جامعة واسط
مجلة كلية التربية

العدد الرابع والأربعون
الجزء الأول / آب /٢٠٢١

أثر اختيار أنموذج الاستجابة للمفردة (1PL,2PL,3PL,4PL) وطرائق التقدير في معالم المفردة والأفراد ودقتها

م.د. هبة عبد اللطيف ضعضع
قسم تربية الطفل– كلية التربية – جامعة حلب– سوريا
Hdodouh1987@gmail.com

ملخص البحث

هدف البحث إلى دراسة أثر اختيار أنموذج الاستجابة للمفردة (أحادي المعلم، ثنائي المعلم، ثلاثي المعلم، رباعي المعلم) وطرائق التقدير (EAP وMAP وMLP) في تقدير معالم المفردة (صعوبة وتمييز وتخمين وعدم اهتمام) والأفراد (القدرة) ودقتها والكشف عن أثر التفاعل بين هذين المتغيرين. لتحقيق أهداف الدراسة اعتمدت الباحثة نتائج اختبار قيد الماجستير المقام في المعهد العالي للغات في جامعة حلب المكون من(٦٠) مفردة من النوع الموضوعي ثنائي الاستجابة (٠ أو١) واستجابتها على(١٠٠٠) متقدم ومتقدمة من الكليات النظرية والتطبيقية.

وبعد التأكد من تحقق افتراضات نظرية الاستجابة للمفردة أحادية البعد (باستخدام التحليل العاملي، تشبع المفردات، معامل الارتباط الثنائي المتسلسل) والاستقلال الموضعي (باستخدام اختبار كاي مربع) والمنحنى المميز للمفردة، قامت الباحثة بالتأكد من ملائمة المفردات للنماذج ثنائية الاستجابة باستخدام برنامج Jmetrik، أظهرت النتائج:

١. لا يوجد فروق ذات دلالة إحصائية في نسبة المفردات الملائمة بين النماذج ثنائي المعلم وثلاثي المعلم ورباعي المعلم، في حين أظهرت النتائج دلالة الفروق في نسبة المفردات الملائمة مع الأنموذج أحادي المعلم.

٢. لا يوجد فروق ذات دلالة إحصائية في معاملات المفردة (صعوبة وتمييز وتخمين) ودقتها تعزى لعامل طريقة التقدير، ووجود فروق ذات دلالة إحصائية في قيمة معامل صعوبة المفردة (لم تظهر فروق دالة بين النموذج الاحادي المعلم والثنائي المعلم) ودقتها والتمييز ودقته وقيمة معامل التخمين، ولم يظهر لتأثير التفاعل بين طريقة التقدير والأنموذج المستخدم دلالة في الفروق.

٣. لا يوجد فروق ذات دلالة إحصائية في قدرة الأفراد تبعاً لطريقة التقدير والأنموذج المستخدم والتفاعل بينهم في حين أظهرت النتائج وجود فروق ذات دلالة احصائية في دقة تقدير قدرة الأفراد تبعاً للأنموذج المستخدم ودقة التقدير والتفاعل بينهما إلا أن دقة تقدير القدرة لم تظهر فروق دالة بين

طريقتي (EAP وMAP)، كما لم يظهر والنموذج الثلاثي الرباعي لم يظهر فروق دالة في دقة تقدير القدرة.

الكلمات المفتاحية: نموذج الاستجابة للمفردة، طرائق التقدير، معالم المفردة والأفراد، دقة التقدير

Abstrct:

The aim of the research is to study the effect of selecting a response model for the item (one parameter, two parameters, three parameters, four parameters) and estimation methods (EAP, MAP and MLP) in estimating the item parameters (difficulty, Discrimination, guess, and lack of interest) and individuals (ability) and its accuracy and detection of the effect of the interaction between These two variables. To achieve the objectives of the study, the researcher adopted the results of the master's entry test held at the Higher Institute of Languages at the University of Aleppo, consisting of (60) subjects of a double-response objective type (0 or 1) and their response to (1000) advanced and advanced from the theoretical and applied colleges.

After making sure that the assumptions of the response theory of the one-dimensional singularity (using factor analysis, item saturation, double-correlation coefficient), local independence (using chi-square test) and characteristic curve of singularity, the researcher made sure of the suitability of the vocabulary for the two-dimensional models using Jmetrik, the results showed:

1. There are no statistically significant differences in the proportion of fitting item between the two- parameter, three parameters and four parameters models, while the results showed significant differences in the proportion of fitting item with the one- parameter model.
2. There are no statistically significant differences in the unitary parameters (difficulty, Discrimination and guss) and their accuracy due to the factor of the estimation method, and the existence of statistically significant differences in the value of the item difficulty factor (no significant differences between the one-parameter model and the two-parameter model appeared) and its accuracy, Discrimination, accuracy, and the value of the two-parameter coefficient. And the effect of the interaction between the estimation method and the model used did not show any indication of differences
3. There are no statistically significant differences in the ability of individuals according to the method of estimation and the model used and the interaction between them, while the results showed that there are statistically significant differences in the accuracy of estimating the ability of individuals

according to the model used and the accuracy of the estimate and the interaction between them. However, the accuracy of estimating the ability did not show significant differences between the two methods. (EAP and MAP) were not shown, and the three-quadruple model did not show significant differences in the accuracy of the power estimation.

key words: irt models, the estimation methods, the parameters of the item and individuals, the accuracy of the estimate.

- **مشكلة البحث:** يتماشى التطور التاريخي لنظرية القياس مع تطور علم النفس كفرع من فروع المعرفة العلمية. وظهرت الدراسات النظرية للقياس النفسي والتربوي من خلال جهود علماء النفس في أوروبا والولايات المتحدة إذ درسوا المشكلات التربوية والنفسية المتنوعة. والتي بدأت في المانيا عام ١٨٧٩م في مختبرات الإدراك في ليبزغ إذ كان ويليم فوندت وإرنست ويبر وغوستاف فخنر من أوائل الذين ثمنوا أهمية الحصول على قياس نفسي ضمن ظروف مضبوطة بدقة، وبنفس الفترة عمل البريطانيون في قياس السمات العقلية وبعكس الألمان حيث انصب اهتمامهم على دراسة الفروق الفردية، ومنها تلتها محاولات فرنسية فعلى الرغم من إسهامات الألمانيين والبريطانيين المهمة في المجال النفسي، إلا أن إنجاز الفرنسيين من خلال أعمال بينه وثيوفيل (١٩٠٥– ١٩٠٨) الذي نقلا دراسة الاختبار العقلي من تدريب أكاديمي إلى أرض الواقع. وأصبح لها تدريبات فورية في غرفة الصف وفي العيادات النفسية وفي أماكن العمل، في حين كانت الولايات المتحدة الأمريكية من أخر الدول تقدماً في مجال القياس النفسي والتربوي وتلتها باقي الدول لتواكب التقدم العلمي في مجال علم النفس والقياس النفسي والتربوي (كروكر والجينا، ٢٠٠٩، ص٢٤–٢٧).

ومع الحركة المتسارعة والتطور في مجال العلوم النفسية والتربوية عامة والقياس النفسي والتربوي الذي توج بداية بنظرية القياس الكلاسيكية التي استخدمت بشكل واسع في إعداد المقاييس النفسية والاختبارات التحصيلية. إلى أن هذه النظرية واجهت بعض نقاط الضعف تتعلق بتقدير معالم المفردة ودقتها وتقديراتها المتعلق بخواص العينة.

الأمر الذي تم عالجته بنظرية الاستجابة التي جاءت كمساهمة لتلافي بعض نقاط ضعف نظرية القياس الكلاسيكية، ومن كون تقدير معاملات المفردة ودقتها من العوامل المؤثرة والمهمة في عملية القياس النفسي والتربوي تتحدد مشكلة البحث في مقارنة تقديرات معاملات المفردة والأفراد بتغير الأنموذج المستخدم (أحادي المعلم، ثنائي المعلم، ثلاثي المعلم، رباعي المعلم) وطريقة التقدير (EAP وMAP وMLP) والكشف عن أثر التفاعل بين هذين المتغيرين، وبذلك تتحدد مشكلة الدراسة بالتساؤل الرئيسي التالي:

ما أثر اختيار نموذج الاستجابة للمفردة (1PL,2PL,3PL,4PL) وطرائق التقدير في تقدير معالم المفردة والأفراد ودقتها؟

يتفرع عنه الأسئلة التالية:

١. ما أثر اختيار نموذج الاستجابة للمفردة (1PL,2PL,3PL,4PL) في نسبة المفردات الملائمة للأنموذج المقترح؟

٢. ما أثر اختيار نموذج الاستجابة للمفردة (1PL,2PL,3PL,4PL) وطرائق التقدير والتفاعل بينهما في تقدير معالم المفردة (صعوبة، تمييز، تخمين) ودقتها ؟

٣. ما أثر اختيار نموذج الاستجابة للمفردة (1PL,2PL,3PL,4PL) وطرائق التقدير ودقتها والتفاعل بينهما في تقدير معالم الأفراد؟

٢- **أهمية البحث:** تبرز أهمية البحث في:

١. الإجابة عن أسئلة الدراسة والتوصل إلى نتائج المقارنة بين نماذج الاستجابة للمفردة الثنائية وطرائق التقدير في تقدير معالم المفردة والأفراد والتفاعل بينهما.

٢. مقارنة النماذج شائعة الاستخدام وهي نموذج أحادي المعلم ونموذج ثنائي المعلم ونموذج ثلاثي المعلم من جهة مع الأنموذج رباعي المعلم وتحديد فاعلية هذا الأنموذج والفروق بينه وبين باقي النماذج ثنائية الاستجابة.

٣. تأتي هذه الدراسة استمرارا لجهود العاملين في مجال القياس والتقويم لتطوير الاختبارات، وانتقاء أفضل الطرائق للكشف عن الأداء التفاضلي للمفردات.

٤. تأتي هذه الدراسة تلبية لتوصيات الدراسات والبحوث السابقة.

٥. إمكانية تعميم نتائج الدراسة واستخدامها مع المقاييس النفسية والاختبارات التحصيلية الأخرى.

٣- **أهداف البحث:** يهدف البحث إلى التعرف:

١. أثر اختيار نموذج الاستجابة للمفردة (1PL,2PL,3PL,4PL) وطرائق التقدير في نسبة المفردات الملائمة للأنموذج المقترح.

٢. أثر اختيار نموذج الاستجابة للمفردة (1PL,2PL,3PL,4PL) وطرائق التقدير والتفاعل بينهما في تقدير معالم المفردة (صعوبة، تمييز، تخمين) ودقتها.

٣. أثر اختيار نموذج الاستجابة للمفردة (1PL,2PL,3PL,4PL) وطرائق التقدير والتفاعل بينهما في تقدير معالم الأفراد ودقتها.

٤- مصطلحات الدراسة:

- **نموذج استجابة للمفردة Item Response Theory Models:** مجموعة من النماذج الرياضية الاحتمالية التي تصف العلاقة بين الاستجابات الملاحظة على مفردات الاختبار والسمات الكامنة غير المباشرة المسببة لها، من خلال نماذج تقوم بالربط بين تلك الدرجات والسمات، وذلك عن طريق تقدير معالم القدرة والمفردة (طومان، ٢٠٠٦، ص٢٨). حيث تناولت الدراسة النماذج ثنائية الاستجابة أحادي المعلم وثنائي المعلم وثلاثي المعلم ورباعي المعلم
- **طرائق التقدير Estimation Methods:** هي مجموعة من الأساليب الرياضية الاحتمالية بغية تقدير معالم المفردة والأفراد في ضوء افتراضات نظرية الاستجابة للمفردة (ضعضع وآخرون، ٢٠٢٠، ص١٣٧). حيث تناولت الدراسة طرائق التقدير (EAP وMAP وMLP)
- **الدقة:** تعبير يشير إلى جودة التقدير لمعالم المفردة والأفراد، والتي تتميز بالاحتمال الكبير في أن يكون التقدير قريب من القيمة الحقيقية للمعلم، وذلك باختيار التقدير غير المتحيز Unbiased Estimator، صاحب أقل تباين بين التقديرات الأخرى غير المتحيزة، وذلك باستخدام الخطأ المعياري في التقدير (الرحيل والدرابسة، ٢٠١٤، ص٢٨).

٥- إطار نظري ودراسات سابقة:

٥-١- إطار نظري:

٥-١-١- نظرية الاستجابة للمفردة:

يعتمد الاتجاه المعاصر في القياس النفسي والتربوي على نظرية الاستجابة للمفردة، التي تفترض أنه يمكن التنبؤ بأداء الأفراد، ويمكن تفسير أدائهم في اختبار نفسي أو تربوي معين، بناء على خاصية أو خصائص معينة مميزة لهذا الأداء تسمى السمات Traits. كما تضيف هذه النظرية مفهوم جديد يدعى (دالة المعلومات)، الذي يعني مقدار الثقة في حصولنا على معلومات تقودنا إلى تقدير قدرة فرد معين أو مجموعة من الأفراد في مستوى قدرة معين، أي أننا نحصل على قيم متعددة لكمية المعلومات التي حصلنا عليها من المفردة أو الاختبار، وهي بذلك المعنى تختلف عن الثبات الذي هو عبارة عن مؤشر واحد فقط للاختبار لكل الأفراد الذين طبق عليهم هذا الاختبار، وتمثل

دالة المعلومات عادة بمنحنى يبين أعلى وأقل كمية معلومات حصلنا عليها عند مستويات القدرة المختلفة.

كما وتلخص الفكرة الأساسية لنماذج الاستجابة للمفردة في محاولة اشتقاق قيم تقديرية للسمات التي تنطوي عليها مجموعة من الاستجابات لمجموعة من المفردات، عادة يفترض أن السمة المقاسة هي قدرة معينة أو خاصية من خصائص الفرد الذي يختبر بها، بحيث لا توجد علاقة منتظمة بين مستويات السمة المقاسة لدى أفراد مختلفين واحتمالات الاستجابة الصحيحة لمفردات مختلفة (علام، ٢٠٠٠، ص٦٨٦). أي تحديد العلاقة بين الاستجابات الملاحظة على الاختبار والسمات غير الملاحظة.

كل هذا دعا إلى استخدام نظرية السمات الكامنة Latent Traits Theory، وهي نظرية حديثة بالمقارنة مع النظرية الكلاسيكية، حيث تعتمد على مجموعة افتراضات قوية، مفادها أن هناك استقلالية للمفردات الاختبارية في قياس القدرات المختلفة، وأن كل مفردة اختبارية لها خصائصها السيكومترية الممثلة في منحنى خاص بها يصف احتمالية الحصول على الإجابة الصحيحة أو الخاطئة بمعرفة قدرة الفرد إضافة لمعالم المفردة (Magno, 2009, P2). وتفترض هذه النظرية وجود واحد أو أكثر من السمات الأساسية التي تحدد استجابات الفرد الملاحظة لبنود اختبار ما. وقد أصطلح على تسميتها بالسمات الكامنة نظراً لعدم أمكانية ملاحظتها أو قياسها بصورة مباشرة، كما تُعين نماذج هذه النظرية العلاقة المتوقعة بين الاستجابات الملاحظة على اختبار ما، والسمات غير الملاحظة التي تفترض أنها تحدد هذه الاستجابات، حيث توفر نماذج السمات الكامنة تقديراً للقدرة مستقلاً عن العينة، كما توفر أيضاً مميزات القياس ذي الفئات المتساوية (Elliott, 1983, P59-60).

٥-١-٢- نماذج الاستجابة للمفردة Item Response Theory model

تهدف هذه النماذج إلى تحديد علاقة بين أداء الفرد في اختبار معين وبين السمات أو القدرات التي تكمن وراء هذا الأداء، وتفسره بدوال رياضية احتمالية يحدد النموذج العلاقة المتوقعة بين الاستجابات الملاحظة على الاختبار والسمات أو القدرات غير الملاحظة التي يفترض أنها تحدد هذه الاستجابات. حيث تمثل السمة بعد كمي يمكن أن يحدد عليه موضع الفرد، أي توفر نماذج الاستجابة للمفردة تقديراً للقدرة مستقلاً عن العينة، ونعني أن يكون القياس متحرراً من العينة Sample – free، أن يعبر عن تقديرات القدرة بوحدات لا تتعلق بصفات أي عينة، أو مجموعة معينة من الأفراد.

ستتوسع الباحثة في النماذج ثنائية الاستجابة وذلك لكونها موضوع الدراسة، وهي ثلاث نماذج يكون منحنى خصائص المفردة حالات خاصة من منحنى خصائص المفردة التالي:

أولاً: نموذج اللوغاريتمي أحادي المعالم (نموذج راش Rasch Model): يعد هذا النموذج أبسط نماذج الاستجابة للمفردة أحادية البعد، ويعرف باسم (نموذج راش Rasch Model) نسبة إلى عالم الرياضيات الدانماركي جورج راش (١٩٦٠)، يكون فيه المنحنى المميز للمفردة ممثل بدالة اللوغاريتمي أحادية البارامترات (One Parameter Logistic Function)، ويعتبر من أبسط نماذج نظرية الاستجابة للمفردة، ويشترط أن تكون هذه المفردات ثنائية الاستجابة، كما يفترض النموذج تساوي معاملات التميّز، وانعدام معامل التخمين، أي يقوم بتقدير صعوبة المفردة فقط. ويعطى بالعلاقة:

$$P_i(\theta_j) = \frac{e^{(\theta_j - b_i)}}{1 + e^{(\theta_j - b_i)}}$$

$P_i(\theta_j)$ هو احتمال الإجابة الصحيحة للفرد الذي قدرته θ_j على المفردة التي صعوبتها b_i.

كما يمكن التعبير عن النموذج اللوغاريتمي أحادي المعلم بالعلاقة التالية:

$$P_i(\theta_j) = \frac{1}{1 + e^{-(\theta_j - b_i)}} = \frac{1}{1 + e^{b_i - \theta_j}}$$

(Hambleton et al,1991,P12)

ثانياً: النموذج اللوغاريتمي ثنائي المعلم (نموذج لوردLord): يتم إضافة معامل جديد إلى نموذج راش وهو معامل التمييز a_i لكل مفردة، الأمر الذي يؤدي إلى تقاطع المنحنيات المميزة لمفردات الاختبار، إي أن هذا النموذج يفترض انعدام معامل التخمين، ويقوم بتقدير معامل التمييز وصعوبة المفردة.

يمكن التعبير عن النموذج اللوغاريتمي ثنائي المعلم بالعلاقة التالية:

$$P_i(\theta_j) = \frac{e^{a_i(\theta_j - b_i)}}{1 + e^{a_i(\theta_j - b_i)}}$$

$P_i(\theta_j)$ هو احتمال الإجابة الصحيحة للفرد الذي قدرته θ_j على المفردة التي صعوبتها b_i وتميزيها a_i.

ثالثاً: النموذج اللوغاريتمي ثلاثي المعلم (نموذج بيرنيوم Birnbaum):

أضاف بيرنيوم معلماً ثالثاً أطلق عليه معامل التخمين ويرمز له بـ C_i. تعطى الدالة الاحتمالية بالعلاقة:

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

$P_I(\Theta_J)$ هو احتمال الإجابة الصحيحة للفرد رقم J الذي قدرته Θ_J على المفردة رقم I التي صعوبتها B_I وتميزيها A_I، ومعامل تخمينها C_I. (Hambleton et al,1991,15)

بالمقارنة بين النماذج الثلاثة نلاحظ أن النموذج ثلاثي المعلم (نموذج بيرنيوم) هو الحالة العامة للدالة احتمال الإجابة الصحيحة على المفردة i، والنموذج ثنائي المعلم هو حالة خاصة من النموذج ثلاثي المعلم وذلك بجعل معامل التخمين مساوياً للصفر، كما أن النموذج أحادي المعلم هو حالة خاصة من نموذج ثنائي المعلم وذلك بجعل معامل التميز ثابت. يشير هامبلتون وسومينان (Hambleton & Swaminathan, 1985) أن قيمة قدرة الفرد θ ومعامل الصعوبة b، تتراوح من ∞+ إلى ∞− نظرياً (عملياً تتراوح قيمتهما بين +٣ و−٣).

رابعاً: الأنموذج رباعي المعلم: وجد أن بعض المفحوصين من ذوي القدرة المرتفعة، لا يجيبون دائماً على المفردة الاختبارية وذلك بسبب قلة اهتمامهم أو بسبب لا مبالاتهم (Hambelton & Swaminathan, 1985, P48). ولتلافي ذلك الأمر تم اقتراح النموذج رباعي المعلم والذي يعبر عنه بالعلاقة الرياضية التالية:

$$P_i(\theta_j) = C_i + (\gamma_i - C_i)\frac{e^{a_i(\theta_j - b_i)}}{1 + e^{a_i(\theta_j - b_i)}}$$

حيث يمثل Γ_i معامل عدم الاهتمام والذي يمثل بالمقارب الأعلى للدالة المميزة للمفردة. (Emons Et Al, 2003, P468)

٥-١-٣- طرائق التقدير:

أ. التقدير باستخدام الأرجحية العظمى (Maximum Likelihood)

ينطلق هذا الاسلوب من إيجاد قيمة θ للقدرة التي تعطي ارجحية عظمى للحصول على نمط استجابات معطى عن الفقرات ذات المعالم المحددة. وهذه الطريقة تنقسم الى عدة طرق منها:

١- **طريقة الارجحية القصوى المشتركة Joint Maximum Likelihood Procedure (JML):** يمكن تطبيق هذه الطريقة في النماذج اللوغاريتمية الاحادية والثنائية والثلاثية ويتم وفقا لهذه

الطريقة تقدير معالم القدرة والفقرة معا(Hambleton & Swaminthan,1985, P36). تسمى هذه الطريقة بتقدير الارجحية القصوى المشتركة، لأنه يتم وفق هذه الطريقة تقدير معالم القدرة والفقرة في آن معاً (Holland, 1990, P593). ويمكن تطبيق هذه الطريقة في النماذج اللوجستية الأحادية والثنائية والثلاثية. ونظرا لان معالم الفقرات غير معلومة في معظم الأحيان، فان تقديرها يمثل احد المشكلات الرئيسة في نماذج الاستجابة للفقرة. واستخدام صيغة دالة الارجحية كما في المعادلة الآتية:

$$L = \prod_{p=1}^{N}\left(\prod_{i=1}^{n} p(X_{ip} = x_{ip}|\theta_p)\right) = \prod_{p=1}^{N}\prod_{i=1}^{n} p_i\,(\theta_p)^{x_{ip}}[1 - p_i(\theta_p)]^{1-x_{ip}} \ldots\ldots(12)$$

اذ ان: ($x_{ip} = 1$) للاستجابة الصحيحة، و ($x_{ip} = 0$) للاستجابة الخطأ (de Gruijter, & van der Kamp, 2005, P106).

وتتم عملية التقدير لمعالم الفقرات في هذه الطريقة في مرحلتين ، تتناول المرحلة الاولى افتراض قيم للقدرة (θ_0)، وتقدير معالم الفقرات وفقا لهذا الافتراض من خلال البدء بمعالم اولية للفقرات ، اما المرحلة الثانية فيتم فيها تقدير القدرة اعتمادا على معالم الفقرات (β_0)، وتتكرر العمليتان حتى يكون هنالك ثبات في عملية التقدير ، بمعنى ان يكون الفرق في القيم المقدرة بين أي دورتين اقل من عدد ثابت (ε_0) يتم اعتماده ، او ان يزيد عدد الدورات عن عدد محدد يتم تحديده للبرنامج (N_0) . ويمثل الشكل المخطط المبدئي في هذه العملية (التقي، ٢٠٠٩، ص١٣٦).

٢- طريقة الارجحية القصوى الهامشية (Marginal Maximum Likelihood (MML:

ويمكن تطبيق هذه الطريقة على النماذج اللوغاريتمية الاحادية والثنائية والثلاثية حيث يتم وفقا لهذه الطريقة ايجاد اقتران الاحتمالية الهامشي (Marginal likelihood function) معالم الفقرة من خلال تكامل اقتران الكثافة الاحتمالية على معلم القدرة ثم ايجاد تقديرات معالم الفقرة وعندما تكون تقديرات معالم الفقرة معروفة فانه يمكن عندئذ تقدير معلم القدرة(Hambleton &Swaminthan, 1985, P36) إنّ تقدير الارجحية القصوى الهامشية وضع من قبل بوك وزملاؤه Bock & Aitken, 1981، وليفين وزملائه Levine,1989 إذ يتعامل هذا الأسلوب مع مستويات القدرة غير المعلومة، وعلى وفق هذه الطريقة لا يحتمل إهمال البيانات الهامشية . (Holland,1990, P593).

وتنطلق طريقة الارجحية القصوى الهامشية من اعتماد توزيع المعلمة التي يراد تقديرها سواء ارتبطت المعلمة بالأفراد أو الفقرات (التقي، ٢٠٠٩، ص١٤٣). وذلك بالتعبير عن احتمالات أنماط الإجابات بتوقعات Expectations من توزيع مجتمع معين. فالبيانات الاختبارية ينظر إليها على أنها عينة عشوائية مستمدة من المجتمع، وتوصل كل من بوك وايتكن إلى برنامج حاسوب يجري تقدير معالم الأفراد والفقرات في مرحلتين: التوقع Expectation، والأخرى تسمى مرحلة التعظيم Maximization، ويرمز بالحرفين (EM). وهذا البرنامج يعتمد تكرار الخطوات Iteration، والذي يؤدي الى تحسين العدد المتوقع للإجابات الصحيحة، ومستويات القدرة (علام، ٢٠٠٥، ص٩٧).

٣- طريقة الارجحية القصوى المشروطة (CML) Conditional Maximum Likelihood:

تنطلق هذه الطريقة في ان هنالك شرطا كافيا لتقدير معالم القدرة عند الافراد هو مجموع الفقرات التي اجاب عنها الفرد ، كما ان هنالك شرطا كافيا لتقدير معالم الفقرات، وهو مجموع الافراد الذين اجابوا عن الفقرة المعنية اجابة صحيحة. ومن هنا فان هذه الطريقة تحذف الفقرات التي تكون اجابات الطلبة عنها اجابة كاملة ، او التي تنعدم فيها اجابة جميع الطلبة ، وكذلك حذف الافراد (الحالات) التي تكون فيها اجابات الطلبة اجابة كاملة او التي تنعدم فيها اجابات الطلبة عن أي من الفقرات (التقي ،٢٠٠٩، ص ١٤٨ – ١٤٩) .

يستخدم هذا الأسلوب مع النموذج الأحادي (راش) ، والنماذج المنبثقة عنه فهو لا يصلح في حالة الانموذج الثنائي المعلم 2PL، او الانموذج ثلاثي المعلم 3PL حيث تختلف معالم تمييز المفردات مما يجعل الدرجات الكلية ، والتي يعتمد عليها الاسلوب في تقدير المعالم احصائيا غير كافي لتقدير مستويات قدرة الافراد ؛ ذلك ان تقديرات المعالم سوف تعتمد على أي المفردات اجاب عليها الفرد اجابة صحيحة . (الزهراني ، ٢٠٠٨، ص ٥٢) .

ب. طرق تعتمد على نظرية بييز :

من الجدير بالذكر أنه توجد اساليب أخرى لتقدير بارامترات نماذج الاستجابة للمفردة تعتمد على نظرية بييز Bayes Theorem التي تربط بين الاحتمالات الشرطية Conditional ، والاحتمالات الهامشية Marginal، وتتطلب هذه الاساليب افتراض احتمالات قبلية Prior Probabilities للبارامترات في ضوء اعتبارات نظرية أو إمبيريقية. وهذا يعني أنه توافرت معلومات عن توزيع القدرة لدى مجموعة من الأفراد المختبرين، فإن الاساليب التي تعتمد على نظرية بييز يمكن أن تساعد في الحصول على تقديرات جيدة لهذه القدرة Ability. إذ يمكن قبل تقدير مستويات قدرة مجموعة معينة

إفتراض أن المعلومات المتعلقة بقدرة أي فرد في المجموعة لا تختلف عن قدرة الأفراد الآخرين. وهذا الافتراض يعني أن القدرات . حيث (j =1 ، 2 ،...، N) يمكن اعتبارها بمثابة عينة عشوائية من مجتمع معين ، وبذلك يؤدي هذا الافتراض إلى أن التوزيع القبلي للقدرات اعتدالي .

ويمكن أيضاً افتراض أشكال أخرى من التوزيعات ، فقد افترض بيرنبوم Birnbaum أن التوزيع القبلي للقدرات يتخذ شكل توزيع الترجيح اللوغاريتمي Prior Logistic Distribution . ودالة الكثافة للتوزيع البعدي Posterior Density يتناسب مع حاصل ضرب دالة الأرجحية Likelihood Function في دالة الكثافة للتوزيع القبلي Prior Density . وعندما يكون لدينا (N) من المختبرين تصبح دالتا الكثافة البعدية والقبلية دالة كثافة مشتركة Joint Densities لمستويات القدرة (Rose,2010, P4). فإذا ما تحدد شكل دالة الكثافة القبلية، واختير أحد نماذج الاستجابة للمفردة، فإن دالة الكثافة المشتركة البعدية للقدرات يمكن تحديدها.

٥-٢-١- دراسات سابقة:

دراسة ضعضع وآخرون (٢٠٢٠) هدفت الدراسة إلى الكشف عن أثر طريقتين لتقدير معالم المفردة (الارجحية العظمى، تقدير بييز، طريقة بروكس) عند أحجام عينة مختلفة (٢٠٠٠، ١٥٠٠، ١٠٠٠، ٥٠٠) على دقة تقدير معالم المفردة (الصعوبة) في ضوء نظرية الاستجابة للمفردة ووفق نموذج راش. لتحقيق لأهداف الدراسة تم توليد عينة عشوائية من (٢٠٠٠) استجابة على (٤٠) مفردة ثنائية استجابة (٠ أو ١)، تم التحقق من افتراضات نظرية الاستجابة للمفردة (أحادية البعد، الاستقلال المحلي، المنحنى المميز للمفردة) والتحقق من ملاءمة المفردات لنموذج راش، نتيجة لذلك حصلنا على (٣١) مفردة ملاءمة لنموذج راش.

أظهرت النتائج وجود فروق ذات دلالة إحصائية في دقة تقدير معالم المفردة وفق طريقة الارجحية العظمى تعزى لحجم العينة، توجد فروق ذات دلالة إحصائية في دقة تقدير معالم المفردة وفق طريقة تقدير بييز تعزى لحجم العينة.تزداد دقة التقدير أي ينخفض الخطأ المعياري كلما زاد حجم العينة في طريقتي الأرجحية العظمى وطريقة بييز، لا يوجد فروق ذات دلالة في دقة تقدير معلم الصعوبة بين طريقة الارجحية العظمى وطريقة تقدير بييز عند أحجام العينة (٢٠٠٠، ١٥٠٠، ١٠٠٠، ٥٠٠٠).

دراسة سرحان (٢٠١٨) تهدف هذه الدراسة إلى المقارنة بين طريقتي الارجحية العظمى والقيمة العظمى للتوزيع البعدي في دقة تقدير القدرة وفقا لنظرية الاستجابة للفقرة، ولتحقيق أهداف الدراسة هذا الهدف أتبع الباحث خطوات علمية في إجراءات تحليل لاختبار تفسير الرموز الصادر من (مركز

التعليم السريع Learning Express ,LLC,2001) على وفق نظرية الاستجابة للفقرة، الذي يضم (٣) اختبارات فرعية تقيس التشابهات الرمزية (٣٠) فقرة، فرز وتصنيف الاشكال (٣٠) فقرة، التفكير في سلسلة (٣٠) فقرة، طبق الاختبار على عينة عشوائية طبقية مكونة من (١٠٠٠) طالب وطالبة. تكون الاختبار في صورته النهائية من (٧٣) فقرة. أشارت النتائج وجود فرق في تقدير القدرة بين الطريقتين ولصالح طريقة القيمة العظمى للتوزيع البعدي من خلال انخفاض قيم الخطأ المعياري لها.

دراسة العبد الله (٢٠١٢) هدفت للكشف عن أثر طريقة التقدير (WML[١]، EAP[٢]، MLE[٣]) على دقة تقدير معلم القدرة لدى الأفراد باستخدام نموذج التقدير الجزئي المعمم GPCM مع مراعاة طول الاختبار وحجم العينة ونوع توزيع البيانات، وتم استخدام بيانات حقيقية ومولدة، حيث تم سحب عينة عنقودية بحجم(٩٠٣) طالب وطالبة من ست مدارس مثلت ثلاث إدارات تربوية (الهرم، الدقي، الجيزة)، كما تم تطبيق اختبار تحصيلي محكي المرجع لوحدتين دراسيتين في اختبار الجبر والإحصاء، حيث تكون من (١٩) مفردة متعددة التدريج ومختلفة في عدد فئات الاستجابة.

كما استخدم بيانات مولدة حاسوبياً باستخدام برنامج WINGEN والذي يمثل نموذج التقدير الجزئي المعمم (GPCM)، حيث تم توليد ثلاثة أحجام مختلفة من البيانات (١٠٠٠، ١٥٠٠، ٢٠٠٠)، واعتماد ثلاثة توزيعات مختلفة لحجم العينة وهي (الطبيعي، المنتظم، بيتا)، واستخدام ثلاثة أطوال مختلفة للاختبار وهي (١٠، ٢٠، ٣٠) مفردة، حيث تكون الاختبار من أربع فئات. أظهرت النتائج بأن ترتيب أفضلية طرق التقدير هو EAP وWML وMLE وذلك عند تغير حجم العينة مع تثبيت كل من طول الاختبار ونوع التوزيع وكذلك عند تغيير محك التقييم RMSE عنها في Bias. ترتيب أفضلية طرق التقدير هو EAP ثم WML ثم MLE وذلك بالنسبة لمحك التقييم RMSE عنها في Bias بغض النظر عن المتغيرات الأخرى المتمثلة بحجم العينة وطول الاختبار ونوع التوزيع. وازدادت دقة تقدير معلم القدرة عند الأفراد بازدياد حجم العينة وطول الاختبار ذلك بالنسبة لمحك التقييم RMSE عنها في Bias بغض النظر عن طريقة التقدير المستخدمة، كما أظهرت أن أفضل النتائج في تقدير معالم القدرة كانت عند استخدام التوزيع الطبيعي للبيانات ثم التوزيع المنتظم وأخيراً توزيع بيتا، وذلك بالنسبة لمحك التقييم RMSE عنها في Bias بغض النظر عن طريقة التقدير المستخدمة.

[١] WML طريقة وورم للأرجحية العظمى الموزونة.
[٢] EAP طريقة التقدير القائمة على نظرية بيز.
[٣] MLE طريقة التقدير القائمة على طريقة الأرجحية العظمى.

التعقيب على الدراسات السابقة:

تتفق وتختلف الدراسة الحالية مع الدراسات السابقة:

١-تتفق الدراسة مع الدراسات السابقة بتناولها لمتغير طرائق التقدير وأثرها في دقة تقدير معالم نموذج الاستجابة للمفردة.

٢-تختلف الدراسة الحالية عن الدراسات السابقة بتناولها أكثر من نموذج، حيث اقتصرت الدراسات السابقة على نموذج واحد في حين تناولت الدراسة الحالية عدة نماذج (أحادي المعلم وثنائي المعلم وثلاثي المعلم وثلاثي المعلم ورباعي المعلم).

٣-ندرة الدراسات التي تناولت النموذج رباعي المعلم على المستوى العربي.

٤-دراسة أثر طرائق التقدير عبر النماذج ومعرفة تأثير اختلاف الأنموذج على معالم المفردة والأفراد ودقتها في ضوء طرائق التقدير.

٦-الدراسة الميدانية:

٦-١- منهج الدراسة: المنهج المستخدم في هذه الدراسة هو المنهج الوصفي التحليلي، الذي يعتبر أسلوباً من أساليب الدراسة العلمية، حيث يعتمد على دراسة الواقع أو الظاهرة كما توجد في الواقع، من حيث وصف الظاهرة وتوضيح خصائصها، أو التعبير عنها تعبيراً كمياً من حيث إعطاء وصفاً رقمياً يوضح مقدار هذه الظاهرة أو حجمها ودرجة ارتباطها مع الظواهر المختلفة، وتحليل هذه النتائج والوصول إلى استنتاجات وتعميمات تساهم في تطوير وفهم الواقع (عبيدات وآخرون، ٢٠٠٠، ٢٤٧). ونظراً لأن الدراسة الحالية تهدف إلى مقارنة تقديرات معالم المفردة والأفراد تبعاً للتغير نموذج الاستجابة للمفردة المستخدم وطريقة التقدير فإنها اندرجت تحت هذا النوع من الدراسات.

٦-٢- مجتمع البحث: جميع الطلاب المتقدمين لامتحان قيد الماجستير المطبق في المعهد العالي للغات في جامعة حلب والمطبق بتاريخ (١٦-٨-٢٠١٩) وذلك لدراسة النماذج ثنائية الاستجابة والبالغ عددهم (١٠٠٠) طالب.

٦-٣- عينة النماذج ثنائية : شملت العينة جميع الطلاب المتقدمين لامتحان قيد الماجستير المطبق في المعهد العالي للغات في جامعة حلب بتاريخ (١٦-٨-٢٠١٩) والتي بلغت (١٠٠٠) متقدم ومتقدمة.

٦-٤- التحقق من افتراضات نظرية الاستجابة للمفردة

أولاً: التحقق من أحادية البعد Investigated of Unidimensionality: وتم التحقق من هذا الافتراض من خلال:

١) التحليل العاملي Factor Analysis: ومنها معامل ألفا كرونباخ، حيث بلغت قيمته (0.845) وهو مؤشر قوي على أحادية البعد. تم التأكد من ملاءمة وكفاية حجم العينة، وذلك من خلال قيمة اختبار Kaiser-Meyer-Olkin (KMO-Test) لكفاية العينة، والتي بلغت (0.888) وهي بحسب محك كايزر Kaiser Test تعد نسبة ممتازة. وفيما يتعلق بالشرط الثاني الذي يتطلب أن يكون اختبار بارتليت Bartlett's Test Of Sphericity دالاً إحصائياً، ودلالته تعني أن مصفوفة الارتباطات ليس مصفوفة الوحدة Identity Matrix أي (خالية من العلاقات بين المفردات)، فقد أظهرت النتائج أن قيمة اختبار بارتليت بلغت (6050.992) وهي دالة إحصائياً وفقاً لقيمة الدلالة .sig=0.000

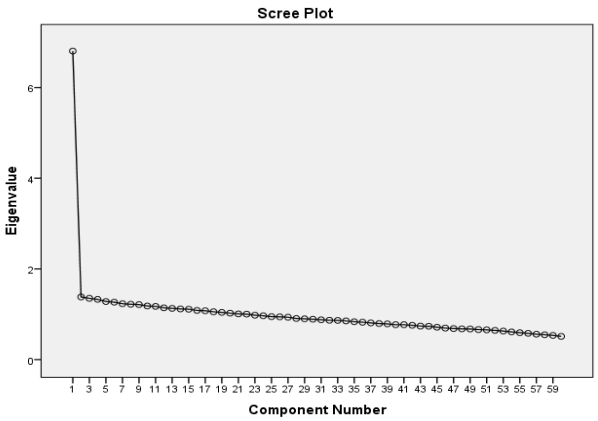
وبعد التأكد من شروط التحليل العاملي قامت الباحثة باستخدام التحليل العاملي الاستكشافي بطريقة المكونات الأساسية Principle Components لاستجابات الطلاب على مفردات الاختبار، وذلك للتحقق ما إذا كان الاختبار يقيس سمة واحدة كما هو مفترض. وقد تم حساب قيمة الجذر الكامن Eigenvalue، ونسبة التباين المفسر Explained Variance، وكذلك التباين المفسر التراكمي لكل عامل من العوامل، نلاحظ أن الجذر الكامن للعامل العام بلغ (٣٠.٠٥٨) وفسر (٤٧.٧١١%) من التباينات وهو أعلى نسبة تباين مفسر قياساً بالعوامل المقترحة الأخرى والذي يعتبر مؤشراً على أن الاختبار أحادي البعد.

٢) **مؤشر دلالة التشبع:** يتضح من جدول (١٥) مدى تشبع مفردات الاختبار بالعامل العام بالاعتماد على نسبة تشبع الاختبار (0.30) فما فوق وفقاً لمعيار جيلفورد Guilford ، كما هي موضحة في الجدول التالي:

الجدول (١). تشبع مفردات الاختبار على العامل العام

المفردة	التشبع	المفردة	التشبع	المفردة	التشبع	المفردة	التشبع
1	0.579	**16**	0.469	**31**	0.440	**46**	0.539
2	0.381	**17**	0.494	**32**	0.597	**47**	0.566
3	0.550	**18**	0.400	**33**	0.554	**48**	0.654
4	0.595	**19**	0.470	**34**	0.542	**49**	0.644
5	0.661	**20**	0.493	**35**	0.622	**50**	0.460
6	0.653	**21**	0.678	**36**	0.415	**51**	0.450
7	0.475	**22**	0.577	**37**	0.512	**52**	0.457
8	0.387	**23**	0.669	**38**	0.487	**53**	0.447
9	0.632	**24**	0.663	**39**	0.560	**54**	0.444
10	0.567	**25**	0.514	**40**	0.551	**55**	0.547
11	0.412	**26**	0.458	**41**	0.423	**56**	0.469
12	0.390	**27**	0.653	**42**	0.512	**57**	0.428
13	0.521	**28**	0.523	**43**	0.445	**58**	0.474
14	0.490	**29**	0.491	**44**	0.516	**59**	0.648
15	0.490	**30**	0.526	**45**	0.551	**60**	0.421

يتضح من الجدول السابق أن تشبع جميع مفردات الاختبار كان تشبعها بالعامل العام أعلى من (0.30)، وعليه لم تستبعد أي مفردة من مفردات المقياس، والتمثيل البياني لقيم الجذور الكامنة للعوامل المكونة للمقياس يوضح ذلك:



يلاحظ من الشكل (١) أن الخط المنحني بدأ يغير ميله بشكل مفاجئ بين النقطتين المناظرتين للعاملين الأول والثاني، كما يلاحظ أن قيم الجذور الكامنة للعوامل بدءاً من العامل الثاني أصبحت متقاربة، مما يُعد مؤشراً لاعتماد العامل الأول واستبعاد بقية العوامل.

٣) **ارتباط درجة المفردة بالدرجة الكلية Item Relation with test score:** هناك مؤشر آخر على أن المفردات تقيس سمة واحدة، حيث يستخدم معامل ارتباط التسلسل الثنائي Point – Biserial Correlations في كثير من الأحيان في تحليل مفردات المقياس لتحديد مدى اتساق درجات الأفراد في كل مفردة مع درجاتهم في الاختبار ككل (علام، ٢٠٠٢، ٣١٦)، والجدول التالي يوضح النتائج:

الجدول(٢). قيم معامل الارتباط المتسلسل الثنائي لمفردات الاختبار

المفردة	معامل الارتباط	المفردة	معامل الارتباط	المفردة	معامل الارتباط	المفردة	معامل الارتباط
1	0.315**	16	0.351**	31	0.424**	46	0.302**
2	0.420**	17	0.237**	32	0.100**	47	0.286**
3	0.380**	18	0.386**	33	0.346**	48	0.207**
4	0.205**	19	0.399**	34	0.277**	49	0.219**
5	0.199**	20	0.355**	35	0.200**	50	0.335**
6	0.211**	21	0.178*	36	0.401**	51	0.392**
7	0.391**	22	0.275**	37	0.291**	52	0.453**
8	0.398**	23	0.205**	38	0.361**	53	0.465**
9	0.167**	24	0.116**	39	0.185**	54	0.427**
10	0.294**	25	0.364**	40	0.187**	55	0.272**
11	0.424**	26	0.514**	41	0.359**	56	0.311**
12	0.442**	27	0.250**	42	0.309**	57	0.492**
13	0.310**	28	0.263**	43	0.432**	58	0.345**
14	0.344**	29	0.446**	44	0.325**	59	0.221**
15	0.383**	30	0.302**	45	0.244**	60	0.510**

يتضح من الجدول السابق أن قيم معامل الارتباط المتسلسل الثنائي إيجابي وذو دلالة إحصائية وهو مؤشر على أحادية البعد.

ثانياً: فرض الاستقلال المحلي Local Item Independence:

يعد هذا الافتراض امتداداً للافتراض الأعم وهو أحادية البعد، وهو يعني أنه إذا تم إزالة أثر العامل أو البعد أو السمة الكامنة خلف المقياس، فلن يوجد أي تغاير منتظم إضافي بين المفردات. وبذلك فإن الاستقلال المحلي يمكن أن يظهر بين المجموعات الجزئية من المفردات التي تسير في اتساق واحد. بحيث يكون بمقدور الباحث أن يختبر إحصائياً أحد الافتراضيين نيابة عن الآخر (زكري، ٢٠٠٩، ٥٤). وأن استقلالية الاستجابات عند الافتراض بوجود بعد واحد هو دليل على صحة الافتراض لكل من أحادية البعد واستقلالية الاستجابات (التقي، ٢٠١٣، ١٥٧). وقد تم التحقق من أحادية البعد وأن مفردات الاختبار تقيس سمة واحدة من خلال مخرجات التحليل العاملي من خلال الجذر الكامن والتباين المفسر، وتشبع المفردات بعامل واحد، فضلاً عن ذلك من خلال

معاملات ارتباط درجة المفردة بالدرجة الكلية، وبهذا يمكن القول بأنه تم التحقق من افتراض الاستقلال المحلي من خلال التحقق من افتراض أحادية البعد.

كما قامت الباحثة بالتحقق من هذا الافتراض من خلال مؤشر (G_2) للكشف عن الفروق بين ما هو متوقع وما هو ملاحظ لزوج من المفردات، وبذلك يمكن عمل جدول توافقي Contingency Table يتم من خلاله تبيان التكرارات الملاحظة والمتوقعة لأي زوج من المفردات، ويمكن تعريف الإحصائي G2 كما يلي:

$$G_2 = -2\sum_{i=1}^{2}\sum_{j=1}^{2}\left(O_{ij}\left(\frac{E_{ij}}{O_{ij}}\right)\right)$$

يشير تومبسون وبومرج (Thomopson & Pommerich, 1996, 3) إلى إمكان حساب المؤشر G_2 من خلال برنامج خاص مصمم للتأكد من هذا الافتراض كما يؤكد شن وأخرون (Chen & Thissen, 1997, 3) أن توزيع G_2 يتوزع توزيع كاي مربع عند درجة حرية مقدارها (١) وأنه إذا زادت نسبة الأزواج التي لا تحقق شرط الاستقلال الموضعي عن (٥%) دل ذلك على انتهاك هذا الافتراض. حيث قامت الباحثة بالتحقق من هذا الافتراض من خلال إضافة EIRT لبرنامج اكسل EXCEL والتي تقوم باختبار ملاءمة كل مفردة اختبارية مع جميع مفردات الاختبار من خلال اختبار كاي مربع عند مستوى دلالة (٠.٠٥)، أظهرت النتائج أن جميع الأزواج البالغ (١٧٧٠) مستقلة ماعدا ٣% من الأزواج، الأمر الذي يعد مؤشراً على الاستقلال الموضعي.

ثالثاً: المنحنى المميز للمفردة: يقترح هامبلتون وآخرون (Hambleton et al., 1991, 10) بأن فحص توزيع معاملات ارتباط مفردات الاختبار بالدرجة الكلية المكتسبة في الاختبار يعطي تصوراً عن مدى تجانس مؤشرات تمييز مفردات المقياس، حيث قامت الباحثة من التحقق من هذا الافتراض من خلال ملائمة المفردة للنماذج المقترحة.

٧-الإجابة عن تساؤلات الدراسة:

التساؤل الأول: ما أثر اختيار نموذج الاستجابة للمفردة (1PL,2PL,3PL,4PL) في نسبة المفردات الملائمة للأنموذج المقترح؟

للإجابة عن هذا التساؤل قامت الباحثة بدراسة ملائمة المفردات للنماذج ثنائية الاستجابة (أحادي المعلم، ثنائي المعلم، ثلاثي المعلم، رباعي المعلم) وذلك من خلال اختبار كاي مربع وباستخدام برنامج Jmtrik، لخصت الباحثة النتائج بالجدول التالي:

الجدول(٣). اختبار كاي مربع لملاءمة المفردات للنماذج ثنائية الاستجابة

	أنموذج أحادي المعلم			أنموذج ثنائي المعلم			أنموذج ثلاثي المعلم			أنموذج رباعي المعلم		
	كاي مربع	Df	SIG	كاي مربع	df	SIG	كاي مربع	df	SIG	كاي مربع	Df	SIG
1	39.437	45	0.706	39.671	44	0.658	38.948	43	0.648	39.539	42	0.580
2	49.376	44	0.267	43.194	44	0.506	42.959	42	0.430	42.154	42	0.464
3	35.808	45	0.835	34.040	43	0.834	33.948	42	0.807	33.671	42	0.817
4	64.426	45	0.030	39.133	44	0.680	39.125	42	0.598	40.377	42	0.542
5	69.139	45	0.012	35.991	43	0.767	36.378	42	0.715	42.330	42	0.457
6	69.150	45	0.012	45.854	44	0.395	44.014	42	0.386	46.582	42	0.289
7	43.047	45	0.555	40.643	43	0.574	40.590	42	0.533	40.205	42	0.550
8	33.710	44	0.869	33.104	43	0.862	38.599	42	0.621	35.270	42	0.759
9	100.058	44	0.000	50.325	43	0.206	53.754	42	0.106	56.705	42	0.064
10	59.772	45	0.069	57.351	44	0.085	57.810	42	0.053	58.079	42	0.050
11	35.122	44	0.828	24.964	42	0.983	24.263	42	0.987	24.235	42	0.987
12	51.421	45	0.237	38.623	43	0.661	38.348	42	0.632	38.508	42	0.625
13	52.459	44	0.179	55.965	44	0.107	55.186	42	0.084	53.714	42	0.106
14	28.841	45	0.971	30.591	44	0.938	30.517	42	0.906	31.377	42	0.885
15	41.582	44	0.576	48.078	44	0.311	46.600	42	0.289	44.612	42	0.363
16	34.506	45	0.872	37.355	44	0.750	26.270	42	0.973	29.261	42	0.932
17	62.801	44	0.033	47.947	43	0.279	48.005	42	0.242	49.535	42	0.198
18	40.281	43	0.590	38.982	43	0.646	37.593	42	0.665	37.347	42	0.675
19	52.046	45	0.219	43.143	44	0.508	42.054	42	0.469	42.994	42	0.428
20	55.052	44	0.123	62.395	43	0.028	57.215	43	0.072	58.470	42	0.047
21	163.246	45	0.000	68.763	44	0.010	66.027	42	0.010	67.734	42	0.007
22	62.771	44	0.033	56.190	44	0.103	51.951	42	0.140	57.653	42	0.054
23	59.753	44	0.057	36.205	44	0.792	35.991	42	0.731	40.669	42	0.529
24	101.176	45	0.000	34.003	43	0.835	36.274	43	0.756	35.762	42	0.740
25	36.880	44	0.768	38.969	44	0.687	37.590	43	0.704	37.114	42	0.685
26	71.261	44	0.006	29.156	43	0.947	29.749	42	0.922	30.033	42	0.916
27	47.261	44	0.341	48.262	44	0.305	45.994	42	0.310	47.168	42	0.269
28	53.465	45	0.181	47.343	43	0.300	44.017	42	0.386	47.542	42	0.257
29	60.425	45	0.062	45.041	43	0.387	43.478	42	0.408	43.929	42	0.390
30	61.113	45	0.055	60.159	43	0.043	49.110	42	0.210	56.511	42	0.067
31	32.279	45	0.922	21.838	44	0.998	21.795	42	0.996	22.164	42	0.995
32	134.927	44	0.000	56.617	44	0.096	53.182	42	0.116	56.249	42	0.070

0.830	42	33.266	0.819	42	33.610	0.816	43	34.605	0.819	44	35.417	33
0.077	42	55.636	0.144	42	51.731	0.052	44	60.251	0.035	44	62.425	34
0.012	42	65.251	0.063	42	56.870	0.231	44	50.550	0.000	44	197.185	35
0.090	42	54.727	0.240	42	48.093	0.054	44	60.019	0.180	44	52.422	36
0.315	42	45.870	0.382	43	45.143	0.408	44	45.538	0.383	45	47.203	37
0.265	42	47.309	0.282	42	46.792	0.826	43	34.292	0.557	45	42.994	38
0.709	42	36.531	0.805	43	34.921	0.842	43	33.792	0.032	44	62.964	39
0.482	42	41.738	0.546	42	40.305	0.589	43	40.296	0.015	45	67.842	40
0.414	42	43.351	0.420	42	43.191	0.443	44	44.692	0.429	44	45.024	41
0.229	42	48.454	0.263	43	48.438	0.264	44	49.467	0.195	43	50.739	42
0.986	42	24.402	0.991	43	24.056	0.972	44	27.926	0.628	44	40.370	43
0.002	42	72.397	0.003	42	72.182	0.004	44	72.468	0.013	44	67.501	44
0.007	42	67.700	0.010	42	66.292	0.015	44	66.836	0.001	45	79.547	45
0.284	42	46.744	0.317	42	45.808	0.355	43	45.854	0.353	44	46.944	46
0.134	42	52.225	0.157	43	52.270	0.168	44	52.913	0.104	43	54.961	47
0.753	42	35.430	0.759	42	35.281	0.814	43	34.662	0.116	45	56.560	48
0.852	42	32.546	0.854	42	32.495	0.863	43	33.056	0.274	45	50.231	49
0.200	42	49.440	0.159	42	51.089	0.173	44	52.695	0.287	44	48.766	50
0.021	42	62.554	0.007	42	67.686	0.003	42	71.359	0.044	45	62.363	51
0.955	42	27.767	0.968	43	27.525	0.971	43	27.213	0.174	44	52.661	52
0.885	42	31.379	0.801	42	34.121	0.786	44	36.365	0.235	45	51.468	53
0.961	42	27.369	0.958	42	27.542	0.969	44	28.218	0.669	45	40.361	54
0.332	42	45.414	0.371	43	45.433	0.431	43	43.949	0.261	45	50.639	55
0.532	42	40.619	0.528	43	41.691	0.753	43	36.358	0.580	45	42.462	56
0.286	42	46.674	0.322	43	46.715	0.355	44	46.880	0.001	44	78.823	57
0.118	42	53.060	0.181	42	50.168	0.173	42	50.509	0.192	44	51.944	58
0.633	42	38.323	0.671	42	37.441	0.707	43	37.534	0.164	45	54.168	59
0.808	42	33.926	0.775	43	35.781	0.757	43	36.253	0.001	45	79.458	60

من الجدول السابق نلاحظ أن:

١. قيمة الدلالة (sig>0.05) وذلك لجميع مفردات الاختبار فيما عدا (٢٠ مفردة) وبالتالي فإن جميع المفردات ملائمة ما عدا (٢٠ مفردة) للنموذج أحادي المعلم وبنسبة (٣٣.٣٣%).

٢. قيمة الدلالة (sig>0.05) وذلك لجميع مفردات الاختبار فيما عدا (٦ مفردات) وبالتالي فإن جميع المفردات ملائمة ما عدا (٦ مفردات) للنموذج ثنائي المعلم وبنسبة (١٠%).

٣. قيمة الدلالة (sig>0.05) وذلك لجميع مفردات الاختبار فيما عدا (٤ مفردة) وبالتالي فإن جميع المفردات ملائمة ما عدا (٤ مفردة) للنموذج ثلاثي المعلم وبنسبة (٦.٦٦%).

٤. قيمة الدلالة (sig>0.05) وذلك لجميع مفردات الاختبار فيما عدا (٥ مفردة) وبالتالي فإن جميع المفردات ملائمة ما عدا (٥ مفردة) للنموذج رباعي المعلم وبنسبة (٨.٣٣%).

٥. نلاحظ أن (٤ مفردات) غير ملائمة في النماذج الأربع (أحادي المعلم، ثنائي المعلم، ثلاثي المعلم، رباعي المعلم).

ولاختبار دلالة الفروق في أثر النموذج في نسبة الفقرات غير الملائمة قامت الباحثة باختبار z لاختبار دلالة نسبة هذه الفقرات، والجدول التالي يوضح النتائج:

الجدول(٤). دلالة الفروق في نسب المفردات الملائمة على وفق الأنموذج ثنائي الاستجابة

المقارنات		*P*	*Q=1-P*	Z المحسوبة	القرار
	ثنائي	٠.١٠	٠.٩٠	3.194	يوجد
	ثلاثي	٠.٠٧	٠.٩٣	3.765	يوجد
	رباعي	٠.٠٨	٠.٩٢	3.567	يوجد
	ثلاثي	٠.٠٧	٠.٩٣	0.590	لا يوجد
	رباعي	٠.٠٨	٠.٩٢	0.383	لا يوجد
	رباعي	٠.٠٨	٠.٩٢	0.199	لا يوجد

نلاحظ من الجدول السابق أن:

١. القيمة المحسوبة لـ z في مقارنة نسبة المفردات غير الملائمة لنموذج أحادي المعلم مع باقي النماذج أكبر من القيمة الجدولية المساوية (١.٩٦) عند مستوى دلالة (٠.٠٥) وبالتالي توجد فروق ذات دلالة إحصائية في نسبة الفقرات غير الملائمة لصالح النموذج أحادي المعلم.

٢. القيمة المحسوبة لـ z في مقارنة نسبة المفردات غير الملائمة لنموذج ثنائي المعلم من جهة والنموذج ثلاثي المعلم والنموذج رباعي المعلم أصغر من القيمة الجدولية المساوية (١.٩٦) عند مستوى دلالة (٠.٠٥) وبالتالي لا توجد فروق ذات دلالة إحصائية في نسبة الفقرات غير الملائمة.

٣. القيمة المحسوبة لـ z في مقارنة نسبة المفردات غير الملائمة لنموذج ثلاثي المعلم من جهة والنموذج رباعي المعلم أصغر من القيمة الجدولية المساوية (١.٩٦) عند مستوى دلالة (٠.٠٥) وبالتالي لا توجد فروق ذات دلالة إحصائية في نسبة الفقرات غير الملائمة.

التساؤل الثاني: ما أثر اختيار نموذج الاستجابة للمفردة (1PL,2PL,3PL,4PL) وطرائق التقدير (EAP وMAP وMLP) والتفاعل بينهما في تقدير معالم المفردة (صعوبة، تمييز، تخمين) ودقتها ؟

ويتفرع عن هذا التساؤل التساؤلات الفرعية التالية:

اولاً: ما أثر اختيار نموذج الاستجابة للمفردة (1PL,2PL,3PL,4PL) وطرائق التقدير (EAP وMAP وMLP) والتفاعل بينهما في تقدير معالم صعوبة المفردة ودقتها ؟

قامت الباحثة بدراسة دلالة الفروق في صعوبة المفردة ودقتها تبعاً للأنموذج المستخدم (أحادي المعلم ثنائي المعلم ثلاثي المعلم ثلاثي المعلم رباعي المعلم) وطريقة التقدير وهي طريقة (EAP وMAP وMLP) والتفاعل بينهما من خلال تحليل التباين المتعدد، وذلك باستخدام برنامج SPSS حيث تم

التأكد من تجانس تباينات المجموعات الفرعية من خلال اختبار ليفين levene's test فبلغت قيمة الدلالة sig=0.000 لكل من صعوبة المفردة ودقتها، وبالتالي فإن المجموعات الفرعية غير متجانسة لصعوبة المفردة ودقتها، وقد أظهرت نتائج تحليل التباين المتعدد ما يلي:

الجدول(٥). نتائج اختبار تحليل التباين المتعدد لدلالة الفروق في صعوبة المفردة ودقتها تبعاً للأنموذج المستخدم وطريقة التقدير

مصدر التباين	معامل/ دقة	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة F	قيمة الدلالة	القرار
طريقة التقدير	صعوبة المفردات	.000	2	.000	.000	1.000	غير دالة
	دقة التقدير	.000	2	.000	.000	1.000	غير دالة
الأنموذج	صعوبة المفردات	93.207	3	31.069	48.104	.000	دالة
	دقة التقدير	13.526	3	4.509	42.365	.000	دالة
طريقة التقدير* الأنموذج	صعوبة المفردات	.000	6	.000	.000	1.000	غير دالة
	دقة التقدير	.000	6	.000	.000	1.000	غير دالة
الخطأ	صعوبة المفردات	457.277	708	.646			
	دقة التقدير	75.350	708	.106			
الكلي	صعوبة المفردات	595.673	720				
	دقة التقدير	133.802	720				

من الجدول السابق نلاحظ أن:

١. قيمة الدلالة (sig>0.05) في دراسة أثر طريقة التقدير على معامل صعوبة الفقرة ودقتها وبالتالي لا يوجد فروق ذات دلالة إحصائية في قيمة معامل صعوبة الفقرة ودقتها تبعاً لطريقة التقدير.

٢. قيمة الدلالة (sig<0.05) في دراسة أثر النموذج المستخدم على معامل صعوبة الفقرة ودقتها وبالتالي يوجد فروق ذات دلالة إحصائية في قيمة معامل صعوبة الفقرة ودقتها تبعاً للنموذج المستخدم.

٣. قيمة الدلالة (sig>0.05) في دراسة أثر التفاعل بين الأنموذج المستخدم وطريقة التقدير على معامل صعوبة الفقرة ودقتها وبالتالي لا يوجد فروق ذات دلالة إحصائية في قيمة معامل صعوبة الفقرة ودقتها تبعاً للتفاعل بين الأنموذج المستخدم وطريقة التقدير.

للكشف عن مصادر الفروق في صعوبة الفقرة ودقتها تبعاً للأنموذج المستخدم قامت الباحثة بإجراء اختبار دانت التتبعي وذلك باستخدام برنامج SPSS، كما هو موضح في الجدول التالي:

الجدول(٦). نتائج اختبار دنت التتبعي لدلالة الفروق في صعوبة المفردة ودقتها تبعاً للأنموذج المستخدم

	المقارنات		متوسط الفروق	الخطأ المعياري	sig
صعوبة المفردة	أحادي المعلم	ثنائي المعلم	.027838	.0847133	.991
		ثلاثي المعلم	-.862244-*	.0847133	.000
		رباعي المعلم	-.364968-*	.0847133	.000
	ثنائي المعلم	ثلاثي المعلم	-.890081-*	.0847133	.000
		رباعي المعلم	-.392806-*	.0847133	.000
	ثلاثي المعلم	رباعي المعلم	.497275*	.0847133	.000
دقة تقدير الصعوبة	أحادي المعلم	ثنائي المعلم	-.097942-*	.0343878	.045
		ثلاثي المعلم	-.310204-*	.0343878	.000
		رباعي المعلم	-.318091-*	.0343878	.000
	ثنائي المعلم	ثلاثي المعلم	-.212262-*	.0343878	.000
		رباعي المعلم	-.220149-*	.0343878	.000
	ثلاثي المعلم	رباعي المعلم	-.007887-	.0343878	.997

نلاحظ من نتائج اختبار دانت في الجدول السابق أن:

أولاً: صعوبة المفردة: لا يوجد فروق ذات دلالة إحصائية بين النموذج أحادي المعلم وثنائي المعلم في معالم صعوبة الفقرة في حين كان هنالك فروق ذات دلالة إحصائية في صعوبة المفردة لصالح النموذج الأكثر معالم.

ثانياً: دقة تقدير صعوبة المفردة: هنالك فروق ذات دلالة إحصائية في دقة تقدير صعوبة المفردة لصالح النموذج الأكثر معالم.

ثانياً: ما أثر اختيار نموذج الاستجابة للمفردة (2PL,3PL,4PL) وطرائق التقدير (EAP وMAP وMLP) والتفاعل بينهما في تقدير معلم تمييز المفردة ودقتها ؟

قامت الباحثة بدراسة دلالة الفروق في تمييز المفردة ودقتها تبعاً للأنموذج المستخدم (ثنائي المعلم، ثلاثي المعلم، رباعي المعلم) وطريقة التقدير وهي طريقة (EAP وMAP وMLP) والتفاعل بينهما من خلال تحليل التباين المتعدد، وذلك باستخدام برنامج SPSS حيث تم التأكد من تجانس تباينات المجموعات الفرعية من خلال اختبار ليفين levene's test فبلغت قيمة الدلالة sig=0.000 لكل من تمييز المفردة ودقتها، وبالتالي فإن المجموعات الفرعية غير متجانسة لتمييز المفردة ودقتها، وقد أظهرت نتائج تحليل التباين المتعدد ما يلي:

الجدول(٧). نتائج اختبار تحليل التباين المتعدد لدلالة الفروق في تمييز المفردة ودقتها تبعاً للأنموذج المستخدم وطريقة التقدير

مصدر التباين	معامل/ دقة	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة F	قيمة الدلالة	القرار
طريقة التقدير	تمييز المفردات	.000	2	.000	.000	1.000	غير دالة
	دقة التقدير	.000	2	.000	.000	1.000	غير دالة
الأنموذج	تمييز المفردات	10.930	2	5.465	39.341	.000	دالة
	دقة التقدير	1.737	2	.869	688.241	.000	دالة
طريقة التقدير* الأنموذج	تمييز المفردات	.000	4	.000	.000	1.000	غير دالة
	دقة التقدير	.000	4	.000	.000	1.000	غير دالة
الخطأ	تمييز المفردات	73.766	531	.139			
	دقة التقدير	.670	531	.001			
الكلي	تمييز المفردات	499.696	540				
	دقة التقدير	14.280	540				

من الجدول السابق نلاحظ أن:

١. قيمة الدلالة (sig>0.05) في دراسة أثر طريقة التقدير على معامل تمييز الفقرة ودقتها وبالتالي لا يوجد فروق ذات دلالة إحصائية في قيمة معامل تمييز الفقرة ودقتها تبعاً لطريقة التقدير.

٢. قيمة الدلالة (sig<0.05) في دراسة أثر النموذج المستخدم على معامل تمييز الفقرة ودقتها وبالتالي يوجد فروق ذات دلالة إحصائية في قيمة معامل تمييز الفقرة ودقتها تبعاً للنموذج المستخدم.

٣. قيمة الدلالة (sig>0.05) في دراسة أثر التفاعل بين الأنموذج المستخدم وطريقة التقدير على معامل تمييز الفقرة ودقتها وبالتالي لا يوجد فروق ذات دلالة إحصائية في قيمة معامل تمييز الفقرة ودقتها تبعاً للتفاعل بين الأنموذج المستخدم وطريقة التقدير.

للكشف عن مصادر الفروق في تمييز الفقرة ودقتها تبعاً للأنموذج المستخدم قامت الباحثة بإجراء اختبار دانت التتبعي وذلك باستخدام برنامج SPSS، كما هو موضح في الجدول التالي:

الجدول(٨). نتائج اختبار دنت التتبعي لدلالة الفروق في تمييز المفردة ودقتها تبعاً للأنموذج المستخدم

	المقارنات		متوسط الفروق	الخطأ المعياري	sig
تمييز المفردة	ثنائي المعلم	ثلاثي المعلم	-.2138-*	.03929	.000
		رباعي المعلم	-.3452-*	.03929	.000
	ثلاثي المعلم	رباعي المعلم	-.1315-*	.03929	.004
دقة تقدير التمييز	ثنائي المعلم	ثلاثي المعلم	-.0792-*	.00374	.000
		رباعي المعلم	-.1385-*	.00374	.000
	ثلاثي المعلم	رباعي المعلم	-.0593-*	.00374	.000

نلاحظ من نتائج اختبار دانت في الجدول السابق أن الفروق في معامل تمييز لصالح النموذج الأكثر معالم.

ثالثاً: ما أثر اختيار نموذج الاستجابة للمفردة (2PL,3PL,4PL) وطرائق التقدير(EAP وMAP وMLP) والتفاعل بينهما في تقدير معلم تخمين المفردة ودقتها ؟

قامت الباحثة بدراسة دلالة الفروق في معامل التخمين المفردة ودقتها تبعاً للأنموذج المستخدم (ثلاثي المعلم، رباعي المعلم) وطريقة التقدير وهي طريقة (EAP وMAP وMLP) والتفاعل بينهما من خلال تحليل التباين المتعدد، وذلك باستخدام برنامج SPSS حيث تم التأكد من تجانس تباينات المجموعات الفرعية من خلال اختبار ليفين levene's test فبلغت قيمة الدلالة sig=0.000 لكل من تخمين المفردة ودقتها، وبالتالي فإن المجموعات الفرعية غير متجانسة لتخميين المفردة ودقته، وقد أظهرت نتائج تحليل التباين المتعدد ما يلي:

الجدول(٩). نتائج اختبار تحليل التباين المتعدد لدلالة الفروق في تخمين المفردة ودقتها تبعاً للأنموذج المستخدم وطريقة التقدير

مصدر التباين	معامل/ دقة	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة F	قيمة الدلالة	القرار
طريقة التقدير	تخميين المفردات	.000	2	.000	.000	1.000	غير دالة
	دقة التقدير	.000	2	.000	.000	1.000	غير دالة
الأنموذج	تخميين المفردات	.045	1	.045	6.782	.010	دالة
	دقة التقدير	.002	1	.002	3.034	.082	غير دالة
طريقة التقدير* الأنموذج	تخميين المفردات	.000	2	.000	.000	1.000	غير دالة
	دقة التقدير	.000	2	.000	.000	1.000	غير دالة
الخطأ	تخميين المفردات	2.345	354	.007			
	دقة التقدير	.210	354	.001			
الكلي	تخميين المفردات	15.740	360				
	دقة التقدير	1.743	360				

من الجدول السابق نلاحظ أن:

٤. قيمة الدلالة (sig>0.05) في دراسة أثر طريقة التقدير على معامل تخمين الفقرة ودقتها وبالتالي لا يوجد فروق ذات دلالة إحصائية في قيمة معامل تخمين الفقرة ودقتها تبعاً لطريقة التقدير.

٥. قيمة الدلالة (sig<0.05) في دراسة أثر النموذج المستخدم على معامل تخمين الفقرة وبالتالي يوجد فروق ذات دلالة إحصائية في قيمة معامل تخمين الفقرة ودقتها تبعاً للنموذج المستخدم، وقيمة الدلالة (sig>0.05) في دراسة أثر النموذج المستخدم على دقة تقدير معامل تخمين الفقرة وبالتالي لا يوجد فروق ذات دلالة إحصائية في دقة تقدير معامل تخمين الفقرة ودقتها تبعاً للنموذج المستخدم.

٦. قيمة الدلالة (sig>0.05) في دراسة أثر التفاعل بين الأنموذج المستخدم وطريقة التقدير على معامل تخمين الفقرة ودقتها وبالتالي لا يوجد فروق ذات دلالة إحصائية في قيمة معامل تخمين الفقرة ودقتها تبعاً للتفاعل بين الأنموذج المستخدم وطريقة التقدير.

رابعاً: ما أثر اختيار نموذج الاستجابة للمفردة (1PL,2PL,3PL,4PL) وطرائق التقدير (EAP وMAP وMLP) ودقتها والتفاعل بينهما في تقدير معالم الأفراد؟

قامت الباحثة بدراسة دلالة الفروق في قدرة الأفراد ودقتها تبعاً للأنموذج المستخدم (أحادي المعلم، ثنائي المعلم، ثلاثي المعلم، رباعي المعلم) وطريقة التقدير وهي طريقة (EAP وMAP وMLP) والتفاعل بينهما من خلال تحليل التباين المتعدد، وذلك باستخدام برنامج SPSS حيث تم التأكد من تجانس تباينات المجموعات الفرعية من خلال اختبار ليفين levene's test فبلغت قيمة الدلالة sig=0.000 لكل من قدرة الأفراد ودقتها، وبالتالي فإن المجموعات الفرعية غير متجانسة لتخمين المفردة ودقته، وقد أظهرت نتائج تحليل التباين المتعدد ما يلي:

الجدول(١٠). نتائج اختبار تحليل التباين المتعدد لدلالة الفروق في قدرة الأفراد ودقتها تبعاً للأنموذج المستخدم وطريقة التقدير

مصدر التباين	معامل/ دقة	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة F	قيمة الدلالة	القرار
طريقة التقدير	قدرة الافراد	5.961	2	2.980	3.090	.046	غير دالة
	دقة التقدير	50.895	2	25.448	99.886	.000	دالة
الأنموذج	قدرة الافراد	1.788	3	.596	.618	.603	غير دالة
	دقة التقدير	69.656	3	23.219	91.138	.000	دالة
طريقة التقدير* الأنموذج	قدرة الافراد	7.131	6	1.189	1.232	.286	غير دالة
	دقة التقدير	37.394	6	6.232	24.463	.000	دالة
الخطأ	قدرة الافراد	11562.786	11988	.965			
	دقة التقدير	3054.123	11988	.255			
الكلي	قدرة الافراد	11578.493	12000				
	دقة التقدير	5009.760	12000				

من الجدول السابق نلاحظ أن:

١. قيمة الدلالة (sig>0.05) في دراسة أثر طريقة التقدير والأنموذج المستخدم والتفاعل بينهما على قدرة الأفراد، وبالتالي لا يوجد فروق ذات دلالة إحصائية في قدرة الأفراد تبعاً لطريقة التقدير والنموذج المستخدم والتفاعل بينهما.

٢. قيمة الدلالة (sig<0.05) في دراسة أثر طريقة التقدير والأنموذج المستخدم والتفاعل بينهما على دقة تقدير قدرة الأفراد، وبالتالي يوجد فروق ذات دلالة إحصائية في دقة تقدير قدرة الأفراد تبعاً لطريقة التقدير والنموذج المستخدم والتفاعل بينهما.

للكشف عن مصادر الفروق في دقة تقدير الفدرة تبعاً للأنموذج المستخدم قامت الباحثة بإجراء اختبار دانت التتبعي وذلك باستخدام برنامج SPSS، كما هو موضح في الجدول التالي:

الجدول(١١). نتائج اختبار دنت التتبعي لدلالة الفروق في دقة تقدير قدرة الأفراد تبعاً لطريقة التقدير

الطريقة		متوسط الفروق	الخطأ المعياري	SIG	القرار
EAP	MAP	.006054	.0112864	.866	لا يوجد
	MLE	-.135024-*	.0112864	.000	يوجد
MAP	MLE	-.141078-*	.0112864	.000	يوجد

من الجدول السابق نلاحظ أنه هنالك فروق ذات دلالة إحصائية في دقة تقدير الأفراد بين طريقة MLE من جهة وطريقتي MAP وMLE من جهة أخرى ولصالحها في حين لم تظهر طريقة MAP , EAP فروق ذات دلالة إحصائية. وللكشف عن مصادر الفروق في دقة تقدير قدرة الأفراد تبعاً للأنموذج المقترح نلاحظ أنه يوجد فروق ذات دلالة إحصائية في دقة تقدير قدرة الأفراد تبعاً لاختلاف الأنموذج المستخدم في الدراسة، قامت الباحثة باستخدام اختبار شيفيه التتبعي للكشف عن مصادر الفروق، والجدول التالي يوضح النتائج:

الجدول(١٢). نتائج اختبار دنت التتبعي لدلالة الفروق في دقة تقدير قدرة الأفراد تبعاً للأنموذج المستخدم

النموذج		متوسط الفروق	الخطأ المعياري	SIG	القرار
أحادي المعلم	ثنائي المعلم	-.095707-*	.0130324	.000	يوجد
	ثلاثي المعلم	-.163988-*	.0130324	.000	يوجد
	رباعي المعلم	-.199981-*	.0130324	.000	يوجد
ثنائي المعلم	ثلاثي المعلم	-.068281-*	.0130324	.000	يوجد
	رباعي المعلم	-.104274-*	.0130324	.000	يوجد
ثلاثي المعلم	رباعي المعلم	-.035993-	.0130324	.054	لا يوجد

من الجدول السابق نلاحظ أن قيمة SIG<0.05 وبالتالي يوجد فروق ذات دلالة إحصائية في دقة تقدير قدرة الأفراد على وفق الأنموذج المستخدم ولصالح انموذج اكثر معالم، إلا أنه ليس هنالك فروق ذات دلالة إحصائية في دقة تقدير قدرة الأفراد بين النموذجين الثلاثي والرباعي المعلم.

الاستنتاجات:

١. يوجد فروق ذات دلالة إحصائية في نسبة المفردات الملائمة بين النماذج ثنائي المعلم وثلاثي المعلم ورباعي المعلم لصالح النموذج ذو المعالم الأكثر، في حين أظهرت النتائج دلالة الفروق في نسبة المفردات الملائمة مع الأنموذج أحادي المعلم.

٢. لا يوجد فروق ذات دلالة إحصائية في معاملات المفردة (صعوبة وتمييز وتخمين) ودقتها تعزى لعامل طريقة التقدير، ووجود فروق ذات دلالة إحصائية في قيمة معامل صعوبة المفردة (لم تظهر فروق دالة بين النموذج الاحادي المعلم والثنائي المعلم) ودقتها والتمييز ودقته وقيمة معامل التخمين لصالح النموذج ذو المعالم الأكثر، ولم يظهر لتأثير التفاعل بين طريقة التقدير والأنموذج المستخدم دلالة في الفروق.

٣. لا يوجد فروق ذات دلالة إحصائية في قدرة الأفراد تبعاً لطريقة التقدير والأنموذج المستخدم والتفاعل بينهم في حين أظهرت النتائج وجود فروق ذات دلالة احصائية في دقة تقدير قدرة الأفراد تبعاً للأنموذج المستخدم ودقة التقدير والتفاعل بينهما إلا أن دقة تقدير القدرة لم تظهر فروق دالة بين طريقتي (EAP وMAP) ولصالحها، كما لم يظهر والنموذج الثلاثي الرباعي لم يظهر فروق دالة في دقة تقدير القدرة.

التوصيات والمقترحات:

١. إن زيادة معاملات المفردة والأفراد ودقتها تبعاً للأنموذج المستخدم ولصالح النموذج ذو المعالم الأكثر، ويمكن المفاضلة في اختيار النموذج تبعاً للمستوى الخطأ المسموح

٢. إذا أهتمت الدراسة بمعاملات صعوبة المفردة فلا فرق بين النموذج أحادي المعلم وثنائي المعلم.

٣. إذا كان الهدف من الدراسة التعامل مع معامل القدرة فليس هنالك فروق في اختيار النموذج ثنائي الاستجابة.

٤. يجب مراعاة اختيار الأنموذج الأنسب في حال تناول دقة تقدير معاملات المفردة والأفراد في النماذج ثنائية الاستجابة، في حين أن تأثير الأنموذج ضعيف عند تناول قيم معاملات المفردات.

٥. تناول النموذج رباعي المعلم والقيام بدراسات مقارنة للتعرف على ما يقدمه هذا الانموذج ومقارنته بالنماذج ثنائية الاستجابة الأخرى.

٦. مراعاة طريقة التقدير الأنسب والأنموذج الأنسب في الدراسات التي تتناول دقة تقدير قدرة الأفراد.

المراجع

١. التقي ، احمد محمد ،(٢٠٠٩)، **النظرية الحديثة في القياس** ، ط١، دار المسيرة ، عمان

٢. **التقي**، احمد محمد .(٢٠١٣). **النظرية الحديثة في القياس**، ط٢، عمان، دار المسيرة للنشر والتوزيع.

٣. **الرحيل**، راتب. **الدرابسة**، رياض. (٢٠١٤). أثر طريقتي التعامل مع القيم المفقودة وطريقة تقدير القدرة على دقة تقدير معالم الفقرات والأفراد. **المجلة الدولية التربوية المتخصصة**. المجلد ٣. العدد ٦. ص ٢٣-٤٧.

٤. **زكري**، علي بن محمد عبد الله. (٢٠٠٩). الخصائص السيكومترية لاختبار أوتيس – لينون للقدرة العقلية مقدرة وفق القياس الكلاسيكي ونموذج راش لدى طلبة المرحلة المتوسطة بمحافظة صبيا التعليمية، **أطروحة دكتوراه غير منشورة**، جامعة أم القرى، كلية التربية.

٥. الزهراني، بندر بن حمدان. (٢٠٠٨): **اثر اختلاف حجم العينة واتساع مدى القدرة على تقدير الدرجة الحقيقة المقدرة باستخدام النظرية التقليدية والنماذج احادية البعد في النظرية الحديثة للقياس**، رسالة دكتورة ، جامعة أم القرى، المملكة العربية السعودية.

٦. سرحان، محمد. (٢٠١٨). المقارنة بين طريقتي الأرجحية العظمى والقيمة العظمى للتوزيع البعدي في تقدير القدرة وفقاً لنظرية الاستجابة للفقرة. **أطروحة دكتوراه غير منشورة**. كلية التربية ابن رشد للعلوم الإنسانية. جامعة بغداد.

٧. **الشرقاوى**، أنور. **الشيخ**، سليمان. **كاظم**، أمينة. **عبد السلام**، نادية. (١٩٩٦): **اتجاهات معاصرة في القياس والتقويم النفسي والتربوى** ، القاهرة ، الانجلو المصرية .

٨. **ضعضع**، هبة. **طومان**، منار. **طيفور**، مصطفى. (٢٠٢٠). أثر حجم العينة وطرائق التقدير في دقة تقدير معالم نموذج راش. **مجلة جرش للبحوث والدراسات**. المجلد ٢١. العدد ١ .ص ص ١٣١– ١٧٠.

٩. **طومان**، منار. (٢٠٠٦). فاعلية استخدام نظرية الاستجابة للمفردة (IRT) في عملية معادلة درجات الاختبارات متعددة الأبعاد والمتغيرات المؤثرة فيها. **أطروحة دكتوراة غير منشورة**. جامعة القاهرة.

١٠.العبد الله، زياد.(٢٠١٢) أثر بعض طرق التقدير على دقة تقدير المعالم ضمن نماذج الاستجابة للمفردة متعددة التدريج. **أطروحة دكتوراه غير منشورة**. معهد الدراسات التربوية. جامعة القاهرة: مصر.

١١.**عبيدات**، ذوقان. **عدس**، عبد الرحمن. **عبد الحق**، كايد. (٢٠٠٠). **البحث العلمي: (مفهومه، أدواته، أساليبه)**. الرياض: دار أسامة للنشر والتوزيع.

١٢.علام، صلاح الدين محمود. (٢٠٠٢).(*القياس والتقويم التربوي والنفسي: اساسياته وتطبيقاته وتوجهاته المعاصرة* .دار الفكر العربي.

١٣.**علام**، صلاح الدين. (٢٠٠٠). **القياس والتقويم التربوي والنفسي – أساسياته وتطبيقاته وتوجهاته المعاصرة**. القاهرة، دار الفكر العربي.

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

١٥.**كروكر**، ليندا، **وألجينا**، جيمز (٢٠٠٩).**مدخل إلى نظرية القياس التقليدية والمعاصرة**، ترجمة، دعنا،زينات يوسف، ط١، دار الفكر ناشرون وموزعون.

16. Chen, W. H., & Thissen, D. (1997). Local dependence indexes for item pairs using item response theory. *Journal of Educational and Behavioral Statistics*, *22*(3), 265-289.

17. **de Gruijter**, D. N. M. & **van der Kamp**, L. J. Th. (2005). **Statistical Test Theory for Education and Psychology**. © D. N. M. de Gruijter & L. J. Th. Van der Kamp.

18. Elliott, C. D. (1983b). *Vertical equating oj unidimensional tests using the Rasch model.* Unpubli shed paper, Department of Education, Uni vers ity of Manchester, England.

19. Emons, W. H.; Glas, C. W.; Meijer, R. R. & Sijtsma, K. (2003): Person Fit in Order-Restricted Latent Class Models. **Applied Psychological Measurement,** Vol(27), No(6), pp.459-478.

20. Hambleton, R. K. & **Swaminathan**, H. (1985). Item Response Theory, Principles and Applications. Bston: Kluwer. **Nijhoff Publishing amembers of the Kluwer Academic Publishers Group.**

21. **Hambleton**, R. K., **Zaal**, J. N., & **Pieters**, J. P. M. (1991). Computerized adaptive testing: Theory**, applications, and standards**. In R. K.

22. Holland, P,W. (1990): **On the Sampling Theory Fondations of Item Response Theory Models**, *PSYCHOMETRIKA*, VOL. 55, NO. 4, 577-601. http://www.psychology.gatech.edu.unfolding.

23. **Magno**, C. (2009). Demonstrating the Difference between Classical Test Theory and Item Response Theory Using Derived Test Data, **The International Journal of Educational and Psychological Assessment** April 2009, Vol. 1, Issue 1, pp. 1-11.

24. Mislevy, R. T. (1990): Modeling Item Response When Different Subjects Employ Different Solations Strategies, **Netherlands: Psychometrica,** Vol. 55, No., PP. 195-215.

25. Rose, Norman .(2010). **Maximum Likelihood and Bayes Ability Estimation in TWO – Parametric IRT Models : Derivations and Implementation** ,Institute of Psychology Friedrich Schiller University Jena.

26. Wainer, H., & Mislevy, R. J. (1990). Item response theory, item calibration, and proficiency estimation. *Computerized adaptive testing: A primer*, *4*, 65-102.