّ، الأردن، عمان.

- ملياني، زياد. (2010). مقارنة أثر اختلاف عدد البدائل في اختبار الاختيار من متعدد على صعوبة الفقرة وقدرة الفرد في ضوء كل من النظرية الكلاسيكية ونموذج راش. أطروحة دكتوراه غير منشورة، جامعة أم القرى، مكة المكرمة.

المراجع الأجنبية:

- Al-Dosari، H (2004). Modern educational measurement and evaluation contemporary principles، applications and issues. Jordan، Amman: Dar al-Fiker for publishing & distributing.

- AL-Kilani، A& Ades، A (1986). Measurement and evaluation in education and psychology. Jordan Book Center، Jordan، Amman.

- Al-Nasraween، M.(2019). Information Function for Item and Test and Reliability When Using Three Forms of Multi-Choice Test Under Item Response Theory، International Journal for Research in Education، 43(3)، 159-181.

- AL- shrifin & Taamnah،E.(2009) conducted a study investigated the effect of multiple choice test number of alternatives on the estimation of a person's ability and the psychometric properties of a test and its items according to Rasch model in item response theory (IRT)،Jordan Journal of Educational Sciences،5(4)،309-335.

- Atalmis، E & Kingston، N .(2017). Three ، four، and none of the above options in multiple-choice items. Turkish Journal of Education، 6(4)، 143-157.

- AL-Zayat، F. (1989). The effect of the dispersal of alternatives in multiple selection questions on the validity and reliability of the test: analytical study، journal of the Faculty of Education، Mansoura University، Volume 11، p. 86-108.

- Ben-simon، A. & Budescu ،D. & Nevo، B. (1997). A comparative study of measures of partial knowledge in multiple – choice tests. Applied psychological measurement، Vol 21(1)، PP.65-88.

- Cerhan، K،Haladyna، T، and Brewer،B. (1993). Use of an inclusive option and the optimal of options for multiple choice items، Educational and psychological Measurement، 53،241-247.

- Crocker، L.& Algina ، J.(1987)، Introduction to Classical and Modern Test Theory. University of Florida، U.S.A.

- Hambleton،R. (1989). Principle and selected application of item response theory، Educational measurement (3rd ed)، 147-201. New York: Amarican council on Education، Macmillan publishing company.

- Hambleton،R.k.، & Swaminathan،H، &Rogers. H.j.(1993). Fundamentals of Item Response Theory: International Educational and Professional. Publisher Newbury park.

- Hambleton،R. k.، & Swaminathan، H. (1985). Item Response Theory: Principles and applications. Boston MA: Kluwer-Nyjhoff.

- Melini، Z.(2010). Comparing the impact of the different number of alternatives in the multiple choice test on the item difficulty and the ability of the individual in the light of both the Calsic theory and the Rasch model. Unpublished Doctoral Thesis، Um al-Qura University، Mecca

- Nwadinigwel، P. & Naibi1، L. (2013). The number of options in a multiple choice test item and the psychometric characteristics. Journal of Education and Practice، 4(28)، 22-34.

- Rich، C، & Johanson، G. (1990). An item-level analysis of "non of the above" paper presented at the annual meeting of the American Educational Research Association، Boston، MA .Retrieved February 15th ،2022، from http://files.eric.ed.gov/fulltext/ED400299.PDF

- Shahin، A & Anil، D. (2017). The effect of the test length and sample size on item parameters in item response theory. Educational sciences: Theory and practice، 17(1) ،321-335.