مجلة كلية التربية

أثر طريقة ترتيب المفردات على دقة تقدير قدرات الأفراد ومعالم المفردات باستخدام نموذج الاستجابة للمفردة رباعي المعلم

The Effect of the Method Arranging the Test Vocabulary on Accuracy of Estimating Persons Ability and Items Parameters using four-Parameter Item Response Theory

إعداد

رحمة احمد الحجاجي

باحثة في القياس والتقويم – جامعة ام القرى

د. ذياب بن عايض المالكي

أستاذ القياس والتقويم المشارك بجامعة أم القرى

المستخلص

هدف البحث الحالي إلى معرفة أثر طريقة ترتيب المفردات على دقة تقدير قدرات الأفراد ومعالم المفردات، تم إعداد اختبار تحصيلي، وتكوين ثلاث نماذج اختبارية تختلف فيما بينها في طريقة ترتيب المفردات. وتكونت عينة البحث من (٣٥٤١) طالبة، وقد أظهرت النتائج أن نموذج الاختبار المرتبة مفرداتها وفق تسلسل المحتوى يعتبر أكثر دقة في تقدير معالم قدرات الأفراد عن نموذج الاختبار التي تم ترتيب مفرداتها وفق تسلسل الصعوبة المعرفية، أو وفق الترتيب العشوائي. وبالنسبة لدقة تقدير معالم الصعوبة، أظهرت النتائج أن النموذج المرتب مفرداته وفق تسلسل المحتوى أو وفق تسلسل الصعوبة المعرفية أكثر دقة في تقدير معالم صعوبة المفردات عن النموذج الاختباري المرتب بشكل عشوائي، وأظهرت النتائج الخاصة بمعلم التمييز، أن النموذج الاختباري المرتب بشكل عشوائي أقل طرق ترتيب المفردات دقة في تقدير معالم تمييز المفردات، وفيما يخص معلم التخمين، أظهرت النتائج عدم وجود فرق دال إحصائياً بين متوسطات الأخطاء المعيارية في تقدير معالم تخمين المفردات تعزى لاختلاف طريقة ترتيب المفردات الاختبارية. وأظهرت النتائج أن أقل طرق ترتيب المفردات دقة في تقدير معالم اللامبالاة لمفردات الاختبار هي طريقة الترتيب العشوائي.

الكلمات المفتاحية: طريقة ترتيب مفردات الاختبار، نموذج الاستجابة للمفردة رباعي المعلم، دقة التقدير، قدرات الأفراد، معالم المفردات.

Abstract

The aim of the current research is to investigate the effect of the method of arranging items on the accuracy of estimating the ability of individuals and the parameters of items. The research sample consisted of 3541 female students. The results showed that the test form with items arranged according to the sequence of content is more accurate in estimating the parameters of individuals' abilities than the form with items arranged according to the sequence of cognitive difficulty or random order. Regarding the accuracy of estimating the difficulty parameters, the results showed that the model with items arranged according to the content sequence or cognitive difficulty sequence is more accurate in estimating the item difficulty parameters than the randomly arranged test model. The results also showed that there was no statistically significant difference between the means of standard errors in estimating item estimation parameters due to the difference in the arrangement of test items. However, the results revealed that the least accurate method of item order in estimating the indifference parameters of the test items was the random order method.

Keywords: method of arranging test items, estimation of a four-parameter model, accuracy of estimation, ability of individuals, parameters of items.

مقدمة:

ما تزال الاختبارات التحصيلية من أدوات القياس المهمة، والتي تستخدم نتائجها في اتخاذ العديد من القرارات التربوية. وقد اعتُرف بالأدوار الكثيرة التي يمكن أن تلعبها الاختبارات التحصيلية في العملية التعليمية منذ مدة طويلة، فمن خلالها يستطيع المعلم الحكم على مدى تحقيق أهداف البرامج التعليمية والتدريبية التي يقوم بتدريسها، كما أنها تفيد الطالب في التعرف على مستوى تحصيله الدراسي من خلال أدائه في الاختبار.

ولما لهذه الاختبارات من أهمية كبيرة في التحصيل الدراسي وأثره على أداء الفرد؛ فإن خبراء القياس والباحثين اهتموا كثيراً بالعوامل المؤثرة بإعداد الاختبار واخراجه بحيث يوفر للطالب القدرة على تحصيل أفضل ما يمكن، وإخراج أكبر ما يمكن من قدراته. فصدق المعلومات التي نحصل عليها من هذه الاختبارات يتوقف على العناية والاهتمام الذين نوفرهما لأعداد هذه الاختبارات.

وقد تكون العوامل التي من الممكن أن تؤثر على الأداء في الاختبارات التحصيلية والتي تختلف عن قدرة الطالب التحصيلية ومدى تمكنه من محتوى المقررات الدراسية، متعلقة بسمات الطلاب أنفسهم سواء المعرفية، أو الوجدانية، أو المهارية، أو متعلقة بالاختبار والظروف الاختبارية، أو ببعض القرارات المتعلقة بمعد الاختبار، وكلما زاد تأثير هذه العوامل على أداء الطلاب في الاختبارات، زادت أخطاء القياس في تقدير تحصيل الطلاب وانخفض تبعاً لذلك صدق هذه الاختبارات، وثباتها أيضاً. (عبد الوهاب،٢٠١٩)

وحجم الأخطاء في عملية القياس النفسي أثناء تطبيق الاختبارات التحصيلية بوجه عام تزداد أو تقل تبعاً لمصدر هذه الأخطاء. (Lord,1980) ويحدد بيمر وقروفز ولايبرج وماثوتز (Biemer; Groves; Lyberg; Mathiowetz (1991 هذه الأخطاء في أربعة مصادر: (أداة القياس، ونموذج جمع البيانات، والمقابلة أو التطبيق،

والمستجيب). وبالنسبة لأداة القياس فهناك مصادر للخطأ تتعلق بها، منها:(صياغة المفردة، وطول المفردة، وطول الأداة، وفئات الاستجابة، وترتب المفردات).

ويشير أيكن وقروث (Aiken; Growth (2006 إلى ان بعض علماء القياس يرون أن الأخطاء المتأتية من أداة القياس تعد من أكثر الأخطاء تأثيراً في تقدير الدرجة للفرد، ويرى حبيب وعزيز (٢٠١٨) أن من المشكلات المعاصرة المهمة التي ترتبط ببناء الاختبارات التحصيلية وكذلك المقاييس النفسية هو ما يتعلق بالخطوات الإجرائية في بناء الاختبارات، وطريقة تنظيمها.

ويعد ترتيب مفردات الاختبار من أحد الخطوات الإجرائية المهمة في بناء الاختبار، إذ يتوجب على واضع الاختبار تبني أسلوباً محدداً لترتيب مفردات الاختبار، وبالرغم من عدم وجود اتفاق في هذا المجال لنظام معين أو أسلوب محدد لترتيب مفردات الاختبار؛ إلا أن هناك عدة وجهات نظر يمكن استخدام أحدها أو اكثر في ترتيب المفردات، فيمكن أن ترتب المفردات وفق سهولة المفردات وصعوبتها، وقد يتم ترتيب المفردات بنفس التسلسل المنطقي لمحتوى المادة الدراسية، أو ترتيب المفردات وفق الأهداف التي تحققها المفردات بتجميع المفردات التي تحقق نوعية معينة من الأهداف وراء بعضها البعض، وغيرها من الطرق المختلفة. (بركات،٢٠١٨)

وقد يلجأ البعض من أجل الوقوف على المستوى الفعلي لأداء الأفراد على الاختبار، إلى إعداد نماذج مختلفة للاختبار من خلال إدارة مجموعة المفردات ذاتها في ترتيبات مختلفة وهي إستراتيجية لمنع النسخ (الغش) ولتعزيز أمن الاختبار، وهذا ما يراه عبد الهادي (٢٠٠١) حيث يذكر "إذا لم توضع أسئلة الاختبار التحصيلي بشكل نماذج مختلفة، فإن هذا بدوره يؤدي إلى الغش، وبذلك تتأثر نتائج الطلبة سلباً". ويبدو ان هذه المعالجة لترتيب المفردات عبر نماذج اختبارية بديلة، قد لا تكون بدون عواقب، فعلى الرغم من تعرض الطلاب لنفس المفردات، قد يقع معد الاختبار– دون دراسة

كافية للنتائج المترتبة على استخدام هذه الاختبارات– في مشكلة التحيز، فقد يحصل بعض الطلاب الذين أجابوا عن نموذج ما، على درجات إضافية لم يحصل عليها أقرانهم الذين أجابوا عن نماذج أخرى من نفس الاختبار. (محمود،٢٠١٥)

وبالنسبة للدراسات القائمة على النظرية الكلاسيكية للاختبارات يذكر كلاً من وين ورونج ويومج (Wen; Rong; Yumg (2012 عند تطبيق الدراسات وفق النظرية الكلاسيكية للاختبارات، في الغالب ما يفترض معد الاختبار أن الأفراد المختبرين سوف يستجيبون على كل مفردة وفقاً لقدرته الحقيقية "أي، أن الفرد سوف يدلي بالإجابة الصحيحة فقط عندما تكون لديه معرفة محدد حول هذه المفردة وبخلاف ذلك سوف يستجيب بشكل خاطئ"، ومع ذلك، هذا الافتراض قد لا يعكس دائماً بشكل صحيح دقة ما يحدث في الممارسة العملية، خاصة في حالة الاختبارات من نوع الاختيار المتعدد، حيث يمكن تصنيف استجابات الأفراد إلى ثلاثة أنواع: (أ): الاستجابات التي تعكس القدرة الحقيقية لدى الفرد المختبر (سواء كانت صحيحة أو خاطئة)، (ب) الاستجابات الصحيحة الناتجة عن التخمين العشوائي، (ج) الاستجابات الخاطئة الناتجة عن القلق أو الإهمال أو اللامبالاة. وقد يتسبب النوعان الأخيران إلى تقدير خاطئ لقدرة الفرد الممتحن الفعلية لأنهما لا يعكسان المعرفة الحقيقية للفرد الممتحن. وبالتالي يمكن القول، بأن تأثير الاستجابات الشاذة في الاختبارات القائمة على النظرية الكلاسيكية للاختبارات محدوداً وذلك بسبب تساوي جميع المفردات في الميزان بشكل عام.

وبما أن البحث السيكومتري قد أهتم بضبط إجراءات وأساليب بناء المقاييس والاختبارات النفسية والتربوية وتحليل مفرداتها والتأكد من صلاحيتها واجراء العمليات الإحصائية المختلفة باستخدام النظرية الحديثة في القياس "نظرية الاستجابة للمفردة الاختبارية -IRT " وما ينبثق عنها من نماذج في محاولة جاهدة للتغلب على عيوب النظرية الكلاسيكية للاختبارات (CTT) والتي سيطرت على منهجيات القياس طوال

القرن الماضي، فالتقدم الناتج عن استخدام نظرية الاستجابة للمفردة في القياس النفسي وما تتمتع به من مزايا في مقابل نظرية القياس التقليدية قد أدى إلى إعادة إجراء كثير من البحوث السيكومترية، ومن منطلق هذا الاتجاه ظهرت الحاجة إلى إجراء هذه البحث.

وفي ضوء مراجعة الأدب التربوي المتعلق بالدراسات التي أجريت حول أثر طريقة ترتيب المفردات على دقة التقدير لقدرات الأفراد ولمعالم المفردات، فإن هذا البحث يأتي استكمالاً لجهود الباحثين حول أثر طرق ترتيب المفردات، وتم استخدام نموذج الاستجابة للمفردة رباعي المعلم وهي من نماذج الاستجابة للمفردة المعلمية أحادية البعد ثنائية الدرجة لمعرفة أثر طريقة ترتيب المفردات على دقة التقدير لقدرات الأفراد ولمعالم المفردات. وقد تم اختيار هذا النموذج لما تتميز به من الواقعية، وملاءمتها لطبيعة مفردات أداة البحث، وإمكانية تنفيذها، وفحص مدى تحقق فرضياتها في البيانات، وسهولة تقدير المعالم، ومنطقية تفسير نتائجه.

مشكلة البحث:

تعتبر إجراءات بناء الاختبار والتي منها ترتيب مفردات الاختبار من الأمور الهامة التي قد تؤثر في خصائص الاختبار السيكومترية وخصائص المفردات، وعلى أداء الأفراد، والتي قد تؤدي إلى خطأ في نتائج أداة القياس.

وعلى الرغم من وجود العديد من الدراسات التي تناولت موضوع أثر طرق ترتيب مفردات الاختبار، إلا أن اغلب هذه الدراسات يظهر بها ثغرة علمية، تتمثل في عدم واقعية الافتراض الضمني الذي قامت عليه، وهو إجابة الطلاب عن المفردات بنفس الترتيب الوارد في الاختبار (عبد الوهاب،٢٠١٩)، فجميع الدراسات السابقة اعتمدت على تطبيق الاختبار بصورة يدوية وليست الكترونية.

فبعض الأساليب الإجرائية تتضمن استخدام تقنيات لا نستطيع تقديمها في الواقع باختبارات الورقة والقلم، وأن دراسة أثر طريقة ترتيب المفردات تتطلب مزيداً من الضبط من الناحية الإجرائية بحيث يقوم المفحوص بالإجابة عن مفردات الاختبار بنفس ترتيبها في الاختبار، وهذا لا يتحقق بدقة عند تطبيق الاختبارات من خلال الصورة الورقية وإنما يمكن ضبطه من خلال التطبيق الإلكتروني للاختبار.

وبما أن الحواسب الالكترونية أسهمت في استكشاف إجراءات ومداخل جديدة في القياس النفسي لدراسة جوانب مختلفة والتي لا يمكن إجرائها بالطريقة التقليدية عند تطبيق الاختبارات والتي تعتمد على "الورقة والقلم" (أناستاسي؛ سوزانا ،٢٠١٥)، ومن هنا ظهرت الحاجة إلى معالجة الثغرة العلمية في الدراسات التي تناولت أثر طرق ترتيب المفردات، من خلال تغيير طريقة تطبيق الاختبار من الطريقة اليدوية إلى الطريقة الإلكترونية التي يمكن من خلالها عرض مفردات الاختبار على المفحوصين بالطريقة التي يريدها الباحث مع إمكانية حرمانهم من العودة إلى المفردات التي قد يتركونها للإجابة عنها في نهاية الاختبار.

وعلى الرغم من انتشار استخدام نظرية الاستجابة للمفردة في شتى مجالات الأبحاث، إلا أنه مع ذلك يوجد اطراد في استخدام إجراءاتها، فأساليب هذه النظرية تم تضمينها سريعاً في الأبحاث السيكومترية التي تتضمن التطبيقات الحديثة لنظرية الاستجابة للمفردة وفي بناء الاختبارات الجديدة.

والبحث العلمي بوجه عام يترتب عليه الحصول على معرفة جديدة، أو تطوير، أو تعديل، أو إضافة معرفة إلى معرفة قائمة، وبناء على ذلك، هناك حاجة إلى إعادة إجراء كثير من البحوث السيكومترية القائمة على النظرية الكلاسيكية للاختبارات وإعادة تقييمها ومراجعتها وفق النظريات الحديثة في القياس النفسي والتربوي.

وهذا جعل هناك حاجة لاستخدام نموذج الاستجابة للمفردة رباعي المعلم (4PL) في هذا البحث، لتدريج مفردات ذات الاستجابة الثنائية، لمعرفة أثر طريقة ترتيب المفردات على دقة تقدير قدرات الأفراد ومعالم المفردات.

وبالتالي فإن مشكلة البحث يمكن صياغتها في التساؤلات الأتية:

- ما مستوى الدلالة الإحصائية للفروق بين متوسطات الأخطاء المعيارية في تقدير قدرات الأفراد المقدرة تعزى لاختلاف طريقة ترتيب المفردات الاختبارية وفق (تسلسل المحتوى، تسلسل الصعوبة المعرفية، الترتيب العشوائي)؟
- ما مستوى الدلالة الإحصائية للفروق بين متوسطات الأخطاء المعيارية في تقديرات معالم صعوبة المفردات تعزى لاختلاف طريقة ترتيب المفردات الاختبارية وفق (تسلسل المحتوى، تسلسل الصعوبة المعرفية، الترتيب العشوائي)؟
- ما مستوى الدلالة الإحصائية للفروق بين متوسطات الأخطاء المعيارية في تقديرات معالم تمييز المفردات تعزى لاختلاف طريقة ترتيب المفردات الاختبارية وفق (تسلسل المحتوى، تسلسل الصعوبة المعرفية، الترتيب العشوائي)؟
- ما مستوى الدلالة الإحصائية للفروق بين متوسطات الأخطاء المعيارية في تقديرات معالم تخمين المفردات تعزى لاختلاف طريقة ترتيب المفردات الاختبارية وفق (تسلسل المحتوى، تسلسل الصعوبة المعرفية، الترتيب العشوائي)؟
- ما مستوى الدلالة الإحصائية للفروق بين متوسطات الأخطاء المعيارية في تقدير معالم اللامبالاة تعزى لاختلاف طريقة ترتيب المفردات الاختبارية وفق (تسلسل المحتوى، تسلسل الصعوبة المعرفية، الترتيب العشوائي)؟

إطار نظري:

أثر ترتيب المفردات على أداء الأفراد

في منتصف الخمسينات، لاحظ نورمان ظهور فرق كبير في الدرجات التي حصل عليها الطلاب في صف علم النفس التمهيدي الذين أخذوا نموذجين مختلفتين من الاختبار، ولقد ابتكر نورمان نموذجين مختلفين فقط في ترتيب المفردات وذلك لتقليل احتمالية الغش، وهي ممارسة لا تزال شائعة لليوم والتي أصبحت ميسرة من خلال توافر برامج شائعة في بناء الاختبارات على نطاق واسع. (Bradburn; (Mason,1964

وتأثير ترتيب مفردة ما يظهر حينما تؤثر استجابة أحد المفردات على المفردة التالية في الاختبار أو المقياس أو على مجموعة من المفردات، ويُدرس هذا النوع من التأثير من خلال فحص التغير في الاستجابات الناتج عن التغيير في موقع المفردة أو التغيير في ترتيب البدائل المتاحة للمستجيبين. (Rose; Nagy; Nagengast; Frey; Becker,2019)

ونقصد بالتأثير وجود تغيرات ذات دلالة إحصائية في مفردات الاختبار عبر صيغ الاختبار المختلفة، وقد يُقدر التأثير كاختلاف في نسبة الإجابات الصحيحة أو في درجة اللوجيت Logit للمفردة، والتي تتغير عبر صيغ وأشكال الاختبار. (Ortenr,2004)

وهناك بعض الأسباب لسلوك الاستجابات المختلفة للأفراد فيما يتعلق بموقع المفردة في الاختبار باختلاف الطريقة المستخدمة للترتيب، وهذا بدوره جعل هناك العديد من التطبيقات لنماذج الاختبار لفحص تأثير ترتيب المفردات على الأداء ضمن التقييم التربوي والنفسي، ومن الجدير أن نأخذ ذلك بالحسبان عند التطرق لهذا الموضوع:

فبالإمكان ترتيب المفردات وفق سهولة المفردة وصعوبتها، وفي حالة ترتيب المفردات من السهل للصعب بمعنى، أن المفردات التي يتم تقديمها في بداية الاختبار تكون في الغالب أسهل من ناحية صعوبتها من تلك التي تقدم في نهاية الاختبار؛ فإن هذا الترتيب يوفر الدافعية لدى الفرد للاستمرار في محاولاته للإجابة عندما يشعر بالتعزيز الفوري بسبب قدرته على الإجابة عن الأسئلة الأولى من الاختبار (الحجاج،٢٠١٣). أما في حالة ترتب المفردات من الصعب إلى السهل بمعنى، أن المفردات التي قدمت في بداية الاختبار قد تكون أكثر صعوبة من تلك التي قدمت في نهاية الاختبار؛ هذا النوع من الترتيب قد يؤدي إلى تكوين اعتقاد أو فكرة لدى الفرد المختبر بأن هناك عموماً إفراط في صياغة المفردات، أو أنها تميل إلى أن تكون غير معقولة وبالتالي تؤثر في سلوك استجابات الأفراد، وقد يكون الفرد المختبر أكثر عرضة للشعور بالتعب أو الملل باتجاه النهاية، والتي قد تسبب بعض النزعات في سلوك الاستجابة، كاللجوء إلى التخمين العشوائي. (Vander Schee,2013)

ومن الأساليب أيضاً لترتيب المفردات وفق السهولة والصعوبة، الترتيب الدوري بالنسبة للصعوبة؛ حيث ترتب الأسئلة بحيث تتدرج من السهل إلى الصعب، ثم تعود بالتدريج من الصعب إلى السهل مرة أخرى، فيبدأ الاختبار وينتهي بأسئلة سهلة، بينما توضع الأسئلة الصعبة في وسط الاختبار تقريباً، وبذلك نشجع التلميذ على المضي في الإجابة وقراءة جميع الأسئلة في الاختبار لتوقعه ان الأسئلة ستتدرج مرة أخرى نحو السهولة. (هاشم،٢٠٠٦)

وقد يتم ترتيب المفردات بنفس التسلسل المنطقي لمحتوى المادة الدراسية؛ وبذلك يتوفر للممتحن فرصة للتركيز الجيد في متابعة المفردات والإجابة عنها بيسر وسلاسة، وتبين أن هذا الترتيب يساعد في عملية التشخيص من خلال معرفة نقاط القوة ونقاط الضعف في استجابات الطلبة وفق المواضيع المختلفة؛ كما تعمل هذه

الطريقة على توجيه تفكير الفرد المختبر وحصر انتباهه حول موضوع محدد وبالتالي تجنبه التشتت في الاستجابة على المفردات الخاصة بموضوع محدد. (بركات،٢٠١٨)

وهناك وجهة نظر تكاملية لترتيب مفردات الاختبار؛ فبالإمكان ترتيب مفردات الاختبار بدمج مميزات مختلفة من وجهات النظر السابقة، بحيث يُصار إلى ترتب مفردات الاختبار مثلاً، وفق تسلسل موضوعها من المحتوى وبنفس الوقت وفق شكلها وصعوبتها. (بركات،٢٠١٨)

نظرية الاستجابة للمفردة الاختبارية

لقد حظيت هذه النظرية باهتمام علماء القياس المعاصر والمؤسسات العالمية المعنية ببناء الاختبارات والمقاييس، وقد تعددت أساليب بناء الاختبارات والمقاييس النفسية والتربوية وطرق تحليل البيانات المستمدة منها وتفسيرها بناءً على الأسس النظرية والنماذج المستخدمة، والآن هي الإطار النظري الرئيسي المستعمل في هذه البحث العلمية.

ونظرية الاستجابة للمفردة هي نظرية إحصائية عامة حول المفردة الاختبارية وأداء الفرد المفحوص على الاختبار وكيفية ارتباط الأداء بالقدرات التي تقيسها مفردات الاختبار. وهي نموذج للقياس العقلي مفاده أن الاستجابات للمفردات في اختبار معين تُعزى إلى سمات كامنة. والسمة الكامنة هي قدرة أو خاصية لفرد معين يُستدل عليها اعتماداً على نظريات للسلوك وكذلك على أدلة إمبريقية، وتفترض أن كل فرد مختبر يمتلك مقداراً معيناً من أي سمة كامنة معلومة، وأن تقديرها لا يعتمد على أي مجموعة معينة من المفردات أو أي إجراء تقييمي. ويمكن أن يكون هناك قدرة واحدة أو العديد من القدرات الكامنة خلف الأداء على الاختبار. (أناستاسي؛ سوزانا،٢٠١٥؛ حبيب، عزيز،٢٠١٨)

وتقدم هذه النظرية عدد من المزايا الممكنة أكثر من النظرية الكلاسيكية للاختبار في تقييم نتائج الاختبارات والمقاييس ونذكر من هذه المزايا:(Hambleton; Swaminathan,1985)

عدم تغير معالم المفردات

أن الاسهام الأكثر شيوعاً لنماذج نظرية الاستجابة للمفردة يتعلق بنتائجها المتحررة من العينة. ويُوصف ذلك من الوجهة الفنية "بالمعالم غير المتغيرة للمفردة".

فمن المفاهيم الأساسية لهذه النظرية أن معالم المفردة ينبغي أن تكون غير متغيرة عند حسابها من مجموعات تختلف في قدرتها، وهذا يعني أنه يمكن تقديم ميزان منتظم للقياس يمكن استخدامه مع مجموعات مختلفة من المختبرين. بمعنى، أنه يمكن اختبار مجموعات وكذلك افراد بمجموعات مختلفة من المفردات الاختبارية المناسبة لمستويات قدرتهم، وسوف تكون درجاتهم قابلة للمقارنة جزئية من المفردات الكلية التي تمت مطابقتها للنموذج.

الدرجات التي تعتمد على نظرية الاستجابة للمفردة

المعيار الذي تقاس عليه الدرجة وفق هذه النظرية هي المفردات التي يتضمنها الاختبار، ويعني ذلك أنها تعتمد على درجات المفردات لا على درجات الاختبار، حيث تتم المقارنة بين مستويات قدرة الفرد، ومعالم المفردات والتي يتحدد موقعها على متصل واحد بالاعتماد على ميزان مشترك (Common Scale)، وبالتالي يمكن تفسير الفرق بين مستوى قدرة الفرد، ومعالم المفردة تفسيراً مباشراً.

الثبات في نظرية الاستجابة للمفردة والأخطاء المعيارية

من المزايا المهمة التي تقدمها نظرية الاستجابة للمفرد، أنها تعالج قضايا الثبات، وخطأ القياس من خلال دالة المعلومات للاختبار.

نظراً لأن دالة المعلومات تتباين بتباين مستوى الدقة، فإن خبراء القياس اقترحوا أن منحنيات معلومات الاختبار تكون بديلاً لتقديرات الثبات المستخدمة في نظرية الكلاسيكية للاختبارات والتي تُقدم تقديراً واحداً فقط للثبات، وللأخطاء المعيارية للقياس المستخدمة في تفسير درجات الاختبارات. (حبيب؛ عزيز،٢٠١٨)، حيث توضح هذه الدالة ثبات القياس عند نقاط مختلفة على مدى التوزيع، بمعنى، أنها تقدم معلومات عن الثبات عند نقاط مختلة على مدى التوزيع، ويعني ذلك أنها توفر تقدير للثبات لكل فرد.

تقدير قدرة الأفراد في النماذج أحادية البعد

يعد تقدير قدرات الأفراد من المكونات الأساسية في الاختبارات، حيث إن جميع استراتيجيات نظرية الاستجابة للمفردة لتقدير الدرجة الاختبارية تحاول تقدير موقع الفرد على متصل السمة وذلك باستعمال أنماط استجابة الفرد للمفردة بالتزامن مع تقديرات معالم المفردة، وهي تعتبر الخطوة الأولى والأكثر أهمية لاستخدام نظرية الاستجابة للمفردة في تصميم الاختبارات والمقاييس وتحليل مفرداتها تتعلق بتقدير المعلمات التي يشتمل عليها النموذج الذي يتم اختياره.

والاستخدام الفاعل لنموذج الاستجابة للمفردة يعتمد على مدى توفر أساليب وإجراءات مناسبة لتقدير هذه المعالم، وهذه الأساليب تهدف لتحديد قيمة القدرة (θ) لكل فرد مُختبر، وكذلك قيم معالم المفردة، من إجابات الأفراد عليها. وهنالك العديد من الأساليب والمنهجيات التي ابتكرها علماء القياس لتيسير تقدير معالم نماذج الاستجابة للمفردة. وتتباين هذه الأساليب في تحقيقها للمحكات التي تستند إليها، في تقييم مدى جودة أسلوب إحصائي معين في المجتمع المستهدف. ومن أهم هذه المحكات: (عدم التحيز، والاتساق، والفعالية النسبية، والكفاية). (علام،٢٠٠٥)

نموذج الاستجابة للمفردة رباعي المعلمة

هذا النموذج يختلف في صيغته عن النموذج ثلاثي المعلم (3PL) في أنه يفترض وجود معلم رابع يؤثر في الاستجابة عن المفردة، وهذا المعلم يأخذ في الحسبان أن هناك احتمالية لعدم استجابة الأفراد من ذوي مستوى القدرة المرتفعة على المفردات إجابة صحيحة أحياناً حتى على المفردات السهلة. وربما كان الأفراد ذوي مستوى المرتفع من القدرة يعتمدوا (اللامبالاة) في اختيار الإجابة، أو في الأوقات الأخرى قد يكون لديهم معلومات غير التي يفترضها مُعد مفردة الاختبار لذلك قد يختار الفرد إجابة غير مرتبطة بالإجابة الصحيحة، أو قد يرتكب الفرد خطأ كتابياً في الإجابة عن المفردات المقالية، أو ربما يعود عدم إجابة الفرد إجابة صحيحة للمفردة إلى صعوبة المفردة.

ويشير المنطق وراء إضافة المعلم الرابع إلى أن الأشخاص ذوي القدرات المرتفعة، قد يرتكبون خطأ في الإجابة على مفردات سهلة نسبياً، لذلك ينخفض منحنى الاستجابة للمفردة بالنهاية العليا والتي تساوي (١) (De Ayala,1993). (ويعني ذلك، أن الأفراد من ذوي القدرة المرتفعة لا يستجيبون لمفردات الاختبار بإجابة صحيحة).

وهذا النموذج يسمح بأن تختلف المفردات في معالم الصعوبة للمفردة (b_i)، والقدرة التمييزية للمفردة (a_i)، ومعلمة التخمين (c_i)، إضافة إلى معلمة اللامبالاة (الحد التقاربي الأعلى d_i)

دراسات السابقة

هدفت دراسة فاندر شي (Vander Schee,2013) لمعرفة تأثير ترتيب المفردات وفق مستوى الصعوبة على أداء الأفراد. حيث تم فحص الدرجات لاختبار من نوع الاختيار من متعدد في مقرر مبادئ التسويق للتحقق من أداء الطلاب في الاختبارات التي تم ترتيب مفرداتها بناء على مستوى صعوبة المفردة. وتم وضع ثلاث نماذج اختبارية،

في النموذج الأول الذي تم تقديمة يحتوي على المفردات الصعبة أولاً (١٧ مفردة) يليها المفردات ذات المستوى المتوسط من الصعوبة (١٦ مفردة) وأخيراً المفردات السهلة (١٧ مفردة). وشمل النموذج الثاني نفس المفردات، ولكن بترتيب عكسي وهي سهلة، متوسطة، صعبة. وتضمن النموذج الثالث أيضاً نفس المفردات، ولكن تم ترتيبها بشكل عشوائي بغض النظر عن مستوى الصعوبة للمفردات. وللإجابة على الأسئلة تم استخدام تحليل التباين لتحديد ما إذا كان أداء الطلاب في اختبار مبادئ التسوق يختلف اختلافاً كبيراً نتيجة التفاعل بين ترتيب مفردات الاختبار بالنسبة لمستوى الصعوبة والجنس أو التخصص. وتشير النتائج إلى عدم وجود فروق دال احصائياً في أداء الطلاب يعود إلى ترتيب مفردات الاختبار وتفاعله مع الجنس أو التخصص. وتوصلت البحث بأن الترتيب بشكل عشوائي لصعوبة المفردات لا يؤثر على الأداء للأفراد بناء على التخصص والجنس.

وقد أجري الشريفين (٢٠١٣) دراسة بهدف الكشف عن أثر موقع المفردات الصعبة في اختبار من نوع الاختيار من متعدد على الخصائص السيكومترية وأداء الطلبة، ولتحقيق ذلك تم بناء اختبار تحصيلي وإعداد أربع نماذج من الاختبار، تختلف في موقع المفردات الصعبة لكل نموذج. ونتائج هذه البحث أوضحت وجود فروق ذات دلالة إحصائية بين متوسطات الصعوبة ومتوسطات التمييز فقط بين النموذجين والذي وضعت المفردات الصعبة ببداية الاختبار والنموذج الذي تم ترتيب مفرداته بشكل عشوائي لصالح النموذج المرتب بشكل عشوائي، وعدم وجود فروقات بين النماذج الأخرى. وبينت النتائج وجود فروق ذات دلالة إحصائية بين متوسطي درجات أداء الطلبة على الاختبار التحصيلي فقط بين النموذجين الذي تم وضع المفردات الصعبة ببداية الاختبار والنموذج المرتب بشكل عشوائي؛ وعدم وجود فروقات بين النماذج الأخرى.

وفي دراسة حول أثر ترتيب مفردات الاختبار على كلاً من أداء الطلاب، وإدراكهم لدرجة صعوبة الاختبار، ومستوى القلق لديهم. طبق كل من بيتجون وساكو (Pettijohn; Sacco,2007) ثلاثة نماذج من نفس الاختبار على مجموعة من الطلاب. ورتبت المفردات في النماذج الثلاثة وفق تسلسل المحتوى (تصاعدياً، تنازلياً، عشوائياً) على الترتيب. إضافة لذلك أجاب كل طالب بعد إنهائه للاختبار على مقياس لدرجة القلق. كما حسب الباحثان الزمن الذي استغرقه كل طالب في الإجابة. وأظهرت النتائج بعدم وجود فروق دالة في درجات الطلاب تعزى لنموذج الاختبار، أو الزمن المستغرق في الإجابة. ولكن ظهرت فروق في درجة إدراك الطلاب لصعوبة الاختبار، حيث أشار الطلاب إلى أن أصعب نموذج هو النموذج ذو الترتيب العشوائي، بينما كان النموذجان الآخران بنفس درجة الصعوبة.

ومن اهم الدراسات التي تطرقت لترتيب المفردات دراسة بلاتش (Blach,1989)، حيث طبق ثلاثة نماذج من اختبارية ورتب المفردات وفق تسلسل الوحدات الدراسية ضمن كل وحدة في النموذج الأول، ورتب المفردات بطريقة عشوائية داخل كل وحدة على حدة في النموذج الثاني، أما في النموذج الثالث رتبت المفردات بطريقة عشوائية بغض النظر عن الوحدة الدراسية. وأظهرت النتائج فروقاً دالة في درجات الطلاب تعزى لنموذج الاختبار وكان أفضل أداء في النموذج الأول، مما يدل على أن الترتيب التسلسلي للمفردات وفق الوحدات الدراسية هو الأفضل.

ومن الدراسات أيضاً دراسة سو (Sue,2009) والتي درست أثر ترتيب مفردات الاختبار على أداء الطلاب، في مقرر مبادئ الإدارة، وكان عدد الملتحقين بالمقرر (أقل من ٣٠ طالب). تم إعداد نموذجين من الاختبار مختلفين من حيث ترتيب المفردات حيث رتبت النموذج الأول وفق تسلسل المحتوى، بينما رتبت النموذج الثاني بطريقة عشوائية. وقد أظهرت نتائج البحث بعد اجراء اختبار مربع كاي: بعدم وجود فروق دالة احصائياً في أداء الطلاب تعزي لترتيب المفردات، مما يدل حسب رأيها

إمكانية المعلم إعداد عدة نماذج من الاختبار بإعادة ترتيب المفردات، دون أن يخشى من وقوعه في التحيز.

وأظهرت بعد ذلك دراسة ناقضت ما ورد في دراسة سو (Sue,2009)، حيث أجرى دورنر وكالهون (Doerner; Calhoun,2009) دراسة حول أثر ترتيب المفردات وفق تسلسل المحتوى على تحصيل الطلاب. وقام بتطبيق: ثلاثة نماذج من اختبارية، وتم ترتيب المفردات في النماذج الثلاثة تبعاً للمحتوى (تصاعدياً، تنازلياً، عشوائياً) على الترتيب. وقد توصلت نتائج البحث بوجود فروق دالة إحصائياً في درجات الطلاب على الاختبار، تعزى لترتيب المفردات. فقد كانت أعلى الدرجات على النموذج الأول، ثم الثاني، ثم الثالث.

إجراءات البحث

منهج البحث

تم الاعتماد على المنهج شبة التجريبي، حيث تم اختيار عينة البحث بأحد أساليب المعاينة العشوائية، والتحكم بالمتغير المستقل عن طريق توزيع النماذج المختلفة للاختبار بالطريقة العشوائية على عينة البحث.

مجتمع البحث: تكون مجتمع البحث من طالبات المرحلة المتوسطة من الصف الثالث متوسط في المدارس الحكومية التابعة للإدارة العاملة للتعليم بمنطقة مكة المكرمة للعام الدراسي (١٤٤٣-١٤٤٤) هـ والبالغ عددهم (١٣٢٥١) طالبة موزعين على (خمس مكاتب للتعليم) حسب إحصائية الإدارة العامة للتعلم بمنطقة مكة المكرمة.

عينة البحث: تكونت العينة من (٣٥٤١) طالبة بالمرحلة المتوسطة من الصف الثالث متوسط. تم اختيارهم بالطريقة الطبقية العشوائية خلال الفصل الدراسي الثاني من العام الدراسي ١٤٣٤/ ١٤٤٤ هـ.

أداة البحث: أداة البحث تمثلت في اختبار تحصيلي في مقرر التفكير الناقد للصف الثالث متوسط (الفصل الدراسي الثاني). وتكون الاختبار من (٥١) مفردة من نوع الاختيار من متعدد، وقد تم تكوين ثلاث نماذج اختبارية بصورة إلكترونية مختلفة فيما بينها فقط في طريقة ترتيب المفردات، بحيث رتبت المفردات في الصورة الأولى وفق تسلسل المحتوى وفي الثانية وفق تسلسل الصعوبة المعرفية وفي الثالثة بطريقة عشوائية.

التطبيق التجريبي للنماذج الاختبارية الثلاث والتي تختلف فيما بينها في طريقة ترتيب المفردات (عينة التحقق من الخصائص السيكومترية): تم التطبيق النماذج الاختبارية الثلاثة على عينة تجريبية مكونة من (٣٥٩)، للتحقق من صدق وثبات درجات الاختبار. حيث تم التحقق من صدق الدرجات المتنوعة التي تم الحصول عليها من عينة التحقق من الخصائص السيكومترية، من خلال التحليلات الإحصائية في ضوء محكات خارجية واقعية يتضمن مزيج من اجراء يعتمد على التحكيم واجراء إحصائي، وذلك من خلال ما يأتي:

صدق المحكمين: وللتحقق من صدق الاختبار تم عرض المفردات والأهداف التعليمة ومستويات الأهداف السلوكية التي تقيسها وجدول مواصفات الاختبار على مجموعة من المحكمين وهم: ثلاث معلمات ممن يدرسون مقرر التفكير الناقد للصف الثالث متوسط أحدهما تحمل شهادة الماجستير في تخصص مناهج وطرق التدريس، ومشرفة تربوية، وأثنان من أعضاء هيئة التدريس بالجامعة تخصص المناهج وطرق التدريس، وثلاثة من أعضاء هيئة تدريس بالجامعة تخصص قياس وتقويم. وأخذت ملاحظاتهم في الاعتبار فقد تم التعديل في مستويات الأهداف السلوكية وتعديل جدول المواصفات واستبدال عدد من المفردات وتعديل صياغة بعض المشتتات لبعض المفردات.

صدق المحك: ولحساب صدق الاختبار المرتبط بمحك حسب معامل ارتباط بيرسون بين درجات الطالبات على النماذج الاختبارية المختلفة فيما بينها في طريقة ترتيب

مفرداتها ودرجات التحصيل الدراسي للطالبات في مقرر التفكير الناقد للفصل الدراسي الأول، والتي كانت متوفرة عند تطبيق الاختبار على عينة التحقق من الخصائص السيكومترية، وكانت قيم معامل الارتباط كما في جدول (١).

جدول (١): معاملات ارتباط درجات الطالبات في النماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها ودرجات التحصيل الدراسي في مقرر التفكير الناقد للفصل الدراسي الأول

نموذج ترتيب مفردات الاختبار	الترتيب وفق تسلسل المحتوى	الترتيب وفق تسلسل الصعوبة المعرفية	الترتيب العشوائي
الدرجات في الفصل الدراسي الأول	٠,٩٢٠	٠,٨٩٣	٠,٨٠٦

ويتضح من جدول (١) أن معاملات الارتباط للنماذج الاختبارية الثلاث المختلفة في طرق ترتيب مفرداتها تراوحت بين (٠,٩٢٠: ٠,٨٠٦) وهي قيم مرتفعة تدل على الصدق التلازمي للنماذج الاختبارية الثلاثة المختلفة في طريقة ترتيبها.

تقدير ثبات درجات الاختبار:

تم حساب معامل ثبات درجات الاختبار باستخدام (معادلة كيودر ريتشاردسون-٢٠ (KR_20)) من خلال برمجية (SPSS Version 26)، للنماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها، وكانت قيم معاملات ثبات درجات الاختبار كما في جدول (٢).

جدول (٢): معاملات ثبات الدرجات للنماذج الاختبارية المختلفة في طريقة ترتيب مفرداتها باستخدام معادلة كيودر ريتشاردسون–٢٠ بعد مطابقة بيانات المفردات لافتراضات نموذج الاستجابة للمفردة رباعي المعلم

نموذج ترتيب مفرد	الترتيب وفق تسلسل المحتوى	الترتيب وفق تسلسل الصعوبة المعرفية	الترتيب العشوائي
معامل الثبات للدرجات الاختبارية KR_20	٠,٩٢٩	٠,٩٢٥	٠,٩٠٩
	٠,٩٠٤	٠,٨٤٥	٠,٨٣٨

ويتضح من جدول (٢) أن جميع قيم معاملات ثبات درجات الاختبار مقبولة للدلالة على ثبات الدرجات للنماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها.

التحقق من مطابقة البيانات لافتراضات نموذج الاستجابة للمفردة أحادية البعد رباعي المعلم، كما يأتي:

أ) التحرر من السرعة

تم تقدير الزمن الأولي للاختبار بعد التجريب على عينة التحقق من الخصائص السيكومترية للاختبار، حيث أعطي الطالبات الوقت الكافي للاستجابة على مفردات الاختبار، وكان زمن (٦٠) دقيقة كافي للإجابة على جميع مفردات الاختبار، وقد أجاب جميع الأفراد على مفردات الاختبار ضمن الزمن المحدد.

ب) التحقق من أحادية البعد

استخدام التحليل العاملي الخطي: تم إجراء أسلوب التحليل العاملي الخطي باستخدام برمجية (SPSS Version 26) بطريقة المكونات الرئيسة (Principal Components) للتحقق من أحادية البعد للبناء العاملي.

تم حساب معاملات ارتباط بيرسون بين المفردات للنماذج الاختبارية الثلاث، ثم حللت عاملياً بطريقة المكونات الأساسية لهوتلنج (Hoteling) وتم اتباع معيار جتمان لتحديد عدد العوامل، حيث يُعد العامل جوهرياً إذا كان جذره الكامن "> (١)"، كما أن محك جيلفورد لجوهرية تشبع الفقرات بالعوامل إذا كانت معاملات الارتباط "≥ (٠,٣٠)". ويوضح جدول (٣). خلاصة نتائج التحليل العاملي الاستكشافي لبيانات النماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها والمكونة من (٥١) مفردة.

جدول (٣): خلاصة نتائج التحليل العاملي الاستكشافي لبيانات النماذج الاختبارية الثلاث المختلفة في طريقة ترتيبها والمكونة من (٥١) مفردة

نموذج ترتيب مفردات الاختبار	رقم العامل	قيمة الجذر الكامن	نسبة التباين المفسر	نسبة التباين المفسر التراكمي	نسبة قيمة هاتي	محك ريكاس
	١	١١,٥٠٤	٢٢,٥٥٨	٢٢,٥٥٨	٦٥,١١٧	٣,٨٩٧٠
الترتيب وفق	٢	٢,٩٥٢	٥,٧٨٩	٢٨,٣٤٧		
تسلسل المحتوى	٣	١,٧٣٦	٣,٤٠٣	٣١,٧٥٠		
	٤	١,٤٧٦	٢,٨٩٣	٣٤,٦٤٣		
	١	٩,٤٩٧	١٨,٦٢١	١٨,٦٢١	٥١,٢٠٥٨	٢,٥٤٢٤
الترتيب وفق	٢	٣,٧٣٤	٧,٣٢٤	٢٥,٩٤٥		
تسلسل	٣	٢,٧٥٠	٥,٣٩٢	٣١,٣٣٨		
الصعوبة	٤	٢,٥٦٤	٥,٠٢٨	٣٦,٣٦٥		
المعرفية						
	١	٦,٢٦٩	١٢,٢٩٢	١٢,٢٩٢	٤٦,١٤١١	٢,٧٩
الترتيب	٢	٣,٠٨٩	٦,٠٥٧	١٨,٣٤٩		
العشوائي	٣	٢,٢٣٧	٤,٣٨٦	٢٢,٧٣٥		
	٤	١,٩٩١	٣,٩٠٤	٢٦,٦٤٠		

يتضح من جدول رقم (٣) أن مصفوفة الارتباطات بين العوامل استخلصت (٤) عوامل جذرها الكامن أعلى من الواحد لجميع النماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها، وقد تم استخدام محك هاتي ومحك ريكاس للحكم على أحادية البعد في النماذج المختلفة، حيث يعد الاختبار أحادي البعد وفق محك هاتي، إذا كانت قيمة التباين الذي يعود للعامل الأول على قيمة التباين الكلي الذي يعود للعوامل أكبر من ٢٠%، ووفق محك ريكاس يعد الاختبار أحادي البعد، عندما تكون قيمة قسمة الجذر الكامن للعامل الأول على قيمة الجذر الكامن للعامل الثاني أعلى من القيمة ٢.

ويلاحظ من الجدول أن النسبة زادت عن (٢٠%) من التباين في جميع النماذج الاختبارية الثلاث، وكذلك نسبة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني في جميع النماذج الاختبارية المختلفة في طريقة ترتيبها زادت عن (٢)، وهذا يعتبر مؤشر على أحادية البعد للاختبار.

ج) التحقق من الاستقلال الموضعي لمفردات الاختبار:

استخدام طريقة الارتباطات البينية: تم حساب معاملات الارتباط بين درجات المفردات بعضها بعضاً في النماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها، وذلك لكامل أفراد المجموعة، ولأفراد المجموعة العليا في المجموعات الثلاث وعددهم (٣١٩،٢٧٨،٣٢٣) طالبة على الترتيب، ولأفراد المجموعة الدنيا في المجموعات الثلاث وعددهم (٣٧٤،٢٧٤،٢٨٩) طالبة على الترتيب، ثم تم تحديد أقل معامل ارتباط وأعلى معامل ارتباط في كل نموذج من النماذج الاختبارية الثلاث، كما تم حساب متوسط معاملات الارتباط بين المفردات في كل النماذج الاختبارية الثلاث، كما في جدول (٤).

جدول (٤): أقل وأعلى ومتوسط معاملات الارتباط بين مفردات النماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها

نموذج	أعلى معامل ارتباط			أقل معامل ارتباط			متوسط معاملات الارتباط		
ترتيب مفردات الاختبار	العينة ككل	العليا	الدنيا	العينة ككل	العليا	الدنيا	العينة ككل	العليا	الدنيا
الترتيب وفق تسلسل المحتوى	٠,٦٢٤	٠,٥١٧	٠,٢٨٢	٠,٣٠٠	٠,٢١٦	٠,٢٣٨	٠,١٩٤	٠,٠٢٣	٠,٠٠٥
الترتيب وفق تسلسل الصعوبة المعرفية	٠,٦٨٧	٠,٥٤٨	٠,٢٦٦	٠,٣٨٢	٠,٣٠٠	٠,١٩٩	٠,١٥٠	٠,٠٢٢	٠,٠٠٥
الترتيب العشوائي	٠,٥٧٦	٠,٣٧٠	٠,٥٥٣	٠,٣٣٧	٠,٢٠٢	٠,٢٠٥	٠,٠٩١	٠,٠٢٩	٠,٠٠٧

ويتضح من جدول (٤) أن متوسط معاملات الارتباط بين المفردات في حالة المجموعتين العليا والدنيا أقل منه في حالة العينة ككل في النماذج الثلاثة المختلفة في طريقة ترتيب مفرداتها، كما أن إلى حد ما معظم متوسطات معاملات الارتباط قريبة من الصفر، وذلك في جميع الحالات، مما يعنى تحقق شرط الاستقلال الموضعي للمفردات (Önder, 20 .07)

التحقق من مطابقة استجابات الأفراد لافتراضات نموذج الاستجابة للمفردة رباعي المعلم

بعد إدخال البيانات إلى برنامج (RStudio Version 3.5.0)، تم إخضاعها لتحليل المطابقة الإحصائية للأفراد وفق نموذج نظرية الاستجابة للمفردة رباعي المعلم بالاستناد إلى إحصائيتين كما يأتي:

١-إحصائية ليفين وروبن لمطابقة الأفراد للنموذج

٢-إحصائية دراجزو لمطابقة الأفراد للنموذج

وأظهرت نتائج الدرجة القاطعة لإحصائية ليفين وروبن (L_0)، وكذلك مستوى الدلالة الإحصائية لهذا المؤشر وفق إحصائية دراجزو (L_z)، عدم مطابقة استجابات عدد من الأفراد لافتراضات نموذج الاستجابة للمفردة رباعي المعلم في النماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها. وكان عدد الأفراد غير المطابقين للنماذج المرتبة وفق (تسلسل المحتوى ١٣، تسلسل الصعوبة المعرفية ١٩، الترتيب العشوائي ٢٨) فرد.

التحقق من مطابقة المفردات لنموذج الاستجابة للمفردة رباعي المعلم

تم إعادة التحليل للنماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها بعد حذف استجابات الأفراد غير المطابقة للتحقق من مدى مطابقة مفردات الاختبار لنموذج الاستجابة للمفردة رباعي المعلم باستخدام برنامج (Exametrika Version 5.5)، وأظهرت نتائج التحليل في المرة الثانية والخاصة بمطابقة المفردات لنموذج الاستجابة للمفردة رباعي المعلم عدم مطابقة عدد من المفردات المرتبة وفق (تسلسل المحتوى خمس من المفردات، تسلسل الصعوبة المعرفية ستة من المفردات، الترتيب العشوائي ثلاث من المفردات).

وبالتالي تم حذف المفردات والإبقاء على (٤٦، ٤٥، ٤٨) مفردة وفق نموذج الاستجابة للمفردة رباعي المعلم للنماذج الاختبارية الثلاث المختلفة في طريقة ترتيب

مفرداتها على التوالي وفق الترتيب (تسلسل المحتوى، تسلسل الصعوبة المعرفية، الترتيب العشوائي).

تحديد الإحصاءات الخاصة بمعالم قدرة الأفراد ومعالم المفردات للنماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها وفق نموذج الاستجابة للمفردة رباعي المعلم:

تم تقدير كلاً من معالم المفردات وفق نموذج الاستجابة للمفردة رباعي المعلم (الصعوبة (b_i)، التمييز (a_i)، التخمين (c_i)، اللامبالاة (d_i))، ومعالم القدرة للأفراد وفق نموذج الاستجابة للمفردة رباعي المعلم باستخدام طريقة الأرجحية القصوى الهامشية (MML) باستخدام برنامج (Exametrika Version 5.5) للنماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها.

وكانت قيم معالم المفردات للنماذج الاختبارية الثلاث كما يلي:

نموذج الترتيب وفق تسلسل المحتوى: أن قيم معلم صعوبة المفردة تتراوح من) – ١,٣٦٢٦) إلى (١,٠٠٢٠)، بمتوسط حسابي (٠,١٠٨٧) وانحراف معياري (٠,٤٨٦٣). وأما القوة التمييزية للمفردات فكانت جميعها موجبة، وتتراوح من (٠,٦٦٢٨) إلى (٢,٨٢٦٦)، بمتوسط حسابي (١,٣٥١٤) وانحراف معياري (٠,٦٠١٧). وتراوحت قيم معلم التخمين من (٠,٠٨٠٨) إلى (٠,٣٦٥٢)، بمتوسط حسابي (٠,٢١٣٦) وانحراف معياري (٠,٠٦٤٩). وقيم معلم اللامبالاة تراوحت من (٠,٧٧٤٣) إلى (٠,٩٨٥٤)، بمتوسط حسابي (٠,٩٢٣٧) وانحراف معياري (٠,٠٥٦٦). وبالنسبة لمقدار معلومات المفردات فتتراوح من (٠,٢١) إلى (٤,٣٠)، بمتوسط حسابي (١,١٩٨) وانحراف معياري (٠,٨٨١٦).

نموذج الترتيب وفق تسلسل الصعوبة المعرفية: أن قيم معلم صعوبة المفردة تتراوح من (–١,١٠٥٣) إلى (١,١٠٦٩)، بمتوسط حسابي (٠١٤٤٨) وانحراف معياري (٠,٥٣٣٧). وأما القوة التمييزية للمفردات فكانت جميعها موجبة، وتتراوح من

(٠,٤٢٣٣) إلى (٢,٨١٢)، بمتوسط حسابي (١,٤٣٠٨) وانحراف معياري (٠,٧٨٣٢). وتراوحت قيم معلم التخمين من (٠,٠٩٣١) إلى (٠,٣٦٣٢)، بمتوسط حسابي (٠,٢٣٣٦) وانحراف معياري (٠,٠٧٢٨). وقيم معلم اللامبالاة تراوحت من (٠,٧٠٠) إلى (٠,٩٩٩٠)، بمتوسط حسابي (٠,٩١١٨) وانحراف معياري (٠,٠٨٧٧). وبالنسبة لمقدار معلومات المفردات فتتراوح من (٠,١٢) إلى (٣,٩٧)، بمتوسط حسابي (١,١٩١) وانحراف معياري (٠,٩٣٩١).

نموذج الترتيب العشوائي: أن قيم معلم صعوبة المفردة تتراوح من (-٢,٤٢٥٦) إلى (٣,١٧٦١)، بمتوسط حسابي (٠,٨٤٨٩) وانحراف معياري (١,٠٠١٦). وأما القوة التمييزية للمفردات فكانت جميعها موجبة، وتتراوح من (٠,٤٠٢٩) إلى (٢,٣٨٤١)، بمتوسط حسابي (١,٢٤٤٩) وانحراف معياري (٠,٦٨٣٠). وتراوحت قيم معلم التخمين من (٠,٠٥١٧) إلى (٠,٣٦٢٦)، بمتوسط حسابي (٠,٢٢٤٣) وانحراف معياري (٠,٠٧٧٢). وقيم معلم اللامبالاة تراوحت من (٠,٧٠٠٠) إلى (٠,٩٨٢٤)، بمتوسط حسابي (٠,٨٩٨١) وانحراف معياري (٠,٠٨٢٥). وبالنسبة لمقدار معلومات المفردات فتتراوح من (٠,١١) إلى (٢,٥٠)، بمتوسط حسابي (٠,٩٠١٧) وانحراف معياري (٠,٦٩٥٠١).

نتائج البحث وتفسيرها

- **السؤال الأول:**

ينص السؤال الأول على: "ما مستوى الدلالة الإحصائية للفروق بين متوسطات الأخطاء المعيارية في تقدير قدرات الأفراد المقدرة تعزى لاختلاف طريقة ترتيب المفردات الاختبارية وفق (تسلسل المحتوى، تسلسل الصعوبة المعرفية، الترتيب العشوائي)؟"

وللإجابة عن هذا السؤال تم تقدير معالم قدرات الأفراد للنماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها باستخدام طريقة الأرجحية القصوى الهامشية (MML) وفق نموذج الاستجابة للمفردة رباعي المعلم باستخدام برنامج (Exametrika Version 5.5)، وقد تم تقدير الخطأ المعياري في تقدير القدرة لكل فرد، ثم تم استخدام تحليل التباين أحادي الاتجاه لمعرفة أثر كل من طريقة ترتيب المفردات الاختبارية على الأخطاء المعيارية في تقدير معالم قدرات الأفراد، وكانت النتائج كما في جدول (٥) وجدول (٦).

جدول (٥): الأوساط الحسابية والانحرافات المعيارية للأخطاء المعيارية في تقدير معالم قدرات الأفراد وفقاً لطرق ترتيب المفردات

طريقة ترتيب المفردات الاختبارية	نوع النموذج	
	رباعي المعلم	
	الوسط الحسابي	الانحراف المعياري
وفق تسلسل المحتوى	٠,٢٩٢١	٠,١٢٨٦
وفق تسلسل الصعوبة المعرفية	٠,٣٠٣٣	٠,١٢٠٤
الترتيب العشوائي	٠,٣٥٥٥	٠,١٢٨٧

ويتضح من جدول (٥) أن أقل وسط حسابي للأخطاء المعيارية للنماذج الاختبارية الثلاث والمختلفة في طريقة ترتيب مفرداتها هي للنموذج الاختباري المرتب وفق تسلسل المحتوى، والتي تدل على الدقة الأكبر في تقدير معالم قدرات الأفراد مقارنة بالنماذج الاختبار الأخرى، وذلك عند التدريج باستخدام نموذج الاستجابة للمفردة رباعي المعلم، وتمت المقارنة بين هذه الأوساط الحسابية للأخطاء المعيارية في التقدير للنماذج الاختبارية الثلاث المختلفة في طرق ترتيب مفرداتها؛ لمعرفة الدلالة الإحصائية للاختلافات الظاهرة كما في جدول (٦).

جدول (٦) نتائج تحليل التباين احادي الاتجاه لأثر طريقة ترتيب المفردات الاختبارية على الأخطاء المعيارية في تقدير معالم قدرات الأفراد

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	مستوى الدلالة	قيمة مربع ايتا	حجم الأثر
طريقة الترتيب	٥,٩٧٥	٢	٢,٩٨٧	١٥٩,٧٤٠	٠,٠٠٠	٠,٠٤٥	صغير
الخطأ	١٢٧,١٩٢	٦٨٠١	٠,٠١٩				
الكلي	١٣٣,٨٢١	٦٨٠٦					

ويتضح من جدول (٦) وجود فرق دال إحصائياً بين متوسطات درجات الأخطاء المعيارية في تقدير معالم قدرات الأفراد يعزى إلى النماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها، وللتعرف على اتجاه الفرق في متوسطات درجات الأخطاء المعيارية في تقدير معالم قدرات أفراد الذي يعزى إلى طريقة ترتيب المفردات تم استخدام اختبار شيفيه للمقارنات البعدية بين كل طريقتين من طرق ترتيب المفردات، وكانت النتائج كما في جدول (٧).

جدول (٧): نتائج اختبار شيفيه للمقارنات البعدية بين متوسطي الأخطاء المعيارية في تقدير معالم قدرات الأفراد بين كل طريقتين من طرق ترتيب مفردات الاختبار

مصدر التباين	الإحصاءة	تسلسل المحتوى	الترتيب العشوائي
تسلسل الصعوبة المعرفية	الفرق بين المتوسطي	٠,٠١٤٤١	٠,٠٥٣٥٥
	مستوى الدلالة	٠,٠٠٢	٠,٠٠٠
الترتيب العشوائي	الفرق بين المتوسطي	٠,٠٦٧٩٦٦	
	مستوى الدلالة	٠,٠٠٠	

ويتضح من جدول (٧) وجود فرق ذا دلالة إحصائية بين متوسطي الأخطاء المعيارية في تقدير معالم قدرات الأفراد بين النموذج الاختباري المرتب وفق تسلسل

المحتوى وكل من النموذجين الاختباريين المرتبة وفق تسلسل الصعوبة المعرفية، ووفق الترتيب العشوائي في اتجاه النموذج الاختباري ذا الترتيب العشوائي في الحالتين، كما يوجد فرق بين متوسطي الأخطاء المعيارية في تقدير معالم قدرات الأفراد بين النموذج الاختباري المرتب وفق تسلسل الصعوبة المعرفية والنموذج الاختباري المرتب بشكل عشوائي في اتجاه النموذج الاختباري ذا الترتيب العشوائي، وهذه النتائج تعنى أن نموذج الاختبار التي تم بها ترتيب المفردات وفق تسلسل المحتوى تعتبر أكثر دقة في تقدير معالم قدرات الأفراد من نموذج الاختبار التي يتم ترتيب مفرداتها وفق تسلسل الصعوبة المعرفية، أو بشكل عشوائي، كما أن نموذج الاختبار المرتب وفق تسلسل الصعوبة المعرفية أكثر دقة في تقدير معالم قدرات الأفراد من نموذج الاختبار المرتب بشكل عشوائي. وهذا يعني أنه أذا روعي ترتيب المفردات وفق تسلسل المحتوى الدراسي كان دقة تقدير معالم قدرات الأفراد أفضل.

ويمكن تفسير ذلك، بأن ترتيب المفردات بنفس التسلسل المنطقي لمحتوى المادة الدراسية يدخل بها جانب الاستدعاء للمفاهيم كما في طريقة تلقيها أثناء التدريس؛ وهذا بدورة قد يوفر للفرد المختبر فرصة للتركيز الجيد في متابعة المفردات والإجابة عنها بيسر وسلاسة. كما تعمل هذه الطريقة على توجيه تفكير الفرد المختبر وحصر انتباهه حول موضوع محدد وبالتالي تجنبه التشتت في الاستجابة على المفردات الخاصة بموضوع محدد بخلاف ترتيب مفردات الاختبار وفق تسلسل الصعوبة المعرفية أو الترتيب العشوائي، والذي قد يُفترض أنها تؤدي إلى تقليل تأثير التهيؤ والاستعداد المسبق لمحتوى مفردات الاختبار مما قد يؤثر على أداء الأفراد في الاختبار. وهذا ما أشار اليه بلاتش في تفسير هذا التأثير "بأن الأسلوب المتخذ في الترتيب لمفردات الاختبار قد يؤثر على دقة عملية التشخيص لقدرات الفرد، وأن المفردات التي تكون بنفس تتسلسل المحتوى الذي يتم عرضة، هي الأفضل في الاسترجاع من الذاكرة وهذا بدورة يجعل

الأفراد يؤدون بشكل أفضل مقارنة بأسلوب الترتيب المعاكس للمحتوى الدراسي أو العشوائي أو المرتب وفق التسلسل المعرفي لمستوى الهدف".

- **السؤال الثاني:**

ينص السؤال الثاني على: "ما مستوى الدلالة الإحصائية للفروق بين متوسطات الأخطاء المعيارية في تقديرات معالم صعوبة المفردات تعزى لاختلاف طريقة ترتيب المفردات الاختبارية وفق (تسلسل المحتوى، تسلسل الصعوبة المعرفية، الترتيب العشوائي)؟"

وللإجابة عن هذا السؤال تم تقدير معالم صعوبة المفردات في النماذج الاختبارية الثلاث المختلفة في طرق ترتيب مفرداتها باستخدام برنامج (Exametrika Version 5.5)، وقد تم تقدير الأخطاء المعيارية في تقدير معالم الصعوبة للمفردات في جميع النماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها، ثم تم استخدام تحليل التباين أحادي الاتجاه لمعرفة أثر كل من طريقة ترتيب المفردات الاختبارية في تقدير معالم صعوبة المفردات، وكانت النتائج كما في جدول (٨) وجدول(٩).

جدول (٨): الأوساط الحسابية والانحرافات المعيارية للأخطاء المعيارية في تقدير معالم صعوبة المفردات وفقاً لطرق ترتيب المفردات

طريقة ترتيب المفردات الاختبارية	نوع النموذج	
	رباعي المعلمة	
	الوسط الحسابي	الانحراف المعياري
وفق تسلسل المحتوى	٠,٠٨٣٥	٠,٠٤٠٤
وفق تسلسل الصعوبة المعرفية	٠,٠٩٥٢	٠,٠٤٨٨
الترتيب العشوائي	٠,١٥٩٤	٠,١٦٥٠

يتضح من جدول (٨) أن أقل متوسط للأخطاء المعيارية للنماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها هي للنموذج الاختباري المرتب وفق تسلسل المحتوى، والتي تدل على الدقة الأكبر في تقدير معالم صعوبة المفردات مقارنة بالنماذج الاختبار الأخرى، وذلك عند التدريج باستخدام نموذج الاستجابة للمفردة رباعي المعلم، وتمت المقارنة بين هذه الأوساط الحسابية المعيارية للأخطاء في التقدير للنماذج الاختبارية الثلاث المختلفة في طرق ترتيب مفرداتها لمعرفة الدلالة الإحصائية للاختلافات الظاهرة كما في جدول (٩).

جدول (٩) نتائج تحليل التباين أحادي الاتجاه لأثر طريقة ترتيب المفردات على الأخطاء المعيارية في تقدير معالم صعوبة المفردات

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	مستوى الدلالة	قيمة مربع ايتا	حجم الأثر
طريقة الترتيب	٠,١٧٠	٢	٠,٠٨٥	١١,٩٠٦	٠,٠٠٠	٠,٠٨٤	متوسط
الخطأ	١,٨٥٩	٢٦١	٠,٠٠٧				
الكلي	٢,٠٦٢	٢٦٦					

يتضح من جدول (٩) أن هناك فرق دال إحصائياً بين متوسطات درجات الأخطاء المعيارية في تقدير معالم صعوبة المفردات يعزى إلى النماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها، ولمعرفة اتجاه الفرق في متوسطات درجات الأخطاء المعيارية في تقدير معالم صعوبة المفردات الذي يعزى إلى طريقة ترتيب المفردات تم استخدام اختبار شيفيه للمقارنات البعدية بين كل طريقتين من طرق ترتيب المفردات، وكانت النتائج كما في جدول (١٠).

جدول (١٠): نتائج اختبار شيفيه للمقارنات البعدية بين متوسطي الأخطاء المعيارية في تقدير معالم صعوبة المفردات بين كل طريقتين من طرق ترتيب مفردات الاختبار

مصدر التباين	الإحصاءة	تسلسل المحتوى	الترتيب العشوائي
تسلسل الصعوبة المعرفية	الفرق بين المتوسطي	٠,٠١٧٠	٠,٠٤٢٦
	مستوى الدلالة	٠,٤١٧	٠,٠٠٤
الترتيب العشوائي	الفرق بين المتوسطي	٠,٠٥٩٥	
	مستوى الدلالة	٠,٠٠٠	

ويتضح من جدول (١٠) وجود فرق ذي دلالة إحصائية بين متوسطي الأخطاء المعيارية في تقدير معالم صعوبة المفردات بين نموذج الاختبار المرتب وفق تسلسل المحتوى ونموذج الاختبار المرتب بشكل عشوائي، ووجود فرق ذا دلالة إحصائية بين نموذج الاختبار المرتب وفق تسلسل الصعوبة المعرفية ونموذج الاختبار المرتب بشكل عشوائي في اتجاه نموذج الاختبار ذا الترتيب العشوائي في الحالتين، أي أن الأخطاء المعيارية الناتجة عنها في تقدير معالم صعوبة المفردات كانت أكبر بشكل دال إحصائياً من النموذجين الاختباريين المرتب وفق تسلسل المحتوى والمرتب وفق تسلسل الصعوبة المعرفية، وبالتالي فهي أقل طرق ترتيب المفردات دقة في تقدير معالم صعوبة المفردات. وعدم وجود فرق بين متوسطي الأخطاء المعيارية في تقدير معالم صعوبة المفردات بين نموذج الاختبار المرتب وفق تسلسل المحتوى ونموذج الاختبار المرتب وفق تسلسل الصعوبة المعرفية، وهذا يعني أن الفرق بين متوسطي طرق ترتيب المفردات وفق تسلسل المحتوى ووفق تسلسل الصعوبة المعرفية يكاد يكون واحد في دقة في تقدير معالم صعوبة المفردات. وهذه النتائج تعنى أن النموذج الاختباري التي يتم بها ترتيب المفردات وفق تسلسل المحتوى أو وفق تسلسل الصعوبة المعرفية تعتبر أكثر دقة في تقدير معالم صعوبة المفردات عن النموذج الاختباري المرتب بشكل عشوائي.

وهذا يمكن أن يفسر بأن ترتيب المفردات الاختبارية وفق تسلسل المحتوى، أو وفق تسلسل الصعوبة المعرفية يوفر لدى الفرد المختبر رغبة في الاستمرار في محاولاته للإجابة على المفردات الاختبارية بسبب قدرته على الإجابة عن الأسئلة الأولى من الاختبار والتي قد تتضمن بعض المفردات الصعبة بين الحين والآخر، بمعنى أن الفرد المختبر قد يشعر بالتعزيز الفوري بسبب قدرته على الإجابة عن الأسئلة الأولى من الاختبار، والتي تسمى بالأسئلة التشجيعية أو الأسئلة الماصة للصدمة، وهذا بدورة ربما يجعل مستوى قدرة معظم الطلبة في النماذج الاختبارية المرتبة وفق تسلسل المحتوى أو فق تسلسل مستوى الصعوبة المعرفية أكبر من مستوى صعوبة المفردة مما يساعدهم في الإجابة عن مفردات الاختبار بشكل أفضل مما يقلل من الخطأ المعياري في التقدير لمعالم صعوبة المفردات، وهذا قد يكون متوقعاً من قبل الفرد المختبر، حيث لديهم قناعة أو معرفة عند ترتيب المفردات الاختبارية وفق تسلسل المحتوى أو وفق تسلسل الصعوبة المعرفية، أن الاختبار لن يكون بنفس المستوى من الصعوبة، لذلك عندما يجيبون عن بعض المفردات، ويواجهون بعض المفردات الصعبة، لن يصلوا إلى مرحلة الإحباط التي قد يصلون إليها في حال كانت المفردات مرتبة بشكل عشوائي.

- **السؤال الثالث:**

ينص السؤال الثالث على: "ما مستوى الدلالة الإحصائية للفروق بين متوسطات الأخطاء المعيارية في تقديرات معالم تمييز المفردات تعزى لاختلاف طريقة ترتيب المفردات الاختبارية وفق (تسلسل المحتوى، تسلسل الصعوبة المعرفية، الترتيب العشوائي)؟"

وللإجابة عن هذا السؤال تم تقدير معالم تمييز المفردات في النماذج الاختبارية الثلاث المختلفة في طرق ترتيب مفرداتها باستخدام برنامج (Exametrika Version

(5.5، وقد تم تقدير الأخطاء المعيارية في تقدير معالم تمييز المفردات في جميع النماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها، ثم تم استخدام تحليل التباين أحدي الاتجاه لمعرفة أثر كل من طريقة ترتيب المفردات الاختبارية وكانت النتائج كما في جدول (١١) وجدول (١٢).

جدول (١١): الأوساط الحسابية والانحرافات المعيارية للأخطاء المعيارية في تقدير معالم تمييز المفردات وفقاً لطرق ترتيب المفردات

طريقة ترتيب المفردات الاختبارية	نوع النموذج	
	رباعي المعلم	
	الوسط الحسابي	الانحراف المعياري
وفق تسلسل المحتوى	٠,١٣٥٢	٠,٠٥٦٦
وفق تسلسل الصعوبة المعرفية	٠,١٥٠١	٠,٠٧٤٦
الترتيب العشوائي	٠,١٩٤٤	٠,١٠٥٩

يتضح من جدول (١١) أن أقل متوسط للأخطاء المعيارية للنماذج الاختبارية الثلاث هي للنموذج الاختباري المرتب وفق تسلسل المحتوى، والتي تدل على الدقة الأكبر في تقدير معالم تمييز المفردات مقارنة بالنماذج الاختبارية الأخرى، وتمت المقارنة بين هذه الأوساط الحسابية المعيارية للأخطاء في التقدير للنماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها لمعرفة الدلالة الإحصائية للاختلافات الظاهرة كما في جدول (١٢).

جدول (١٢) نتائج تحليل التباين أحادي الاتجاه لأثر طريقة ترتيب المفردات على الأخطاء المعيارية في تقدير معالم تمييز المفردات

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	مستوى الدلالة	قيمة مربع ايتا	حجم الأثر
طريقة الترتيب	٠,٠٧١	٢	٠,٠٣٦	٤,٥٤٥	٠,٠١١	٠,٠٣٤	صغير
الخطأ	٢,٠٤٩	٢٦١	٠,٠٠٨				
الكلي	٢,٤٧١	٢٦٦					

يتضح من جدول (١٢) أن هناك فرق دال إحصائياً بين متوسطات درجات الأخطاء المعيارية في تقدير معالم تمييز المفردات يعزى إلى النماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها، وللتعرف على أتجاه الفرق في متوسطات درجات الأخطاء المعيارية في تقدير معالم تمييز المفردات الذي يعزى إلى طريقة ترتيب المفردات تم استخدام اختبار شيفيه للمقارنات البعدية بين كل طريقتين من طرق ترتيب المفردات، وكانت النتائج كما في جدول (١٣).

جدول (١٣): نتائج اختبار شيفيه للمقارنات البعدية بين متوسطي الأخطاء المعيارية في تقدير معالم تمييز المفردات بين كل طريقتين من طرق ترتيب مفردات الاختبار

مصدر التباين	الإحصاءة	تسلسل المحتوى	الترتيب العشوائي
تسلسل الصعوبة المعرفية	الفرق بين المتوسطي	٠,٠١٠٠	٠,٠٢٦٥
	مستوى الدلالة	٠,٧٥٧	٠,١٣٨
الترتيب العشوائي	الفرق بين المتوسطي	٠,٠٣٦٥	
	مستوى الدلالة	٠,٠٢٣	

ويتضح من جدول (١٣) عدم وجود فرق ذات دلالة إحصائية بين متوسطي الأخطاء المعيارية في تقدير معالم تمييز المفردات بين نموذج الاختبار المرتب وفق تسلسل المحتوى ونموذج الاختبار المرتب وفق تسلسل الصعوبة المعرفية، وأيضاً عدم وجود فرق بين متوسطي الأخطاء المعيارية في تقدير معالم تمييز المفردات بين

نموذج الاختبار المرتب وفق تسلسل الصوبة المعرفية ونموذج الاختبار المرتب بشكل عشوائي، وهذا يعني أن الفرق بين طريقتي ترتيب المفردات وفق تسلسل المحتوى ووفق تسلسل الصعوبة المعرفية، والفرق بين طريقتي ترتيب المفردات وفق تسلسل الصعوبة المعرفية ووفق الترتيب العشوائي تكاد تكون واحد في دقة في تقدير معالم تمييز المفردات. ووجود فرق دال إحصائياً بين نموذج الاختبار المرتب وفق تسلسل المحتوى ونموذج الاختبار المرتب بشكل عشوائي في اتجاه نموذج الاختبار ذا الترتيب العشوائي، وهذه الدلالة تعني أن متوسطات الأخطاء المعيارية الناتجة من النموذج المرتب بشكل عشوائي في تقدير معالم تمييز المفردات كانت أكبر بشكل دال إحصائياً من النموذج الاختباري المرتب وفق تسلسل المحتوى، وبالتالي فهي أقل طرق ترتيب المفردات دقة في تقدير معالم تمييز المفردات.

ويمكن تفسير ذلك، أن جميع الأفراد المختبرين–على اختلاف مستوياتهم– قد يشعرون بالعجز عن الإجابة على النموذج المرتب بشكل عشوائي، بسبب غياب عامل التعزيز الفوري في النموذج الاختباري والتي بدورها قد يؤثر بشكل سلبي من حيث شعور الطلبة بعدم وجود نتائج إيجابية لدراستهم، حيث أشار فاند شي (Vander Schee, 2013) "بأن الترتيب العشوائي يؤدي إلى تكوين اعتقاد أو فكرة لدى الفرد المختبر بأن هناك عموماً إفراط في صياغة المفردات"، وهذا قد يكون كفيلاً لخلف إحباط لدى الأفراد ذوي التحصيل المرتفع، وعاملاً مثيراً لعدم إكمالهم الاختبار بدافعية مرتفعة بسبب شعورهم بأن نتائجهم ستكون متقاربة مع غيرهم من الأفراد ذوي التحصيل المتدني، والذين لا يكونون قد بذلوا جهداً في الاستعداد للاختبار، مما يؤثر في طريقة إكمالهم الاختبار. وهذا التأثير قد لا يكون موجوداً في كلا النموذجين الاختباريين المرتبة وفق تسلسل المحتوى ووفق تسلسل الصعوبة المعرفية.

- **السؤال الرابع:**

ينص السؤال الرابع على: "ما مستوى الدلالة الإحصائية للفروق بين متوسطات الأخطاء المعيارية في تقديرات معالم تخمين المفردات تعزى لاختلاف طريقة ترتيب المفردات الاختبارية وفق (تسلسل المحتوى، تسلسل الصعوبة المعرفية، الترتيب العشوائي)؟"

وللإجابة عن هذا السؤال تم تقدير معالم تخمين المفردات في النماذج الاختبارية الثلاث المختلفة في طرق ترتيب مفرداتها باستخدام برنامج (Exametrika Version 5.5)، وقد تم تقدير الأخطاء المعيارية في تقدير معالم تخمين المفردات في جميع النماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها، ثم تم استخدام تحليل التباين أحادي الاتجاه لمعرفة أثر كل من طريقة ترتيب المفردات الاختبارية، وكانت النتائج كما في جدول (١٤) وجدول (١٥).

جدول (١٤): الأوساط الحسابية والانحرافات المعيارية للأخطاء المعيارية في تقدير معالم تخمين المفردات وفقاً لطرق ترتيب المفردات

طريقة ترتيب المفردات الاختبارية	نوع النموذج	
	رباعي المعلم	
	الوسط الحسابي	الانحراف المعياري
وفق تسلسل المحتوى	٠,٠٣١١	٠,٠١١٣
وفق تسلسل الصعوبة المعرفية	٠,٠٣٥١	٠,٠١٥٧
الترتيب العشوائي	٠,٠٣١٠	٠,٠١٧٠

يتضح من جدول (١٤) أن أقل وسط حسابي للأخطاء المعيارية للنماذج الاختبارية الثلاث والمختلفة في طريقة ترتيب مفرداتها هي للنموذج الاختباري المرتب وفق تسلسل المحتوى، والتي تدل على الدقة الأكبر في تقدير معالم قدرات الأفراد مقارنة بالنماذج الاختبار الأخرى، وتمت المقارنة بين هذه الأوساط الحسابية للأخطاء

المعيارية في التقدير للنماذج الاختبارية الثلاث المختلفة في طرق ترتيب مفرداتها لمعرفة الدلالة الإحصائية للاختلافات الظاهرة كما في جدول (١٥).

جدول (١٥). نتائج تحليل التباين أحادي الاتجاه لأثر طريقة ترتيب المفردات على الأخطاء المعيارية في تقدير معالم التخمين المفردات

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	مستوى الدلالة
طريقة الترتيب	٠,٠٠١	٢	٠,٠٠١	٢,٢٨٣	٠,١٠٤
الخطأ	٠,٠٥٩	٢٦٢	٠,٠٠٠		
الكلي	٠,٠٦٠	٢٦٧			

يتضح من جدول (١٥) عدم وجود فرق بين متوسطات درجات الأخطاء المعيارية في تقدير معالم تخمين المفردات يعزى إلى النماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها، ويمكن تفسير هذه النتيجة على أساس أن تقديرات تخمين المفردة تعتمد بدرجة كبيرة على الأفراد وما يمتلكونه من مهارات وكفايات وكيفية توظيفه لمعارفه أثناء استجابته عن المفردة، وكون الاختبار من نوع الاختيار من متعدد ومن المعروف أن مشتتات مفردات الاختيار من متعدد تبنى من أجل الكشف عن عامل التخمين وأخذه بعين الاعتبار، وبسبب أن جميع النماذج الاختبارية تتضمن نفس المفردات الاختبارية بالتالي فإن فاعلية المشتتات والتي تقليل من عامل التخمين من قبل الأفراد المستجيبين واحدة؛ مما ترتب عليه التوازي في دقة تقدير معالم تخمين المفردات في النماذج الاختبارية الثلاث المختلفة فيما بينها في طريقة ترتيب مفرداتها.

- **السؤال الخامس:**

ينص السؤال الخامس على: "ما مستوى الدلالة الإحصائية للفروق بين متوسطات الأخطاء المعيارية في تقدير معالم اللامبالاة تعزى لاختلاف طريقة ترتيب المفردات الاختبارية وفق (تسلسل المحتوى، تسلسل الصعوبة المعرفية، الترتيب العشوائي)؟"

وللإجابة عن هذا السؤال تم تقدير معالم اللامبالاة للمفردات في النماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها باستخدام برنامج (Exametrika Version 5.5) مع نموذج الاستجابة للمفردة رباعي المعلم، وقد تم تقدير الأخطاء المعيارية في تقدير معالم اللامبالاة للمفردات في جميع النماذج الاختبارية المختلفة في طريقة ترتيب مفرداتها، ثم تم استخدام تحليل التباين أحادي الاتجاه لمعرفة أثر طريقة ترتيب المفردات الاختبارية الثلاث المختلفة في طريقة ترتيب مفردتها على الأخطاء المعيارية في تقدير معالم اللامبالاة للمفردات وفق نموذج الاستجابة للمفردة رباعي المعلم، وكانت النتائج كما في جدول (١٦) وجدول(١٧).

جدول (١٦): الأوساط الحسابية والانحرافات المعيارية للأخطاء المعيارية في تقدير معالم اللامبالاة للمفردات باختلاف طريقة ترتيب المفردات الاختبارية

طريقة ترتيب المفردات الاختبارية	تسلسل المحتوى	تسلسل الصعوبة المعرفية	الترتيب العشوائي
الوسط الحسابي	٠,٠٢١٢	٠,٠١٩٩	٠,٠٣١٧
الانحراف المعياري	٠,٠٠٧٨	٠,٠١٠٢	٠,٠١٢٣

ويتضح من جدول (١٦) أن أقل وسط حسابي للأخطاء المعيارية لتقدير معالم اللامبالاة للمفردات باختلاف طريقة ترتيب المفردات الاختبارية هي للنموذج المرتب وفق تسلسل الصعوبة المعرفية، والتي تدل على الدقة الأكبر في تقدير معالم اللامبالاة، وتمت المقارنة بين هذه المتوسطات كما في جدول (١٧)

جدول (١٧): تحليل التباين أحادي الاتجاه لأثر طريقة ترتيب المفردات على الأخطاء المعيارية في تقدير معالم اللامبالاة للمفردات

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف	مستوى الدلالة	قيمة مربع ايتا	حجم الأثر
طريقة الترتيب	٠,٠٠٤	٢	٠,٠٠٢	١٧,٩٥٢	٠,٠٠٠	٠,٢٢٢	كبير
الخطأ	٠,٠٠١٤	١٣٢	٠,٠٠٠				
الكلي	٠,٠١٨	١٣٤					

ويتضح من جدول (١٧) أنه يوجد فرق دال إحصائياً بين متوسطات الأخطاء المعيارية في تقدير معالم اللامبالاة للمفردات يعزى إلى طريقة ترتيب المفردات، ولمعرفة اتجاه الفرق الذي يعود إلى طريقة ترتيب مفردات الاختبار تم استخدام اختبار شيفيه للمقارنات البعدية بين كل متوسطين من متوسطات الأخطاء المعيارية في تقدير معالم اللامبالاة للمفردات باستخدام نموذج الاستجابة للمفردة رباعي المعلم في النماذج الاختبارية الثلاث المختلفة في طريقة ترتيب مفرداتها، كما في جدول (١٨).

جدول (١٨): نتائج اختبار شيفيه للمقارنات البعدية بين متوسطي الأخطاء المعيارية في تقدير معالم اللامبالاة للمفردات بين كل طريقتين من طرق ترتيب مفردات الاختبار

مصدر التباين	الإحصاءة	تسلسل المحتوى	الترتيب العشوائي
الصعوبة المعرفية	الفرق بين المتوسطي	٠,٠٠١٣٧	٠,٠١١٨٣
	مستوى الدلالة	٠,٨٢٢	٠,٠٠٠
الترتيب العشوائي	الفرق بين المتوسطي	٠,٠١٠٤	
	مستوى الدلالة	٠,٠٠٠	

ويتضح من جدول (١٨) وجود فرق دال إحصائياً بين متوسطي الأخطاء المعيارية في تقدير معالم اللامبالاة للمفردات الاختبارية بين نموذج الاختبار المرتب وفق تسلسل المحتوى ونموذج الاختبار ذا الترتيب العشوائي، ووجود فرق دال إحصائياً

بين نموذج الاختبار المرتب وفق تسلسل الصعوبة المعرفية ونموذج الاختبار ذا الترتيب العشوائي وفي اتجاه نموذج الاختبار ذا الترتيب العشوائي في الحالتين، وهذا يدل على أن الأخطاء المعيارية في تقدير معالم اللامبالاة لمفردات الاختبار ذا الترتيب العشوائي كانت أكبر بشكل دال إحصائياً من النموذجين الاختباريين المرتب وفق تسلسل المحتوى والمرتب وفق تسلسل الصعوبة المعرفية، وبالتالي فهي أقل طرق ترتيب المفردات دقة في تقدير معالم اللامبالاة لمفردات الاختبار. وأظهرت نتائج التحليل عدم وجود فرق بين متوسطي الأخطاء المعيارية في تقدير معالم اللامبالاة لمفردات الاختبار بين نموذج الاختبار المرتب وفق تسلسل المحتوى ونموذج الاختبار المرتب وفق تسلسل الصعوبة المعرفية، وهذا يعني أن الفرق بين متوسطي طرق ترتيب المفردات وفق تسلسل المحتوى ووفق تسلسل الصعوبة المعرفية يكاد يكون واحد في دقة في تقدير معالم اللامبالاة للمفردات. وهذه النتائج تعنى أن النموذج الاختباري التي يتم بها ترتيب المفردات وفق تسلسل المحتوى أو وفق تسلسل الصعوبة المعرفية تعتبر أكثر دقة في تقدير معالم اللامبالاة للمفردات عن النموذج الاختباري المرتب بشكل عشوائي. ويمكن تفسير ذلك، بأنه قد تنعكس طرق ترتيب المفردات وفق تسلسل المحتوى أو وفق تسلسل الصعوبة المعرفية على أداء الأفراد المختبرين عند الاستجابة على المفردات من حيث التعامل معها بشكل فاعل، فربما يجعل هذا النوع من طرق ترتيب المفردات مستوى قدرة الفرد المختبر أكبر من مستوى صعوبة المفردة مما يظهر تأثير احتمالية اللامبالاة في اختيار الإجابة الصحيحة عن المفردات السهلة، وقد يكون ذلك من بسبب الابداع الاستثنائي على المفردات، وهو قيام الفرد من ذوي القدرة المرتفعة بتفسير المفردات السهلة بطريقة إبداعية غير تقليدية؛ مما يجعله يجيب عن المفردة إجابة خاطئة، أما بالنسبة للأفراد ذوي مستوى القدرة المتدنية فربما تزيد طرق ترتيب المفردات وفق تسلسل المحتوى أو وفق تسلسل الصعوبة المعرفية دافعيتهم وقدرتهم على الإجابة عن المفردات السهلة، والتي تسمى بالأسئلة التشجيعية أو

الأسئلة الماصة للصدمة، والتي بدورها توفر الرغبة للاستمرار في المحاولة في الإجابة على بعض المفردات السهلة بين الحين والأخر بسبب قدرته على الإجابة عن الأسئلة الأولى من الاختبار المرتب وفق تسلسل المحتوى أو وفق تسلسل الصعوبة المعرفية.

المراجع:

الحجاج، أيمن عودة. (٢٠١٣). أثر ترتيب الفقرات حسب خصائصها السيكومترية على أداء الأفراد والخصائص السيكومترية للاختبار. رسالة ماجستير غير منشورة. عمادة الدراسات العليا، جامعة مؤتة.

الخرشة، هناء عوض. (٢٠١٥). أثر ترتيب فقرات الاختبار على دقة تقدير معالم الفقرات وقدرات الأفراد والخصائص السيكومترية للاختبار. رسالة ماجستير منشورة، عمادة الدراسات العليا، جامعة مؤتة.

الزبون، حابس سعد؛ الصرايرة، راجي عوض.(٢٠١٧). أثر ترتيب فقرات اختبار الاختيار من متعدد وفقاً لمعامل الفقرة على تقدير قدرة المفحوصين ودالة المعلومات للاختبار وفقاً للنموذج ثلاثي المعلمة. مجلة جامعة الحسين بن طلال للبحوث، المجلد١، العدد٣.

الشريفين، علي.(٢٠١٣). أثر موقع الفقرات الصعبة في اختبار من متعدد على خصائص السيكومترية وأداء الطلبة عليه. رسالة ماجستير منشورة، جامعة اليرموك.

أناستاسي، أنا؛ سوزانا، يورين.(٢٠١٥). القياس النفسي. ترجمة: أ.د. صلاح الدين محمود علام. عمان: دار الفكر ناشرون وموزعون.

بركات، زياد.(٢٠١٨). القياس والتقويم النفسي والتربوي بين النظرية والتطبيق. رام الله: جامعة القدس المفتوحية، عمادة البحث العلمي.

حبيب، صفاء طارق؛ عزيز، صادق عبدالنور.(٢٠١٨). بناء المقاييس النفسية وفقاً لنظرية الاستجابة للفقرة باستعمال الأنموذج الكشفي المعمم. عمان: الدار المنهجية للنشر والتوزيع.

رمضان، رمضان صالح.(٢٠٠٧). أثر ترتيب فقرات الاختبار حسب درجة صعوبتها على تحصيل طلاب المرحلة الثانوية لمادة الرياضيات، بحث منشور، مجلة التربية، جامعة الأزهر– كلية التربية،(٨)، ١٧٦–١٥٧.

عبد الهادي، نبيل.(٢٠٠١). القياس والتقويم التربوي واستخدامه في مجال التدريس الصفي. عمان: دار وائل للنشر

عبد الوهاب، محمد محمود محمد.(٢٠١٩). أثر طريقة ترتيب مفردات الاختبار وطريقة تطبيقه على تقدير قدرات طلاب كلية التربية وفق النموذج اللوغاريتمي ثلاثي البارامتر. بحث منشور، رسالة الخليج العربي، المجلد ٤٢، (١٥٩).

علام، صلاح الدين محمود.(٢٠٠٥).نماذج الاستجابة للمفردة الاختبارية "أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي". القاهرة: دار الفكر العربي.

محمود، فريال محمود الحاج.(٢٠١٥). أثر كل من ترتيب فقرات الاختبار ومستوى القلق على أداء الطلاب في اختبار من نوع الاختيار من متعدد. بحث منشور، مجلة العلوم التربوية والنفسية، جامعة البحرين– مركز النشر العلمي، المجلد ١٦، (٢).

هاشم، كمال الدين محمد.(٢٠٠٦). التقويم التربوي مفهومة، أساليبه، مجالاته، توجهاته الحديثة. الرياض: مكتبة الرشد للنشر والتوزيع.

Aiken, L.R.; Growth-Marnat, G.(2006). Psychological Testing and Assessment (12th ed.). Boston, MA: Allyn and Bacon.

Balch, William. R.(1989). Item Order Effects Performance on Multiple-Choice Exams. The Pennsylvania State University, Altoona. Teaching of psychology, (16),75-77.

Barton, M. A.; Lord, F. M.(1981). An upper asymptote for the three-parameter logistic item response model. ETS Research Report Series, (1), 1–8.

Biemer, paul. P.; Groves, Robert; Lyberg, Lars; Mathiowetz, Nancy. A.(1991). Measurement Errors in Surveys. Technimetrics, vol 35 (1).

Bradburn, Norman. M.; Mason, William. M.(1964). The Effect of Question Order on Responses. Journal of Marketing Research, (4), 57-61.

Chen, Pei-Hua.(2010). Item Order Effect on Attitude Measures. Theses of Doctoral, The Morgridge College of Education, University of Denver. ProQuest Dissertation & Theses Database.

De Ayala, R,J.(1993). The Influence of Multidimensionality on the Graded Response Model. Paper Presented at the Annual Meeting of the National Council on Measurement in Education (Atlanta, GA, April 13-15).

Doerner, W.; Calhoun, P.(2009). The impact of the order of test questions in introductory economics. (working paper).

Flaugher, Ronald. L.; Melton, Richard. S; Myers, Charles.(1968). Item Rearrangement Under Typical Test Conditions. Research Article, (3),813-824.

Hambleton, R. K.; Traub, R. E. (1974). The effects of item order on test performance and stress. Journal of Experimental Education, (1), 40–46.

Hambleton, R.; Swaminathan, H.(1985). Item Response Theory: Principles and Applications, Boston, MA: Kluwer- Nijhoff.

Lord, M. F.(1980). Application of Item Response Theory to Practical Testing Problems. Hillsdale, NJ: Erlbaum.

Ortenr, Tuulia M.(2004). On changing the position of items in personality questionnaires Analyzing effects of item sequence using IRT. Psychology Science,(4), 466-476.

Pettijohn, T.; Sacco, M.(2007). Multiple-choic exam question order influences on student performance, Completion time, and Perceptions, Journal of Instructional Psychology,(3), 142-149.

Rose, Norman; Nagy, Gabriel; Nagengast, Benjamin; Frey, Andreas; Becker, Michael.(2019). Modeling Multiple Item Context Effects With Generalized Linear Mixed Models. Quantitative Psychology and Measurement.

Vander Schee, B. A.(2013). Test item order, level of difficulty, and student performance in marketing education. *Journal of Education for Business,* (1), 36–42.

Weinstein, Yana; Bard, Gabriele.(2017). The Effect of Question Order on Evaluation of Test Performance: Can The Bias Dissolve?. QJ Exp Psychol (Hove), (10), 2130-2140.

Wen, Wei Liao; Rong, Guey Ho; Yung, Chin Yen.(2012). The Four-Parameter Logistic Item Response Theory Model As a Robust Method of Estimating Ability Despite Aberrant Responses. Social Behavior and Personality An International Journal. Vol 40, (10).