أثر النموذج اللوجستي وطرق تقدير الأرجحية العظمى على تحيز قيم معالم الفقرة وقدرة الأفراد

The effect of Maximum Likelihood Estimation Methods and the Logistic Model on the Bias of Item Parametres Values and the Ability of Individuals

أ.م.د. وليد خالد عبدالكريم بابان

كلية التربية/ جامعة صلاح الدين–أربيل

Assis. Prof.Dr. Waleed Khalid Abdulkareem Baban 1

Salahaddin University/Erbil

waleed.baban@su.edu.krd

ملخص البحث:

استهدف البحث الحالي المفاضلة في تقدير بارامترات الفقرة وقدرة الأفراد وفقا لطريقة التقدير والنموذج اللوجستي باستخدام مؤشري التحيز (BAIS)، و (Akaiki)، بوصفهما مؤشرين على تحيز التقديرات، لذلك استخدم الباحث المنهج التجريبي وفق تصميم القياسات المتكررة، إذ طبق أداة البحث والمتكونة من اختبار الذكاء المجرد لـ روبرت ثورندايك والقائمة على القدرة لمعالجة الألفاظ والرموز وكيفية التعامل معها، حيث طبق الاختبار على عينة بلغت (1072) راشدا وراشدة تم الحصول عليهم من طلبة جامعة صلاح الدين/أربيل، ولغرض تحليل البيانات التي تم الحصول عليها، لجأ الباحث إلى ثلاثة من النماذج الثنائية التدريج لنظرية الاستجابة للفقرة وبعد الاطمئنان لمطابقة البيانات المتولدة من النماذج اللوجستية الثلاث (راش، لورد، بيرنبوم)، حيث تم استخدام برنامج(Bilog–MG3) لتقدير معالم الفقرة كل من (الصعوبة، التمييز، التخمين) وتبعاً للنماذج المذكورة آنفًا.

وكان الباحث قد استخدم كل من طريقتي الأرجحية العظمى المشتركة والهامشية، لتقدير معالم الفقرات، وبعد ذلك أجرى الباحث مفاضلة في تقدير بارامترات الفقرة وفقا لطريقة التقدير والنموذج اللوجستي باستخدام مؤشر التحيز (BAIS)، وتبين له أن تقدير بارامترات الفقرة وفق طريقة الأرجحية العظمى المشتركة أقل انحرافًا عن متوسط

المعالم، وهو ما يؤشر إلى أن هذه الطريقة أقل تحيزا من تقدير البارامترات وفق طريقة الأرجحية العظمى الهامشية.وهو ما يؤشر إلى أن طريقة الارجحية المشتركة أكثر دقة في تقدير البارامترات الصعوبة والتمييز والتخمين وفق نموذج بيرنبوم ثم يليه نموذج لورد وبعدها نموذج راش وحسب خصوصية النموذج في تقدير البارامترات، مقارنة بطريقة الأرجحية الهامشية. وبعد تقدير بارامترات القدرة وفق النماذج اللوجستية الثلاث تم إجراء مفاضلة لتقدير القدرة وفق مؤشر التحيز لطريقتي الأرجحية العظمى المشتركة والهامشية تبعاً لنموذج بيرنبوم كانت أقل من قيم الوسط الحسابي لنموذج لورد وكذلك بالنسبة لنموذج راش، وهو ما يؤشر أفضلية نموذج بيرنبوم على النموذجين الآخرين لتقليل تحيز تقدير قدرة الأفراد يليه نموذج لورد ثم نموذج راش، كما جاءت نتائج حساب التحيز وفق معيار أكايكي للمعلومات مطابقة للنتائج التي تم الحصول عليها وفق مؤشر التحيز (BAIS). وبناء على النتائج التي توصل إليها البحث الحالي فإن الباحث أوصى باستخدام نموذج بيرنبوم وتبعا لطريقة الأرجحية العظمى المشتركة في تقدير معالم الفقرة وقدرة الأفراد، وذلك لتحاشي تحيز التقديرات والحصول على نتائج أكثر دقة في التقدير. كما اقترح الباحث إلى إجراء المزيد من الدراسات كمقارنة النماذج المتدرجة لنظرية الاستجابة للفقرة في الكشف عن النموذج الأكثر دقة وأقل تحيزا في تقدير معالم الفقرة والأفراد.

الكلمات المفتاحية: النموذج اللوجستي، الأرجحية، التحيز، معالم الفقرة، قدرة الأفراد.

Abstract

The current research aims to compare the estimation of Item parameters and the ability of individuals according to the estimation method and the logistic model using the bias indices (BAIS) and (Akaiki), as indicators of bias estimates. Therefore, the researcher used the experimental method according to the repeated measurements design, where he applied the research tool consisting of Robert Thorndike's abstract intelligence test based on the ability to process words and symbols and how to deal with them. The test was applied to a sample of (1072) adults, male and female, obtained from students at the University of Salahaddin/Erbil. In order to analyze the data obtained, the researcher resorted to three of the binary gradual models of the Item response theory. After ensuring that the data generated from the three logistic models (Rasch, Lord, Birnbaum) matched, the (Bilog–MG3) program was used to estimate the Item parameters (difficulty, discrimination, guessing) according to the aforementioned models.

The researcher used both the joint maximum likelihood and marginal likelihood methods to estimate the Item parameters. After that, the researcher made a comparison in estimating the Item parameters according to the estimation method and the logistic model using the bias index (BAIS). It became clear to him that estimating the Item

parameters according to the joint maximum likelihood method is less deviant from the average of the parameters, which indicates that this method is less biased than estimating the parameters according to the marginal maximum likelihood method. This indicates that the joint likelihood method is more accurate in estimating the difficulty, discrimination and estimation parameters according to the Birnbaum model, followed by the Lord model, then the Rasch model, and according to the specificity of the model in estimating the parameters, compared to the marginal likelihood method. After estimating the ability parameters according to the three logistic models, a comparison was made to estimate the ability according to the bias index for the joint maximum and marginal odds methods according to the Birnbaum model, which was less than the arithmetic mean values for the Lord model as well as for the Rasch model, which indicates the superiority of the Birnbaum model over the other two models in reducing the bias in estimating the ability of individuals, followed by the Lord model and then the Rasch model.The results of calculating bias according to Akaike Information Criterion were consistent with the results obtained according to the bias index (BAIS). Based on the results reached by the current research, the researcher recommended using the Birnbaum model and according to the joint maximum likelihood method in estimating the Item parameters and the ability of individuals, in order to avoid bias in the estimates and obtain more accurate results in the estimation. The researcher also suggested conducting more studies such as comparing the graded models of the Item response theory in revealing the most accurate and least biased model in estimating the Item parameters and individuals.

Keywords: Logistic model, likelihood, Bias, Item parameters, individual ability.

مشكلة البحث:

لاشك أن توخي الدقة بتقدير قدرة الأفراد وموضوعية القياس بات هدف الباحثين وعلماء القياس التربوي والنفسي، من أجل الحصول على أعلى مراتب الدقة المتوخاة من تقدير السمات والخصائص النفسية والتربوية، وفي ظل التطورات التي يشهدها ميدان القياس النفسي والتربوي، فقد ارتبطت الدقة والموضوعية بنظرية الاستجابة للفقرة، لكونها تغلبت على العديد من المشكلات التي يعاني منها الاتجاه الكلاسيكي والمتمثل بنظرية القياس الكلاسيكية (علام، ٢٠٠٥، صفحة ٩١). لذلك سعى الكثير من علماء القياس النفسي والتربوي إلى تحقيق الموضوعية والدقة في تقدير سلوك الأفراد، باعتباره خطوة هامة لفهم الظاهرة السلوكية وضبطها ثم التحكم فيها، وبغرض الحصول على تقدير للقدرات أو السمات التي تتسم بأعلى درجات الدقة، جاءت نظرية الاستجابة للفقرة (IRT)، التي تعد تطورا مهما في ميدان القياس النفسي والتربوي، لكونها تعالج الكثير من قضايا القياس بنحو أكثر فعالية من النظرية الكلاسيكية (Hambleton & Jones, 1993, p. 40)(CTT).

وتعد عملية تقدير المعالم وفق طرق تقدير الأرجحية العظمى عملية محفوفة بالحذر لأنه كل طريقة أو أسلوب قد تمارس دورًا في تضخيم أو تدني قيم التقدير، لذلك فإن هذا التذبذب الراجع لطبيعة طريقة التقدير قد يؤدي إلى تحيز قيم التقدير سواء على صعيد معالم الفقرات أو قدرة الأفراد، وهو ما قد ينعكس بنحو مباشر على الدقة والموضوعية في القياس. وقد تتفاقم تلك المشكلة إذا ما أخذنا بالحسبان طبيعة النموذج اللوجستي المعتمد في التقدير واستخلاص درجات المفحوصين بالاعتماد على خصائص الفقرات التي يضمها المقياس لذلك فإن عملية التقدير ليست مستقلة عن تلك الفقرات. وبناء على ذلك فإننا أمام مشكلة واجبة الدراسة والتحقق من آثارها على تحيز قدرة الأفراد وقيم معالم الفقرة المتمثلة بـ(صعوبة، تمييز، تخمين)، إذ إنّ تباين أخطاء القياس قد يكون هنا مختلفا عما عهدناه في مبادئ القياس النفسي والتربوي والذي تم حصر مصادرها بـ(الأداة، المفحوصين، القائم بعملية القياس، ظروف التطبيق لأداة القياس)، إذ إن تباين طرق تقدير الأرجحية العظمى وتعددها والنموذج اللوجستي قد يكونان هما مصادر للخطأ والذي قد تتباين بدورها تبعا لطريقة التقدير الأرجحية ونماذج نظرية الاستجابة للفقرة.

وتُعتمد عملية تقدير معالم الفقرات وتقديرات معلمة القدرة للمفحوصين على عوامل عدة في مقدمتها اختيار الطريقة المناسبة للتقدير. فقد أشار بعض الباحثين مثل هامبلتون، ولورد إلى أنه تزداد الدقة في تقدير القدرة كلما قل الخطأ المعياري للتقدير (Standard Error of Estimate) (Luc & Adams, 2013, p. 72).كما ويرى لندن وهامبلتون (Linden & Hambleton, 1997) أن وجود أخطاء في تقدير معالم الفقرات ينعكس بنحو سلبي على دقة تقدير تلك المعالم، كما يؤثر على دالة معلومات الاختبار(Linden & Hambleton, 1997, p. 39).

ومن هنا استشعر الباحث بخطورة هذه المشكلة وبلورها في إطار هذا البحث، إذ سيخضع طريقتي الأرجحية العظمى المشتركة (Joint Maximum likelihood)، وطريقة الأرجحية العظمى الهامشية (Marginal) Maximum likelihood، والنماذج اللوجستية الثنائية الدرجة والمتمثلة بـ(راش، لورد، بيرنبوم) للبحث والتمحيص،وبناءً على ماسبق فإن الباحث ومن خلال البحث الحالي يسعى إلى الإجابة على تساؤل رئيسي مفاده:

- هل هناك أثر للنموذج اللوجستي على تحيز كل من قيم تقدير معالم الفقرة وقدرة الأفراد؟
- هل هناك أثر لطريقة التقدير الأرجحية العظمى على تحيز كل من قيم تقدير معالم الفقرة وقدرة الأفراد؟
- هل هناك أفضلية للنموذج اللوجستي وطريقة التقدير الأرجحية العظمى على تحيز كل من قيم تقدير معالم الفقرة وقدرة الأفراد؟

أهمية البحث:

أثبتت البحوث التجريبية بأنها تحقق الدقة والموضوعية التي تهدف العلوم النفسية والتربوية بلوغها. يتطلب الأمر استخدام التوجهات الحديثة في القياس، حيث يعد الكشف عن أداء الفرد على فقرات المقاييس والأختبارات أمرًا مهمًا، لما له من أهمية بالغة، حيث إن هذا الأداء يتمثل في أنه قد يكون مؤشرًا على تحيز هذه الفقرات، وحينئذٍ يؤدي إلى عدم تحقق عدالة الاختبار، كما أنه قد يؤثر في صدق الاختبارات وثباتها (Salubayba, 2013, p. 43).

فقد يحدث في كثير من الأحيان مقارنة شخصين أو أكثر بناء على درجاتهم الخام في نماذج مختلفة من اختبار معين، وهذا فيه شيء من الظلم والإجحاف، ولهذا فإن نظرية القياس الحديثة توفر أسلوبا أفضل للمقارنة، فهي لا توازن الدرجات الخام وإنما يكون الهدف هو معرفة موقع القيمة الرقمية لقدرة الفرد على مقياس القدرة، الأمر الذي يؤدي لتحيز التقديرات، سواء بالمغالاة أو خفض التقديرات، من هنا تتضح الحاجة للمزيد من الجهد البحثي في هذا المضمار، الذي يتبنى الحكم على تقديرات الفقرات وقدرة الأفراد المفحوصين، ومدى تحيز تلك التقديرات، ولاسيما وأن تلك التقديرات متعلقة بطرائق تقديرها والنماذج اللوجستية التي يوظفها الباحثون في هذا الصدد، لاسيما وأن الكشف عن التحيز في الاختبار (Test bias) ، يعد من تطبيقات نظرية الاستجابة للفقرة،والتي من خلالها يمكن للكشف عن تحيز الاختبار (جرادات، ٢٠١٢، صفحة ٤٤).

بيد أن استخدام المنحى أو الاتجاه الحديث في القياس النفسي والتربوي، أو ما أشيع تسميته بنظرية الاستجابة للفقرة، وما تتطلبه من أساليب إحصائية كطرق تقدير الأرجحية العظمى لتقدير البيانات أو لتحليلها، ليست بالأمر الهين أو البسيط، وذلك لضرورة ملائمة تلك الأساليب مع النماذج المستخدمة لتقدير معالم الفقرة، فضلاً عن ملائمة هذه النماذج مع الفقرات المراد استخراج معالمها، مما يحتم على الباحثين بذل جهد مضاعف بغية التأكد من ذلك، لذلك حاول علماء القياس المعاصر الاستفادة من التقدم التكنولوجي في التوصل إلى طرق قياسية، جديدة تنبثق عنها نماذج عرفت بنماذج الاستجابة للفقرة، فباستخدام هذه النماذج يمكن الحصول على تقديرات إحصائية لمعالم الفقرات مستقلة عن خصائص المفحوصين، كما أن الدرجات تعبر عن درجة امتلاك الفرد للسمة قيد القياس، والتي تكون مستقلة عن الخصائص القياسية للفقرة.

وفي ذات السياق فإن استخدام نظرية الاستجابة للفقرة، يقترن مع تبني نموذج لوجستي، يوفر الأساس لربط فقرات الاختبار بمستويات القدرة للأفراد، إذ إن أكثر ميزة في هذه النظرية يمكن أن تجتذب الباحثين إليها هي تعدد

نماذجها لتقدير معالم (بارامترات) الفقرة، مما تكسبها سمة المرونة الفائقة في التعامل وتناول الكثير من قضايا ومجالات القياس النفسي والتربوي.

وفي هذا الصدد يشير (Hambleton & Jones, 1993)، في دراسة أجرياها للمقارنة بين النظريتين، إلى تفوق النظرية الحديثة على النظرية التقليدية في التغلب على جوانب القصور التي تعاني منها النظرية التقليدية في مواجهة المسائل الاختبارية مثل معادلة درجات الاختبار، الكشف عن تحيز الفقرات، تدريج المقاييس(Hambleton & Jones, 1993, p. 51).ومن هنا تبرزت الحاجة وأهمية البحث الحالي الذي يرمي إلى تركيز الجهود للتثبت من صلاحية النموذج الذي يتم تبنيه، وتوخي دقة تقديراتها لمعالم الفقرة، وقدرة عينة المفحوصين من جراء تطبيق أداة البحث عليهم، لاسيما وأن الأخطاء المعيارية في نظرية الاستجابة للفقرة، تتأثر بمعالم الفقرة والنموذج المستخدم.

أهداف البحث:

يسعى الباحث من خلال البحث الحالي إلى تحقيق الأهداف التالية:

١. تقدير معالم الفقرة وفق النماذج اللوجستية تبعا لطريقة الأرجحية العظمى المشتركة والهامشية.
٢. المفاضلة في تقدير بارامترات الفقرة وفقا لطريقة التقدير والنموذج اللوجستي باستخدام مؤشر التحيز.
٣. المفاضلة في تقدير قدرة الأفراد وفقا لطريقة التقدير والنموذج اللوجستي باستخدام مؤشر التحيز.
٤. المفاضلة في تقديرقدرة الأفراد وفقا لطريقة التقدير والنموذج اللوجستي وفق معيار "أكايكي" للمعلومات.

تحديد المصطلحات:

- **النموذج اللوجستي:** محموعة من الدوال الرياضية، التي تفسر أداء الفرد في اختبار ما، والسمات التي تكمن خلف الأداء (ضعضع، ٢٠٢١، صفحة ٤٥٤).
- **الأرجحية العظمى Maximum Likelihood methods:**وهي تقدير للمعالم من خلال تعظيم الاحتمالية للمعلمة المراد تقديرها،وطرق تقدير الأرجحية القصوى لمعالم الفقرات تختلف في كيفية التعامل مع مستويات القدرة غير المعلمة(Embretson & Reise, 2000, pp. 217-218).
- **تحيز التقدير** Bias:وهو الفرق بين المعلمة المقدرة والمعلمة المولدة، وهو مؤشر لدقة تقدير المعالم (Brian & Diane, 1997, p. 57).

ويعرفها الباحث إجرائيًا، بأنها الفرق بين تقديرات معالم الفقرات وقدرة الأفراد المقدرة من النماذج اللوجستية الأحادية والثنائية والثلاثية المعلمات وبين المعالم الحقيقية للفقرات وقدرة الأفراد تبعا لطريقة تقدير الأرجحية العظمى.

- **طرق التقديرEstimation Methods** :هي مجموعة من الأساليب الرياضية الاحتمالية تستخدم بهدف تقدير معالم الفقرة والأفراد في إطار افتراضات نظرية الاستجابة للفقرة (ضعضع، طومان، و طيفور، ٢٠٢٠، صفحة ١٣٧).
- **معالم الفقرة (Item Parametres)**:هي دوال لوغارتمية للأداء على الاختبار باستخدام المجموعة المرجعية والمستهدفة من خلال معلم قدرة الأفراد والتي تستخدم معلم ($B_{uncorected}$) الذي يتم تقديره من خلال التحيز.(Engelhard & Wind, 2018, p. 121)(BIAS)

ويعرفها الباحث إجرائيًا، بأنها ما يتمخض من النماذج اللوجستية الأحادية والثنائية والثلاثية المعلم لفقرات أداة البحث، تبعا لطريقتي تقدير الأرجحية العظمى المشتركة والهامشية.

- **معلم القدرةParameter Ability** :هي درجة المفحوص وتعبر عن مقدار ما يمتلكه من السمة المقاسة بأداة القياس، ويعبر عنها بموقع المفحوص على متصل السمة قيد القياس (Hambleton & Swaminathan, 1985).

ويعرفها الباحث إجرائيًا، بأنها قيم إحصائية يتم تقديرها في هذا البحث باستخدام طريقتي تقدير الأرجحية العظمى (المشتركة، الهامشية)، و وفق كل من النماذج (راش ولورد وبيرنبوم) الذي يضم معالم (الصعوبة،التمييز،التخمين).

الإطار النظري

- **النماذج اللوجستية ModelsLogistics**:

تضم نظرية الاستجابة للفقرة من مجموعة من النماذج التي تهدف إلى تحديد العلاقة بين أداء الفرد على الاختبار والذي يمكن ملاحظته بصورة مباشرة من خلال الأداء الأفراد، وبين مجموعة السمات أو القدرة التي تكمن وراء هذا الأداء، في صورة دالة رياضية احتمالية تقوم على افتراض أساسي أي أنّه كلما ازدادت قدرة الفرد على صعوبة الفقرة ازدادت احتمالية حدوث الاستجابة الصحيحة للفقرة والعكس الصحيح. إذ إنّ هذه النماذج احتمالية Probabilistic models، إذ يمكن الاعتماد عليها واستعمالها بدرجة أقرب إلى الثقة والموضوعية في المقاييس النفسية(Linden & Hambleton, 1997, p. 10).

وتفترض هذه نماذج أنه يمكن التنبؤ بأداء المفحوصين من خلال واحدة أو أكثر من السمات التي تميز الأداء والتي تكمن خلفها القدرة، كما يجب أن تقدر تلك السمات أو القدرات من خلال الأداء الملاحظ لمجموعة المفحوصين على مجموعة فقرات الاختبار(Hambleton & Swaminathan, 1985, pp. 9-13). إذ إنّ هذه النماذج تحاول اشتقاق قياسات أو قيم تقديرية للسمة أو السمات التي تنطوي عليها مجموعة من الاستجابات لمجموعة من المفردات الاختبارية، أي أن هذه النماذج تفسر استجابة الفرد للفقرة الاختبارية تنطوي على السمة التي تقيسها الفقرة،كما وأن جميع النماذج تستند، نماذج رياضية، بوصفها قيمة عامة للتوزيع المتصل للقدرة، وتستند هذه النماذج على عدد من الفروض التي يجب أن تتوافر في البيانات المستمدة من الاختبارات (Elliott, 1983, p. 14).

بيد أن اختلاف عدد معالم الفقرة يرجع إلى اختلاف منظور علماء القياس فيما يتعلق بنمذجة العملية الاختبارية، فمنهم من يرى أنه ينبغي أن تكون للفقرة معلم واحد هو الصعوبة (Item Diffuculty Parameter)، وبينهم من يضيف إلى ذلك معلم آخر هو تمييز الفقرة (ItemDiscrimination Parameter)، وآخرون يرون أنّه إلى جانب المعلمين السّابقين من الأفضل إضافة معلم ثالث يأخذ في الاعتبار عامل تخمين الاستجابة وخاصة في فقرات الأختيار من متعدد (Guessing Parameter) لذلك تعددت النماذج المتعلقة بهذه النظرية وتباينت أشكال المنحنيات تبعًا لاختلاف عدد معالم المفردات، فنموذج راش(Rasch Model)هو نموذج أحادي البارامتر أو المعلم، ويفترض هذا النموذج توازي المنحنيات المميزة لجميع مفردات الاختبار (تختلف في صعوبتها)، بينما نموذج لورد (Lord Model) هو نموذج ثنائي المعلم، ويتحرى تمييز الفقرة، أما نموذج بيرنبوم (Birnbum Model) هو نموذج ثلاثي المعلم، يهتم بتخمين الفقرة(علام، ٢٠٠٥، الصفحات ٤٧-٦١).

- **طريقة الأرجحية العظمى (Maximum Likelihood methods):**

ويتم تقدير المعالم (Parametres Estimation)فيها من خلال تعظيم الاحتمالية للمعلمة (Maximumming Likelihood Parameter)المراد تقديرها، إذ إنّ هناك ثلاث طرق لتقدير معالم الفقرات ضمن طريقة الأرجحية العظمى وهذه الطرق تختلف في كيفية التعامل مع مستويات القدرة غير المعلومة، وهي كالآتي:

أ. طريقة الأرجحية العظمى المشتركة (Joint Maximum Likelihood):

ويتم تقدير معالم الفقرات وفق هذه الطريقة على مرحلتين حيث يتم بدأً افتراض قيم للقدرة (θ)كقيم مبدئية لمعالم الأفراد ومن خلال هذه القيم المفترضة يتم تقدير معالم الفقرات من خلال البدء بمعالم أولية للفقرات ليتم في

المرحلة الثانية تقدير القدرة أو ما تسمى بمعالم الأفراد اعتمادا على معالم الفقرات التي تم تقديرها في المرحلة الأولى وبتكرارهذه العملية كما في المرحلتين السابقتين حتى نحصل على قيم تقديرية للمعالم لا تختلف بين خطوتين متتالتين بأقل من مقدار ثابت معتمد مسبقا محك التقارب أي هناك نوع من الثبات في تقديرات المعالم (علام، ٢٠٠٥، صفحة ٩٦).ونلاحظ في هذه الطريقة أن استراتيجية التقدير لكل من معلم قدرة الأفراد أو معالم الفقرات يتم في آن واحد من خلال رفع الحد الأعلى لدالة الأرجحية المشتركة(De Ayala, 2009, p. 41).

وفي هذه الطريقة يتم النظر إلى أن فقرات الأداة والأفراد على أنهما ثابتين، أي ينصب اهتمامها فقط على مفردات الأداة فقط ولا تتجاوز إلى تعميم مفردات أخرى ربما قد تحتويها هذه الأداة ونفس الشيء بالنسبة للأفراد فالاهتمام لا يتجاوز عينة التدريج المستخدمة.

ب. طريقة الأرجحية العظمى الهامشية (Marginal Maximum Likelihood):

تم تطوير هذه الطريقة من قبل بوك واينكن(Bock & Aitkin, 1981) حيث استخدمت في تقدير معالم فقرات نماذج الاستجابة للمفردة الأحادية والمتعددة الأبعاد، وهي فعالة سواء كانت عدد الفقرات قليلة أو كثيرة (علام، ٢٠٠٥، صفحة ٩٨)، ويتم في هذه الطريقة الفصل بين عمليتي تقدير قدرة الأفراد وتقدير معالم الفقرة، حيث يمكن الحصول على تقديرات معالم الفقرات في الأول وبعد تحقق من مطابقة البيانات للنموذج يتم الانتقال إلى تقدير قدرة الأفراد.

وفي هذا الصدد تذكر(De Ayala, 2009) أن عملية الفصل بين تقديرات الأفراد والتي هي معالم عرضية (Incidental parameters)، وبين معالم المفردات التي هي معالم بنيوية أو هيكلية (Structural Parameters)،ستؤدي إلى نتائج أقل تحيزا في تقديرات قدرة الأفراد ومعالم المفردات من طريقة الأرجحية العظمى المشتركة في التقدير ولاسيما عند تدريج أدوات قياس تتكون من عدد فقرات قليلة (١٥ فقرة أو أقل) (De Ayala, 2009, p. 69).ولكون هذه الطريقة تتطلب عمليات حسابية كثيرة لذلك تعد غير عملية إلا في الاختبارات القصيرة جدا، ولكن فيما بعد ومع تطوير إجرائها من طرف بوك واينكن (Bock & Aitkin, 1981) من خلال خوارزمية توقع القيمة العظمى أدى إلى أن يكون إجراؤها مقبولا نظريا ومتاحة بنحو عملي وأكثر بساطة من الناحية الحسابية.(Embretson & Reise, 2000, p. 170)

والافتراض هو أن الأفراد يتم انتقاؤهم بطريقة عشوائية من مجتمع يتبع التوزيع الطبيعي (The normal distribution)وحينئذ يتم الحصول على معلومات المجتمع في تقدير معالم المفردات دون

الحاجة إلى تقدير قدرات الأفراد و يتم الفصل بين تقدير معالم المفردات وتقديرات قدرات الأفراد (Baker & Kim, 2004, p. 158).

ج. طريقة الأرجحية العظمى الشرطية (conditional Maximum Likelihood):

قدم هذه الطريقة راش (Rach, ١٩٦٠) ثم قام أندرسون (Andersen, 1973) فيما بعد بتطوير هذه الطريقة لتقدير معالم النموذج الأحادي المعلمة. علاوة على ذلك فإنه يمكن تطبيقها على النماذج المنبثقة عن نموذج راش. ففي هذه الطريقة يتم اشتراط دالة الأرجحية بعدد الإجابات الصحيحة التي أجاب عنها الأفراد، وأن الدرجة الكلية للفرد تمثل مجموع الفقرات التي أجاب عنها الفرد لذلك تعد إحصاءً كافيا لتقدير قدرة الفرد وليس هناك حاجة إلى معلومات أخرى من البيانات لتقدير المعلمة. وحينئذٍ فإن في هذه الطريقة يتم حذف الفقرات التي يجيب عنها جميع الأفراد إجابة صحيحة أو إجابة خاطئة كما يحذف الأفراد الذين تكون إجاباتهم كلها صحيحة أو كلها خاطئة.(Engelhard & Wind, 2018, p. 238)

وقد أشار هاميلتونوسوامينثان (Hambleton & Swaminathan, 1985, p. 139)بأن هذا الأسلوب في تقدير يكون فعالا في تقدير معالم الفقرة لغاية(٤٠) فقرة، وتكون أقل سرعة عند (٦٠) فقرة أو أكثر وعند عدد الفقرات تتراوح بين (٨٠) (١٠٠) تصبح غير مناسبة. تهدف طريقة الأرجحية العظمى الشرطية CML إلى نمذجة احتمالات نمط استجابة المفردة ما مشروط بإجمالي درجات اختبار الأفراد، ولتوضيح ذلك نفترض المصفوفة.

ويشير أمبريستون وريزر(Embretson & Reise, 2000) بأن طريقة الأرجحية العظمى المشروطة في التقدير تتميز بمجموعة من المميزات وهي:

- لا تعتمد على أي افتراض حول اي توزيع محدد المستويات القدرة أو السمة الكامنة في تقدير معلمة الصعوبة.
- تحقق مبدأ اللا تغاير (Invariance)في تقدير معالم الفقرة (صعوبة الفقرات) عند اختلاف الأفراد بسبب عدم تأثر هذه التقديرات بمستوى قدرات عينة الأفراد.
- تتميز التقدير المحصل عليها بهذه الطريقة بخصائص جيدة مثل الاتساق(Consistency) حيث كلما كبر حجم العينة فإن التقديرات تقترب من القيم الحقيقية، التوزيع الاعتدالي، والفعالية (Efficiency)أي أن لهذه التقديرات أقل تباينا(Embretson & Reise, 2000, pp. 217-218).

وقدم(Embretson & Reise, 2000) بعض الانتقادات لهذه الطريقة في التقدير يمكن إيجازها كما يلي:

- اقتصار استخدام هذه الطريقة في النموذج الأحادي المعلم (نموذج راش) وتعميماته فقط، إذ إنه في النماذج الأخرى (الثنائي والثلاثي المعلم) الدرجات الكلية في الاختبار لا تعد إحصاء كافيا لتقدير مستويات القدرة للأفراد. (ولهذا السبب فإن الباحث سيستبعد تقدير معالم الفقرات وقدرة الأفراد وفق طريقة الأرجحية العظمى الشرطية (conditional Maximum Likelihood)، لكون البحث الحالي سيستند إلى استخدام نموذج الأحادي إلى جانب النموذجين اللوجستيين الثنائي والثلاثي المعلم، وذلك لتعذر استخدام النموذجين اللوجستيين الأخيرين.
- طريقة الأرجحية العظمى المشروطة لا تقوم بالتقدير للعلامات الكاملة أو الصفرية سواء كانت للأفراد أو المفردات وحينئذٍ تؤدي إلى فقدان بعض المعلومات.
- توجد صعوبات في التحليل العددي عندما يكون عند البنود أو المفردات في الاختبار كبيرا مما يؤدي إلى فشل في استكمال إجراءات التقدير لبرامج الحاسوب ولاسيما إذا فاق عدد الفقرات ١٠٠ فقرة.

- **مؤشر التحيز للمعلمة:**

عندما تبدي فقرة ما خصائص إحصائية مختلفة في مجموعات مختلفة ظاهرة للعيان بعد معادلة المجموعات على مقياس للقدرة، عندها يظهر التحيز كفرق بين احتمالات شرطية لاستجابة واحدة في مجموعتين مختلفتين باختلاف طريقة تقديرها (Angel,1993:351).لذلك فإن الفرق بين المعلمة المقدرة (estimated) والمعلمة المولدة (الحقيقية) والمعالم المولدة مؤشر على تحيز التقدير، والتي يمكن استخلاصها باستخدام المعادلة الآتية:

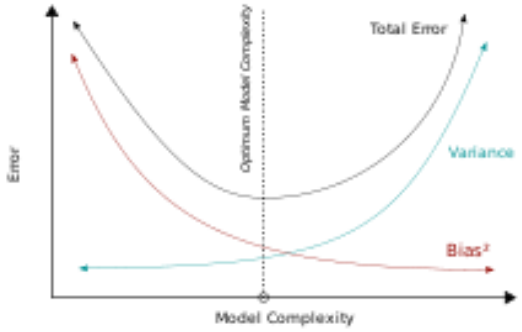
$$BIAS_x = \frac{\sum_i^n = 1 \left(\hat{X}_{estimated} _ X_{true}\right)}{n}$$

إذ إن:

$X_{estimatal}$ القيمة المقدرة للمعلمة

X_{true} القيمة الحقيقية للمعلمة

nعدد المعالم المقدرة.



شكل (١) يوضح مخطط التحيز

وفي هذا الصدد يوصي "كاميلي"، و "شبارد" (Camilli & Shepard, 1994) باستبعاد الفقرات التي يظهر فيها اختلاف أكبر لكونها متحيزة عن تقدير المتغير الشرطي، وعندما يعاد تحليلها باستخدام عملية تكرار الإعادة هذه (iterative process)، تتم "تنقية " المتغير الشرطي. والافتراض في هذا الإجراء أن استعمال المتغير الشرطي "النقي" سيكشف الفقرات التي تبدي أداء تفاضليا ذا دلالة، وفي الوقت نفسه يساعد في الكشف عن فقرات أخرى فيها أداء متباين كانت محجوبة بالمتغير الشرطي الملوث.

ويؤكد "هولاند وثاير (Holland and Thayer 1988) على أن الفقرات المتحيزة تقلص من قوة الاختبار، ومن منظور نظرية الاستجابة للفقرة تعني أن تقديرات معالم تلك الفقرة لا توصف باللا تغاير في المجموعات الظاهرة أي أنها غير مطابقة لبيانات الفقرة. وهذا الافتقار إلى "اللا تغاير" الذي يفسر كدليل على وجود تحيز في قيمها(Fu, 2010, p. 46). كما أن المؤشرات التي تستند إليها نظرية الاستجابة للفقرة للكشف عن الأداء المتحيز وعن الافتقار إلى اللا تغاير، يعتمد على طريقة نسبة الأرجحية (TSW-AG) وطريقة مربع كاي وطريقة (Exact Signed Area & Statistic)، والنموذج الخطي اللوجستي، والانحدار اللوجستي. وطريقة الإحصائي اللا معلمي مربع كاي لـ مانتل – هانسزل (Mantel-Haenszel Chi-Square,MH). (Raju, 1990)(Thissen, & Wainer, 1993).

- **تقدير معالم الفقرات:**

تعد تقدير وتحليل معالم الفقرات عند استخدام نظرية الاستجابة للمفردة، من الخطوات الهامة، والتي يطلق عليها مصطلح التعيير Calibration(علام، ٢٠٠٥، صفحة ٩١)، وفي الواقع العملي فإننا لا نعلم قيمة معالم الفقرة (Items Parametres Values) (الصعوبة Difficulty،التمييز Discrimination ،التخمين Guessing ولا معالم القدرة الأفراد,Values Ability Parametres of Individualsوليس بين أيدينا سوى استجابات الأفراد

على مفردات المقياس أو الاختبار ومن خلال هذه الاستجابات نقوم بتحديد قيمة القدرة لكل فرد اجاب على فقرات الاختبار، وبما أن نماذج الاستجابة للمفردة هي نماذج غير خطية Nonlinear تستخدم أساليبا وطرقا معينة لتقدير معالم نماذج الاستجابة للمفردة في هذا الجانب هناك اتجاهان، الأول يعتمد على المعلومات الظاهرة المدروسة فقط في تقدير معالم النماذج، أما الاتجاه الثاني فيعتمد على دمج معلومات مسبقة مع المعلومات الحالية للظاهرة قيد الدراسة والقياس، ليتكون لدينا معالم النموذج في شكل توزيع احتمالي Prior Distribution على وفق تقديرات نظرية بييز (Baye's Theorem) نسبة إلى العالم توماس بييز (Embretson & Reise, 2000, p. 21).

وما يهمنا في البحث الحالي هو الاتجاه الأول الذي يعتمد على طرق الأرجحية العظمى الذي سنتطرق لطرقه فيما بعد.

- **تقدير قدرة الأفراد:**

عند استخدام بعض طرق الأرجحية العظمى في تقدير معالم المفردات (صعوبة تمييز، تخمين مثل طريقة الأرجحية العظمى الهامشية(MML)فإننا نمر إلى تقدير قدرات الأفراد وهنا توجد عدة طرق من بينها طريقة الأرجحية العظمى ((ML) طريقة توقع التوزيع البعدي (EAP) وطريقة القيمة العظمى التوزيع البعدي (MAP)، إلا أن الباحث سيكتفي بتسليط الضوء على طرق الأرجحية العظمى، لكونها تقع في دائرة التحري للبحث الحالي.

منهجية البحث وإجراءاته:

سيلجأ الباحث للمنهج التجريبي، إذ سيعمد ومن خلال استخدام تصميم القياسات المتكررة (Repeated Measurements Desing) ذات المجموعات المرتبطة (غير المستقلة)،إلى تطبيق الاختبار على مجموعة من الأفراد، ثم سيخضع استجاباتهم لتحليلات متعددة، تختلف باختلاف النموذج اللوجستي المستخدم،(راش ولورد وبيرنبوم)، لتوليد بيانات تقدير المعالم أو باختلاف طريقة الأرجحية العظمى وهما طريقة الأرجحية المشتركة والهامشية والمعتمدتان لتقدير معالم الفقرات والأفراد الخاضعين للاختبار ثم حساب تحيز تلك التقديرات وفق مؤشر التحيز(BIAS)، ومعيار أكايكي (Akaike)للمعلومات، كمؤشرين للاستدلال وفقهما إلى تحيز التقديرات وأفضلية طريقة تقدير الأرجحية العظمى والنموذج اللوجستي لتقليل تحيز قيم معالم الفقرة وقدرة الأفراد.

مجتمع البحث:

يتكون مجتمع البحث من كل الطلبة الذكور فقط، من المستمرين بالدوام الصباحي، في الكليات والأقسام في تخصصات العلوم الإنسانية والطبيعية بجامعة صلاح الدين/ أربيل للعام الدراسي (2023 – 2024).وقد بلغ المجموع الكلي لمجتمع البحث(11802) طلاب، في جميع المراحل الدراسية.

عينة البحث:

تم اختيار عينة ممثلة لمجتمع البحث، للحصول على استجابات هؤلاء الأفراد على أداة البحث للتحقق من افتراضات نماذج نظرية الاستجابة للفقرة المستخدمة لأغراض البحث الحالي، ووفق طريقتي تقدير الأرجحية العظمى المعتمدة، حيث تألفت عينة البحث من طلبة الجامعة صلاح الدين/ أربيل،فقد تم اختيار (1072) طالبا وفق أسلوب المعاينة الطبقية العشوائية، وبأسلوب الاختيار المتساوي، حيث حرص الباحث على أن يتم الاستعانة بالذكور فقط، وذلك لعدم إتاحة الفرصة للفروق الديموغرافية الراجعة للجنس أو ما شابهها، للتأثير على نتائج البحث، إذ شكلوا نسبة (9%) من حجم مجتمع البحث، وتم إجراء التحليل الإحصائي عليهم كما سيتم توضيحه فيما بعد.

أداة البحث:

وقع اختيار الباحث على استخدام اختبار القدرة المنطقية لدى الراشدين أو ما يسمى بالذكاء المجرد لـ روبرت ثورندايك والقائمة على القدرة لمعالجة الألفاظ والرموز وكيفية التعامل معها، حيث يقيس الاختبار قدرة الراشدين على التفكير المنطقي المنهجي والقدرة على الاستنتاج، والاختبار مستقل ومتحرر من العوامل الثقافية والاجتماعية للمستجيبين (المفحوصين)، لكون فقراته قائمة على التحليل المنطقي والتي لا تختلف باختلاف المجتمع الذي تطبق فيه، وتضم هذه الأداة القائمة على أقصى الأداء (٢٥) فقرة تركز على قياس القدرة الرياضية المنطقية، يتم تطبيقها بنحو فردي أو جماعي، أما معايير تصحيح فقراته فتعطى درجة واحدة لكل جواب صحيح، فيما يعطى المستجيب صفر عن كل إجابة خاطئة.

صدق الظاهري:

أشار ابيل (Ebel) إلى أن الصدق الظاهري وسيلة جيدة للتأكد من صلاحية الفقرات من خلال قيام عدد من الخبراء المختصين بتقرير صلاحيتها لقياس الصفة التي وضعت من أجلها (Robert & Ebel, 1972, p.

(140. لذلك اعتمد الباحث أسلوب الصدق الظاهري من خلال عرض الاختبار على (١٩) خبيرًا ومحكمًا من الخبراء والمتخصصين في العلوم النفسية والقياس النفسي، لإصدار أحكامهم على مدى صلاحية الفقرات، وصلاحية البدائل المعتمدة في الاستجابة على كل فقرة، وللتعرف على مدى اتفاق السادة الخبراء والمحكمين عن ذلك، تم حساب هذا النمط من الصدق باستخدام طريقتين هما النسبة المئوية واختبار مربع كاي لحسن المطابقة، حيث اعتمد الباحث نسبة اتفاق (89.4 %) فأكثر بين المحكمين وقيمة مربع كاي تبلغ (٣.٨٤) فأكثر للإبقاء على الفقرة، وعلى ضوء استجابة السادة الخبراء والمحكمين لم تحذف أي فقرة من فقرات الاختبار. والجدول (١) يوضح نتائج أحكام السادة الخبراء والمحكمين على فقرات المقياس:

جدول (١) يوضح نتائج الصدق الظاهري لأداة البحث

تسلسل الفقرات	الموافقون		غير الموافقين		قمة مربع كاي	
	العدد	النسبة المئوية	العدد	النسبة المئوية	المحسوبة	الجدولية
١، ٢، ٣، ٥، ٧، ٨، ٩، ١١، ١٢، ١٣، ١٤، ١٥، ١٦، ١٧، ١٩، ٢٠، ٢٢، ٢٣، ٢٥	١٩	١٠٠٪	٠	٠٪	١٩	٣.٨٤
٤، ٦، ١٨، ٢١، ٢٤	١٨	%94.7	١	%5.3	15.21	
١٠	١٧	%89.4	٢	%10.6	11.84	
الفقرات المستبعدة	**لاتوجد أية فقرة مستبعدة من جراء حساب الصدق الظاهري**					

التحقق من افتراضات النماذج الثلاثة:

تقوم نظرية الاستجابة للفقرة افتراض أحادية البعد، والاستقلال المحلي للفقرات، والمطابقة لمنحنى خصائص الفقرة، والتحرر من السرعة. والتحقق من هذه الافتراضات أمر ضروري قبل استخدام أي النموذج لوجستي في تحليل الإحصائي لمعالم الفقرة والقدرة. وقد تم ذلك وكما يأتي:

أ- **التحقق من افتراض أحادية البعد:**

تم التحقق من هذا الافتراض في البحث الحالي على مؤشر التحليل العاملي كأسلوب إحصائي يعتمد على ارتباطات ودرجة تشبع الفقرة بالعامل العام، ولذلك تم استخلاص المحكات والمؤشرات الإحصائية الدالة على أحادية البعد لاختبارالقدرة المنطقية لدى الراشدين، وفق كل نموذج من النماذج الثلاث المعتمد وهي: (راش، لورد، بيرنبوم)، على حدة،إذ قام الباحث باستخدام جميع الإجابات لعينة البحث والبالغة(1072)إجابة لغرض إدخالها في التحليل العاملي الاستكشافي بطريقة المكونات الأساسية(Principle Component) لهوتيللنج (Hottelling, 1933)،وباستخدام الحقيبة الإحصائية للعلوم الاجتماعية(SPSS). فقد اعتمد الباحث على محك كايزر، إذ إن هذا المحك يتطلب مراجعة الجذر الكامن للعوامل الناتجة وعلى أن تقبل العوامل التي يزيد جذرها الكامن (Eigenvalue)عن الواحد الصحيح وتعد عوامل عامة. ويبدو هذا الأسلوب صالحا ومناسبا على وجه الخصوص لطريقة المكونات الأساسية(Hambleton, Swaminathan, & Rogers, 1991, p. 9).

وبعد إجراء التحليل العاملي الاستكشافي، وباستخدام الحقيبة الإحصائية (SPSS) لكل نموذج على حدة أفرز التحليل بعد تدوير العامل على محاور متعامدة بطريقة الفاريمكس (Varimax) لكايزر Kaiser، بهدف الحصول على عامل عام واحد لكل نموذج منها، يكون الجذر الكامن أكبر من (1) وتفسرون جميع العوامل معا نسبة (100%) من التباين.ففي هذا الصدد يشير ريكيس (Reckase, 1979) إلى أن تكون النسبة على الأقل (20%) لكي يفسر المكون الأول من نسبة التباين كدليل لأحادية البعد (تيغزة، ٢٠١٢، صفحة ٨٩). ويتفق هامبلتون (Hambleton, 2004) مع ريكيس في اعتماد هذه النسبة، ثم يمكن اعتباره عاملا سائدا. وهذا دليل إحصائي على أحادية البعد حيث كانت النتائج كما يلي:

١. أحادية البعد لنموذج راش Rasch Model:

كانت قيم الجذر الكامن للعامل الأول(21.56) ويفسر(21.838%) من التباين. فيما فسر العاملين الآخرين معا نسبة(78.162%) من التباين. والجدول رقم (٢) يوضح تلك النتائج:

جدول رقم (٢)يوضح قيم الجذر الكامن ونسبة التباين المفسر ونسبة التباين المفسر التراكمي لنموذج راش

العامل	الجذر الكامن	نسبة التباين المفسر	نسبة التباين المفسر التراكمي
1	21.560	21.838	**21.838**
2	28.974	29.083	**50.921**
3	48.971	49.079	**100.000**

٢. أحادية البعد لنموذج لورد Lord Model:

كانت قيم الجذر الكامن لعامل الأول(26.001) ويفسر(28.153%) من التباين. فيما فسر العاملين الآخرين معًا نسبة(71.847%) من التباين. والجدول رقم (٣) يوضح تلك النتائج:

جدول رقم (٣) يوضح قيم الجذر الكامن ونسبة التباين المفسر ونسبة التباين المفسر التراكمي لنموذج لورد

العامل	الجذر الكامن	نسبة التباين المفسر	نسبة التباين المفسر التراكمي
1	26.001	28.153	**28.153**
2	31.347	31.992	**60.145**
3	39.520	39.855	**100.000**

٣. أحادية البعد لنموذج بيرنبوم Birnbum Model:

كانت قيم الجذر الكامن للعامل الأول(30.692) ويفسر(30.918%) من التباين. فيما فسر العاملين الآخرين معا نسبة(69.082%) من التباين. والجدول رقم (٤) يوضح تلك النتائج:

جدول رقم (٤) يوضح قيم الجذر الكامن ونسبة التباين المفسر ونسبة التباين المفسر التراكمي لنموذج بيرنبوم

العامل	الجذر الكامن	نسبة التباين المفسر	نسبة التباين المفسر التراكمي
1	30.692	30.918	**30.918**
2	33.512	33.770	**64.688**
3	35.026	35.312	**100.000**

ويتبين من الجداول أعلاه أن جميع النماذج قد حققت افتراض أحادية البعد. وهو ما يؤشر إلى ملاءمة النماذج الثلاثة (راش، لورد، بيرنبوم) لنظرية الاستجابة للفقرة، مع البيانات المولدة من العينة الخاضعة لأداة البحث الحالي.

ب. التحقق من افتراض الاستقلال الموضعي Local Independence:

ويشير هذا الافتراض إلى طبيعة المنحنى أو الدالة المميزة لكل فقرة التي تصف العلاقة بين القدرة والأداء على الفقرة. ويعتمد شكل المنحنى المميز للفقرة على معالم الفقرة من صعوبة (β)، وتمييز (a)، وتخمين (c). وقدرة الأفراد (θ)، وهذا الافتراض يكافئ افتراض أحادية البعد والذي تمت الإشارة إليه آنفًا، ولذلك اكتفى الباحث بالتحقق من أحادية البعد في البيانات لكل نموذج لوجستي، فعندما يتم تحقيق فرض أحادية البعد فإن افتراض الاستقلال الموضعي فإن ذلك يشير إلى أن الاختبار يقيس قدرة عامة واحدة، فمن الملاحظ أن افتراض أحادية البعد

وافتراض الاستقلال الموضعي متلازمان، بمعنى أنه إذا تحقق الافتراض الأول فإن الافتراض الثاني يتحقق أيضًا، ومنطقيا سيكون هذان المفهومان متكافئين(Lord, 1980, p. 402).

د- التحرر من السرعة Speediness:

اهتم الباحث بهذا الافتراض، حيث تمت مراعاة الزمن اللازم لإجابة الاختبار، مع ترك فسحة إضافية من الوقت حتى لا يترك المستجيب إكمال الإجابة عن بعض الفقرات بداعي عدم كفاية الوقت، حيث بلغ الوقت الإضافي (10) دقائق، يتم إتاحتها أمامهم بالإضافة لـ (35) دقيقة المخصصة للإجابة ليكون مجمل الزمن المتاح أمام المستجيب (45) دقيقة.

حيث تم تخصيص الوقت اللازم للإجابة بناء على معطيات استخدام معادلة حساب الوقت اللازم للإجابة، لذلك فإن إخفاق أي مستجيب للاختبار يرجع إلى ضعف القدرة وليس لعدم كفاية الوقت أو لعدم تحرر الاختبار من عامل السرعة.وبذلك يكون الباحث قد تحقق من افتراضات نظرية الاستجابة للفقرة، وهو ما جعلهما يتأكدا من أن البيانات المستمدة من أداة البحث مناسبة للتحليل وفقا لنظرية الاستجابة للفقرة.

الوسائل الإحصائية:

تحقيقا لأغراض البحث استخدم الباحث الوسائل الإحصائية الآتية:

- اختبار مربع كاي لحساب الصدق الظاهري لأداة البحث.
- التحليل العاملي الاستكشافي للتعرف على مدى تحقق افتراضات نظرية الاستجابة للفقرة في نماذج المعتمدة بالبحث الحالي.
- المتوسط الحسابي للتعرف على متوسط تقديرات بارامترات الفقرة وقدرة الأفراد وفق الطريقة التقدير والنموذج اللوجستي.
- الانحراف المعياري للتعرف على انحراف تقديرات بارامترات الفقرة وقدرة الأفراد عن متوسطاتها وفق الطريقة التقدير والنموذج اللوجستي.
- معادلة التحييز للتعرف على تقدير بارامترات الفقرة وقدرة الأفراد وفق الطريقة التقدير والنموذج اللوجستي.
- اختبار تحليل التباين الثنائي بتفاعل للتعرف على الفروق في طريقة تقدير الأرجحية العظمى والنموذج اللوجستي

على تحيز قيم تقدير قدرة الأفراد بالاستناد على مؤشر التحيز.

نتائج البحث:

تحقيقا لأهداف البحث فقد قام الباحث بحساب الوسائل الإحصائية التي تتلاءم مع كل هدف من أهداف البحث، وتمخض عن ذلك النتائج الآتية:

١. **تقدير معالم الفقرة وفق نماذج اللوجستية تبعا لطريقة الأرجحية العظمى المشتركة والهامشية:**بهدف تقدير معالم الفقرة وفق نماذج اللوجستية الثلاث (راش، لورد، بيرنبوم) تبعا لطريقة الأرجحية العظمى المشتركة والهامشية، فقد تم استخدام برنامج(Bilog–MG3)لتحديد تلك الدوال. والجدول (٦) يوضح ذلك:

جدول رقم (٦) يوضح تقدير معالم الفقرة وفق نماذج اللوجستية تبعا لطريقة الأرجحية العظمى المشتركة والهامشية

الفقرات	تقدير معالم الفقرات وفق النموذج اللوجستي تبعاً لطريقة الأرجحية العظمى المشتركة						تقدير معالم الفقرات وفق النموذج اللوجستي تبعاً لطريقة الأرجحية العظمى الهامشية					
	راش		لورد		بيرنبوم		راش		لورد		بيرنبوم	
	الصعوبة	التمييز	الصعوبة	التمييز	الصعوبة	التمييز	التخمين	الصعوبة	التمييز	الصعوبة	التمييز	التخمين
١	-.227	.189	–.283	.223	1.059	.062	-.314	-.319	1.001	-.307	1.216	.114
٢	.241	.187	.224	.204	.502	.078	.263	.231	.604	.231	.593	.821
٣	.220	.207	.257	.251	.628	.116	.279	.266	.611	.255	.773	.131
٤	-.240	.224	–.210	.238	.764	.050	-.281	-.223	.733	-.217	.943	.072
٥	.280	.172	.205	.216	1.422	.085	.328	.291	1.183	.280	1.500	.911
٦	.267	.183	.203	.242	.454	.152	.305	.298	.408	.284	.462	.163
٧	-.250	.212	–.227	.211	.462	.216	-.265	-.247	.434	-.228	.488	.225
٨	.221	.211	–.235	.203	.784	.099	.247	-.240	.720	.233	.957	.128
٩	-.211	.220	.143	.223	1.002	.048	-.241	.187	.925	.180	1.115	.052

الفقرات	تقدير معالم الفقرات وفق النموذج اللوجستي تبعاً لطريقة الأرجحية العظمى المشتركة						تقدير معالم الفقرات وفق النموذج اللوجستي تبعاً لطريقة الأرجحية العظمى الهامشية					
	راش	لورد		بيرنبوم			راش	لورد		بيرنبوم		
	الصعوبة	الصعوبة	التمييز	الصعوبة	التمييز	التخمين	الصعوبة	الصعوبة	التمييز	الصعوبة	التمييز	التخمين
١٠	.236	.248	.206	.235	1.209	.173	.250	.262	1.113	.226	1.429	.183
١١	-.227	.272	.210	.202	.316	.196	-.257	.288	.216	.200	.621	.212
١٢	.230	-.292	.241	.229	.814	.127	.261	-.337	.754	.226	1.022	.147
١٣	-.270	.193	.230	.235	1.583	.130	-.312	-.220	1.248	.230	1.559	.166
١٤	.222	.258	.221	.202	.395	.065	.238	.263	.527	.216	.533	.088
١٥	.240	-.187	.201	.234	1.565	.077	.252	-.196	1.708	.228	1.602	.093
١٦	-.225	.278	.237	.229	.845	.236	-.239	.302	.939	.211	1.086	.251
١٧	.203	.257	.163	.211	1.014	.158	.209	.275	1.266	.204	1.231	.167
١٨	-.210	.216	.166	.203	.857	.223	-.223	.234	1.172	-.205	1.035	.233
١٩	.230	-.252	.215	.262	.274	.082	.237	-.258	1.223	.220	.319	.099
٢٠	.238	.218	.188	.272	.348	.093	.255	.222	.496	.227	.372	.121
٢١	-.262	.255	.175	.230	.963	.181	-.271	.261	1.186	.223	1.238	.206
٢٢	.236	.194	.209	.222	.554	.112	.242	.246	.911	.217	.621	.152
٢٣	.240	.237	.244	.266	.735	.147	.258	.253	.958	.253	.883	.173
٢٤	.245	-.224	.226	.248	.177	.043	.240	-.266	.429	.225	.330	.074
٢٥	-.251	.297	.193	.241	.377	.090	-.272	.311	.701	.236	.442	.118

٢. المفاضلة في تقدير بارامترات الفقرة وفقا لطريقة التقدير والنموذج اللوجستي باستخدام مؤشر التحيز:

اعتمد الباحث على تقديرات كل نموذج من النماذج الثلاث، وفق كل من طريقتي تقدير الأرجحية العظمى الهامشية والمشتركة، حيث تم حساب مؤشر التحيز(BIAS) كأحد المؤشرات الإحصائية التي تشير إلى دقة التقديرات بين التقديرات المقدرة والحقيقية، كما وتم حساب المتوسطات الحسابية لمؤشر التحيز وانحرافاتها المعيارية. وتمخض عن ذلك المعادلة القيم الواردة في الجدول (٧) مايلي:

جدول (٧) يوضح المتوسطات الحسابية والانحرافات المعيارية لقيم مؤشر التحيز لتقدير بارامترات الفقرة وفقا لطريقة التقدير والنموذج اللوجستي

الانحراف المعياري	المتوسط الحسابي	المعالم	النموذج اللوجستي	الأرجحية العظمى
.065	.248	الصعوبة	راش	الأرجحية العظمى المشتركة
.036	.234	الصعوبة	لورد	
.023	.205	التمييز		
.020	.229	الصعوبة		
.015	.243	التمييز		
.017	.221	التخمين	بيرنبوم	
.015	.230	المجموع		
.069	.261	الصعوبة	راش	الأرجحية العظمى الهامشية
.037	.259	الصعوبة	لورد	
.035	.258	التمييز		
.027	.220	الصعوبة		
.041	.264	التمييز		
.020	.234	التخمين	بيرنبوم	
.017	.246	المجموع		

ومن القيم أعلاه يتبين لنا أن تقدير بارامترات الفقرة (Estimation the Item Parametres)، وفق طريقة الأرجحية العظمى المشتركة أقل انحرافا عن متوسط المعالم، وهو ما يؤشر إلى أن هذه الطريقة أقل تحيزًا من

تقدير البارامترات وفق طريقة الأرجحية العظمى الهامشية.وهو ما يؤشر إلى أن طريقة الأرجحية المشتركة أكثر دقة في تقدير بارامترات الصعوبة والتمييز والتخمين وفق نموذج بيرنبوم ثم يليه نموذج لورد ثم نموذج راش مقارنة بطريقة الأرجحية الهامشية.

٣. المفاضلة في تقدير قدرة الأفراد وفقا لطريقة التقدير والنموذج اللوجستي باستخدام مؤشر التحيز:

ولتحقيق هذا الهدف اعتمد الباحث على تقدير قدرة الأفراد لكل نموذج من النماذج الثلاث، وفق كل من طريقتي تقدير الأرجحية العظمى الهامشية والمشتركة، إذ تم حساب مؤشر التحيز(BIAS)، وكذلك حساب المتوسطات الحسابية لمؤشر التحيز وانحرافاتها المعيارية،

كما قام الباحث بحساب الفروق في القدرة في ضوء كل من طريقتي تقدير الأرجحية العظمى والنموذج اللوجستي على تحيز قدرة الأفراد من خلال اختبار تحليل التباين الثنائي بتفاعل (Two Way Interaction ANOVA). وتمخض عن ذلك النتائج الموضحة في الجداول (٨) و (٩):

جدول (٨)يوضح قيم المتوسطات الحسابية والخطأ المعياري لمؤشر التحيز لتقديرقدرة الأفراد وفقا لطريقة التقدير والنموذج اللوجستي

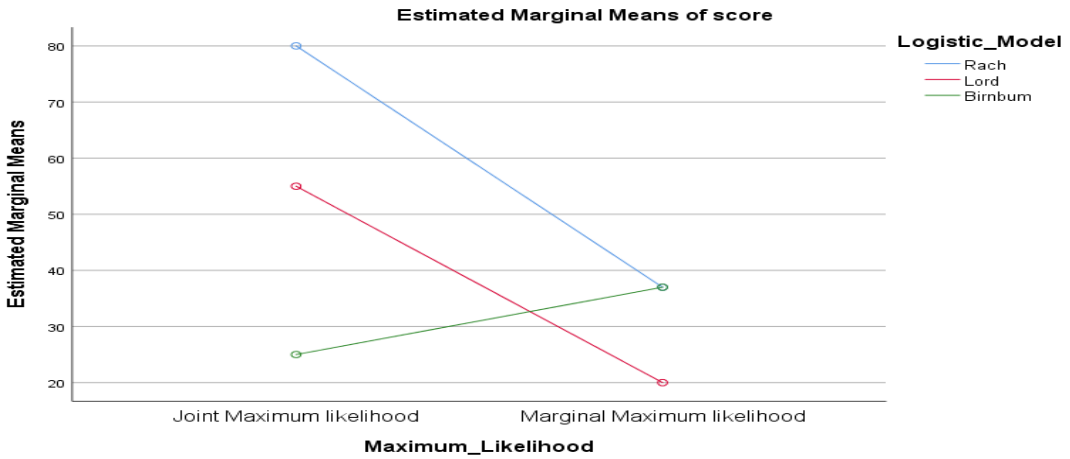
الأرجحية العظمى		النموذج اللوجستي	القيمة الدنيا	القيمة القصوى	المتوسط الحسابي	الانحراف المعياري
الأرجحية العظمى المشتركة		راش	-.17	.27	.220	3.100
		لورد	-.15	.24	.200	2. 701
		بيرنبوم	-.14	.22	.170	2.280
		المجموع	-.46	.73	.196	8.081
الأرجحية العظمى الهامشية		راش	-.16	.25	.190	2.643
		لورد	-.13	.23	.188	1.899
		بيرنبوم	-.11	.20	.160	1.810
		المجموع	-.40	.68	.176	6.352

ويلاحظ من القيم الواردة في الجدول أعلاه أن قيم المتوسط الحسابي والتي تمثل قيمة تقدير القدرة وفق مؤشر التحيز لطريقتي الأرجحية العظمى المشتركة والهامشية تبعا لنموذج بيرنبوم كانت أقل من قيم الوسط الحسابي

لنموذج لورد وكذلك بالنسبة لنموذج راش، وهو ما يؤشر أفضلية نموذج بيرنبوم على النموذجين الآخرين لتقليل تحيز تقدير قدرة الأفراد يليه نموذج لورد ثم نموذج راش، كما أوضحت النتائج أفضلية طريقة الأرجحية العظمى المشتركة على طريقة الأرجحية العظمى الهامشية في خفض مؤشر التحيز ودقة تقدير بارامتر قدرة الأفراد وفق البيانات التي تم الحصول عليها من الأفراد لقياس القدرة المنطقية المجردة.

جدول (٩) يوضح نتائج تحليل التباين الثنائي بتفاعل للفروق بين طريقة تقدير الأرجحية العظمى والنموذج اللوجستي على تحيز قيم تقدير قدرة الأفراد بالاستناد على مؤشر التحيز(BIAS)

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	القيمة الفائية	الدلالة الإحصائية
النموذج المصحح	3161.667[a]	5	632.333	3.293	.090
التفاعل	17595.273	1	17595.273	91.642	.000
الاحتمالية القصوى	1188.000	1	1188.000	6.188	.047
النموذج اللوجستي	1404.000	2	702.000	3.656	.092
أقصى احتمالية والنموذج اللوجستي	1857.333	2	928.667	4.837	.056
درجة الخطأ	1152.000	6	192.000		
المجموع	22722.000	12			
المجموع المصحح	4313.667	11			



شكل (٢) يوضح تقدير قدرة الأفراد وفقا لطريقة التقدير الأرجحية العظمى والنموذج اللوجستي بالاستناد على مؤشر التحيز(BIAS)

٤. المفاضلة في تقدير قدرة الأفراد وفقا لطريقة التقدير والنموذج اللوجستي وفق معيار أكايكي للمعلومات:

إلى جانب مؤشر التحيز(BIAS) والذي تم عرضه آنفًا، فقد اعتمد الباحث أيضًا على معيار أكايكي للمعلومات(Akaike Information Criterion)، كمؤشر أو معيار إضافي آخر يتم الاستناد إليه كمؤشر الإحصائي يتم الاستدلال وفقه إلى دقة التقديرات المقدرة والحقيقية، وذلك بغية التأكد أكثر من أفضلية النموذج اللوجستي في تقليل درجة تحيز تقدير بارامتر قدرة الأفراد، وعليه فقد لجأ الباحث إلى هذا المعيار الذي يوصف بأنه ذات موثوقية في تحديد النموذج الأمثل في تقدير القدرة لدى المفحوصين، حيث يعمل على ترجيح كفة النماذج التي تحقق توافق أحسن بين درجة تعقيد النموذج و قوته لوصف القدرة، كما تعد تقديرًا نسبيًا لكمية المعلومات الضائعة أو المحجوبة عند نمذجة البيانات وفق نموذج محدد(Akaike, 1970, pp. 203-217).

وتأسيسا على ذلك وبهدف المقارنة بين النماذج اللوجستية الثلاث فقد استخدم الباحث معيار أكايكي(AIC)، والجدول (١٠) يوضح نتائج قيم أفضلية النماذج اللوجستية لتقليل تحيز تقدير قدرة الافراد:

جدول (١٠) يوضح قيم التحيز للنماذج اللوجستية وفق معيار أكايكي للمعلومات

النموذج اللوجستي	قيم التحيز وفق معيار أكايكي للمعلومات	ترتيب التحيز
راش	704.208	1
لورد	581.334	2
بيرنبوم	506.791	3

ويتبين منالجدول أعلاه أن النتائج كانت مطابقة للنتائج التي تم الحصول عليها وفق مؤشر التحيز في تقدير القدرة لدى الأفراد المفحوصين، حيث تظهر لنا قيم معيار أكايكي للمعلومات أن نموذج بيرنبوم هو أقل النماذج اللوجستية تحيزًا وأكثرها دقة في تقدير قدرة الأفراد، ثم يليها نموذجي لورد وراش على الترتيب.

و يفسر الباحث تلك النتائج إلى انخفاض مؤشر التحيز في التقدير لمعلم القدرة وفق نموذج بيرنبوم مقارنة بالنموذجين الآخرين لورد وراش، يرجع لطبيعة النموذج الثلاثي الذي يستند على معالم الصعوبة، التمييز، التخمين. إذ إن أخذه بالحسبان عدد أكثر من المعالم سواء في النموذج الثنائي أو الأحادي مما يتيح المجال بشكل

أوسع لانتقاء وتدريج الفقرات والأفراد الأمر الذي يضيق المجال أمام احتمالية الحصول على درجة الخطأ، وهو ما ينعكس وبصورة إيجابية على تقليل نسبة التحيز في تقدير معالم القدرة والفقرة.

وبناء على نتائج البحث الحالي يوصي الباحث باستخدام نموذج بيرنبوم وتبعا لطريقة الأرجحية العظمى المشتركة في تقدير معالم الفقرة وقدرة الأفراد، وذلك لتحاشي تحيز التقديرات والحصول على نتائج أكثر دقة في التقدير.

كما ويقترح على الباحثين إجراء دراسات مماثلة للبحث الحالي للتعرف على أفضلية النماذج اللوجستية المتدرجة كنموذج سلالم التقدير، والتقدير الجزئي.. ألخ. في الحصول على تقديرات أقل تحيزا في تقدير معالم الفقرة وقدرة الأفراد.

المصادر:

Baker, F., & Kim, S. (2004). *Item response theory: Parameter estimation techniques.* CRC Press.

Brian, B., & Diane, R. (1997). Educational assessment of mathematics and abilities. *Journal of Learning Disabilities*.

De Ayala, R. (2009). *The theory and practice of item response theory.* Guilford Publications.

Elliott, C. (1983). *British ability scales manual1:Introductory Handbook.* Windsor, England : National foundation for educational Research.

Engelhard, J., & Wind, S. (2018). *(Invariant measurement with raters and rating scales) Rasch models for rater-mediated assessments.* Routledge.

Fu, Q. 3. (2010). *Comparing 59TA Accuracy59T of 59TParameter Estimation59T Using IRT Models in the Presence of Guessing.* USA: Illinois University.

Hambleton, R., & Jones, L. (1993). Latent trail Models and Their use in the Analysis of Educational teat data. *Journal of Educational Measurement*.

Hambleton, R., & Swaminathan, H. (1985). *Item Response Theory : Principles and Application.* Kluwer, Boston: Nijhoff Publishing.

Hambleton, R., Swaminathan, H., & Rogers, H. (1991). *Fundamentals of Item Response Theory.* Newbury Park CA: Sage Publications.

Hirotugu Akaike .(١٩٧٠) .Statistical Predictor Identification .*Ann. Inst. Statist*.

Linden, W., & Hambleton, R. (1997). *Item response theory, brief history , common models and extension.* New york: sprmger.

Lord, F. M. (1980). *Application of item response theory to practical testing problems.* Hillsadle, New Jersey: Lawrence Erlbaum Associates. Psychometrica.

Luc, L., & Adams, R. (2013). *Evaluate Rasch item parameter recovery in MML and JML estimations by ACER ConQuest.* Camberwell, Australia: University of Melbourne Australian Council for Educational Research.

Robert, L., & Ebel. (1972). *Essential of Education measurement* (Vol. 2). New Jersey.

S.E Embretson و ،S.P Reise .(٢٠٠٠) .*Item response theory for psychologists. Maheah*.New Jersey: Lawrence Erlbaum Associates, Publishers.

Salubayba, T. (2013). Differential Item Functioning Detection in Reading Comprehension Test Using Mantel-Haenszel, Item Response Theory, and Logical Data Analysis. *The International Journal of Social Sciences*(1).

أمحمد بوزيان تيغزة. (٢٠١٢). *التحليل العاملي الاستكشافي والتوكيدي مفاهيمها ومنهجها بتوظيف SPSS وليزرل LISREL.* عمان: دار مسافة للنشر والتوزيع والطباعة حزمة.

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

محمد جرادات. (٢٠١٢). *http://www.asfbs.com/fb/archive/index.bhb/t-16354.html*. (بناء الاختيار الحديث الخيار، المنتج) تم الاسترداد من موقع أكاديمي لعلم النفس.

هبة ضعضع، منار طومان، و مصطفة طيفور. (٢٠٢٠). أثر حجم العينة وطرائق التقدير في دقة تقدير معالم نموذج راش. *مجلة جرش للبحوث والدراسات*(١).

هبة عبداللطيف ضعضع. (٢٠٢١). أثر اختيار إنموذج الاستجابة للمفردة (PL,2PL,3PL,4PL١) وطرائق التقدير في معالم المفردة والأفراد ومعالمها. *مجلة كلية التربية*(الجزء الأول).