أثر عدد البدائل والفقرات في تقدير معالم الأفراد والفقرات بإستخدام النموذج رباعي المعلم البارامتري

The Effect of the Number of Alternatives and Items in Estimating the Parameters of Individuals and Items Using the Four-Parameter Parametric Model

محمد فايق مرشود
أستاذ مساعد / جامعة عمان الأهلية / الأردن

Mohammed Faieq Marshoud
Assistant Professor / Al-Ahliyya Amman University / Jordan
m.marshoud@ammanu.edu.jo

تاريخ الاستلام: 30/ 10/ 2019م، **تاريخ القبول:** 23/ 5/ 2020م.
Received: 30/ 10/ 2019, **Accepted:** 23/ 5/ 2020.
DOI: 10.33977/1182-011-032-006
https: //journals.qou.edu/index.php/nafsia
E-ISSN: 2307-4655
P-ISSN: 2307-4647

الملخص:

هدفت الدراسة للكشف عن أثر عدد البدائل والفقرات في تقدير معالم الأفراد والفقرات باستخدام النموذج رباعي المعلم البارامتري، بإستخدام المنهج الوصفي المسحي، وتألف مجتمع الدراسة من جميع الطلبة الدراسيين لمادة علم النفس والحياة، والبالغ عددهم ما يقارب (245) طالباً وطالبة، وعينة الدراسة من (237)، منهم (156) طالباً و(81) طالبة، بأسلوب العينة القصدية، فيما تكونت أداة الدراسة من ثلاثة نماذج إختبارية، وجرى التحقق من صدق المحتوى والبناء والثبات بطريقة الإتساق الداخلي.

وقد توصلت النتائج لمطابقة الأفراد والفقرات للنموذج رباعي المعلم بإستخدام مؤشرات المطابقة، كما تم تقدير معالم الأفراد والفقرات والخطأ المعياري في النظرية الكلاسيكية (الصعوبة، التمييز) ونظرية إستجابة الفقرة (الصعوبة، التمييز، التخمين، اللامبالاه)، وتحقق إفتراضات (أحادية البعد، إطرادية السمة، منحنى خصائص الفقرة، دالة معلومات الفقرة والإختبار).

وقد توصلت النتائج في النظرية الكلاسيكية لوجود فروق بين متوسطي معلم (التمييز) لفقرات البديلين والثلاثة، وبين البديلين والأربعة، وذلك لصالح الثلاثة والأربعة بدائل، ووجود فروق في نظرية الإستجابة للفقرة بين متوسطي معلم (اللامبالاه) لعدد فقرات (9) و(15)، و(9) و(24)، وذلك لصالح (15، 24) فقرة.

الكلمات المفتاحية: عدد البدائل، عدد الفقرات، تقدير معالم الأفراد والفقرات، النموذج رباعي المعلم البارامتري.

Abstract:

The study sought to reveal the effect of the number of alternatives and items in estimating the parameters of individuals and items using the four-parameter parametric model, using the descriptive survey methodology. The study population consisted of all the students who study psychology and life sciences subject, it included 245 male and female students. The study sample consisted of 237 (156 male students and 81 female students) who were selected using purposive sample. The study instrument consisted of three experimental models. The validity and reliability of their content and structure were verified using internal consistency.

Individuals and items of the four-parameter model were matched using matching indicators. The parameters of individuals, items, and standard errors were estimated in classical theory (difficulty, discrimination) and item response theory (difficulty, discrimination, guessing, upper asymptote). Moreover, assumptions were achieved (unidimensionality, monotone homogeneity, item characteristic curve, item information function and test).

The results according to the classical theory have found differences between the mean parameters (discrimination) of the two and three variants, as well as between the two alternatives and the four, in favor of the three and four. Moreover, there were differences in the theory of response to the item between the parameter means (upper asymptote) for the number of items 9 and 15, and 9 and 24, in favor of 15, 24 items.

***Key words**: Alternatives Number, Item Number, Estimation of Persons and Items Parameters, Four Parameter Logistic Model.*

مقدمة:

يرى الباحث أن الإختبارات تلعب دوراً أساسياً في جميع مراحل الطالب التعليمية، وبناءاً على نتائجها يتحدد مستقبله، وتتخذ بحقه الكثير من القرارات، ويرى (عودة، 2014) بأنها تعدّ الوسيلة الأساسية في تقويم العملية التربوية التعليمية التعلمية، بهدف التعرف على ما حققه الطالب من أهداف تعليمية.

وتمتاز الإختبارات بالعديد من الأشكال، كإختبارات المقال والإجابة القصيرة، والتي يعطي فيها الطالب إجابة بلغته الخاصة، ومنها ما يطلب منه إختيار الإجابة الصحيحة من بين عدد من البدائل كإختبارات المطابقة، والإختيار من متعدد، والصواب والخطأ، والتكميل ذات الإجابة القصيرة والمحددة (عودة، 2014؛ العيسوي، 2003)، إذ يؤثر شكل الفقرة في مستوى صعوبتها بحيث تصبح سهلة أو صعبة جداً، أو متوسطة حسب ما تتطلبه الفقرة من قدرة للإجابة عنها بشكل صحيح، مما يؤدي إلى حصول الطالب على علامات لا تعكس قدرته الحقيقية، كما تؤثر على خصائص الإختبار السيكومترية (الصدق والثبات) (المصري، 2009).

وتعد فقرات الإختيار من متعدد من أفضل الفقرات الموضوعية وأكثرها إنتشاراً وكفاءة وسهولة في تصحيحها، كما أنها تمتاز بثبات درجة الطالب فيها، وعدم تأثرها بذاتية المصحح (دبك، 1998)، ويرى (علام، 2016) إلى أن أبسط صيغ فقرات الإختيار من متعدد هي أسئلة الصواب والخطأ، كما يمكن أن يكون لكل فقرة ثلاثة أو أربعة أو خمسة بدائل، فزيادتها يقلل من التخمين، فإحتمالية التخمين لفقرات البديلين (50 %)، وللثلاثة (33 %)، وللأربعة (25 %)، وللخمسة (20 %).

وقد جرى إستخدام النظرية الكلاسيكية في بناء الإختبارات إلا أن الفوائد المنتظرة منها كانت محدودة، لإعتمادها في تحليل الفقرات على خصائص عينة الأفراد المطبق عليهم الإختبار مما قلل من إتساقها وفاعليتها ومرونتها، لذلك تم الإنتقال إلى استخدام نظرية الإستجابة للفقرة بهدف مواجهة جوانب قصورها (علام، 2005)، وتهتم نظرية الإستجابة للفقرة بالربط بين إستجابة الفرد لفقرة إختبار ذات خصائص معينة وقدرته (الشافعي، 2008)، كما تركز على تحديد مواقع الأفراد بالنسبة للاختبارات، التي تتمثل في مجموعة من النماذج الرياضية والإحصائية المستخدمة في تحليل الفقرات، لذلك تعتبر نموذجاً رياضياً يركز على مجموعة من الإفتراضات للعلاقة بين الأداء الملاحظ للفرد على الاختبار

والسمات الكامنة، هذه الإفتراضات ينبغي تحققها في البيانات لكي تؤدي إلى نتائج يمكن الوثوق بها، ويشير إفتراض أحادية البعد لوجود سمة أو قدرة واحدة تفسر أداء الأفراد على الاختبار، بمعنى أن جميع فقرات الإختبار تقيس بعداً واحداً، فيما يقصد بافتراض الاستقلال الموضعي المكافئ لافتراض أحادية البعد، أن إستجابات الفرد للفقرات المختلفة في الإختبار مستقلة إحصائياً، أي أن إستجابة الفرد على فقرة ما لا تؤثر إيجاباً أو سلباً في استجابته على فقرة أخرى، ويشير افتراض منحنى خصائص الفقرة إلى أن إحتمال إجابة الفرد على فقرة إجابة صحيحة يزداد بازدياد قدرته والذي يأخذ الشكل اللوجستي متمثلاً بشكل الحرف (S)، ويشير إفتراض السرعة في الإجابة إلى أن عامل السرعة لا يلعب دوراً في الإجابة على فقرات الاختبار، بمعنى أن إخفاق الأفراد في إجابة فقرات الاختبار يرجع إلى إنخفاض قدرتهم، وليس إلى تأثير عامل السرعة على إجاباتهم، بينما يشير إفتراض دالة معلومات الفقرة لكمية المعلومات التي تقدمها الفقرة عند مستوى القدرة التي تقيسها، وذلك بتحديد أقصى ارتفاع لمنحنى دالة معلومات الفقرة عند مستوى معين للقدرة، وبهذا يمكن تحديد أي الفقرات الاختبارية تقيس المتغير المراد قياسه بدرجة أفضل عند مستويات محددة للقدرة (Hambleton & Swaminathan, 1985).

وتعتبر نماذج إستجابة الفقرة دوالاً رياضية احتمالية تختلف عن بعضها باختلاف عدد المعالم المتعلقة بتقدير احتمالية الإجابة الصحيحة على الفقرة، كذلك من حيث الصيغة الرياضية (Baker & Kim, 2004)، وتركز الدراسة الحالية على النموذج رباعي المعلم، إذ يمكن التعرف على إحتمالية إستجابة الفرد على الفقرة في النموذج الرباعي من خلال أربع معالم (Magis & Raiche, 2012).

(**الصعوبة b**) والتي تعبر عن النقطة على متصل القدرة التي تكون عندها احتمالية استجابة الفرد الصحيحة عند تلك الفقرة مساوية لـ(50 %)، عندما تكون نقطة تقاطع منحنى خصائص الفقرة مع المحور الصادي تساوي صفر تقريباً، (أي أن معلم التخمين هنا يساوي صفراً، كما هو في النموذج أحادي وثنائي المعلم) أما إذا كانت قيمة المقطع الصادي للمنحنى أكبر من صفر فإن صعوبة الفقرة هي القدرة الممثلة على محور السينات التي تقابل إحتمالية الإجابة الصحيحة في منتصف المسافة بين تقاطع المنحنى مع المحور الصادي والقيمة واحد (كما هو في النموذج اللاثي ورباعي المعلم)، (Baker, 2001; Hambleton & Swaminathan, 1985 ; Magis & Raich, 2012).

(**التمييز a**) وتعبر قيم هذا المعلم عن مدى قدرة الفقرة على التمييز بين الأفراد ذوي القدرات التي تقع دون موقع الفقرة على متصل القدرة وأولئك ذوي القدرات التي تقع فوق موقع هذه الفقرة، حيث ترتبط قيم معلم التمييز بميل منحنى خصائص الفقرة، وبزيادة ميل المنحنى تزداد قيمة معلم التمييز عند نقطة الإنعطاف (Hambleton & Swaminathan, 1985).

(**التخمين c**) يمثل معلم التخمين المقطع الصادي لمنحنى خصائص الفقرة، وتعبر قيم هذا المعلم عن احتمالية الاستجابة الصحيحة للفقرة لفرد ذي قدرة متدنية على الفقرات الصعبة جداً أو متوسطة الصعوبة، وتعتبر الفقرة جيدة عندما يقترب الخط التقاربي الأدنى من الصفر أي أن قيمة التخمين صفر (Baker, 2001; Hambleton & Swaminathan, 1985).

وبإضافة **معلم رابع وهو** (**d**) الذي يعبر عن الخط التقاربي الأعلى لمنحنى خصائص الفقرة (اللامبالاه)، أي أن الأفراد من ذوي القدرة المرتفعة لا يستجيبوا لفقرات الإختبار بإجابة صحيحة، وهو يفترض قيمة أقل من (1) (Hambleton & Swaminathan, 1985)، والذي يمثل المقطع الصادي لمنحنى خصائص الفقرة، وتعبر قيم هذا المعلم عن إحتمالية الاستجابة الصحيحة للفقرة لفرد ذي قدرة عالية، وتعتبر الفقرة جيدة عندما يقترب من الواحد الصحيح أي أن قيمته (1) (Loken & Rulison, 2010).

ويشير المنطق وراء المعلم الرابع إلى أن الأشخاص ذوي القدرات المرتفعة جداً، قد يرتكبون خطأ في الإجابة على فقرات سهلة نسبياً، لذلك ينخفض منحنى الإستجابة للفقرة بالنهاية العليا والتي تساوي (1) (de Ayala, 2003).

وتعد نظرية استجابة الفقرة اللابارمترية كعائلة لنماذج قياس إحصائية والتي تعتمد على مجموعة صغيرة من الافتراضات الضرورية للحصول على فوائد قياسات الأفراد والفقرات (Stochl, 2007)، فقد لعبت دوراً هاماً في تطوير نظرية إستجابة الفقرة، فهي تمثل الحالة العامة لها والتي تشكل النماذج البارامترية حالة خاصة منها (Sijtsma & Hemker, 2000).

ولقد شاع إستخدام نماذج نظرية إستجابة الفقرة البارامترية من قبل الباحثين على الرغم من موانعها في تحليل البيانات ذات المستوى الرتبي، إلا أن مصداقية النتائج التي تتوصل لها النماذج البارامترية يكون موضع تساؤل عندما لا يتحقق فرض وقوع البيانات على مستوى القياس الفئوي، الأمر الذي يبرر إستخدام نماذج إستجابة الفقرة اللابارامترية والتي لا تضع قيود حول شكل دالة استجابة الفقرة (IRF)، مما يثير التساؤل حول مدى مطابقة النوعين للبيانات التحصيلية التي يعتبرها بعضهم رتبية وبعضهم الآخر شبه فئوية، ومدى دقة النتائج التي تفرزها مثل هذه الاختبارات (Liang, 2010) .

وتعرف دالة (IRF) بالعلاقة بين احتمالية حصول الفرد على الدرجة (1) الاستجابة الصحيحة لفقرة ثنائية التدريج، والسمة الكامنة أو القدرة للفرد (Θ)، وتفترض النماذج البارامترية قيوداً لشكل هذه الدالة إلا أن نظيرتها اللابارامترية لا تفترض شكلاً معيناً لها، فشرطها الوحيد ألا تكون متناقصة مع إزدياد مستوى القدرة وما عدا ذلك فجميع الأشكال مقبولة (Sijtsma,1998).

ويعتبر تقدير معالم الفقرات من الإجراءات المهمة التي تستخدم نظريات إحصائية تفسر كيفية الاستجابة للفقرات الاختبارية بتحديد الفقرات المطابقة للنموذج المستخدم، والتي تقدر وفق عدة طرق تستخدم أساليب التحليل العددي من خلال إستخدام برامج حاسوبية معينة (Hambelton, Swaminathan, Rogers, 1991)

وينظر للهدف من تقدير معالم الأفراد (قدرتهم) هو تقييم المفحوص تبعاً لقدرته الحقيقية التي تكمن وراء استجابته عن فقرات الإختبار (Baker, 2001)، إذ تساعد الاستجابات التي يقدمها المفحوص في التعبير عن موقعه على متصل القدرة التي يقيسها ذلك الاختبار مع احتمال وجود خطأ في تقديرها، بحيث تعبر قيمة هذا

الخطأ عن مقدار التفاوت بين علامة المفحوص الحقيقية وعلامته الملاحظة (Hambleton, Swaminathan & Rogers, 1991).

وتحاول جميع النماذج تقدير موقع المفحوص على متصل القدرة بإستخدام نمط إستجابات المفحوصين، وأن أداء المفحوص على مجموعة من الفقرات يرتبط بقدرته، والتي تعبر عن كمية لا يمكن ملاحظتها ولكن يمكن تقديرها، وتعتبر أهم ميزات تقدير القدرة، تحررها من خصائص عينة الفقرات التي إستخدمت في تقديرها، وتحررها من أداء باقي المفحوصين، وهذا يسهل مقارنة الفقرات ومقارنة الاختبارات بالإضافة إلى مقارنة أداء المفحوصين (Embretson & Reise, 2000).

وقد تناولت العديد من الدراسات تقدير معالم الأفراد والفقرات وفق متغيرات الدراسة، فهدفت دراسة دبك (1998) لمعرفة أثر شكل الفقرات (تكميل، اختيار من متعدد) وعدد البدائل (3، 4) في الخصائص السيكومترية للإختبار (الصدق، الثبات)، والخصائص السيكومترية للفقرة (الصعوبة والتمييز)، وتألف مجتمع الدراسة من جميع طلبة الصف التاسع الأساسي في مديرية تربية الكورة للعام (1996 /1997)، فيما تألفت العينة من (300) طالباً، باستخدام إختبار تحصيلي في مادة الفيزياء، تكون من (49) فقرة بثلاث نماذج، نموذج من نوع التكميل، ونموذجين وفق الاختيار من متعدد (3، 4) بدائل، وقد توصلت النتائج لعدم وجود فروق في معاملات الصعوبة تعزى لعدد البدائل (3، 4) بدائل، بينما توجد فروق في معاملات التمييز لعدد البدائل (3، 4) بدائل ولصالح (3) بدائل.

بينما أشارت دراسة **فرانسيسكو وأوليا وبونسودا (Fran-cisco, Olea & Ponsoda, 2001)** لتحليل عدد البدائل الأمثل في فقرات الاختيار من متعدد وفق النظرية الكلاسيكية والنموذج ثلاثي المعلم، وقد تألفت عينة الدراسة من (452) فرداً من طلبة المدارس الثانوية وطلبة المرحلة الجامعية ومعلمي الجامعات في إسبانيا، وذلك ببناء إختبار لمفردات اللغة الإنجليزية، مكون من (221) فقرة لها (5) بدائل، وبشكل متتال تم إعادة تدوير إجاباتهم لأسوأ البدائل الثلاثة (1، 2، 3) عشوائياً لتوليد إجابات افتراضية لفقرات البدائل (2، 3، 4)، وقد توصلت النتائج لتحقق افتراضات نظرية الاستجابة للفقرة، ومطابقة (113) من أصل (221) بنسبة (51.13 %)، فيما أشارت إلى أن أسوأ النتائج كانت للفقرات ذات البديلين، في حين أن النتائج للفقرات ذات (3، 4) بدائل لا تكاد تختلف عن الفقرات ذات (5) بدائل، وقد إعتبرت فقرات (3) بدائل أفضل خيار.

وقام **التعمري** بدراسة **(2003)** هدفت لتطبيق النموذج اللوغاريتمي ذي الثلاث معالم لتقدير قدرات الأفراد ومعالم الفقرات في اختبار الاختيار من متعدد باختلاف عدد البدائل، إذ تألف مجتمع الدراسة من جميع طلبة الصف العاشر الأساسي في مديرية عمان الرابعة للعام (2001 /2002) والبالغ عددهم (10245) طالباً وطالبة، فيما تألفت العينة من (1200) طالباً وطالبة، التي اختيرت باستخدام العينة العشوائية البسيطة، وذلك ببناء اختبار تحصيلي في مادة الرياضيات، تكون من (35) فقرة من نوع الإختيار من متعدد بثلاث نماذج (3، 4، 5) بدائل، وقد توصلت النتائج لتحقق إفتراضات نظرية الاستجابة للفقرة، وعدم وجود فروق في قدرات الأفراد على اختبار الاختيار من متعدد بإختلاف عدد البدائل وفق النظرية الحديثة في القياس، بينما توجد فروق وفق النظرية الكلاسيكية في القياس ولصالح الاختبار ذي (5) بدائل، كما وجدت فروق في معالم الصعوبة والتمييز والتخمين وفق النظرية الحديثة في القياس لصالح الاختبار ذي (3) بدائل.

وفي دراسة **الشريفين وطعامنة (2009)** والتي هدفت للكشف عن أثر عدد البدائل (3، 4، 5) في تقديرات القدرة للأفراد لاختبار الإختيار من متعدد، وتقديرات الصعوبة للفقرات وفق نموذج راش أحادي المعلم، وتألف مجتمع الدراسة من جميع طلبة الصف العاشر الأساسي في مديرية أربد الأولى للعام (2006 /2007)، فيما تألفت العينة من (600) طالباً وطالبة، منهم (300) طالباً و(300) طالبة، وذلك بإعداد اختبار تحصيلي في مادة الرياضيات، مكون من (45) فقرة من نوع الاختيار من متعدد بثلاثة نماذج مختلفة في عدد بدائلها (3، 4، 5)، وقد توصلت النتائج لتحقق إفتراضات نظرية الاستجابة للفقرة، ومطابقة (37) فقرة بنسبة (92.5 %) لـ(3) بدائل، و(38) فقرة بنسبة (95 %) لـ(4) بدائل، و(33) فقرة بنسبة (82.5 %) لـ(5) بدائل، ولعدم وجود فروق بين متوسط الخطأ المعياري لمعلم الصعوبة ومعاملات ثبات الأفراد تعزى لعدد البدائل، ووجود فروق بين متوسط الخطأ المعياري في معالم القدرة للأفراد، والتي تمتعت بدقة أكثر للاختبار من (3) و(4) بدائل مقارنة بـ(5) بدائل، بينما لم تختلف بين (3) و(4) بدائل، كما أعطى الاختبار من (3) بدائل كمية معلومات أكبر من إختبار (4) و(5) بدائل عند مستويات القدرة المنخفضة، بينما أعطى الإختبار من (5) بدائل كمية معلومات أكبر عند مستوى القدرة المتوسطة والمرتفعة من (3) و(4) بدائل.

في حين كشفت دراسة **الزواهره (2014)** عن أثر عدد البدائل في تقدير معالم الأفراد والفقرات بإستخدام نموذج بارامتري (رباعي المعلم) ونموذج لابارامتري (موكن)، وذلك بتوليد بيانات ذات استجابة ثنائية (0، 1) حيث بلغ عدد الأفراد (4500) فرد بواقع (1500) لكل بديل (2، 3، 4)، وعدد فقرات الاختبار (45) فقرة، وقد توصلت النتائج لتحقق افتراضات نظرية الاستجابة للفقرة، ومطابقة جميع فقرات الاختبار باختلاف عدد البدائل، كما أن الاختبارات من (2، 3، 4) بدائل تعطي معلومات أكثر عند الأفراد ذوي القدرات دون المتوسط، كذلك وجود فروق في التخمين للنموذج الرباعي المعلم تعزى لمتغير عدد البدائل (3، 4) بدائل ولصالح (3) بدائل؛ ووجود فروق في معلمة التمييز عندما يكون عدد البدائل (4 ثم 3) مقارنة بنظيرتهما عندما يكون عدد البدائل (2)، ووجود فروق في تقديرات القدرة تعزى لمتغير عدد البدائل (2) مقارنة بعدد البدائل (3، 4).

بينما هدفت دراسة **أبو مسلم (2016)** للكشف عن أثر كل من عدد وترتيب البدائل لفقرات اختبار اختيار من متعدد تحصيلي على الخصائص السيكومترية لفقراته وفق النظرية الكلاسيكية، وتألف مجتمع الدراسة من جميع الطلبة الدارسين في برنامج الدبلوم العام في التربية بمعهد البحوث والدراسات التربوية في جامعة القاهرة للعام (2012 /2013)، فيما تألفت العينة من (113) طالباً وطالبة، منهم (51) طالباً و(62) طالبة، وذلك بإعداد إختبار تحصيلي في مقرر علم النفس التربوي، مكون من (50) فقرة من نوع الاختيار من متعدد، وقد توصلت النتائج إلى أنه في اختبار (3) بدائل يزداد الصدق والثبات عما لو كان (4، 5) بدائل، وذلك عندما

يتوزع البديل الصحيح بشكل عشوائي، كما توصلت إلى أن الاختبار يكون أكثر صعوبة وتمييزاً كلما زاد عدد البدائل، وذلك عندما يتوزع البديل الصحيح عشوائياً.

وقام **علاونة (2016)** بدراسة للكشف عن أثر حجم العينة ونمط الفقرة وجنس الطلبة على الخصائص السيكومترية لاختبار تحصيلي، فقد تألف مجتمع الدراسة من جميع طلبة الصف الحادي عشر للعلوم الإنسانية بمدينة نابلس والبالغ عددهم (3500) طالباً وطالبة، فيما تألفت العينة من (3356) طالباً وطالبة، والتي اختيرت بإستخدام العينة العشوائية البسيطة، وذلك ببناء إختبار تحصيلي في مادة الرياضيات، تكون من (72) فقرة بنموذجين ((36) فقرة اختيار من متعدد بأربع بدائل، (36) فقرة الصح والخطأ)، وقد توصلت النتائج لوجود فروق في ثبات الإختبار التحصيلي لصالح فقرات الاختيار من متعدد، بينما لا توجد فروق في الصعوبة والتمييز وفق نمط الفقرة (اختيار من متعدد، الصح والخطأ).

وفي دراسة **العنزي (2017)** والتي هدفت إلى تعرف أثر شكل فقرات الإختيار من متعدد والصواب والخطأ على معالم الفقرة والخصائص السيكومترية للاختبار وفق النموذج ثنائي المعلم، وقد تألف مجتمع الدراسة من جميع طلبة الصف الأول الثانوي بمدينة تبوك والبالغ عددهم (3268) طالباً، فيما تألفت العينة من (520) طالباً، والتي اختيرت بإستخدام العينة العشوائية العنقودية، وذلك ببناء اختبار تحصيلي في مادة الحاسب وتقنية المعلومات، تكون من (24) فقرة بنموذجين (اختيار من متعدد، الصواب والخطأ)، وقد توصلت النتائج لتحقق افتراضات نظرية الاستجابة للفقرة، ومطابقة (19) فقرة بنسبة (79.2 %) لشكلي الفقرات، ولعدم وجود فروق في الصعوبة والثبات بين شكلي الاختبار، بينما وجد فروق في التمييز لصالح اختبار الصواب والخطأ.

ومن خلال الإطلاع على الدراسات السابقة يلاحظ أنها تباينت واختلفت في العديد من القضايا التي استهدفتها بالدراسة، فمنها ما ركزت على النظرية الكلاسيكية كدراسات دبك (1998) وفرانسيسكو وأوليا وبونسودا (Francisco, Olea & Ponsoda, 2001) والتعمري (2003) وأبو مسلم (2016) وعلاونة (2016)، ومنها ما ركزت على نمط فقرات الإختبار متجاهلة عدد بدائلها كدراستي العنزي (2017) وعلاونة (2016)، وكذلك ما ركزت على نماذج إستجابة الفقرة رباعي المعلم كدراسة الزواهره (2014) والتي كانت على بيانات مولدة وليست حقيقية، وثلاثي المعلم كدراستي فرانسيسكو وأوليا وبونسودا (Francisco, Olea & Ponsoda, 2001) والتعمري (2003)، وثنائي المعلم كدراسة العنزي (2017)، وأحادي المعلم كدراسة الشريفين وطعامنة (2009)، لكن الدراسة الحالية تتشابه مع الدراسات التي تطرقت للنظرية الكلاسيكية، كذلك الدراسات التي ركزت على نمط الفقرات لكن تختلف عنها بدراسة عدد الفقرات والبدائل والتفاعل بينهما، كذلك تشابهها مع دراسة الزواهره (2014) بدراسة النموذج رباعي المعلم لكن تختلف معها في أن بيانات الدراسة الحالية حقيقية.

وباستعراض الدراسات السابقة يلاحظ أن أغلبها أشارت إلى أن أفضل عدد بدائل كانت لنموذج الاختبار من (3) بدائل في تقدير معالم الصعوبة والتمييز والتخمين والقدرة كدراسات دبك (1998) وفرانسيسكو وأوليا وبونسودا (Francisco, Olea & Pon-soda, 2001) والتعمري (2003) والشريفين وطعامنة (2009)، كذلك زيادة الصدق والثبات كدراسة أبو مسلم (2016)، كما أن أغلبها أشارت لعدم وجود فروق بين عدد البدائل (3، 4) في تقدير معالم الصعوبة والتمييز والتخمين والقدرة كدراسات دبك (1998) وفرانسيسكو وأوليا وبونسودا (Francisco, Olea & Ponsoda, 2001) والشريفين وطعامنة (2009)، وتتشابه الدراسة الحالية مع الدراسات التي تطرقت لعدد البدائل لكن تختلف عنها بدراسة عدد الفقرات والبدائل والتفاعل بينهما في نظريتي القياس الكلاسيكية ونظرية الاستجابة للفقرة.

كما أن جميع الدراسات تجاهلت متغير عدد فقرات الاختبار وركزت على عدد البدائل ونمط أو شكل الفقرة على الرغم من أهميته في تقدير قدرة الأفراد، وعدم تناولها حجم الفقرات المشكلة للاختبار ومدى تأثيرها في خصائص الإختبار، وتركيزها على عدد الفقرات المرتفع (49، 221، 35، 40، 45، 50، 36) فقرة باستثناء دراسة العنزي (2017) والتي ركزت على (24) فقرة وكان أدنى عدد فقرات في الدراسات، والذي تتشابه فيه مع الدراسة الحالية، التي ركزت على عدد فقرات (9، 15، 24) بمجموع (48) فقرة، كذلك ندرة الدراسات التي تناولت الطلبة الجامعيين وتركيزها على طلبة المرحلة الأساسية والثانوية بإستثناء دراسة فرانسيسكو وأوليا وبونسودا (Francisco, Olea & Ponsoda, 2001)، كذلك ندرة الدراسات التي تتناول مواد أكاديمية تلامس واقع الحياة كمواد علم النفس والاتصال والتواصل، وتركيزها على مواد أكاديمية بحثية تقيس قدرات مهارية كالرياضيات والفيزياء والحاسوب واللغة الإنجليزية.

وتسعى الدراسة الحالية لسد الفجوة ودراسة ما تم تجاهله في الدراسات السابقة، بحيث جمعت ما بين عدد البدائل والفقرات والنموذج رباعي المعلم على بيانات حقيقية في تقدير معالم الأفراد والفقرات، وهو ما لم يجر تناوله في أي من الدراسات السابقة، بهدف الكشف عن أفضل عدد من البدائل، كذلك عن أفضل عدد من الفقرات، والتفاعل بينهما في تقدير معالم الأفراد والفقرات وفق استخدام النموذج رباعي المعلم من خلال استخدام بيانات حقيقية، بحيث يمثل ذلك إضافة علمية لم يتم التحقق منها من خلال الدراسات السابقة في ضوء اطلاع الباحث.

مشكلة الدراسة وأسئلتها:

يعد القياس عملية مهمة في الوقت الحاضر من أجل الكشف عن كمية ما يمتلكه الأفراد من مهارات ومعلومات بشتى الجوانب، ومن أجل أن تتم عملية القياس فلا بد من استخدام أدوات لذلك، وترتكز الدراسة الحالية على الإختبارات كأداة قياس لكونها تحقق أهداف الدراسة، كما أنها تعتبر من أفضل الأدوات لقياس التحصيل، فتستخدم بشتى أشكالها وأنواعها من أجل التعرف على مقدار ما يمتلكه الأفراد من السمة المقيسة، ولأنها أداة فاعلة ومهمة كان لا بد وأن يتم بنائها على أسس سليمة وموضوعية، بالإضافة لأن تمتع بدرجة عالية من الخصائص السيكومترية المحكمة والموثوقة بحيث تحقق الهدف المنشود منها، وهو إستخدام فقراتها عند إعداد اختبارات ذات مواصفات محددة، لذلك أتجه معدو الإختبارات إلى

إعدادها بأقل وقت وجهد ومال، بحيث تتكون من عدد مناسب من الفقرات والبدائل وتتمكن من قياس سمة واحدة محددة بأقصى درجة من الدقة والموضوعية، إضافة إلى إمتلاكها خصائص سيكومترية جيدة، ونظراً لأهمية بناء الإختبارات وضرورة مراعاتها لعدد الفقرات المشكلة للاختبار أو عدد بدائلها، جاءت هذه الدراسة لتكشف عن أفضل عدد مناسب لتشكيل اختبار تحصيلي من بين ثلاث مستويات من عدد الفقرات (9، 15، 24)، إضافة لأخذ عدد البدائل في الاعتبار بتناولها ثلاثة أنماط (2، 3، 4)، **والتفاعل بينهما، والتي أدرجت وفق التساؤلات التالية:**

1. ما مدى مطابقة الأفراد والفقرات للنموذج رباعي المعلم وفق عدد البدائل والفقرات؟
2. ما تقدير معالم فقرات الإختبار التحصيلي والخطأ المعياري في نظريتي الكلاسيكية واستجابة الفقرة وفق عدد البدائل والفقرات؟
3. ما مدى تحقق افتراض أحادية البعد للاختبار التحصيلي وفق عدد البدائل والفقرات؟
4. ما مدى تحقق افتراض اطرادية السمة للاختبار التحصيلي وفق عدد البدائل والفقرات؟
5. ما مدى تحقق دالة المعلومات للفقرات والاختبار التحصيلي وفق عدد البدائل والفقرات؟
6. ما تقديرات القدرة لدى أفراد الدراسة على الاختبار التحصيلي وفق عدد البدائل والفقرات؟
7. ما أثر عدد البدائل والفقرات على التوافق مع معالم الفقرات في النظرية الكلاسيكية؟
8. ما أثر عدد البدائل والفقرات على التوافق مع معالم الفقرات في نظرية الاستجابة للفقرة؟

أهمية الدراسة:

تكمن أهمية هذه الدراسة في جانبين فمن الناحية النظرية تحاول تسليط الضوء على كيفية بناء إختبار على درجة عالية من الفاعلية والكفاءة من حيث الخصائص السيكومترية لفقراته (الصعوبة، التمييز، التخمين، اللامبالاه) بحيث يأخذ بالاعتبار عدد الفقرات وبدائلها، كذلك تناول النموذج رباعي المعلم ومعالمه وإفتراضاته.

أما من الناحية العملية فتحاول إفادة العاملين في المجال التربوي والنفسي ومطورو الإختبارات بمحتوى الإختبار الجيد، كذلك يتوقع منها أن تقدم عدد الفقرات والبدائل الأنسب لقياس التحصيل، إضافة لمقارنتها بين معالم الفقرات في نظريتي القياس الكلاسيكية والحديثة، والإسهام بتناول العديد من المفاهيم الحديثة (كمؤشرات مطابقة الفقرات وإفتراض الإطرادية).

أهداف الدراسة:

تهدف الدراسة الحالية للكشف عن أثر عدد البدائل والفقرات في تقدير معالم الأفراد والفقرات بإستخدام النموذج رباعي المعلم البارامتري.

التعريفات الاصطلاحية والإجرائية:

عدد البدائل: وهي عدد الاستجابات الممكنة للفقرة، والتي تكون إحداهما صحيحة والأخرى خاطئة تسمى بالمموهات (المشتتات)، وفي هذه الدراسة تم استخدام عدد البدائل (2، 3، 4).

عدد الفقرات: وهي عدد الأسئلة التي تشكل الاختبار، وفي هذه الدراسة تم إستخدام عدد الفقرات (9، 15، 24).

تقدير معالم الأفراد: وهي قيمة تدريجية تنتج من تطبيق الدالة الرياضية لنموذج استجابة الفقرة عند إستجابة المفحوص على مجموعة من الفقرات المشكلة للاختبار، والتي تقيس مدى إمتلاكه للسمة المراد قياسها، إذ يكون لكل فرد معلم يسمى معلم الفرد أو القدرة.

تقدير معالم الفقرات: وهي قيمة تدريجية تنتج من تطبيق الدالة الرياضية لنموذج إستجابة الفقرة عند استجابة المفحوص على مجموعة من الفقرات المشكلة للاختبار، والتي تقيس الخصائص السيكومترية المتعلقة بالفقرة كالصعوبة والتمييز والتخمين واللامبالاه لمنحنى خصائص الفقرة.

النموذج رباعي المعلم البارامتري: وهو أحد نماذج استجابة الفقرة البارامترية ثنائية التدريج، والذي يتم من خلاله معرفة احتمالية استجابة الفرد على الفقرة من خلال أربع معالم (الصعوبة b، التمييز a، التخمين c، الحد التقاربي الأعلى d).

حدود الدراسة:

1. الحدود الموضوعية: اقتصرت الدراسة على بناء إختبار تحصيلي نصفي في مادة علم النفس والحياة، كما تتحدد أداة الدراسة بمدى مطابقتها للنموذج رباعي المعلم البارامتري ومدى توافر دلالات إحصائية مقبولة عن الخصائص السيكومترية لأداة الدراسة، كذلك تعتمد النتائج التي تم التوصل إليها على البرامج الحاسوبية التي أمكن توفيرها أثناء تحليل البيانات التي جمعت في هذه الدراسة.
2. الحدود الزمانية: اقتصرت الدراسة على مادة علم النفس والحياة في الفصل الدراسي الأول من العام الجامعي (2017/ 2018) والتي تمثل فيها مجتمع الدراسة.
3. الحدود المكانية: اقتصرت الدراسة على عينة من طلبة جامعة عمان الأهلية في الفصل الدراسي الأول من العام الجامعي (2017 / 2018).

محددات الدراسة:

1. يقتصر تعميم نتائج الدراسة بالقدر الذي يكون فيه هناك عينات مماثلة للعينة التي تم إختيارها من الطلبة.
2. أقتصرت أداة الدراسة على بناء فقرات ثنائية التدريج من نوع (الاختيار من متعدد بـ(2، 3، 4) بدائل) وبأعداد (9، 15، 24) فقرة، في مادة علم النفس والحياة، والذي كان محتواه في وحدتين دراسيتين هما (ماهية علم النفس والتعلم، من كتاب علم النفس

العام لمؤلفه معاوية أبو غزال 2015).

3. تتحدد أداة الدراسة في مدى مطابقتها للنموذج رباعي المعلم البارامتري ومدى توافر دلالات إحصائية مقبولة عن الخصائص السيكومترية لأداة الدراسة .

4. تعتمد النتائج التي تم التوصل إليها على البرامج الحاسوبية التي أمكن توفيرها أثناء تحليل البيانات التي جمعت في هذه الدراسة.

منهج البحث:

استخدم المنهج الوصفي المسحي والذي يعتبر من أكثر المناهج استخداماً في مثل هذه الدراسات، وهذا ينسجم مع طبيعة الدراسة وأهدافها، وذلك باستخدام الاختبار كأداة لجمع البيانات اللازمة.

مجتمع الدراسة وعينتها:

تألف مجتمع الدراسة من جميع طلبة جامعة عمان الأهلية الدارسيين لمادة علم النفس والحياة في الفصل الدراسي الأول من العام الجامعي (2017 / 2018)، والبالغ عددهم (245) طالباً وطالبة، أما عينة الدراسة فقد تكونت من (237) طالباً وطالبة، منهم (156) طالباً و(81) طالبة، وقد شكلت العينة ما نسبته (96.7 %) من مجتمع الدراسة الأصلي، والتي أخذت بأسلوب العينة القصدية، لكون الباحث هدف لإجراء دراسته على طلبة المساق الذي يقوم بتدريسه، وذلك لسهولة الحصول على البيانات وعلى أذن من الجامعة لجمعها، ولكون إعداد نماذج الاختبار تمت وفق أسس القياس والتقويم، كذلك لمناسبة عدد الأفراد لإجراء دراسة علمية.

أداة الدراسة:

قام الباحث بكتابة إختبار تحصيلي نصفي في مادة علم النفس والحياة، والذي كان محتواه في وحدتين دراسيتين هما (ماهية علم النفس والتعلم، من كتاب علم النفس العام لمؤلفه معاوية أبو غزال 2015) والذي أعتبر أداة الدراسة، فقد تألفت من ثلاثة نماذج إختبارية، فكل نموذج تكون من (48) فقرة، النموذج الأول تكون من (9 فقرات من (4) بدائل، 15 من بديلين، 24 من (3) بدائل)، والنموذج الثاني تكون من (9 فقرات من (3) بدائل، 15 من (4) بدائل، 24 فقرة من بديلين)، والنموذج الثالث تكون من (9 فقرات من بديلين، 15 من (3) بدائل، 24 من (4) بدائل).

كما استخدم تدريج ثنائي بحيث تعطى الدرجة (1) للإجابة الصحيحة والدرجة (0) للإجابة الخاطئة، بحيث تتراوح الدرجة الكلية على الإختبار بين (0– 48).

صدق أداة الدراسة وثباتها:

صدق الأداة: يشير الصدق إلى ما إذا كان المقياس المعد يقيس بالفعل ما وضع لقياسه ولا يقيس شيء أخر، وقد أعتمد الباحث على الصدق الظاهري من أجل التحقق من صدق الأداة المصممة من خلال عرض النماذج الاختبارية على زملائه مدرسي المادة، والبالغ عددهم (4)، والذين اعتبروا كمحكمين لتمتعهم بالإختصاص والخبرة، وقد جرى الأخذ برأيهم حول مدى مناسبة الأداة التي وُضعتْ من حيث ملائمتها للهدف المراد قياسه والدقة والصياغة والسلامة اللغوية ووضوحها وخلوها من الإشارات التوجيهية ومناسبتها، وقد قام الباحث في ضوء ملاحظات المحكمين بتعديل الصياغة اللغوية لثلاث فقرات، وحذف فقرتين، وتعديل (10) بدائل بحيث تكون مناسبة.

كما قام الباحث بالتحقق من الصدق البنائي على العينة الأساسية من خلال التحقق من إيجاد معامل الإرتباط بين الفقرة والدرجة الكلية على الإختبار الفرعي، والتي يعرضها الجدول (1)،

الجدول (1)

معاملات الإرتباط بين الفقرة والدرجة الكلية على الإختبار الفرعي بإختلاف عدد البدائل والفقرات

النموذج	ن	الفقرات	معاملات الارتباط		عدد المعاملات		
			أدنى	أعلى	**	*	غير دال
الأول	79	9 فقرات أربع بدائل	.162	.624	7	1	1
		15 فقرة بديلين	.016	.553	11	2	2
		24 فقرة ثلاث بدائل	.188	.550	21	1	2
		كل الفقرات	.054	.546	30	9	9
الثاني	78	9 فقرات ثلاث بدائل	.211	.493	7	1	1
		15 فقرة أربع بدائل	.124	.511	11	3	1
		24 فقرة بديلين	.020	.530	14	4	6
		كل الفقرات	.055	.484	25	6	17

النموذج	ن	الفقرات	معاملات الإرتباط أدنى	أعلى	عدد المعاملات **	*	غير دال
الثالث	80	9 فقرات بديلين	.188	.590	6	2	1
		15 فقرة ثلاث بدائل	.225	.615	13	2	0
		24 فقرة أربع بدائل	.173	.670	21	2	1
		كل الفقرات	.005	.664	37	3	8
عدد الفقرات باختلاف عدد البدائل	237	9 فقرات	.271	.480	9	0	0
		15 فقرة	.297	.507	15	0	0
		24 فقرة	.126	.526	23	0	1
		كل الفقرات	.145	.499	46	2	0

* دالة إحصائياً عند مستوى الدلالة (α = .05)

** دالة إحصائياً عند مستوى الدلالة (α = .01)

حيث أظهرت النتائج أن معامل الإرتباط بين الفقرة والدرجة الكلية على الإختبار الفرعي دالة إحصائياً عند مستوى الدلالة (.01) و(.05)، بإستثناء بعض الفقرات جاءت غير دالة إحصائيا عند مستوى الدلالة (.01) و(.05) والتي ينصح بتعديلها أو حذفها، وهذا مؤشر على أن المقياس يتمتع بمؤشرات صدق بنائي مرتفعة تدلل على مدى صلاحيته لقياس ما وضع لقياسه ومناسبتها لتحقيق غرض الدراسة.

الثبات: استخدمت معادلة كرونباخ الفا للتحقق من دلالات ثبات أداة الدراسة، والتي تعمل على فحص الإتساق الداخلي لاستجابات المفحوصين على فقرات الأداة، وقد بلغت قيم معامل الثبات، والتي يعرضها الجدول (2).

الجدول (2)

قيم معاملات الثبات على الإختبار الفرعي بإختلاف عدد البدائل والفقرات

نموذج الإختبار	معاملات الثبات 9 فقرات	15فقرة	24 فقرة	كل الفقرات
الأول	أربع بدائل	بديلين	ثلاث بدائل	الكلي
	.416	.785	.431	.827
الثاني	ثلاث بدائل	أربع بدائل	بديلين	الكلي
	.174	.513	.610	.760
الثالث	بديلين	ثلاث بدائل	أربع بدائل	الكلي
	.189	.748	.773	.850

يلاحظ من الجدول أعلاه أن قيم معامل الثبات لـ (9 فقرات (2، 3) بدائل) منخفضة، بينما كانت معاملات الثبات الأخرى جيدة وهذا يعبر عن مدى صلاحية الأداة ومناسبتها لتحقيق غرض الدراسة.

إجراءات الدراسة:

وقد أتبع الباحث الإجراءات التالية من أجل جمع البيانات، إذ قام بمخاطبة رئاسة قسمه من أجل تسهيل مهمته بجمع البيانات بجميع شعب مادة علم النفس والحياة، ومن ثم جرى إبلاغ مدرسي المادة بذلك، وقام الباحث لكونه منسق المادة في حينه، بتحليل محتوى المادة الدراسية المشمولة بالدراسة، وكذلك إعداد جدول مواصفات للاختبار، ومن ثم أعد الاختبار بنماذجه الثلاثة، وقام بعرضها على مدرسي المادة الذين اعتبروا كمحكمين، وقام بأخذ ملاحظاتهم في الاعتبار عند إعداد الصورة النهائية للاختبار، ومن ثم جرى تحديد مكان وزمان الاختبار وكيفية توزع الطلبة والمراقبين على القاعات التدريسية، وقام بجمع أوراق الاختبار بعد تطبيقه، وأعد مفتاح مثقب للتصحيح بالتعاون مع مدرسي المادة، وقاموا بتصحيحها، ومن ثم قام الباحث بتجهيز الإستجابات من أجل إدخالها على البرامج الإحصائية المستخدمة في الدراسة لتحليلها إحصائياً والحصول على النتائج المتعلقة بالإجابة على أسئلة الدراسة.

المعالجات الإحصائية:

جرى التحقق من الصدق البنائي باستخدام معامل الارتباط بيرسون، والثبات وفق معادلة كرونباخ الفا بإستخدام برنامج (spss)، كما جرى مطابقة الأفراد والفقرات وتقدير معالم الفقرات للنموذج رباعي المعلم، كذلك التحقق من إفتراضي (أحادية البعد، اطرادية السمة) ودالة معلومات الفقرة والاختبار وتقديرات القدرة باستخدام برنامج (Exametrika v.5.3)، كما تم الكشف عن أثر عدد البدائل والفقرات في نظريتي القياس الكلاسيكية والاستجابة للفقرة باستخدام برنامج (spss).

نتائج الدراسة ومناقشتها

يتناول هذا الجزء عرضاً لنتائج الدراسة التي تحصل عليها الباحث، وفقاً للتسلسل الآتي:

◄ أوَّلاً: للإجابة عن سؤال الدراسة الأوَّل والذي ينص على:

"ما مدى مطابقة الأفراد والفقرات للنموذج رباعي المعلم وفق عدد البدائل والفقرات؟".

1. **التحقق من مطابقة الأفراد**:

تمت مطابقة إستجابات (237) فرداً؛ منهم (79) فرد طبق عليهم النموذج الأول، و(78) الثاني، و(79) الثالث، حيث لم يتم حذف أي منهم لمطابقتهم للنموذج رباعي المعلم، وذلك من خلال إستخدام برنامج (Exametrika v.5.3).

2. **التحقق من مطابقة الفقرات**:

تم التحقق من مطابقة الفقرات للنموذج رباعي المعلم وفق عدد البدائل والفقرات لبيانات أفراد الدراسة، من خلال برنامج (Exametrika v.5.3).

3. **عدد الفقرات** (9) **وبإختلاف عدد البدائل**:

وقد توصلت النتائج وفق مؤشر ($\chi 2$) المحسوبة أحد مؤشرات جودة المطابقة المطلقة لفقرات الإختبار أن (8 فقرات) بنسبة (88.9 %) للبديلين والأربع بدائل، و(9 فقرات) بنسبة (100 %) للثلاثة بدائل كانت مطابقة بدرجة ممتازة، والتي تشير إلى أن قيمة مؤشر ($\chi 2$) المحسوبة بين (0 إلى 2) تؤشر على مطابقة ممتازة، بينما كانت (فقرة واحدة) بنسبة (11.1 %) مطابقة بشكل جيد للبديلين والأربع بدائل، والتي تشير إلى أن قيمة مؤشر ($\chi 2$) المحسوبة بين (أعلى من 2 إلى 3) تؤشر على مطابقة جيدة، كذلك وفق مؤشري (CFI, RMSEA) إحدى مؤشرات جودة المطابقة غير المركزية لفقرات الإختبار أن جميع الفقرات (بديلين، ثلاث، أربع بدائل) كانت مطابقة بنسبة (100 %) بدرجة ممتازة، والتي تشير إلى أن قيمة مؤشر (CFI) بين (0.97 إلى 1) ومؤشر (RMSEA) بين (0 إلى 0.05) تؤشر على مطابقة ممتازة (Karin, Helfried & Hans, 2003).

4. **عدد الفقرات** (15) **وبإختلاف عدد البدائل**:

كما توصلت النتائج وفق مؤشر ($\chi 2$) المحسوبة أن (11 فقرة) بنسبة (73.3 %) للبديلين، و(12 فقرة) بنسبة (80 %) للثلاث بدائل، و(15 فقرة) بنسبة (100 %) للأربع بدائل مطابقة بدرجة ممتازة، والتي تشير إلى أن قيمة مؤشر ($\chi 2$) المحسوبة بين (0 إلى 2) تؤشر على مطابقة ممتازة، بينما كانت (فقرة واحدة) بنسبة (6.7 %) مطابقة بشكل جيد للثلاث بدائل، والتي تشير إلى أن قيمة مؤشر ($\chi 2$) المحسوبة بين (أعلى من 2 إلى 3) تؤشر على مطابقة جيدة، فيما كانت (4 فقرات) بنسبة (26.7 %) للبديلين، و(فقرتين) بنسبة (13.3 %) للثلاث بدائل غير مطابقة للنموذج، والتي تشير إلى أن قيمة مؤشر ($\chi 2$) المحسوبة أعلى من (3) تؤشر على عدم مطابقة الفقرات للنموذج، كذلك وفق مؤشري (CFI, RMSEA) أن جميع الفقرات كانت مطابقة بنسبة (100 %) بدرجة ممتازة، والتي تشير إلى أن قيمة مؤشر (CFI) بين (0.97 إلى 1)، ومؤشر (RMSEA) بين (0 إلى 0.05)، تؤشر على مطابقة ممتازة (Karin, Helfried & Hans, 2003).

5. **عدد الفقرات** (24) **وبإختلاف عدد البدائل**:

كما توصلت النتائج وفق مؤشر ($\chi 2$) المحسوبة لفقرات الإختبار أن (9 فقرات) بنسبة (37.5 %) للبديلين، و(17 فقرة) بنسبة (70.8 %) للثلاث بدائل، و(19 فقرة) بنسبة (79.2 %) للأربع بدائل مطابقة بدرجة ممتازة، والتي تشير إلى أن قيمة مؤشر ($\chi 2$) المحسوبة بين (0 إلى 2) تؤشر على مطابقة ممتازة، بينما كانت (7 فقرات) بنسبة (29.2 %) للبديلين، و(5 فقرات) بنسبة (20.8 %) للثلاث بدائل، و(3 فقرات) بنسبة (12.5 %) للأربع بدائل مطابقة بشكل جيد، والتي تشير إلى أن قيمة مؤشر ($\chi 2$) المحسوبة بين (أعلى من 2 إلى 3) تؤشر على مطابقة جيدة، فيما كانت (8 فقرات) بنسبة (33.3 %) للبديلين، و(فقرتين) بنسبة (8.3 %) للثلاث والأربع بدائل غير مطابقة للنموذج، التي تشير إلى أن قيمة مؤشر ($\chi 2$) المحسوبة أعلى من (3) تؤشر على عدم مطابقة الفقرات للنموذج، كذلك وفق مؤشري (CFI, RMSEA) أن جميع الفقرات كانت مطابقة بنسبة (100 %) بدرجة ممتازة، والتي تشير إلى أن قيمة مؤشر (CFI) بين (0.97 إلى 1)، ومؤشر (RMSEA) بين (0 إلى 0.05) تؤشر على مطابقة ممتازة (Karin, Helfried & Hans, 2003).

كما توصلت النتائج أن جميع الفقرات (9، 15، 24) من (بديلين، ثلاث بدائل، أربع بدائل) مطابقة للنموذج رباعي المعلم، إذ لم تستبعد أي فقرة غير مطابقة لجميع مؤشرات المطابقة، وهذا يشير إلى تحقق إفتراضي مطابقة الفقرات ضمن مؤشر أو أكثر من مؤشرات المطابقة.

وتتفق نتيجة هذه الدراسة مع دراسة الزواهره (2014) بمطابقة جميع فقرات الاختبار باختلاف عدد بدائلها، كما تتفق مع دراسة العنزي (2017) في عدد ونسبة الفقرات المطابقة للنموذج (4) بدائل، لكن تختلف معها في نموذج البديلين، كما تختلف مع دراسة فرانسيسكو وأوليا وبونسودا (Francisco, Olea & Ponsoda, 2001) والشريفين وطعامنة (2009) في عدد ونسبة الفقرات المطابقة للنموذج باختلاف عدد بدائلها.

ويعزو الباحث هذه النتيجة لدقة ومتانة الإختبار في قياس سمة التحصيل، وهذا يؤكد مدى سلامة الإجراءات المتبعة في بنائه، والتي أرتسمت بتحقق خصائص الاختبار الجيد وهي الصدق والثبات والموضوعية، ومدى صحة وسلامة البيانات التي جرى جمعها والتي أشرت على جودة وملائمة الفقرات للنموذج المستخدم، وتمتعها بمستوى مناسب من الصعوبة وملائمتها لمستويات القدرة المختلفة للأفراد الخاضعين للاختبار.

◄ **ثانياً**: للإجابة عن سؤال الدراسة الثاني والذي ينص على: "ما تقدير معالم فقرات الاختبار التحصيلي والخطأ المعياري في نظريتي الكلاسيكية واستجابة الفقرة وفق عدد البدائل والفقرات؟".

6. **عدد الفقرات** (9) **بإختلاف عدد البدائل**:

تم تقدير معالم فقرات الاختبار التحصيلي والخطأ المعياري في نظريتي الكلاسيكية واستجابة الفقرة وفق عدد الفقرات (9) بإختلاف عدد البدائل، والتي يعرضها الجدول (3).

الجدول (3)

إحصائيات معالم فقرات الإختبار التحصيلي والخطأ المعياري في نظريتي الكلاسيكية وإستجابة الفقرة وفق عدد الفقرات (9) باختلاف عدد البدائل

عدد البدائل	إحصائيات	النظرية الكلاسيكية		نظرية إستجابة الفقرة							
		معلمة التمييز	معلمة الصعوبة	a	S.E	b	b	S.E	c	S.E	d	d	S.E
بديلين	minimum	.363	.188	.262	.145	-2.455	.365	.103	.078	.717	.051		
	maximum	.888	.590	1.328	.537	.701	1.564	.518	.136	1.000	.114		
	mean	.631	.360	.616	.302	-.709	.894	.271	.113	.867	.080		
	std	.177	.130	.344	.125	.993	.428	.138	.019	.094	.021		
ثلاث بدائل	minimum	.359	.211	-.248	.104	-5.566	.262	.104	.076	.619	.022		
	maximum	.846	.493	1.818	.597	1.748	111.449	.662	.135	1.000	.121		
	mean	.584	.361	.486	.335	-.382	22.357	.265	.114	.837	.091		
	std	.160	.087	.727	.182	2.171	39.332	.173	.018	.127	.034		
أربع بدائل	minimum	.266	.162	.160	.139	-2.288	.363	.082	.054	.729	.051		
	maximum	.810	.624	1.196	.556	1.410	2.537	.455	.127	1.000	.138		
	mean	.533	.420	.770	.358	-.090	.882	.215	.093	.855	.093		
	std	.188	.139	.372	.126	1.098	.770	.122	.025	.100	.028		

يتضح من الجدول أعلاه أن تقديرات معالم الفقرات في النظرية الكلاسيكية لفقرات البديلين كانت على الشكل التالي، إذ تراوحت معلمة التمييز بين (.363 – .888) بمتوسط (.631) وإنحراف (.177)، وللثلاثة بدائل تراوحت بين (.359 – .846) بمتوسط (.584) وإنحراف (.160)، وللأربع بدائل تراوحت بين (.266 – .810) بمتوسط (.533) وانحراف (.188).

بينما تراوحت معلمة الصعوبة لفقرات للبديلين بين (.188 – .590) بمتوسط (.360) وإنحراف (.130)، وللثلاثة بدائل تراوحت بين (0.211 – 0.493) بمتوسط (.361) وانحراف (.087)، وللأربع بدائل تراوحت بين (.162 – .624) بمتوسط (.420) وإنحراف (.139).

كما يتضح من الجدول أعلاه أن تقديرات معلم التمييز في نظرية إستجابة الفقرة لفقرات البديلين تراوحت بين (.262 – 1.328) بمتوسط (.616) وانحراف (.344)، وللثلاث بدائل بين ((.248 –) – 1.818) بمتوسط (.486) وانحراف (.727)، وللأربع بدائل بين (.160 – 1.196) بمتوسط (.770) وانحراف (.372).

بينما تراوحت معلمة الصعوبة لفقرات البديلين بين ((–2.455)– (.701)) بمتوسط (.709 –) وإنحراف (.993)، وللثلاث بدائل بين ((5.566–)– (1.748)) بمتوسط (.382 –) وإنحراف (2.171)، وللأربع بدائل بين ((2.288–)– (1.410)) بمتوسط (.090 –) وإنحراف (1.098).

بينما تراوحت معلمة التخمين لفقرات البديلين بين (0.103 – .518) بمتوسط (.271) وإنحراف (.138)، وللثلاث بدائل بين (.104 – .662) بمتوسط (.265) وإنحراف (.173)، وللأربع بدائل بين (.082 – .455) بمتوسط (.215) وإنحراف (0.122).

بينما تراوحت معلمة اللامبالاه لفقرات البديلين بين (.717 – 1.000) بمتوسط (.867) وإنحراف (.094)، وللثلاث بدائل بين (.619 – 1.000) بمتوسط (.837) وإنحراف (.127)، وللأربع بدائل بين (.729 – 1.000) بمتوسط (.855) وإنحراف (.100).

7. **عدد الفقرات (15) بإختلاف عدد البدائل:**

تم تقدير معالم فقرات الاختبار التحصيلي والخطأ المعياري في نظريتي الكلاسيكية واستجابة الفقرة وفق عدد الفقرات (15) باختلاف عدد البدائل، والتي يعرضها الجدول (4).

الجدول (4)

إحصائيات معالم فقرات الإختبار التحصيلي والخطأ المعياري في نظريتي الكلاسيكية وإستجابة الفقرة وفق عدد الفقرات (15) باختلاف عدد البدائل

عدد البدائل	إحصائيات	النظرية الكلاسيكية		نظرية إستجابة الفقرة							
		معلمة التمييز	معلمة الصعوبة	a		b		c		d	
				a	S.E	b	S.E	c	S.E	d	S.E
بديلين	minimum	.253	.016	.029	.040	-8.213	.416	.052	.076	.694	.040
	maximum	.911	.553	1.153	.465	.921	12.168	.375	.175	1.000	.108
	mean	.717	.327	.635	.270	-1.158	1.439	.266	.133	.908	.068
	std	.177	.125	.266	.096	2.055	2.974	.088	.022	.079	.021
ثلاث بدائل	minimum	.513	.225	.436	.162	-1.915	.345	.152	.100	.813	.030
	maximum	.938	.615	1.410	.580	.022	.957	.488	.174	1.000	.088
	mean	.748	.466	.951	.369	-.805	.525	.272	.131	.932	.051
	std	.131	.113	.261	.105	.635	.170	.106	.022	.053	.016
أربع بدائل	minimum	.410	.124	.016	.069	-4.836	.402	.108	.075	.801	.011
	maximum	.949	.511	1.166	.496	1.082	23.483	.505	.194	1.000	.109
	mean	.697	.362	.717	.312	-.914	2.246	.325	.120	.930	.069
	std	.146	.108	.346	.108	1.464	5.889	.130	.030	.060	.031

يتضح من الجدول أعلاه أن تقديرات معالم الفقرات في النظرية الكلاسيكية لفقرات البديلين كانت على الشكل التالي، حيث تراوحت معلمة التمييز بين (.253 – .911) بمتوسط (.717) وانحراف (.177)، وللثلاثة بدائل تراوحت بين (.513 – .938) بمتوسط (.748) وانحراف (.131)، وللأربع بدائل تراوحت بين (.410 – .949) بمتوسط (.697) وانحراف (.146).

بينما تراوحت معلمة الصعوبة لفقرات للبديلين بين (.016 – .553) بمتوسط (0.327) وانحراف (.125)، وللثلاثة بدائل تراوحت بين (–0.225 0.615) بمتوسط (.466) وانحراف (.113)، وللأربع بدائل تراوحت بين (.124 – .511) بمتوسط (.362) وإنحراف (.108).

كما يتضح من الجدول أعلاه أن تقديرات معلم التمييز في نظرية استجابة الفقرة لفقرات البديلين تراوحت بين (.029 – 1.153) بمتوسط (.635) وإنحراف (.266)، وللثلاث بدائل بين ((.436 – 1.410) بمتوسط (0.951) وانحراف (.261)، وللأربع بدائل بين (.016 – 1.166) بمتوسط (.717) وانحراف (.346).

بينما تراوحت معلمة الصعوبة لفقرات البديلين بين ((–8.213)– (.921)) بمتوسط (–1.158) وإنحراف (2.055)، وللثلاث بدائل بين ((–1.915)– (.022)) بمتوسط (.805 –) وإنحراف (.635)، وللأربع بدائل بين ((–4.836)– (1.082)) بمتوسط (.914 –) وإنحراف (1.464).

بينما تراوحت معلمة التخمين لفقرات البديلين بين (.052 – .375) بمتوسط (.266) وانحراف (.088)، وللثلاث بدائل بين (.152 – .488) بمتوسط (0.272) وانحراف (.106)، وللأربع بدائل بين (.108 – .505) بمتوسط (.325) وانحراف (.130).

بينما تراوحت معلمة اللامبالاه لفقرات البديلين بين (0.694 – 1.000) بمتوسط (.908) وانحراف (0.079)، وللثلاث بدائل بين (.813 – 1.000) بمتوسط (.932) وانحراف (.053)، وللأربع بدائل بين (.801 – 1.000) بمتوسط (.930) وانحراف (.060).

8. **عدد الفقرات (24) بإختلاف عدد البدائل:**

تم تقدير معالم فقرات الإختبار التحصيلي والخطأ المعياري في نظريتي الكلاسيكية واستجابة الفقرة وفق عدد الفقرات (24) باختلاف عدد البدائل، التي يعرضها الجدول (5).

الجدول (5)

إحصائيات معالم فقرات الإختبار التحصيلي والخطأ المعياري في نظريتي الكلاسيكية وإستجابة الفقرة وفق عدد الفقرات (24) باختلاف عدد البدائل

عدد البدائل	إحصائيات	النظرية الكلاسيكية		نظرية إستجابة الفقرة							
		معلمة التمييز	معلمة الصعوبة	a		b		c		d	
				a	S.E	b	S.E	c	S.E	d	S.E
بديلين	minimum	.253	.016	.029	.040	-8.213	.416	.052	.076	.694	.040
	maximum	.911	.553	1.153	.465	.921	12.168	.375	.175	1.000	.108
	mean	.717	.327	.635	.270	-1.158	1.439	.266	.133	.908	.068
	std	.177	.125	.266	.096	2.055	2.974	.088	.022	.079	.021
ثلاث بدائل	minimum	.513	.225	.436	.162	-1.915	.345	.152	.100	.813	.030
	maximum	.938	.615	1.410	.580	.022	.957	.488	.174	1.000	.088
	mean	.748	.466	.951	.369	-.805	.525	.272	.131	.932	.051
	std	.131	.113	.261	.105	.635	.170	.106	.022	.053	.016
أربع بدائل	minimum	.410	.124	.016	.069	-4.836	.402	.108	.075	.801	.011
	maximum	.949	.511	1.166	.496	1.082	23.483	.505	.194	1.000	.109
	mean	.697	.362	.717	.312	-.914	2.246	.325	.120	0.930	.069
	std	.146	.108	.346	.108	1.464	5.889	.130	.030	.060	.031

يتضح من الجدول أعلاه أن تقديرات معالم الفقرات في النظرية الكلاسيكية لفقرات البديلين كانت على الشكل التالي، إذ تراوحت معلمة التمييز بين (.253 – .911) بمتوسط (.717) وإنحراف (0.177)، وللثلاثة بدائل تراوحت بين (.513 – .938) بمتوسط (0.748) وانحراف (.131)، وللأربع بدائل تراوحت بين (.410 – .949) بمتوسط (.697) وانحراف (.146).

بينما تراوحت معلمة الصعوبة لفقرات للبديلين بين (.016 – .553) بمتوسط (0.327) وإنحراف (.125)، وللثلاثة بدائل تراوحت بين (.225 – .615) بمتوسط (0.466) وانحراف (0.113)، وللأربع بدائل تراوحت بين (.124 – .511) بمتوسط (.362) وانحراف (.108).

كما يتضح من الجدول أعلاه أن تقديرات معلم التمييز في نظرية إستجابة الفقرة لفقرات البديلين تراوحت بين (.029 – 1.153) بمتوسط (.635) وإنحراف (.266)، وللثلاث بدائل بين (.436 – 1.410) بمتوسط (.951) وانحراف (.261)، وللأربع بدائل بين (.016 – 1.166) بمتوسط (.717) وإنحراف (.346).

بينما تراوحت معلمة الصعوبة لفقرات البديلين بين (8.213– – 0.921) بمتوسط (–1.158) وانحراف (2.055)، وللثلاث بدائل بين (1.915–)– 0.022) بمتوسط (.805 –) وانحراف (.635)، وللأربع بدائل بين (4.836–)(–1.082) بمتوسط (.914 –) وانحراف (1.464).

بينما تراوحت معلمة التخمين لفقرات البديلين بين (.052 – .375) بمتوسط (.266) وانحراف (.088)، وللثلاث بدائل بين (.152 – .488) بمتوسط (.272) وانحراف (.106)، وللأربع بدائل بين (.108 – .505) بمتوسط (.325) وانحراف (.130).

بينما تراوحت معلمة اللامبالاه لفقرات البديلين بين (.694 – 1.000) بمتوسط (.908) وانحراف (.079)، وللثلاث بدائل بين (.813 – 1.000) بمتوسط (.932) وانحراف (.053)، وللأربع بدائل بين (.801 – 1.000) بمتوسط (.930) وانحراف (.060).

وتتفق نتيجة هذه الدراسة مع دراسة دبك (1998) بوجود فروق في معلم التمييز لصالح (3) بدائل، لكن تختلف معها بمعلم الصعوبة والتي تشير الدراسة لأفضلية (3) بدائل، كما تتفق مع دراسة فرانسيسكو وأوليا وبونسودا (Francisco, Olea & Ponsoda, 2001) في أن أسوا النتائج كانت للبديلين وأفضلها لـ(3) بدائل، لكن تختلف معها في معلم التمييز لعدد فقرات (9) والذي كان أسوا لفقرات (3) بدائل في النموذج الثلاثي، ولفقرات (4) بدائل في النظرية الكلاسيكية بإختلاف عدد الفقرات، كما تتفق مع دراسة التعمري (2003) بوجود فروق في معالم الصعوبة والتمييز والتخمين لصالح فقرات (3) بدائل، وتختلف مع دراسة الشريفين وطعامنة (2009) بوجود فروق بين متوسط الخطأ المعياري لمعلم الصعوبة ولصالح فقرات (3) بدائل.

كما تتفق مع دراسة الزواهره (2014) بوجود فروق في التخمين لعدد البدائل (3، 4) ولصالح (3) بدائل عند عدد الفقرات (15، 24) لكن تختلف عند عدد الفقرات (9) ولصالح (4) بدائل، كما تتفق معها بوجود فرق في معلمة التمييز لعدد البدائل (3، 4) مقارنة بعدد البدائل (2) عند عدد الفقرات (15، 24) لكن تختلف عند عدد الفقرات (9) فكانت أفضل ببدائل (4، 2) من (3) بدائل، كما تختلف مع دراسة أبو مسلم (2016) في أن الإختبار يكون

أكثر صعوبة وتمييزاً كلما زاد عدد البدائل، فكان أفضل عند (3، 2، 4) بدائل لمعلم التمييز و(3، 4، 2) لمعلم الصعوبة لعدد فقرات (15، 24)، كذلك في معلم التمييز لعدد فقرات (9) جاءت النتيجة معاكسة فكلما زاد عدد البدائل قل التمييز، كما اتفقت معها في معلم الصعوبة لعدد الفقرات (9).

وتختلف مع دراسة علاونة (2016) بوجود فروق في الصعوبة والتمييز وفق نمط الفقرة أو عدد البدائل (2، 3، 4)، كما تختلف مع دراسة العنزي (2017) بوجود فروق في الصعوبة بين شكلي الاختبار (اختيار من متعدد، الصواب والخطأ)، كذلك بوجود فروق في التمييز لصالح إختبار الصواب والخطأ.

ويعزو الباحث هذه النتيجة لمدى دقة وحرص الطلبة في الإجابة على فقرات الاختبار ضمن ما توفرت لديهم من معلومات دون اللجوء إلى التخمين، لكون الإختبار المصمم اختبار تحصيلي يحقق أهداف المادة التي تم دراستها، كذلك إلى مناسبة فقرات الإختبار بقدرتها على قياس التحصيل الدراسي، ومدى شموليتها ومراعاة الأسس السليمة في كتابتها وتحقيقها لأهداف المادة الدراسية.

◄ **ثالثاً: للإجابة عن سؤال الدراسة الثالث والذي ينص على:» ما مدى تحقق إفتراض أحادية البعد للاختبار التحصيلي وفق عدد البدائل والفقرات ؟ «.**

جرى التحقق من افتراض أحادية البعد للاختبار التحصيلي وفق عدد البدائل والفقرات باستخدام مؤشرات جودة المطابقة، والتي يعرضها الجدول (6).

الجدول (6)

مؤشرات التحقق من إفتراض أحادية البعد للإختبار التحصيلي وفق عدد البدائل والفقرات

مؤشرات مطابقة الإختبار	9 فقرات			15 فقرة			24 فقرة		
	بديلين	ثلاث بدائل	أربع بدائل	بديلين	ثلاث بدائل	أربع بدائل	بديلين	ثلاث بدائل	أربع بدائل
كا2 المحسوبة	5.902	.787	5.804	38.149	22.691	8.790	70.146	45.890	36.513
CFI	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
RMSEA	.000	.000	.000	.000	.000	.000	.000	.000	.000

يلاحظ من الجدول أعلاه وفق مؤشر (χ2) المحسوبة أن الإختبار التحصيلي وفق عدد البدائل والفقرات غير مطابقة للنموذج رباعي المعلم، والتي تشير إلى أن قيمة مؤشر (χ2) المحسوبة أعلى من (3) تؤشر على عدم مطابقة الفقرات للنموذج، بينما طابق الإختبار التحصيلي من (9 فقرات بثلاثة بدائل) النموذج رباعي المعلم، والتي تشير إلى أن قيمة مؤشر (χ2) المحسوبة أقل من (2) تؤشر على مطابقة الفقرات للنموذج بدرجة ممتازة.

كما يلاحظ من الجدول أعلاه وفق مؤشري (CFI, RMSEA) أن جميع الاختبارات التحصيلية وفق عدد البدائل والفقرات كانت مطابقة بنسبة (100 %) بدرجة ممتازة، والتي تشير إلى أن قيمة مؤشر (CFI) بين (97. إلى 1)، ومؤشر (RMSEA) بين (0 إلى05.) تؤشر على مطابقة ممتازة.

كما يلاحظ من الجدول أعلاه أن الاختبار التحصيلي وفق عدد البدائل والفقرات، مطابقة للنموذج رباعي المعلم، فلم يتم استبعاد أي إختبار غير مطابق لجميع مؤشرات المطابقة، وهذا يشير إلى تحقق افتراضي مطابقة فقرات الاختبار التحصيلي ضمن مؤشر أو أكثر من مؤشرات المطابقة، وإفتراض أحادية البعد، الذي يقود تتحقيق افتراض الاستقلال الموضعي، وفق ما أكد (Hamble-،1991 ton , Swaminathan & Rogers) أنه يعد مكافئاً لإفتراض أحادية البعد.

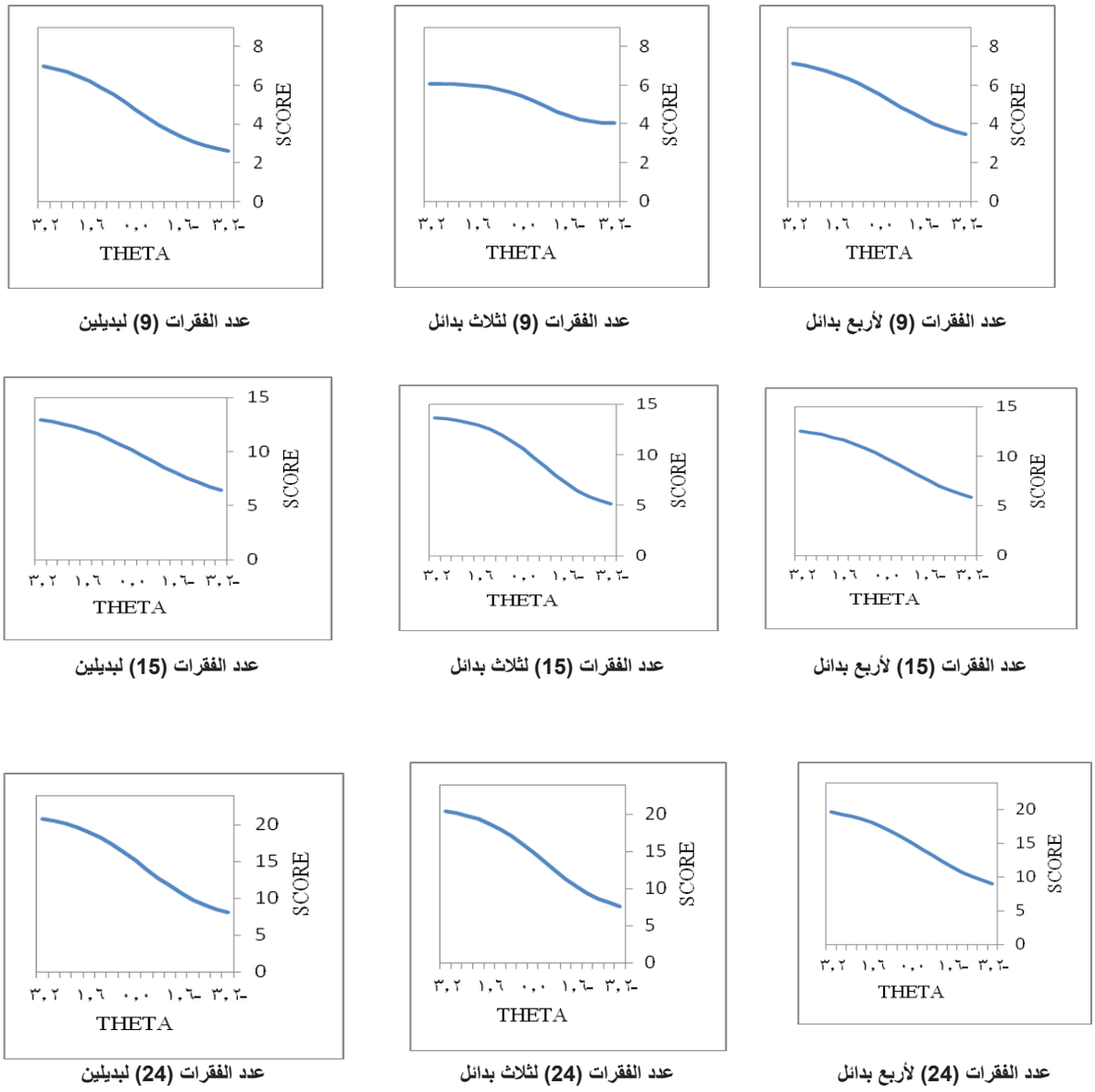
وتتفق هذه النتيجة مع نتيجة دراسة فرانسيسكو وأوليا وبونسودا (Francisco, Olea & Ponsoda, 2001) والتعمري (2003) والشريفين وطعامنة (2009) والزواهره (2014) والعنزي (2017) بتحقق افتراضات نظرية الاستجابة للفقرة.

ويعزو الباحث هذه النتيجة لاستقلالية الفقرات المشكلة للاختبار وعدم ارتباط إجابة فقرة بالفقرات الأخرى، مما يؤشر على أن كل فقرة تقيس هدف معين من المحتوى التدريسي، وهذا يدل على مدى تمثيل جدول مواصفات الإختبار لمحتوى المادة، والذي يؤشر على دقة ومتانة الاختبار وسلامة الإجراءات المتبعة في بنائه، مما يؤشر على تمتع النماذج الاختبارية بصدق بنائي مرتفع، وهذا بدوره يؤشر على أن هناك قدرة واحدة فقط تفسر أداء الفرد في الاختبار وهي تحصيله الأكاديمي في مادة علم النفس والحياة.

◄ **رابعاً: للإجابة عن سؤال الدراسة الرابع والذي ينص على:» ما مدى تحقق إفتراض إطرادية السمة للإختبار التحصيلي وفق عدد البدائل والفقرات؟ «.**

تمّ التحقق من إفتراض اطرادية السمة لفقرات الاختبار التحصيلي وفق عدد البدائل والفقرات، وذلك برسم منحنيات الخصائص لكل فقرة من فقرات الإختبار التحصيلي وفق عدد البدائل والفقرات، والتي تشير إلى أن جميع الفقرات تحقق إفتراض إطرادية السمة؛ حيث أن إحتمال الإجابة عن الفقرة الواحدة إجابة صحيحة يتزايد وتيرياً تبعاً لزيادة مستوى القدرة.

أما للتحقق من افتراض اطرادية السمة للاختبار التحصيلي ككل وفق عدد البدائل والفقرات؛ فقد تم رسم منحنى خصائص الاختبار التحصيلي وفق عدد البدائل والفقرات، كما هو مبين في الشكل (1).



عدد الفقرات (9) لأربع بدائل | عدد الفقرات (9) لثلاث بدائل | عدد الفقرات (9) لبديلين

عدد الفقرات (15) لأربع بدائل | عدد الفقرات (15) لثلاث بدائل | عدد الفقرات (15) لبديلين

عدد الفقرات (24) لأربع بدائل | عدد الفقرات (24) لثلاث بدائل | عدد الفقرات (24) لبديلين

الشكل رقم (1):

منحنى خصائص الإختبار التحصيلي وفق عدد البدائل والفقرات

يتضح من الشكل (1) أن افتراض إطرادية السمة متحقق للاختبار التحصيلي ككل وفق عدد البدائل والفقرات للنموذج رباعي المعلم؛ إذ أن احتمال الإجابة عن الفقرات في الإختبار إجابة صحيحة يتزايد وتيرياً تبعاً لزيادة مستوى القدرة.

وتتفق هذه النتيجة مع نتيجة دراسة الزواهره (2014) بتحقق افتراض الاطرادية أحد افتراضات النموذج رباعي المعلم.

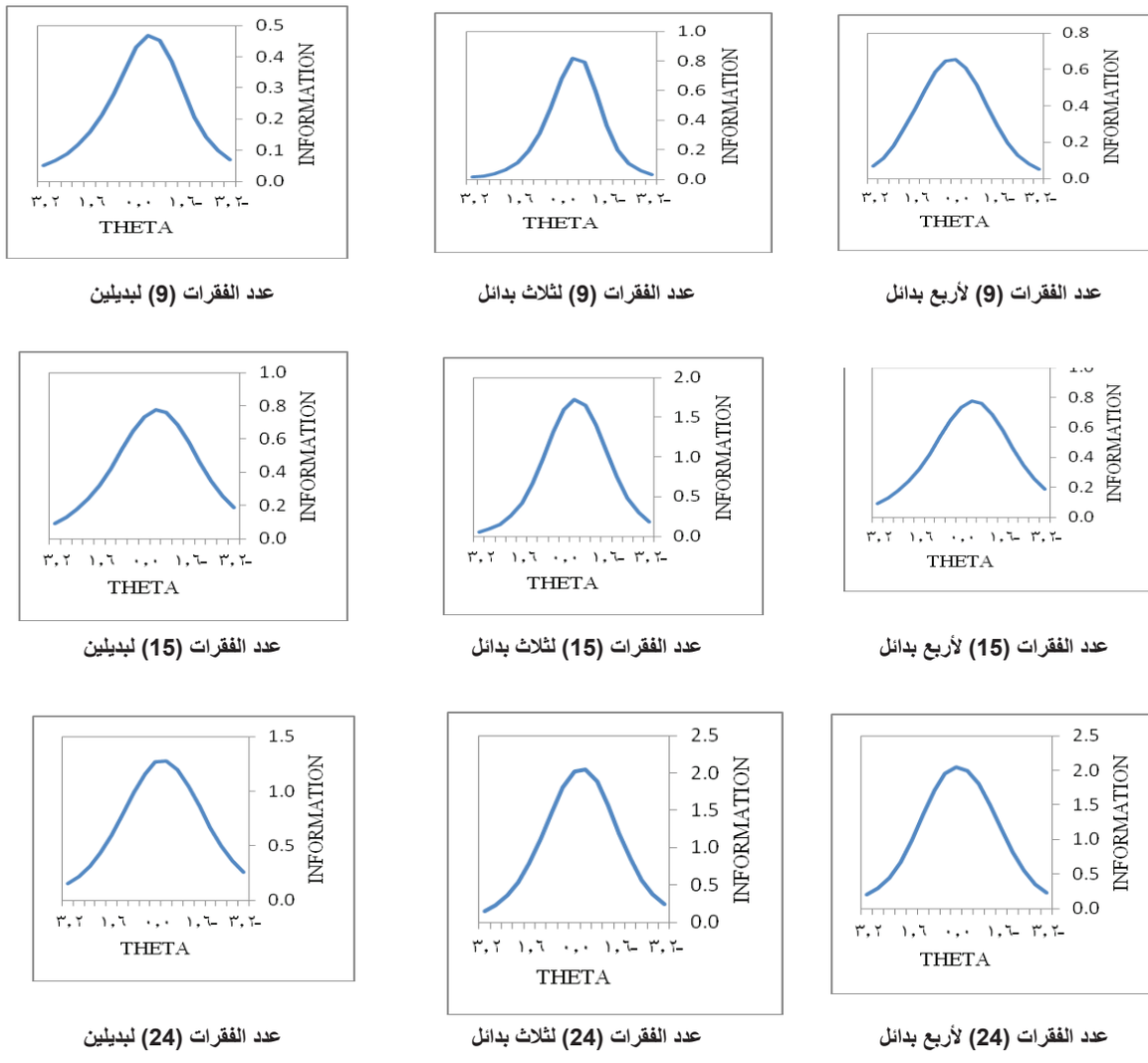
ويعزو الباحث هذه النتيجة إلى مدى وجود فروق فردية ما بين الطلبة والتي تمثلت بتوزعهم على متصل السمة، من القدرات المنخفضة إلى المرتفعة، والتي توضح مدى إرتفاع فرصة الأفراد ذوي القدرات المرتفعة في الإجابة على فقرات الاختبار بشكل صحيح دون تخمين كلما ارتفعت قدرتهم، ومن ثمّ، تؤشر على مدى امتلاكهم للسمة المدروسة بشكل أفضل وهي التحصيل الدراسي.

◄ **خامساً: للإجابة عن سؤال الدراسة الخامس والذي ينص على: «ما مدى تحقق دالة المعلومات للفقرات والإختبار التحصيلي وفق عدد البدائل والفقرات؟»**.

تمّ التحقق من دالة المعلومات لفقرات الاختبار التحصيلي وفق عدد البدائل والفقرات والتي تؤشر على ثبات فقرات الإختبار، وذلك برسم منحنيات دالة معلومات كل فقرة من فقرات الاختبار التحصيلي وفق عدد البدائل والفقرات والتي تتناسب عكسيا مع الخطأ المعياري للتقدير، إذ يتضح من منحنيات دوال المعلومات أن قيم دالة المعلومات المتعلقة بالفقرات تكون أكبر ما يمكن عندما يقترب مستوى القدرة من الصفر؛ أي أن هذه الفقرات تناسب الأفراد ذوي القدرة المتوسط أو دونها بقليل أو أعلى منها بقليل، بينما تكون قيم دالة المعلومات التي تعطيها هذه الفقرات أقل ما يمكن عند مستويات القدرة العالية والمتدنية.

بينما للتحقق من افتراض دالة معلومات الإختبار التحصيلي ككل وفق عدد البدائل والفقرات؛ فقد تم رسم دالة معلومات الاختبار التحصيلي وفق عدد البدائل والفقرات، كما هو مبين في الشكل (2).

الشكل رقم (2):



عدد الفقرات (9) لأربع بدائل | عدد الفقرات (9) لثلاث بدائل | عدد الفقرات (9) لبديلين

عدد الفقرات (15) لأربع بدائل | عدد الفقرات (15) لثلاث بدائل | عدد الفقرات (15) لبديلين

عدد الفقرات (24) لأربع بدائل | عدد الفقرات (24) لثلاث بدائل | عدد الفقرات (24) لبديلين

دالة معلومات الإختبار التحصيلي وفق عدد البدائل والفقرات

يلاحظ من الشكل (2)، أن قيم دالة معلومات الاختبار التحصيلي وفق عدد البدائل والفقرات تكون أكبر ما يمكن عند مستوى القدرة (0.000)؛ بمعنى أن الاختبار يعطي معلومات أكثر فاعلية عن الأفراد ذوي القدرة المتوسط، بينما تكون قيم دالة المعلومات التي يقدمها الإختبار أقل ما يمكن عند مستويات القدرة المتدنية والمرتفعة، وهذا يعني أن الاختيار يعطي معلومات قليلة عن الأفراد ذوي القدرات المتدنية والمرتفعة.

وتختلف هذه النتيجة مع نتيجة دراسة الشريفين وطعامنة (2009) بإعطاء الإختبار من (3) بدائل كمية معلومات أكبر من إختبار (4) بدائل عند مستويات القدرة المنخفضة، كما تختلف عن دراسة الزواهره (2014) في أن الاختبارات من (2، 3، 4) بدائل تعطي معلومات أكثر عند الأفراد ذوي القدرات دون المتوسط، بينما تتفق مع دراسة العنزي (2017) بعدم وجود فروق بين شكلي الاختبار (اختيار من متعدد، الصواب والخطأ).

ويعزى الباحث هذه النتيجة لمدى تمثيل الاختبار لمستويات القدرة المختلفة، والتي تؤشر على توزعها طبيعياً وفقاً لمستوى تحصيل الطلبة، وذلك لكون المادة الدراسية المشمولة بالدراسة تعتبر مادة اختيارية للطلبة فبعض الطلبة يسعى لتحقيق علامة مرتفعة فيها سعياً منه لرفع معدله التراكمي، مما يسهم ببذله جهداً أعلى، ومن ثمّ ترتفع قدرته فيها، والبعض الأخر يسعى لأخذ المادة أعتقاداً منه بأنها سهلة وليست بحاجة إلى جهد كبير، أما غالبية الطلبة وهم فئة متوسطي القدرة فيعتبرونها مادة دراسية كالمواد الأخرى، ومن ثمّ يبذلون مجهوداً متوسطاً فيها، وهذا ما يفسر توزع تحصيلهم الدراسي (قدرتهم) توزعاً طبيعياً على متصل السمة، والتي تنعكس على مستوى ثبات الإختبار المتمثلة بدالة المعلومات ومدى كمية المعلومات التي يوفرها الاختبار عند مستويات القدرة المختلفة.

◄ **سادساً: للإجابة عن سؤال الدراسة السادس والذي ينص على:» ما تقديرات القدرة لدى أفراد الدراسة على الإختبار التحصيلي وفق عدد البدائل والفقرات؟ «.**

جرى حساب كافة الإحصاءات الوصفية لكل من تقديرات القدرة لأفراد الدراسة على الاختبار التحصيلي للنموذج رباعي المعلم وفق عدد البدائل والفقرات، والتي يعرضها الجدول (7).

الجدول (7)

الإحصاءات الوصفية لكل من تقديرات القدرة لأفراد الدراسة على الإختبار التحصيلي للنموذج رباعي المعلم وفق عدد البدائل والفقرات

إحصاءات وصفية للقدرة	9 فقرات			15 فقرة			24 فقرة		
	بديلين	ثلاث بدائل	أربع بدائل	بديلين	ثلاث بدائل	أربع بدائل	بديلين	ثلاث بدائل	أربع بدائل
N-person	80	78	79	79	80	78	78	79	80
N-item	9	9	9	15	15	15	24	24	24
minimum	-2.470	-2.355	-2.220	-2.636	-2.720	-2.630	-2.391	-2.927	-2.342
maximum	2.093	1.736	2.472	2.490	2.355	2.472	2.875	2.971	2.987
median	.408	.545	-0.019	.652	.761	.205	.474	.405	.309
mean	.153	.014	-0.007	.556	.544	.253	.585	.431	.393
variance	1.265	1.682	1.697	1.395	2.144	1.975	1.949	2.372	2.506
std	1.125	1.297	1.303	1.181	1.464	1.406	1.396	1.540	1.583

يتضح من الجدول أعلاه أن تقديرات القدرة للاختبار التحصيلي (9) فقرات لبديلين تراوحت بين ((–2.470) – 2.093) بمتوسط (153.) وإنحراف (1.125)، وللثلاثة بدائل تراوحت بين ((–2.355) – 1.736) بمتوسط (0.014) وانحراف (1.297)، وللأربع بدائل تراوحت بين ((–2.220) – 2.472) بمتوسط (–0.007) وانحراف (1.303).

كما يتضح من الجدول أعلاه أن تقديرات القدرة للاختبار التحصيلي (15) فقرة لبديلين تراوحت بين ((–2.636) – 2.490) بمتوسط (556.) وانحراف (1.181)، وللثلاثة بدائل تراوحت بين ((–2.720) – 2.355) بمتوسط (0.544) وانحراف (1.464)، وللأربع بدائل تراوحت بين ((–2.630) – 2.472) بمتوسط (253.) وانحراف (1.406).

يتضح من الجدول أعلاه أن تقديرات القدرة للإختبار التحصيلي (24) فقرة لبديلين تراوحت بين ((–2.391) – 2.875) بمتوسط (0.585) وانحراف (1.396)، وللثلاثة بدائل تراوحت بين ((–2.927) – 2.971) بمتوسط (405.) وانحراف (1.540)، وللأربع بدائل تراوحت بين ((–2.342) – 2.987) بمتوسط (393.) وإنحراف (1.583).

كما يلاحظ من الجدول أعلاه أن مدى تقديرات القدرة لدى أفراد الدراسة لـ(9) فقرات لبديلين بلغ (4.563)؛ ولثلاث بدائل (4.091)، وللأربع بدائل (4.692)، ولـ(15) فقرة بديلين بلغ (5.126)؛ ولثلاث بدائل (5.075)، وللأربع بدائل (5.102)، ولـ(24) فقرة بديلين بلغ (5.266)؛ ولثلاث بدائل (5.898)، وللأربع بدائل (5.329)، بما يفيد بوجود فروق فردية بين تقديرات القدرة لدى أفراد الدراسة، وهذا يعني أن قدرات الأفراد تتوزع توزيعاً ملتوياً نحو اليسار على متصل السمة بإستثناء الاختبار (9 فقرات لأربع بدائل) (ذلك بالنظر إلى قيمة المتوسط)؛ كما أن قدرات الأفراد غير متجانسة كون الانحراف المعياري لها مرتفع، وأن الأفراد الذين قدراتهم العالية أكثر عدداً من الأفراد الذين قدراتهم متدنية (ذلك بالنظر إلى قيمة الوسيط) باستثناء الإختبار (9 فقرات لأربع بدائل).

وتتفق نتيجة هذه الدراسة مع دراسة الزواهره (2014) والشريفين وطعامنة (2009)، لكن تختلف مع دراسة التعمري (2003) بوجود فروق في تقديرات القدرة.

ويعزى الباحث هذه النتيجة لوجود فروق فردية ما بين الطلبة في تحصيلهم ومدى تمثيل الاختبار لمستويات القدرة المختلفة، كذلك لتشابه محتوى المادة الدراسية والإختبار وظروف تطبيق الاختبار بين جميع الطلبة الذين خضعوا للاختبار باستثناء اختلافهم بمدرس المادة ونموذج الاختبار الذي خضعوا له.

◄ **سابعاً: للإجابة عن سؤال الدراسة السابع والذي ينص على: "ما أثر عدد البدائل والفقرات على التوافق مع معالم الفقرات في النظرية الكلاسيكية؟"**.

يلاحظ من الجداول (4، 5، 6) أنه توجد فروق ظاهرية بين المتوسطات الحسابية لمعالم فقرات الاختبار التحصيلي في النظرية الكلاسيكية وفقاً لمتغيري (عدد البدائل والفقرات والتفاعل بينهما)، ولمعرفة هل توجد فروق دالة إحصائيا أجري تحليل التباين الثنائي (Two Way ANOVA)، والتي يعرضها الجدول (8).

الجدول (8):

تحليل التباين الثنائي لمعالم الفقرات بإستخدام النظرية الكلاسيكية (التمييز، والصعوبة) وفقاً لمتغيري عدد البدائل والفقرات والتفاعل بينهما

معلم الفقرة	مصدر التباين	SS	DF	MSS	F	sig
تمييز	عدد فقرات	.003	2	.001	.095	.910
	عدد بدائل	.134	2	.067	4.590	.012*
	تفاعل	.099	4	.025	1.699	.154
	الخطأ	1.968	135	.015	5.450	

معلم الفقرة	مصدر التباين	SS	DF	MSS	F	sig
صعوبة	عدد فقرات	.322	2	.161	.708	.005*
	عدد بدائل	.042	2	.021	.371	.494
	تفاعل	.044	4	.011		.829
	الخطأ	3.988	135	.030		

يلاحظ من الجدول أعلاه أنه توجد فروق دالة إحصائياً عند مستوى الدلالة (α = .05) بين المتوسطات الحسابية لمعلم التمييز وفقاً لمتغير عدد البدائل، وبين المتوسطات الحسابية لمعلم الصعوبة وفقاً لمتغير عدد الفقرات، ولمعرفة لمن تؤول الفروق أُجريتْ المقارنات البعدية باستخدام اختبار (Scheffe)، والجدولين (9، 10) يبينان ذلك.

الجدول (9)

نتائج إختبار Scheffe للمقارنات البعدية لمعلم التمييز وفقاً لمتغير عدد البدائل

عدد البدائل	التمييز	بديلين	ثلاث بدائل	أربع بدائل
	متوسط	.328	.416	.393
بديلين	0.328		.088-*	.065-*
ثلاث بدائل	0.416	.088*		.023
أربع بدائل	0.393	.065*	.023	

يتضح من نتائج الجدول أعلاه أنه توجد فروق دالة إحصائيا عند مستوى الدلالة (α = .05) بين متوسطي معلم (التمييز) لفقرات البديلين وللثلاث بدائل، وبين البديلين والأربع بدائل، وذلك لصالح الثلاث والأربع بدائل كون المتوسط الحسابي لها أعلى من المتوسط الحسابي للبديلين.

الجدول (10)

نتائج إختبار Scheffe للمقارنات البعدية لمعلم الصعوبة وفقاً لمتغير عدد الفقرات

عدد الفقرات	الصعوبة	9 فقرات	15 فقرة	24 فقرة
	متوسط	.583	0.721	0.669
9 فقرات	0.583		.138-*	.087-
15 فقرة	0.721	.138*		.052
24 فقرة	0.669	.087	.052	

يتضح من نتائج الجدول أعلاه أنه توجد فروق دالة إحصائياً عند مستوى الدلالة (α = .05) بين متوسطي معلم (الصعوبة) لعدد فقرات (9) و(15 فقرة)، وذلك لصالح (15) فقرة كون المتوسط الحسابي لها أعلى من المتوسط الحسابي لـ(9 فقرات).

وتتفق هذه النتيجة مع دراسة فرانسيسكو وأوليا وبونسودا (Francisco, Olea & Ponsoda, 2001) في أن أسوأ النتائج كانت للفقرات ذات البديلين، كما تتفق مع دراسة أبو مسلم (2016) في أن الاختبار يكون أكثر تمييز كلما زاد عدد البدائل للبديل مقارنة بـ (3، 4) لكن تختلف معها عند مقارنة (3) و(4) بدائل، لكن تختلف مع دراسة دبك (1998) بوجود فروق في التمييز لعدد البدائل (3، 4) بدائل ولصالح (3) بدائل، وعلاونة (2016) بعدم وجود فروق في التمييز وفق نمط الفقرة (اختيار من متعدد، الصح والخطأ).

كما يعزو الباحث هذه النتيجة إلى أن عدد الفقرات المرتفع أو المنخفض يؤثر في مستوى صعوبة الفقرة بسبب أن العدد المنخفض من الفقرات يكون دقيق جداً، وبالتالي يحتوي على فقرات صعبة، فيما العدد المرتفع قد يحتوي أيضاً فقرات صعبة بسبب عدم كفاية الوقت للإجابة عليها أو عدم تمعن الطالب بفهم محتوى الفقرة بشكل دقيق، ومن ثمّ ترتفع فرصة التخمين، لذلك يفضل أن يكون الاختبار مشكل من عدد متوازن من الفقرات بحيث يراعي محتوى المادة ومستوى صعوبة الفقرات ومدى كفاية الوقت للإجابة عليها وتهيئة الظروف الفيزيقية لذلك.

◄ **ثامناً: للإجابة عن سؤال الدراسة الثامن والذي ينص على:"ما أثر عدد البدائل والفقرات على التوافق مع معالم الفقرات في نظرية الإستجابة للفقرة ؟".**

يلاحظ من الجدول (4، 5، 6) أنه توجد فروق ظاهرية بين المتوسطات الحسابية لمعالم فقرات الاختبار التحصيلي في نظرية الإستجابة للفقرة وفقاً لمتغيري (عدد البدائل والفقرات والتفاعل بينهما)، ولمعرفة هل توجد فروق ذات دلالة إحصائية، تم إجراء تحليل التباين الثنائي (Two Way ANOVA)، والتي يعرضها الجدول (11).

الجدول (11):

تحليل التباين الثنائي لمعالم الفقرات بإستخدام نظرية الإستجابة للفقرة وفقاً لمتغيري عدد البدائل والفقرات والتفاعل بينهما

معلمة الفقرة	مصدر التباين	SS	DF	MSS	F	sig
a	عدد فقرات	.452	2	.226	1.899	.154
	عدد بدائل	.579	2	.289	2.433	.092
	تفاعل	.974	4	.244	2.048	.091
	الخطأ	16.057	135	.119		
b	عدد فقرات	15.575	2	7.788	1.916	.151
	عدد بدائل	.324	2	.162	.040	.961
	تفاعل	7.562	4	1.891	.465	.761
	الخطأ	548.634	135	4.064	1.724	
c	عدد فقرات	.044	2	.022	.074	.182
	عدد بدائل	.002	2	.001	.906	.929
	تفاعل	.046	4	.012		.462
	الخطأ	1.714	135	.013		
d	عدد فقرات	.086	2	.043	6.508	.002*
	عدد بدائل	.003	2	.002	.239	.788
	تفاعل	.010	4	.002	.380	.823
	الخطأ	.889	135	.007		

يلاحظ من الجدول أعلاه أنه توجد فروق دالة إحصائياً عند مستوى الدلالة (α = .05) بين المتوسطات الحسابية لمعلم اللامبالاه وفقاً لمتغير عدد الفقرات، ولمعرفة لمن تؤول الفروق أجريتْ المقارنات البعدية بإستخدام إختبار (Scheffe)، التي يعرضها الجدول (12).

الجدول (12)

نتائج إختبار Scheffe للمقارنات البعدية لمعلم اللامبالاه وفقاً لمتغير عدد الفقرات

عدد الفقرات	d	9 فقرات	15 فقرة	24 فقرة
	متوسط	.583	.721	.669
9 فقرات	.583		.070-*	.052-*
15 فقرة	.721	.070*		.018
24 فقرة	.669	.052*	.018	

يتضح من نتائج الجدول أعلاه أنه توجد فروق دالة إحصائياً عند مستوى الدلالة (α = .05) بين متوسطي معلم (اللامبالاه) لعدد فقرات (9) و(15 فقرة)، وبين عدد فقرات (9) و(24 فقرة)، وذلك لصالح (15، 24) فقرة كون المتوسط الحسابي لها أعلى من المتوسط الحسابي لـ(9 فقرات).

وتختلف نتيجة هذه الدراسة مع دراسة التعمري (2003) بعدم وجود فروق في معالم الصعوبة والتمييز والتخمين وفق النظرية الحديثة، ودراسة الزواهره (2014) والعنزي (2017) بوجود فروق في التمييز والتخمين تعزى لمتغير عدد البدائل (3، 4)، وتتفق مع دراسة الشريفين وطعامنة (2009) والعنزي (2017) بعدم وجود فروق لمعلم الصعوبة تعزى لعدد البدائل.

ويعزو الباحث هذه النتيجة إلى أنه كلما قل أو زاد عدد الفقرات عن العدد المناسب يسهم بزيادة مستوى اللامبالاه لدى الطلبة من ذوي القدرات المرتفعة، بحيث يجيبون على فقرات سهلة إجابة خاطئة، وتعتبر الفقرة جيدة عندما تقترب قيمة (d) من (1) (Loken & Rulison, 2010)، وبالنظر إلى متوسط قيمة (d) لوحظ أنه كان أفضل لعدد فقرات (15) مقارنة بـ(9، 24)، ومن ثمّ يجب على المدرسين بناء اختبارات متوازنة من حيث عدد فقراتها ومستوى صعوبتها، بحيث تسهم بأن يجيب عليها الطلبة إجابة صحيحة بثقة ودقة مع مراعاة محتوى المادة ومدى تمثيله بالشكل الصحيح.

التوصيات:

وفي ضوء النتائج التي توصل لها الباحث من خلال إجراءه للدراسة، فإنّه يوصي بما يلي:

1. الكشف عن أثر عدد البدائل والفقرات في مواد دراسية أخرى وفق نظريات القياس المختلفة والمتعددة (نظرية القياس الكلاسيكية، ونظرية الاستجابة للفقرة بنماذجها البارامترية واللابارامترية).
2. العمل على استخدام برمجيات مختلفة للتحقق من الخصائص السيكومترية وتقدير معالم الأفراد.
3. تقدير معالم الأفراد والفقرات بإستخدام الأنواع المختلفة من الفقرات والتي لم تتناولها الدراسة من مثل الفقرات ذات الإجابة المنتقاة (التكميل، الإنشائية المحددة)، والفقرات ذات الإجابة المصوغة.
4. تقدير معالم الأفراد والفقرات في الإختبارات النفسية، مما يمكننا من تفسير مدى إمتلاك الفرد للسمة المراد قياسها من خلال العلامة الكلية التي يحصل عليها الفرد.

المصادر والمراجع العربية:

- أبو مسلم، مايسة. (2016). أثر كل من عدد البدائل وترتيبها لمفردات اختبار تحصيل دراسي من نوع الاختيار من متعدد على الخصائص السيكومترية للاختبار ومفرداته، مجلة إتحاد الجامعات العربية للتربية وعلم النفس، كلية التربية، جامعة دمشق، 14(2):154 – 187.
- التعمري، بسمة محمود. (2003). تطبيق النموذج اللوغاريتمي ذي الثلاث معالم في تقدير قدرة الفرد ومعالم الفقرات لاختبار الإختيار من متعدد، (رسالة ماجستير غير منشورة)، جامعة مؤتة، الكرك، الأردن.
- دبك، أحمد رشيد. (1998). أثر شكل الفقرات وعدد البدائل في الخصائص السيكومترية للإختبار وفقراته، (رسالة ماجستير غير منشورة)، جامعة اليرموك، إربد، عمان.
- الزواهره، ريما فايز. (2014). أثر عدد البدائل في تقدير معالم الأفراد والفقرات بإستخدام نموذج بارامتري ونموذج لابارامتري (أطروحة دكتوراه غير منشورة)، كلية التربية، جامعة اليرموك، إربد، الأردن.
- الشافعي، محمد منصور. (2008). تأثير انتهاك افتراض أحادية البعد واستقلالية المحل في تدريج بنك الأسئلة ودقة معادلة درجات الإختبارات البنكية المسحوبة، ورقة مقدمة إلى الندوة الإقليمية لعلم النفس بعنوان: علم النفس وقضايا التنمية الفردية والمجتمعية، جامعة الملك سعود، الرياض27 – 29 ت أول، 2008م.
- 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 الأردنية في العلوم التربوية، عمادة البحث العلمي، جامعة اليرموك، 5(4):309 – 335.
- علام، صلاح الدين محمود. (2005). نماذج الإستجابة للمفردة الإختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي. القاهرة: دار الفكر العربي.
- علام، صلاح الدين محمود. (2016). الإختبارات والمقاييس التربوية والنفسية، عمان: دار الفكر.
- علاونة، معزوز جابر. (2016). أثر حجم العينة ونمط الفقرة وجنس الطلبة على الخصائص السيكومترية لإختبار تحصيلي في الرياضيات لدى طلبة الصف الحادي عشر في المدارس الحكومية الفلسطينية، مجلة جامعة المدينة العالمية (مجمع)، جامعة المدينة العالمية، (17):426 – 473.
- العنزي، محمد عويض. (2017). أثر شكل فقرة الإختبار من متعدد والصواب والخطأ على الخصائص السيكومترية وفق نظرية استجابة الفقرة لإختبار مادة الحاسب الآلي للصف الأول ثانوي في مدينة تبوك، (رسالة ماجستير غير منشورة)، جامعة تبوك، تبوك، السعودية.

- عودة، أحمد سليمان. (2014). القياس والتقويم في العملية التدريسية. إربد: دار الأمل للنشر والتوزيع.
- العيسوي، عبد الرحمن محمد. (2003). القياس والمقاييس النفسية والعقلية. الإسكندرية: منشأة المعارف.
- المصري، أحمد. (2009). أثر شكل فقرات الإختبار في تقدير قدرات المفحوصين ومعالم الفقرات وفق نظرية استجابة الفقرة (أطروحة دكتوراه غير منشورة)، كلية التربية، جامعة اليرموك، إربد، الأردن.

المصادر والمراجع العربية مترجمة:

- *Abu Muslim, M. (2016). The Effect of Both the Number Alternatives and Arrangement of Vocabulary for A Multiple Choice Study Achievement Test on the Psychometric Properties of the Test and its Vocabulary. Association of Arab Universities Journal for Education and Psychology, Faculty of Education, Damascus University, 14 (2): 154-187.*
- *Al-Taamari, B. (2003). Application of the Three Dimensional Logarithmic Model in Estimating the Ability of the Individual and Item Parameters to Test Multiple Choice. Unpublished Master Thesis. Mutah University, Karak, Jordan.*
- *Dabak, A. (1998). The Effect of the Shape of the Items and the Number of Alternatives on the Psychometric Properties of the Test and its items. Unpublished Master Thesis, Yarmouk University, Irbid, Amman.*
- *Zawahreh, R.(2014). The Effect of Number Alternatives on the Estimation of Persons and Items Parameters Using Parametric Model and Non-parametric Model. Unpublished Doctoral Thesis, Faculty of Education, Yarmouk University, Irbid, Jordan.*
- *Shafei, M. (2008). the effects of violating some IRT assumption on the calibration of item bank and equating scores of drone tests. Paper presented to the regional symposium of psychology Entitled: Psychology and Individual Community Development Issues. King Saud University, Riyadh, 27-29 October, 2008.*
- *Al-Shraaifin, N. & Ta'mnah, E. (2009). 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, Scientific Research & Graduate Studies Deanship, Yarmouk University, 5(4): 309-335.*
- *Allam, S. (2005). Response Models for Undimensional and Multidimensional Experimental Items and their Applications in Psychological and Educational Measurement. Cairo: Dar Al-Fikr Al-Arabi.*
- *Allam, S. (2016). Educational and Psychological Tests and Measurement. Amman: Al-Fikr house.*
- *Alawne, M. (2016). The effect of Sample Size, Type of Test Items and Student's Gender on the Psychometric Characteristics of A mathematical Achievement Test of 11th Grade Students in the Palestinian Governmental Schools. Al-Madinah International University (MEDIU) journal, Al-Madinah International University, (17): 426-473.*
- *Alenezi, M. (2017). The Effect of Item Format the Multiple Choice and True/false on Psychometrics Properties According to Item Response Theory (IRT) of the Computer Test for the First Secondary Grade in The City of Tabuk. (Unpublished Master Thesis), University of Tabuk, Tabuk, Kingdom of Saudi Arabia.*
- *Odeh, A. (2014). Measurement and Evaluation in the Teaching Process. Irbid: Al-Amal for Publishing and Distribution house.*
- *Issawi, A. (2003). Measurement and Mental and Psychological Parameters. Alexandria: Knowledge Facility.*
- *Al-Masry, A. (2009). The Effect of the form of the Test Items on Estimating the Ability of the Subjects and the Parameters of the Items According to the Item response theory. Unpublished Doctoral Thesis, College of Education, Yarmouk University, Irbid, Jordan.*

المصادر والمراجع الأجنبية:

- *Baker, Frank B. (2001). The Basic Item Response Theory. ERIC Clearinghouse on Assessment and Evaluation .*
- *Baker, Frank B., & Kim, Seock-Ho. (2004). Item Response Theory : Parameter Estimation Techniques. (2 ed), New York: Revised and Expanded. Marcel Dekker, INC.*
- *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, 1993).*
- *Francisco, A., & Olea ., J & Ponsoda., V. (2001). Analysis of the Optimum Number Alternatives Form the Item Response Theory. Psicothema, 13(1): 152-158.*
- *Hambleton, Ronald K., & Swaminathan, Hariharan. (1985). Item Response Theory : Principles and Applications. Boston : Kluwer.*
- *Hambleton, Ronald K., Swaminathan, Hariharan., & Rogers, Jane H. (1991). Fundamentals of item response theory. Newbury perk California: stage publications.*
- *Karin, S., Helfried, M., & Hans, M. (2003). Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. Methods of Psychological Research, 8(2): 23-74.*
- *Liang, T. (2010). An Assessment of the Nonparametric Approach of Evaluating the Fit of Item Response Model. Unpublished Doctor Dissertation, The Graduate School, School of Education, University of Massachusetts Amherst, Massachusetts.*
- *Loken, Eric., & Rulison, Kelly L. (2010). Estimation of A Four-Parameter Item Response Theory Model. British Joumal of Mathematical and Statistical Psychology, 63(3): 509-525.*
- *Magis, D., & Raiche, G. (2012). Random Generation of Response Patterns Under Computerized A Daptive Testing With The R Package CatR. Journal of Statistical Software, 48(8): 1-31.*
- *Sijtsam, K. (1988). Contributions to Mokken's Nonparametric Item Response Theory. Amesterdan: Free University Press.*
- *Sijtsma, K., & Hemker, B. (2000). A Taxonomy of IRT Models For Ordering of Persons and Items Using Simple Sum Scores. Journal of Educational and Behavioral Statistics, 25(4): 391-415.*
- *Stochl, J. (2007). Nonparametric Extension Of Item Response Theory Models And Its Usefulness For Assessment Of Dimensionality Of Motor Tests. Acta Universitatis Carolinae, 42(1): 75-94.*