OPEN ACCESS
تاريخ الاستلام: 23/ 05/ 2025م
تاريخ القبـول: 29/ 06/ 2025م

مجلة الآداب
للدراسات النفسية والتربوية

تأثير عدد بارامترات النموذج وطريقة التقدير وحجم العينة وشكل توزيع القدرة على دقة تقدير قدرات الأفراد دراسة محاكاة

أ. شكري أحمد خُجا بخاري*
Dr.shokribukhari@gmail.com

د. أحمد علي حسن المعمري**
a.almaamari@qu.edu.sa

الملخص

هدفت الدراسة إلى التعرف على تأثير كل من النموذج (أحادي–ثنائي–ثلاثي)، وشكل توزيع القدرة (طبيعي–ملتوي موجب–ملتوي سالب)، وحجم العينة (100–500–1000)، وطريقة التقدير (ML–MP–EP)، بالإضافة إلى التفاعلات الثنائية والثلاثية والرباعية بينها على دقة تقدير قدرات الأفراد، واعتمد الباحثان على المنهج التجريبي وأسلوب المحاكاة لتوليد البيانات وتحليلها باستخدام برنامج (R) وحزمة [MIRT]، مع استخدام SPSS للتحليلات الإحصائية المكمّلة، وExcel لحفظ البيانات وفرزها، وتم تطبيق تحليل التباين العاملي Factorial Anova (3×3×3×3)، مع حساب حجم التأثير (مربع إيتا)، واستخدام اختبار Bonferroni للمقارنات البعدية (Post Hoc)، وأظهرت النتائج وجود تأثير للنموذج وحجم العينة وشكل توزيع القدرة وطريقة التقدير والتفاعلات الثنائية والثلاثية والرباعية بينهما على دقة تقدير قدرات الأفراد إلا أن هذا التأثير لم يرق إلى الدلالة العملية.

الكلمات المفتاحية: بارامترات النموذج، توزيع القدرة، تقدير القدرة، دراسة المحاكاة

* طالب الدكتوراة، قياس وتقويم تربوي، قسم علم النفس، كلية اللغات والعلوم الإنسانية، جامعة القصيم.

** أستاذ علم النفس التربوي المشارك، قسم علم النفس، كلية اللغات والعلوم الإنسانية، جامعة القصيم.

للاقتباس: بخاري، شكري أحمد خُجا؛ المعمري، أحمد علي حسن. (2025). تأثير عدد بارامترات النموذج وطريقة التقدير وحجم العينة وشكل توزيع القدرة على دقة تقدير قدرات الأفراد دراسة محاكاة ، *مجلة الآداب للدراسات النفسية والتربوية*، 7(3)، 113-149.

© نُشر هذا البحث وفقًا لشروط الرخصة Attribution 4.0 International (CC BY 4.0)، التي تسمح بنسخ البحث وتوزيعه ونقله بأي شكل من الأشكال، كما تسمح بتكييف البحث أو تحويله أو الإضافة إليه لأي غرض كان، بما في ذلك الأغراض التجارية، شريطة نسبة العمل إلى صاحبه مع بيان أي تعديلات أُجريت عليه.

OPEN ACCESS
Received: 23-05-2025
Accepted: 29-06-2025

مجلة الآداب
للدراسات النفسية والتربوية

Impact of Number of Model Parameters, Estimation Method, Sample Size and Ability Distribution Shape on Individuals' abilities Estimation Accuracy: A Simulation Study

Shokri Ahmed Khoga Bukhari* Dr. Ahmed Ali Hasan Al-Maamari**

Dr.shokribukhari@gmail.com a.almaamari@qu.edu.sa

Abstract:

This study aimed to identify the impact of the model (unidimensional–bidimensional–tridimensional), ability distribution shape (normal–positively skewed–negatively skewed), sample size (100–500–1000), and estimation method (ML–MP–EP), in addition to the two-way, three-way, and four-way interactions among them, on the accuracy of estimating individuals' abilities. The comparative descriptive approach was employed, using simulation methods to generate and analyze data using the R software and the [MIRT] package, along with SPSS for supplementary statistical analyses and Excel for storing, sorting, and organizing the data. A factorial ANOVA (3×3×3×3) was applied, along with the calculation of effect size (Eta squared), and the use of Bonferroni post hoc tests for multiple comparisons. The results showed that the model, sample size, ability distribution shape, estimation method, and their various interactions had a statistically significant effect on the accuracy of ability estimation. However, this effect did not reach practical significance.

Keywords: Model Parameters, Ability Distribution, Ability Estimation, Simulation Study.

* Ph.D. candidate in Educational Assessment and Measurement, Department of Psychology, College of Languages and Humanities, Qassim University

** Associate Professor of Educational Psychology, , Department of Psychology, College of Languages and Humanities, Qassim University.

Cite this article as:: Bukhari, Shokri Ahmed Khoga. & Al-Maamari, Ahmed Ali Hasan. (2024). Impact of Number of Model Parameters, Estimation Method, Sample Size and Ability Distribution Shape on Individuals' abilities Estimation Accuracy: A Simulation Study . *Journal of Arts for Psychological & Educational Studies* 7(2) 113-149

© This material is published under the license of Attribution 4.0 International (CC BY 4.0), which allows the user to copy and redistribute the material in any medium or format. It also allows adapting, transforming or adding to the material for any purpose, even commercially, as long as such modifications are highlighted and the material is credited to its author.

مقدمة

قبل ظهور نماذج الاستجابة للمفردة، اعتمدت النظرية الكلاسيكية على مؤشرات مثل معامل الثبات والصدق الكلي دون التمييز بين خصائص المفردات أو مستوى دقة تقدير معالم الأفراد، لكن مع مرور الوقت، ظهرت الحاجة إلى نماذج تركز على المفردة الواحدة بدلاً من النظر للاختبار ككل (Lord & Novick, 1968).

في أوائل الستينيات، بدأ ظهور أول نموذج في إطار نظرية الاستجابة للمفردة وهو نموذج راش (Rasch model)، الذي يفترض أن جميع المفردات لها القدرة التمييزية نفسها، مما يعني أن له معلمة واحدة فقط (الصعوبة) (Wright & Stone, 1979).

عند استخدام نموذج راش في تطبيقات ميدانية، لاحظ الباحثون أن المفردات تختلف في قدرتها على التمييز بين الأفراد، ما أدى إلى ظهور نموذج ثنائي المعلمة (2PL) ومن ثم ثلاثي المعلمة (3PL).

تم التمهيد لذلك من خلال دراسات بيرنباوم (1968) Birnbaum، الذي أوضح رياضيًا أن إدخال معلمات إضافية (مثل التمييز والتخمين) يؤدي إلى تحسين دقة التقدير، خاصة في الاختبارات ذات الطبيعة المتنوعة أو التي تحتوي على احتمالات عالية للتخمين.

لورد قام بتطبيق النماذج متعددة المعلمات في السياقات العملية، وأكد من خلال محاكاة إحصائية أن زيادة عدد معلمات المفردة (من معلمة واحدة إلى ثلاث معلمات) يؤدي إلى زيادة دقة واستقرار تقديرات معالم الأفراد، بشرط توافق النموذج مع البيانات وتوفر عينة كافية (Lord, 1980).

وبدأت النظرية بالتطبيق في مجالات متعددة مثل تصميم الاختبارات القياسية (SAT وGRE) وإنشاء بنوك المفردات واختبارات التكيف الحاسوبي (Computerized Adaptive Testing) CAT، مما أسهم في تقديم قياسات دقيقة تتسم بالاستقلال عن العينة والاختبار (Embretson & Reise, 2000).

وتُعد دراسة العوامل المؤثرة على دقة تقدير قدرات الأفراد والمفردات في نظرية الاستجابة للمفردة (IRT) موضوعًا رئيسيًا منذ منتصف القرن العشرين، إذ بدأت هذه الدراسات بالتزامن مع

تطوير النماذج الاحتمالية وتطبيقها على اختبارات القياس النفسي والتربوي، إذ أصبح من الضروري تقييم تأثير المتغيرات المختلفة مثل حجم العينة، وطرق التقدير، وشكل توزيع القدرة، وعدد البارامترات المستخدمة في النماذج.

ونجد أن بداية فكرة تأثير عدد معلمات المفردة على دقة التقدير كانت من ملاحظة محددات نموذج راش، الذي اعتمد على معلمة واحدة، ومع تعقيد التطبيقات التربوية والنفسية، ظهرت نماذج ذات معلمتين وثلاث معلمات، بفضل أعمال بيرنباوم Birnbaum ولورد Lord التي أثبتت عبر نماذج محاكاة وتحليل رياضي أن زيادة عدد المعلمات يؤدي إلى تقديرات أكثر دقة واستقرارًا لقدرات الأفراد، خاصة في حالة تنوع خصائص المفردات أو وجود احتمالات للتخمين (Birnbaum, 1968; Lord, 1980).

وكانت بداية دراسة أثر حجم العينة على دقة تقدير معالم الأفراد في عام 1904، إذ نشر تشارلز سبيرمان Charles Spearman دراسة رائدة بعنوان: General intelligence objectively determined and measured إذ قدم مفهوم العامل العام (g factor) من خلال تحليل التغاير بين اختبارات القدرات العقلية، ورغم أن سبيرمان Spearman لم يناقش حجم العينة بشكل مباشر، فإن استخدامه لأساليب إحصائية معقدة آنذاك استدعى الحاجة إلى عينات كافية لضمان دقة النتائج.

وفي عام 1910، ركز ثورندايك Thorndike على أهمية الاتساق في تقييم الأداء في دراسته التي تدور حول تقييم جودة الكتابة اليدوية وهي بعنوان: " The measurement of the quality of handwriting"، إذ أشار إلى أن التقييمات قد تختلف باختلاف العينات، مما يبرز أهمية حجم العينة في تحقيق تقديرات دقيقة (Thorndike,1910).

وفي عام 1925 ناقش رونالد فيشر Ronald Fisher في كتابه: " Statistical methods for research workers"، أهمية حجم العينة في تقليل الخطأ المعياري وزيادة دقة التقديرات، وأوضح أن التباين في التقديرات يتناقص مع زيادة حجم العينة، مما يعزز من موثوقية النتائج (Fisher, 1925).

وفي عام 1935 قدم ثيرستون Thurstone في كتابه "The vectors of the mind" منهجية تحليل العوامل لتفسير التباين في الاختبارات النفسية، إذ أشار إلى أن استقرار العوامل المستخرجة

يعتمد على حجم العينة، إذ إن العينات الصغيرة قد تؤدي إلى نتائج غير مستقرة Thurstone, (1935).

فيمكن القول بأن دراسة أثر حجم العينة على دقة تقدير معالم الأفراد بدأت في أوائل القرن العشرين، إذ كان سبيرمان (Spearman, 1904) وثورندايك (Thorndike, 1914) من أوائل الباحثين الذين تناولوا تأثير حجم العينة على دقة التقديرات في مجالات القياس، وقام فيشر (Fisher, 1922) بتقديم مفهوم المعلومات الإحصائية وأثر حجم العينة على دقة تقدير المعالم، موضحًا أن حجم العينة يزيد من دقة التقديرات، كما لعب ثورستون (Thurstone, 1931) دورًا كبيرًا في تطوير النماذج الإحصائية لفهم العلاقة بين حجم العينة ودقة التقديرات في مجالات متعددة مثل القياس النفسي والتربوي.

أما عن بدايات دراسة أثر شكل توزيع القدرة على دقة تقدير معالم الأفراد، فكان رونالد فيشر أحد أبرز العلماء الذين أسسوا للكثير من المبادئ الإحصائية التي نستخدمها اليوم، وخاصة في تقدير المعالم من خلال البيانات التجريبية، ففي نظرية فيشر تم تقديم المعلومات الإحصائية كمقياس لحساسية تقدير المعالم بالنسبة لتغيرات في التوزيع الاحتمالي للبيانات، عندما يكون التوزيع أكثر توافقًا مع الافتراضات (مثل التوزيع الطبيعي)، تكون المعلومات الإحصائية أكثر دقة، مما يؤدي إلى تقديرات أدق، وفي أبحاثه الشهيرة التي نشرت في عام 1922 تحت عنوان: " On the Mathematical Foundations of Theoretical Statistics"، الذي وضح فيه فيشر Fisher كيف أن دقة التقديرات تعتمد على مدى توافق التوزيع الفعلي للبيانات مع الافتراضات المعتمدة في النماذج الإحصائية، واقترح فيشر أن الشكل الطبيعي للتوزيع يُعد الأكثر فاعلية في تقدير المعالم بدقة، بينما التوزيعات الأخرى، مثل تلك التي تتسم بالانحراف أو الملتوية، يمكن أن تؤدي إلى تقليل دقة التقديرات بسبب انخفاض المعلومات التي تقدمها هذه التوزيعات (Fisher, 1922).

ومع تطور علم القياس النفسي في النصف الثاني من القرن العشرين، بدأ الباحثون في التوسع في تطبيقات فيشر على مختلف مجالات القياس، خاصة في اختبارات القدرات.

فرّق لورد (Lord, 1953) في دراساته المبكرة بين التوزيع الطبيعي المفترض وتوزيعات غير طبيعية لقدرات الأفراد، وبدأ يشير إلى أن توزيع القدرة غير الطبيعي (مثل التوزيع الملتوي أو ثنائي المنوال) يمكن أن يضعف دقة التقدير في اختبارات القدرات، وافترض أن التوزيع الطبيعي للقدرات

هو الأساس الذي يتم من خلاله تقدير معالم الأفراد، وقد لا يكون - دائمًا - صالحًا في البيئات التعليمية الحقيقية (Lord, 1953).

توسع لورد ونوفيك في كتابهما الشهير Statistical Theories of Mental Test Scores في الحديث عن أثر توزيع القدرة، وأكّدا أن النموذج الكلاسيكي يفترض توفر توزيع طبيعي، وأن هذا الافتراض قد لا يتحقق - دائماً - في العينات الحقيقية، وأشار بيرنباوم Birnbaum - ضمن إسهامه في الكتاب نفسه - إلى أن النماذج ثلاثية المعلمة تصبح أكثر حساسية لشكل التوزيع، خاصة إذا كان هناك انحراف أو تطرف في توزيع القدرة (Lord & Novick, 1968).

قام بيرنباوم Birnbaum (1968) بتوسيع هذا الفهم من خلال تطوير النماذج الحديثة لاستجابة الأفراد للمفردات (IRT)، وتمكّن من تقديم طُرق أكثر دقة لاحتساب المعلومات الإحصائية التي توفرها في اختبارات القدرات، مع مراعاة التأثيرات الناتجة عن الانحرافات في توزيع القدرات بين الأفراد (Birnbaum, 1968).

ومع تطور التقنيات الإحصائية، أصبح من الواضح أن شكل التوزيع لا يؤثر فقط على دقة تقديرات المعالم، بل يؤثر أيضًا على استقرار هذه التقديرات، خاصةً مع ظهور أساليب مثل التحليل متعدد المتغيرات (دراسة أكثر من متغير واحد في وقت واحد) لفهم العلاقات بين المتغيرات المتعددة، إذ أصبح من الضروري فهم العلاقة بين الشكل الاحتمالي للتوزيع ودقته في القياس.

هذه التطورات في دراسة أثر شكل التوزيع على تقدير المعالم لم تكن مستقلة عن الأفكار التي بدأها فيشر، بل كانت توسيعًا وتطويرًا لها في إطار النماذج الحديثة، إذ كانت المعلومات الإحصائية التي قدمها فيشر Fisher في عام (1922) هي الأساس لفهم تأثير الشكل التوزيعي على دقة التقدير، ووسّع لورد ونوفيك Lord & Novick (1968) هذه الفكرة لتشمل تطبيقات عملية في القياس النفسي واختبارات القدرة، وجاء بيرنباوم Birnbaum (1968) ليُظهر كيف يمكن أن تؤثر التوزيعات غير الطبيعية في الاستقرار والدقة في تقديرات معالم الأفراد.

ومن الدراسات الحديثة التي تناولت أثر هذه العوامل المختلفة على دقة التقدير:

دراسة شرفاوي وبن نابي (2022) التي هدفت إلى تقصي أثر طرق تقدير القدرة ومعالم المفردة على دقة التقدير باستخدام نظرية الاستجابة للمفردة، ولتحقيق أهداف الدراسة تم الاعتماد على أسلوب المحاكاة للحصول على بيانات مولدة ثنائية الاستجابة (0،1) اختبارين الأول (20) مفردة

والثاني (40) مفردة وفق النموذج الأحادي المعلم (نموذج راش)، إذ تم توليد استجابات الأفراد بواقع حجم عينة (250) فردًا ، 500 فرد، 1000 فرد، باستخدام برنامج (WinGen3) تبعًا للتوزيع الطبيعي للقدرة، التوزيع موجب الالتواء، توزيع بيتا = 4، والتوزيع سالب الالتواء توزيع بيتا = 2 وبالاعتماد على كل من: طرق تقدير القدرة، طريقة الأرجحية العظمى (ML)، طريقة توقع التوزيع البعدي (EAP)، وطريقة تعظيم الاقتران البعدي (MAP)، وطرق تقدير معالم المفردة، طريقة الأرجحية العظمى المشتركة (JML)، طريقة الأرجحية العظمى الهامشية (MML)، وطريقة الأرجحية العظمى الشرطية (CML)، اعتمدا على الخطأ المعياري للتقدير كمؤشر على دقة التقدير، وقد توصلت النتائج إلى أن طريقة الأرجحية العظمى الشرطية (CML) تقديرات أكثر دقة لتقدير المعلم صعوبة المفردة عند جميع مستويات حجم العينة وطول الاختبار في حالة شكل التوزيع الاعتدالي للقدرة، و أنتجت طريقة الأرجحية العظمى الهامشية (MML) تقديرات أكثر دقة لتقدير معلم صعوبة المفردة عند جميع مستويات حجم العينة وطول الاختبار في حالة شكل التوزيع الموجب والسالب الالتواء للقدرة.

أما دراسة Al-Tarawnah & Al- Qahtani (2022) التي هدفت إلى مقارنة تأثير طول الاختبار على درجة تقدير معلمة القدرة في النموذج اللوجستي ثنائي المعلمة وثلاثي المعلمات، باستخدام الطريقة البايزية للوضع المسبق المتوقع والاحتمال الأقصى، وتم اتباع المنهج التجريبي باستخدام طريقة مونت كارلو في المحاكاة، يتكون مجتمع الدراسة من جميع المواد ذات مستوى القدرة المحدد، وتشمل الدراسة عينات عشوائية من المواضيع والعناصر، أظهرت النتائج أن دقة تقدير معلمة القدرة في النموذج اللوجستي ثنائي المعلمة وفق طريقة الاحتمالية القصوى وطريقة بايزي تزداد مع زيادة عدد فقرات الاختبار، وتشير النتائج إلى أن زيادة معلمة القدرة في النموذج اللوجستي ثلاثي المعلمات تزداد مع زيادة عدد عناصر الاختبار، تتفوق طريقة بايزي فيما يتعلق بدقة التقدير في جميع ظروف حجم العينة، بينما في الاختبارات الطويلة تتفوق طريقة الاحتمالية القصوى في جميع الظروف.

كما هدفت دراسة الخرشة (2023) إلى تقصي أثر حجم العينة وعدد مفردات الاختبار على دقة تقدير معالم المفردات وقدرات الأفراد، في ضوء عدد نقاط التربيع في برنامج بايلوج (3) Bilog Mg، ولتحقيق هدف الدراسة تم توليد بيانات ثنائية التدريج من خلال النموذج اللوجستي ثلاثي المعلمة لثلاثة اختبارات طول كل منها (20 ، 30 ، 60) مفردة، وأربعة مستويات لحجم العينة (200،

400، 800، 1200) باستخدام برنامج (WINGEN)، وتم تحليل البيانات المولدة تبعاً لاختلاف عدد نقاط التربيع (15 ، 20 ، 25 ، 30) من خلال برمجية Bilog Mg3، أظهرت نتائج الدراسة وجود أثر ذي دلالة إحصائية لكل من طول الاختبار، وحجم العينة، والتفاعل بينهما في دقة تقديرات معالم المفردات (الصعوبة، التمييز، التخمين) ومعلمة القدرة للأفراد، وكانت تزداد دقة تقديرات معالم المفردات ومعلمة القدرة للأفراد بزيادة طول الاختبار وحجم العينة، كما أظهرت النتائج عدم وجود أثر ذي دلالة إحصائية لعدد نقاط التربيع في دقة تقديرات القدرة للأفراد، وبشكل عام تزداد دقة تقديرات القدرة للأفراد بزيادة عدد نقاط التربيع.

أما بحث العنزي (2023) فهدف إلى التعرف على أثر اختيار نموذج نظرية الاستجابة للمفردة، وطرق التقدير والتفاعل بينهما في دقة تقدير بارامترات مفردات اختبار كاتل للذكاء وقدرات الأفراد. تكونت عينة البحث من (1104) مشاركين من طلاب وطالبات جامعة الحدود الشمالية بالمملكة العربية السعودية، وبلغ المتوسط الحسابي لأعمارهم الزمنية (80.20) بانحراف معياري (1.31)، وتمثلت أداة البحث في اختبار كاتل للذكاء (المتحرر من أثر الثقافة) المقياس الثالث الصورة (أ)، واعتمد الباحث على البرامج الإحصائية (SPSS), (BILOG-MG) واستخدم اختبار تحليل التباين ثنائي الاتجاه، واختبار "ت" لعينتين مستقلتين، واختبار شيفيه، وتوصل البحث إلى وجود فروق ذات دلالة إحصائية في دقة تقدير بارامتر الصعوبة، ترجع لاختلاف النموذج المستخدم لصالح النموذج الأحادي، وعدم وجود فروق ذات دلالة إحصائية في دقة تقدير بارامتر الصعوبة، ترجع لاختلاف طريقة التقدير، أو التفاعل بين النموذج المستخدم وطريقة التقدير، ووجود فروق ذات دلالة إحصائية في دقة تقدير بارامتر التمييز، ترجع لاختلاف النموذج المستخدم لصالح النموذج الثنائي، وعدم وجود فروق ذات دلالة إحصائية في دقة تقدير بارامتر التمييز، ترجع لاختلاف طريقة التقدير، أو التفاعل بين النموذج المستخدم وطريقة التقدير، ووجود فروق ذات دلالة إحصائية في دقة تقدير قدرات الأفراد، ترجع لاختلاف النموذج وطريقة التقدير والتفاعل بينهما، وقد أوصى البحث بضرورة مراعاة اختيار نموذج نظرية الاستجابة للمفردة الملائم للبيانات في ضوء افتراضات كل نموذج للحصول على دقة أعلى في تقدير بارامترات المفردات وقدرات الأفراد، وفق طرق التقدير المتعددة.

كما هدفت دراسة كريشان والعناتي (2023) إلى التحقق من أثر شكل التوزيع لدرجات الطلبة على الاختبارات الوطنية في دقة تقدير دالة معلومات المفردة وثبات الاختبار في ضوء نظرية الاستجابة للمفردة، اتبعت الدراسة المنهج الوصفي المقارن، ولتحقيق هدف الدراسة تم اختيار عينة عشوائية مكونة من (12000) طالب وطالبة من مجتمع الدراسة، تم استخدام البرنامج الإحصائي (BILOG_MG3) لتحليل البيانات في ضوء نظرية الاستجابة للمفردة حسب النموذج ثلاثي المعلمة، إذ تم استخراج دالة معلومات الفقرة والخطأ المعياري لها باختلاف شكل التوزيع للبيانات، كما تم استخراج معاملات الثبات النظري والتجريبي، وأظهرت النتائج أن قيم متوسط الخطأ المعياري للتقدير لدالة معلومات المفردة الاختبارات الوطنية؛ كانت الأعلى في البيانات ذات الالتواء الموجب، وكانت قيم الثبات النظري والتجريبي للاختبارات الوطنية منخفضة في البيانات ذات الالتواء الموجب، كما كانت الفروق بين متوسطات الخطأ للبيانات ذات التوزيع المعتدل مع البيانات موجبة وسالبة الالتواء دالة إحصائيًا في مواد اللغة العربية، واللغة الانجليزية، والرياضيات، في حين كانت الفروق غير دالة في مادة العلوم.

وهدفت دراسة النفيعي (2024) إلى مقارنة قوة طرق اختيار نموذج الاستجابة للمفردة عديدة الاستجابة الأكثر مطابقة للبيانات في ضوء عدد من الشروط التجريبية، ولتحقيق ذلك تم تصميم دراسة محاكاة من خلال تصميم بحثي عاملي متقدم تم من خلاله دراسة أثر التفاعل بين طرق اختيار النموذج وكل من اختلاف عدد رتب الاستجابة (أربع رتب، وخمس رتب) ، وشكل توزيع السمة المقاسة (ملتوي سالب واعتدالي وملتوي موجب)، وحجم العينة (500,1000) وذلك من خلال فحص معدلات قوة الاختبار، وتم التوصل إلى النتائج الآتية: جميع معدلات القوة لجميع الطرق كانت كبيرة جداً وفقاً لمعيار كوهين المعدل، إذ كانت أكبر من القيمة (0.80) مما يدل على أن جميع الطرق لديها قدرة كبيرة جداً على تحديد نموذج نظرية الاستجابة للمفردة عديدة الاستجابة الرتبية الأكثر ملاءمة، كما أن معدلات قوة الاختبار الإحصائي لا تختلف باختلاف التفاعل الثنائي بين طرق اختيار النموذج الأكثر مطابقة للبيانات وعدد رتب الاستجابة، بينما اختلفت باختلاف التفاعل الثنائي بين طرق اختيار النموذج الأكثر مطابقة للبيانات وكل من شكل توزيع السمة المقاسة وحجم العينة.

أما دراسة شكار وبن نابي (2024)، فهدفت إلى التعرف على مختلف طرق تقدير معالم المفردة والقدرة في نظرية الاستجابة للمفردة وفق نموذج راش، بالإضافة إلى تحديد طرائق التقدير الأكثر

دقة في التقدير تحت ظروف مختلفة في عدد الفقرات وحجم العينة والنموذج اللوغاريتمي المستخدم، وتوصلت الدراسة إلى أنه توجد طرق مختلفة وعديدة، من أهمها طريقة الأرجحية العظمى بأساليبها الثلاثة، والطريقة "البيزية" بأسلوبيها، وتوصل - من خلال إجراء المقارنة بين طرائق التقدير، واستنادًا إلى مختلف الدراسات السابقة العربية منها والأجنبية - إلى تفوق الطرق "البيزية" على طريقة الأرجحية العظمى، خاصة في العينات الصغيرة، أو عند استخدام النموذج اللوغاريتمي الثنائي، كما أنها أعطت تقديرات أقل خطأ مقارنة مع طريقة الأرجحية العظمى، بالإضافة إلى أنها تتمتع بثبات أفضل من تقديرات المعالم والأخطاء المعيارية التي يتم الحصول عليها من خلال طريقة الأرجحية العظمى، خاصة عند تقدير القدرات على أطراف متصل القدرة، كذلك بزيادة عدد الفقرات وحجم العينة.

مشكلة الدراسة

تُعَدّ عمليات التقدير الدقيقة لقدرات الأفراد من الأهداف الأساسية في مجالات القياس النفسي والتعليمي ونظرية الاستجابة للمفردة (IRT)، وقد حظيت باهتمام كبير من الباحثين نظرًا لدورها الحاسم في اتخاذ القرارات التعليمية والمهنية وتقييم الأداء، وتعتمد دقّة تقدير قدرات الأفراد على مجموعة من العوامل المؤثرة، مثل: عدد بارامترات النموذج المستخدم، وطرق التقدير، وشكل توزيع القدرة، وحجم العينة، وتؤثر هذه العوامل بشكل مباشر على مخرجات التقدير.

ركزت الدراسات الأولى في أثر عدد البارامترات على دقة التقدير على مقارنة النماذج الأحادية مع النماذج الثنائية والثلاثية، وأوضحت النتائج أن زيادة عدد البارامترات تحسّن من ملاءمة النموذج للبيانات، لكنها تزيد أيضًا من متطلبات حجم العينة (Lord & Novick, 1968)، تلتها مجموعة من الدراسات التي أظهرت أن اختلاف النماذج يؤثر على تقدير القدرات (Embretson & Reise, 2000) مما يعكس أهمية اختيار النموذج المناسب حسب الغرض من التقدير. (Samejima, 1969; Choi et al., 1997)

أما فيما يتعلق بشكل توزيع القدرة فتم تناوله في الدراسات نظراً لتأثيره على دقة الطرق المستخدمة لتقدير معالم الأفراد، فهو يغير افتراضات النموذج الإحصائي، ويحدد كفاءة دوال المعلومات (Information Functions) للنموذج، وقد يؤثر على الأخطاء المعيارية ودقة التقدير، ومن الدراسات المبكرة في شكل توزيع القدرة دراسة Samejima (1969) التي أكدت على أهمية التوزيع

الطبيعي لتحقيق دقة عالية في التقدير، تلتها مجموعة من الدراسات مثل دراسة (Hambleton et al., 1979) التي أظهرت أن التوزيع غير الطبيعي يمكن أن يؤدي إلى أخطاء في التقدير مثل التمييز واحتمالية التخمين في النموذج الثلاثي (3PL)، وبينت دراسة Choi et al., (1997) أن الانحراف الشديد يقلل دقة التقدير ويزيد الأخطاء، وأظهرت دراسة Dai et al., (2021) أن تأثير التوزيع المنحرف يمكن تقليله بزيادة حجم العينة، أو باستخدام نماذج متعددة البارامترات.

أما عن الدراسات المبكرة لتأثير حجم العينة، دراسة Lord (1983) التي هدفت إلى تحديد الحد الأدنى من عدد المفحوصين اللازم للحصول على تقديرات دقيقة لبارامترات المفردات والقدرات، وأوضح أن العينات الصغيرة قد تؤدي إلى تقديرات غير مستقرة، خاصة عند استخدام نماذج متعددة المعايير مثل نموذج (3PL)، وتلت دراسته مجموعة من الدراسات في حجم العينة، إذ أوضحت دراسة Choi et al. (1997) أن كبر حجم العينة يقلل من الخطأ المعياري لجذر المتوسط (RMSE) في تقدير القدرات، إذ يتناقص RMSE بشكل كبير حتى يصل إلى حجم عينة معين، ثم يبدأ معدل التحسن في الدقة بالانخفاض، كما أكدت على ذلك دراسة Dai et al. (2021) التي أوضحت أنه يمكن تحسين التقدير بشكل ملحوظ قبل الوصول إلى 1000 مفحوص، ولكن بعده تكون الفوائد محدودة في دقة التقدير.

ومن المهم أن ندرك أن دراسة أثر هذه المتغيرات وتفاعلاتها مهمة جداً لرفع كفاءة أدوات القياس من خلال تحسين دقة التقدير، إذ أوصت الدراسات السابقة بدراسة أثر تفاعلات هذه المتغيرات، على سبيل المثال لا الحصر، دراسة كلٍ من (شرفاوي وبن نابي، 2022؛ القضاة، 2020؛ محمود، 2017) التي أوصت بدراسة طرق التقدير باستخدام النماذج اللوجستية الثنائية، ودراسة القضاة (2020) التي أوصت بدراسة أثر النماذج الثلاثية على دقة تقدير قدرة الأفراد، ودراسة عبدالوهاب (2019) التي أوصت بدراسة أثر نموذج الاستجابة للمفردة على دقة تقدير معالم الأفراد، أما الدراسات التي أوصت بدراسة دقة التقدير تحت ظروف مختلفة من أشكال التوزيع (الحواري، 2015؛ الخرشة، 2018؛ شرفاوي وبن نابي ، 2022؛ الحمدانية والنصراوين، 2020)، أما دراسة (كريشان والعناتي، 2023) فأوصت بدراسة أثر شكل التوزيع على دالة المعلومات في ضوء النموذج اللوجستي (أحادي المعلمة، ثنائي المعلمة، ثلاثي المعلمة)، أما عن الدراسات التي أوصت بدراسة أثر حجم العينة بأحجام مختلفة، دراسة كلٍ من (البادية وآخرون، 2018؛ شرفاوي وبن

نابي ، 2022؛ ضعضع وآخرون، 2020؛ عبدالوهاب، 2019؛ الحمدانية والنصراوين، 2018؛ Fathi & Al_Alem, 2022)، وأوصت بعض الدراسات السابقة بدراسة طرق التقدير باستخدام البيانات المولدة مثل دراسة (الخوالدة والنصراوين، 2020؛ الصبح، 2021؛ الحمدانية والنصراوين، 2018؛ Fathi & Al_Alem, 2022) مما يشكل الحاجة الملحّة للقيام بهذه الدراسة التي تهدف إلى تحديد أثر عدد بارامترات النموذج (أحادي، ثنائي، ثلاثي) وشكل توزيع القدرة (طبيعي، ملتو موجب، ملتو سالب)، وحجم العينة (صغيرة، متوسطة، كبيرة)، والتفاعلات الثنائية والثلاثية والرباعية بينهما على دقة تقدير قدرات الأفراد كمتغير تابع، وذلك من خلال دراسة محاكاة، إذ تشكل هذه الدراسة إضافة علمية للميدان؛ فهي تسلط الضوء على العلاقة بين هذه العوامل وأثرها التفاعلي، مما يُمكِّن الباحثين والممارسين من فهم التفاعل بين العوامل المختلفة، واختيار النماذج الأكثر ملاءمة وفقًا لطبيعة البيانات وظروف التطبيق، وسيسعى الباحثان - من خلال الدراسة الحالية – إلى الإجابة عن السؤال الرئيس الآتي:

هل تختلف تقديرات قدرات الأفراد باختلاف عدد بارامترات النموذج، وطرق التقدير، وشكل توزيع القدرة، وحجم العينة، والتفاعلات المشتركة بينها؟

ويتمثل الهدف الرئيسي للدراسة الحالية في الكشف عن مدى تأثر تقدير قدرات الأفراد باستخدام نظرية الاستجابة للمفردة في ضوء عدد من المتغيرات (عدد البارامترات، وطرق تقدير القدرة، وشكل توزيع القدرة، وحجم العينة) والتفاعلات المشتركة بينها.

أهمية الدراسة

تتمثل الأهمية النظرية في أنها تعدّ إضافة نوعية للمكتبة العلمية في مجال القياس والتقويم، إذ تتميز بأهميتها النظرية العميقة التي تسهم في توسيع الفهم العلمي للعوامل المؤثرة في دقة تقدير قدرات الأفراد، وتتناول الدراسة مجموعة من العوامل الإحصائية والمنهجية، مثل حجم العينة، وعدد البارامترات، وطرق التقدير، وشكل توزيع القدرة، بالإضافة إلى التفاعلات المشتركة بينها ضمن إطار نماذج نظرية الاستجابة للمفردة، كما تُثري هذه الدراسة الأدبيات العلمية من خلال اختبار الفروق بين النماذج التقديرية المختلفة في سياقات متنوعة من حيث خصائص البيانات، ما يوفر إطارًا معرفيًا متقدمًا يدعم الأسس النظرية للقياس، ويُسهم في تطوير نماذج أكثر دقة وفاعلية

لتقدير قدرات الأفراد، ويُعد هذا الإسهام النظري مرجعًا مهمًا للباحثين الساعين إلى الارتقاء بممارسات القياس والتقويم، وتطوير نماذج قياس أكثر دقة وفعالية.

وتتمثل الأهمية التطبيقية لهذه الدراسة في تقديم معلومات دقيقة وعملية تُمكّن المختصين في مجالي القياس والتقويم النفسي والتربوي من اختيار أكثر الأساليب والنماذج ملاءمةً لتقدير قدرات الأفراد، وذلك بما يتوافق مع ظروف الدراسة وخصائص البيانات المتاحة، كحجم العينة وشكل توزيع القدرة. وتسهم الدراسة في توجيه الممارسين نحو اتخاذ قرارات أكثر كفاءة فيما يتعلق بتصميم الاختبارات، وتحليل نتائجها، وتفسير مستويات الأداء بدقة أعلى، كما تُوفّر نتائج الدراسة أساسًا تجريبيًا يمكن الاعتماد عليه في تطوير الاختبارات بما يضمن أعلى درجات الدقة والموضوعية في تقدير القدرات.

حدود الدراسة

الحدود الموضوعية: اقتصرت الدراسة الحالية على توظيف بيانات مُولّدة باستخدام أسلوب المحاكاة (Simulation) من خلال برنامج R (الإصدار 4.4.2) The R Project for Statistical Computing، وذلك وفقًا لثلاثة نماذج لوجستية: النموذج الأحادي البارامتر (1PLM)، والنموذج الثنائي البارامتر (2PLM)، والنموذج الثلاثي البارامتر (1PLM)، وقد تم تنفيذ المحاكاة باستخدام ثلاثة مستويات لحجم العينة (100، 500، 1000)، وثلاثة أساليب تقدير إحصائي هي: أسلوب الاحتمال الأقصى (Maximum Likelihood - ML)، والتقدير البعدي المتوقع (Expected A Posteriori - EP)، والتقدير البعدي الأقصى (Maximum A Posteriori - MP). كما تم أخذ ثلاثة أشكال لتوزيع القدرة في الاعتبار: التوزيع الطبيعي، والتوزيع الموجب الالتواء، والتوزيع السالب الالتواء.

الحدود الزمانية: العام الدراسي 1445-1446هـ/2024-2025 م.

مصطلحات الدراسة

بارامترات النموذج (Model Parameters):

عرفها Hambelton & Swaminathan (1985) بأنها القيم التي تمثل الخصائص الداخلية لنموذج معين وتستخدم لتحديد سلوك أو أداء النموذج في تحليل البيانات، كما عرفها Kim (2004)

Baker & بأنها المتغيرات التي تتحكم في أداء النموذج الرياضي أو الإحصائي ويتم تقديرها من البيانات للحصول على توقعات دقيقة.

توزيع القدرة (Ability Distribution):

عرفه Lord (1980) بأنه التوزيع الإحصائي الذي يمثل تنوّع مستويات القدرة بين الأفراد في عينة ما، كما عرفه Embretson & Reise (2000) بأنه الكيفية التي توزع بها القدرة أو الكفاءة بين الأفراد في نموذج القياس النفسي، ويستخدم في تحديد مستويات أداء الأفراد.

تقدير القدرة (Ability Estimation):

هي عملية تكرارية تستخدم لحساب احتمالية الاستجابة الصحيحة للفرد على مفردات الاختبار للحصول على تقدير أولي لقدرة المفحوص، ومن ثم يتم التعديل على هذا التقدير بحيث يصبح التغيير في موقع المفحوص أقل ما يمكن (بيكر وفرانك، 2010).

دراسة المحاكاة (Simulation Study):

عرفها Ripley (1987) بأنها طريقة بحث تعتمد على إنشاء بيانات محاكاة لتقدير الأداء المحتمل لنماذج أو فرضيات معينة، كما عرفها (Law & Kelton, 2000) بأنها أسلوب بحثي يُستخدم فيه نموذج حاسوبي لتمثيل سيناريوهات افتراضية واختبار الفرضيات عن طريق تكرار التجارب عبر بيانات محاكاة، مما يسمح بتحليل سلوك الأنظمة المعقدة في بيئات مُتحكم بها.

منهج الدراسة

تم الاعتماد على المنهج التجريبي للإجابة عن تساؤل الدراسة، إذ تم توليد البيانات باستخدام أسلوب المحاكاة (بيانات مولدة)، تتميز البيانات المولدة عن البيانات الواقعية في أنها تضبط المتغيرات الخارجية، وتقدم قيمًا حقيقية للمعالم، مما يمكن مقارنتها مع القيم المقدرة لفحص دقة التقدير بالاعتماد على البرامج المختلفة للتقدير، وتسهم البيانات المولدة في التخلص من تأثير بعض العوامل على استجابات الأفراد مثل الغش والتخمين، وتساعد أيضاً على ضبط الموقف الاختباري من خلال معالجة مشكلة اللامبالاة وترك الإجابة عن بعض المفردات، وتقدم نتائج نظرية يمكن الاعتماد عليها ميدانياً، وتسهّل عملية جمع البيانات والمعلومات، والتكلفة المادية.

إجراءات الدراسة

1. تم توليد بيانات ثنائية الاستجابة (0،1) باستخدام حزمة mirt من برنامج (R) الإصدار (4.2.2) لواحد وثمانين نموذجًا للبيانات وفقاً لمتغيرات الدراسة: عدد بارامترات النموذج (النموذج اللوجستي الأحادي، النموذج اللوجستي الثنائي، النموذج اللوجستي الثلاثي)، وشكل توزيع القدرة (ملتوي لليمين، طبيعي، ملتوي لليسار)، حجم العينة (100، 500، 1000)، وطرق التقدير (الأرجحية العظمى (ML)، التقدير البعدي المتوقع (EP)، التقدير البعدي الأقصى (MP)).

2. استخراج ملفات البيانات في صيغة Excel.

3. تهيئة البيانات المولدة على برنامج SPSS لمعالجتها إحصائياً.

4. إجراء تحليل البيانات لمعالجة البيانات المولدة، والإجابة عن سؤال الدراسة عن طريق استخدام برنامج (R).

5. تفسير ومناقشة النتائج.

6.تقديم التوصيات والمقترحات في ضوء نتائج الدراسة.

نتائج الدراسة وتفسيراتها

قبل التطرق للإجابة عن سؤال الدراسة يجب ذكر الآتي:

1-المتغيرات المستقلة الأربعة المتضمنة في كل سؤال (عدد بارامترات النموذج، وحجم العينة، وشكل توزيع القدرة، وطريقة التقدير)، تُعد متغيرات بين المجموعات Between Subjects أو قياسات مستقلة، لأن عملية المحاكاة مبنية على أساس إنتاج بيانات جديدة data sets في كل عملية محاكاة، وكل مجموعة بيانات تجرى عليها وحدة تكرار replication، التي تمثله الحالة case، وتتكرر وحدات التكرارات طبقاً لعدد التكرارات المطلوبة في كل موقف بحثي (مثلاً استجابات 100 فرد افتراضي على اختبار مكون من 50 مفردة افتراضية ثنائية الاستجابة، بطريقة ML وتوزيع القدرة طبيعي وفقاً للنموذج أحادي المعلم)، فإذا تم تحديد عدد التكرارات=1000، سنجد 1000 تكرار replications ببيانات جديدة في كل موقف بحثي، وبالتالي لا توجد قياسات مكررة وإنما قياسات مستقلة، للمتغيرات المستقلة موضوع الدراسة.

2-يُعد تحليل التباين العاملي (Factorial Anova) الأسلوب الإحصائي المناسب لمعالجة السؤال إذ أنه يستخدم عند دراسة تأثير متغيرين مستقلين أو أكثر على متغير تابع كمّي، إذ يُمكن من خلاله التعرف على التأثيرات الرئيسية لكل متغير مستقل على حدة، بالإضافة إلى التأثيرات التفاعلية بين هذه المتغيرات، ولا يمكن هنا بحث إمكانية اللجوء لبديل لا بارامتري بعد فحص الافتراضات اللازمة، نظراً لكبر حجم العينة (81000 حالة)، وهذا العدد كبير جداً يجعل هناك مناعة إحصائية ضد انتهاك افتراض الاعتدالية، كما أن تساوي أحجام الخلايا نظراً لأن كل موقف بحثي يحتوي على 1000 تكرار يجعل افتراض التجانس متحقق (Field et al., 2012).

ينص سؤال الدراسة على الآتي:

هل تختلف تقديرات قدرات الأفراد باختلاف عدد بارامترات النموذج وحجم العينة وشكل توزيع القدرة وطريقة التقدير والتفاعلات المشتركة بينها؟

وقبل الإجابة عن السؤال يجب عرض الإحصائيات الوصفية (المتوسطات والانحرافات المعيارية)، لكافة المواقف البحثية التي يتضمنها السؤال، ويوضح جدول1 ذلك.

جدول 1

المتوسطات والانحراف المعيارية لقدرات الأفراد وفقًا لعدد بارامترات النموذج وحجم العينة وشكل توزيع القدرة وطريقة التقدير والتفاعلات المشتركة بينها.

النموذج	حجم العينة	طريقة التقدير	توزيع القدرة	المتوسط الحسابي	الانحراف المعياري	حجم العينة
1PL	100	ML	Normal	-.000559203	.1820305433	1000
			p_skewed	-.029966587	.2345842836	1000
			n_skewed	.040665207	.2367697864	1000
			Total	.003379806	.2210936400	3000
		MP	Normal	.000059440	.0028398613	1000
			p_skewed	-.007700205	.0020805202	1000
			n_skewed	.007777276	.0020344903	1000
			Total	.000045504	.0067413635	3000
		EP	Normal	.000014353	.0005371908	1000
			p_skewed	-.000112903	.0000523577	1000
			n_skewed	.000114377	.0000537114	1000
			Total	.000005276	.0003265804	3000
		Total	Normal	-.000161803	.1050739534	3000

النموذج	حجم العينة	طريقة التقدير	توزيع القدرة	المتوسط الحسابي	الانحراف المعياري	حجم العينة
			p_skewed	-.012593232	.1359891051	3000
			n_skewed	.016185620	.1377863444	3000
			Total	.001143528	.1277035359	9000
		ML	Normal	.007093108	.1428578392	1000
			p_skewed	-.036473601	.1664957678	1000
			n_skewed	.037125117	.1652029877	1000
			Total	.002581541	.1613596012	3000
		MP	Normal	.000030053	.0024755054	1000
			p_skewed	-.007821175	.0017038625	1000
			n_skewed	.007846819	.0016866132	1000
			Total	.000018565	.0066995562	3000
	500	EP	Normal	.000004992	.0005185330	1000
			p_skewed	-.000100405	.0000406997	1000
			n_skewed	.000101564	.0000423404	1000
			Total	.000002050	.0003122817	3000
		Total	Normal	.002376051	.0825318771	3000
			p_skewed	-.014798394	.0973653409	3000
			n_skewed	.015024500	.0966774287	3000
			Total	.000867386	.0932389654	9000
1PL		ML	Normal	.005085106	.1445021471	1000
			p_skewed	-.042700799	.1557060686	1000
			n_skewed	.047003283	.1541402388	1000
			Total	.003129197	.1558514445	3000
		MP	Normal	.000101286	.0023545063	1000
			p_skewed	-.007804295	.0016299664	1000
			n_skewed	.007829786	.0016781134	1000
	1000		Total	.000042259	.0066650190	3000
		EP	Normal	-.000010417	.0004976977	1000
			p_skewed	-.000095940	.0000330724	1000
			n_skewed	.000094895	.0000330249	1000
			Total	-.000003821	.0002988875	3000
		Total	Normal	.001725325	.0834459444	3000
			p_skewed	-.016867011	.0917641861	3000
			n_skewed	.018309321	.0913081503	3000
			Total	.001055878	.0900652099	9000

النموذج	حجم العينة	طريقة التقدير	توزيع القدرة	المتوسط الحسابي	الانحراف المعياري	حجم العينة
	المجموع	ML	Normal	.003873004	.1574867316	3000
			p_skewed	-.036380329	.1888604185	3000
			n_skewed	.041597869	.1889354841	3000
			Total	.003030181	.1818307600	9000
		MP	Normal	.000063593	.0025642439	3000
			p_skewed	-.007775225	.0018157179	3000
			n_skewed	.007817960	.0018070208	3000
			Total	.000035443	.0067013182	9000
		EP	Normal	.000002976	.0005179863	3000
			p_skewed	-.000103083	.0000433689	3000
			n_skewed	.000103612	.0000445741	3000
			Total	.000001168	.0003127755	9000
		Total	Normal	.001313191	.0909454644	9000
			p_skewed	-.014752879	.1101434829	9000
			n_skewed	.016506480	.1105534272	9000
			Total	.001022264	.1050571696	27000

النموذج	حجم العينة	طريقة التقدير	توزيع القدرة	المتوسط الحسابي	الانحراف المعياري	حجم العينة
2PL	100	ML	Normal	.000662129	.0129255474	1000
			p_skewed	.053868724	.0259786275	1000
			n_skewed	-.055294547	.0286626682	1000
			Total	-.000254565	.0504114920	3000
		MP	normal	.000054672	.0048398383	1000
			p_skewed	-.016985712	.0068641254	1000
			n_skewed	.017545309	.0069511169	1000
			Total	.000204756	.0154402860	3000
	100	EP	normal	.000020165	.0004088612	1000
			p_skewed	.000019458	.0001839926	1000
			n_skewed	-.000048958	.0013025136	1000
			Total	-.000003112	.0007957075	3000
		Total	normal	.000245655	.0079748498	3000
			p_skewed	.012300823	.0339553309	3000
			n_skewed	-.012599399	.0354073339	3000
			Total	-.000017640	.0304403350	9000

النموذج	حجم العينة	طريقة التقدير	توزيع القدرة	المتوسط الحسابي	الانحراف المعياري	حجم العينة
2PL	500	ML	normal	-.000085110	.0054969181	1000
			p_skewed	.035895508	.0163415259	1000
			n_skewed	-.037782336	.0157999170	1000
			Total	-.000657313	.0329754601	3000
		MP	normal	.000078245	.0036564427	1000
			p_skewed	-.013115210	.0042567902	1000
			n_skewed	.012913573	.0043544634	1000
			Total	-.000041131	.0113915268	3000
		EP	normal	.000006416	.0004372105	1000
			p_skewed	-.000032764	.0000946120	1000
			n_skewed	.000030193	.0000989936	1000
			Total	.000001282	.0002656979	3000
		Total	normal	-.000000150	.0038192954	3000
			p_skewed	.007582511	.0229014398	3000
			n_skewed	-.008279524	.0235053689	3000
			Total	-.000232387	.0201429863	9000
	1000	ML	normal	.000107533	.0048270637	1000
			p_skewed	.032311851	.0164480899	1000
			n_skewed	-.032572341	.0164606490	1000
			Total	-.000050985	.0298336236	3000
		MP	normal	-.000028553	.0032406084	1000
			p_skewed	-.012324581	.0040535314	1000
			n_skewed	.012276094	.0040461766	1000
			Total	-.000025680	.0107389144	3000
		EP	normal	.000000761	.0001911127	1000
			p_skewed	-.000025927	.0000279130	1000
			n_skewed	.000025864	.0000277085	1000
			Total	.000000233	.0001145827	3000
		Total	normal	.000026581	.0033578916	3000
			p_skewed	.006653781	.0212153087	3000
			n_skewed	-.006756794	.0213087598	3000
			Total	-.000025477	.0183044655	9000
	المجموع	ML	normal	.000228184	.0085778871	3000
			p_skewed	.040692027	.0222007049	3000
			n_skewed	-.041883075	.0232716690	3000

النموذج	حجم العينة	طريقة التقدير	توزيع القدرة	المتوسط الحسابي	الانحراف المعياري	حجم العينة
2PL	المجموع		Total	-.000320954	.0388069184	9000
		MP	normal	.000034788	.0039694629	3000
			p_skewed	-.014141834	.0055994430	3000
			n_skewed	.014244992	.0057775998	3000
			Total	.000045982	.0126941467	9000
		EP	normal	.000009114	.0003627581	3000
			p_skewed	-.000013078	.0001227009	3000
			n_skewed	.000002366	.0007549686	3000
			Total	-.000000532	.0004887829	9000
		Total	normal	.000090695	.0054612935	9000
			p_skewed	.008845705	.0267419583	9000
			n_skewed	-.009211906	.0275567090	9000
			Total	-.000091835	.0235747994	27000

النموذج	حجم العينة	طريقة التقدير	توزيع القدرة	المتوسط الحسابي	الانحراف المعياري	حجم العينة
3PL	100	ML	normal	-.090586941	.1001303626	1000
			p_skewed	-.063303308	.0680701217	1000
			n_skewed	-.115922462	.0729045946	1000
			Total	-.089937570	.0843541582	3000
		MP	normal	.029846459	.0288659569	1000
			p_skewed	.065957307	.0348424880	1000
			n_skewed	.062490039	.0273621085	1000
			Total	.052764602	.0345841989	3000
		EP	normal	-.018233261	.0277253436	1000
			p_skewed	-.028521554	.0260489599	1000
			n_skewed	-.016665992	.0207824745	1000
			Total	-.021140269	.0255661359	3000
		Total	normal	-.026324581	.0795257403	3000
			p_skewed	-.008622518	.0718165046	3000
			n_skewed	-.023366138	.0865631211	3000
			Total	-.019437746	.0798973516	9000
		ML	normal	-.008867334	.0163078792	1000
			p_skewed	-.014952805	.0221238962	1000
			n_skewed	-.046362504	.0235139700	1000
			Total	-.023394214	.0265678253	3000

النموذج	حجم العينة	طريقة التقدير	توزيع القدرة	المتوسط الحسابي	الانحراف المعياري	حجم العينة
3PL	500	MP	normal	.025393300	.0084805989	1000
			p_skewed	.090562822	.0176958441	1000
			n_skewed	.049533202	.0161244597	1000
			Total	.055163108	.0306400199	3000
		EP	normal	-.000368725	.0028826561	1000
			p_skewed	-.001456008	.0028663240	1000
			n_skewed	-.000336051	.0003199837	1000
			Total	-.000720261	.0024103671	3000
		Total	normal	.005385747	.0180990334	3000
			p_skewed	.024718003	.0496886160	3000
			n_skewed	.000944882	.0424831483	3000
			Total	.010349544	.0404965179	9000
	1000	ML	normal	-.001852570	.0055028487	1000
			p_skewed	-.018504834	.0203663519	1000
			n_skewed	-.031640331	.0160681107	1000
			Total	-.017332579	.0195674055	3000
		MP	normal	.020468139	.0071249151	1000
			p_skewed	.080339805	.0134319875	1000
			n_skewed	.028331021	.0103037268	1000
			Total	.043046321	.0286062277	3000
	1000	EP	normal	-.000012857	.0000623613	1000
			p_skewed	-.000331016	.0014119233	1000
			n_skewed	-.000067944	.0001032707	1000
			Total	-.000137273	.0008295749	3000
		Total	normal	.006200904	.0113742580	3000
			p_skewed	.020501318	.0452207858	3000
			n_skewed	-.001125752	.0268618614	3000
			Total	.008525490	.0323378351	9000
		ML	normal	-.033768949	.0711432516	3000
			p_skewed	-.032253649	.0482597769	3000
			n_skewed	-.064641766	.0582419457	3000
			Total	-.043554788	.0617746158	9000
		MP	normal	.025235966	.0182512026	3000
			p_skewed	.078953311	.0258980214	3000
			n_skewed	.046784754	.0238680751	3000

النموذج	حجم العينة	طريقة التقدير	توزيع القدرة	المتوسط الحسابي	الانحراف المعياري	حجم العينة
			Total	.050324677	.0318061043	9000
		EP	normal	-.006204948	.0181993250	3000
			p_skewed	-.010102859	.0199831053	3000
			n_skewed	-.005689996	.0142891700	3000
			Total	-.007332601	.0177591098	9000
		Total	normal	-.004912643	.0498928863	9000
	المجموع		p_skewed	.012198934	.0586759511	9000
			n_skewed	-.007849003	.0588236942	9000
			Total	-.000187571	.0566455548	27000
			p_skewed	-.013133724	.1501000432	3000
			n_skewed	-.043517267	.1577203695	3000
			Total	-.028937443	.1461906046	9000
		MP	normal	.009986857	.0220299689	3000
			p_skewed	.013757130	.0424121647	3000
			n_skewed	.029270874	.0288914768	3000
			Total	.017671620	.0333020533	9000
		EP	normal	-.006066248	.0181729273	3000
			p_skewed	-.009538333	.0201565985	3000
			n_skewed	-.005533525	.0143677253	3000
			Total	-.007046035	.0178159664	9000
		Total	normal	-.008746910	.0772187373	9000
			p_skewed	-.002971642	.0915712734	9000
			n_skewed	-.006593306	.0975741428	9000
			Total	-.006103953	.0892263291	27000
	500	ML	normal	-.000619778	.0833037752	3000
			p_skewed	-.005176966	.1020151576	3000
			n_skewed	-.015673241	.1037554038	3000
			Total	-.007156662	.0969961216	9000
		MP	normal	.008500532	.0131599086	3000
			p_skewed	.023208812	.0488367161	3000
			n_skewed	.023431198	.0209504916	3000
			Total	.018380181	.0323672530	9000
	500	EP	normal	-.000119106	.0017182728	3000
			p_skewed	-.000529726	.0017805103	3000
			n_skewed	-.000068098	.0002733666	3000
			Total	-.000238977	.0014519123	9000

النموذج	حجم العينة	طريقة التقدير	توزيع القدرة	المتوسط الحسابي	الانحراف المعياري	حجم العينة
		Total	normal	.002587216	.0488762224	9000
			p_skewed	.005834040	.0664733474	9000
			n_skewed	.002563286	.0631842347	9000
			Total	.003661514	.0600169945	27000
		ML	normal	.001113356	.0835585287	3000
			p_skewed	-.009631260	.0963424780	3000
			n_skewed	-.005736463	.0973759033	3000
			Total	-.004751456	.0927351618	9000
		MP	normal	.006846957	.0107264659	3000
			p_skewed	.020070310	.0434358837	3000
			n_skewed	.016145634	.0109233491	3000
	1000		Total	.014354300	.0271589967	9000
		EP	normal	-.000007504	.0003098544	3000
			p_skewed	-.000150961	.0008256649	3000
			n_skewed	.000017605	.0000928799	3000
			Total	-.000046953	.0005172774	9000
		Total	normal	.002650937	.0487259155	9000
			p_skewed	.003429363	.0622551720	9000
			n_skewed	.003475592	.0573196855	9000
			Total	.003185297	.0563772728	27000
		ML	normal	-.009889253	.1013121538	9000
			p_skewed	-.009313984	.1186611631	9000
			n_skewed	-.021642324	.1236664353	9000
			Total	-.013615187	.1150824292	27000
		MP	normal	.008444782	.0161071678	9000
			p_skewed	.019012084	.0451497341	9000
			n_skewed	.022949235	.0222046881	9000
	المجموع		Total	.016802034	.0311089537	27000
		EP	normal	-.002064286	.0109127163	9000
			p_skewed	-.003406340	.0124703550	9000
			n_skewed	-.001861339	.0086929536	9000
			Total	-.002443988	.0108251014	27000
		Total	normal	-.001169586	.0600310136	27000
			p_skewed	.002097253	.0746545561	27000
			n_skewed	-.000184809	.0749650686	27000
			Total	.000247619	.0702425279	81000

جدول 2

نتائج تحليل التباين العاملي لتأثير كل من عدد بارامترات النموذج وحجم العينة وشكل توزيع القدرة وطريقة التقدير والتفاعلات المشتركة بينها على قدرات الأفراد

	مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة ف	مستوى الدلالة	حجم التأثير مربع إيتا η2
القدرة	النموذج	.024	2	.012	3.099	.045	.000
	حجم العينة	1.637	2	.818	207.678	.000	.005
	الطريقة	12.784	2	6.392	1621.876	.000	.039
	توزيع القدرة	.152	2	.076	19.240	.000	.000
	النموذج *-حجم العينة	3.381	4	.845	214.497	.000	.010
	النموذج * والطريقة	27.621	4	6.905	1752.128	.000	.080
	النموذج * توزيع القدرة	7.824	4	1.956	496.349	.000	.024
	حجم العينة *الطريقة	1.928	4	.482	122.287	.000	.006
	حجم العينة *توزيع القدرة	.069	4	.017	4.394	.001	.000
	الطريقة* توزيع القدرة	1.745	4	.436	110.697	.000	.005
	النموذج *حجم العينة *الطريقة	3.897	8	.487	123.615	.000	.012
	النموذج *حجم العينة *توزيع القدرة	.209	8	.026	6.623	.000	.001
	النموذج *الطريقة *توزيع القدرة	17.630	8	2.204	559.176	.000	.052
	حجم العينة *الطريقة *توزيع القدرة	1.332	8	.167	42.255	.000	.004
	النموذج*حجم العينة*الطريقة* *توزيع القدرة	.513	16	.032	8.130	.000	.002
	الخطأ	318.904	80919	.004			

يتضح من الجدول (2) وجود تأثيرات دالة إحصائيًا لكل من: النموذج، حجم العينة، الطريقة، شكل توزيع القدرة، بالإضافة إلى التفاعلات الثنائية والثلاثية والرباعية بينها، على القدرة المقدّرة، ومع ذلك، فإن معظم أحجام التأثير كانت لا تذكر (أقل من 0.01)، وتشمل تأثيرات كل من: النموذج، حجم العينة، توزيع القدرة، التفاعل بين حجم العينة والطريقة، بين حجم العينة وتوزيع القدرة، بين الطريقة وتوزيع القدرة، بين النموذج وحجم العينة وتوزيع القدرة، بين حجم العينة والطريقة وتوزيع القدرة، وأخيرًا التفاعل الرباعي بين النموذج، حجم العينة، الطريقة، وتوزيع القدرة.

أما التأثيرات التي جاءت بأحجام صغيرة (بين 0.01 وأقل من 0.06)، فقد شملت: الطريقة، التفاعل بين النموذج وحجم العينة، بين النموذج وتوزيع القدرة، بين النموذج وحجم العينة والطريقة، وبين النموذج والطريقة وتوزيع القدرة، في حين سُجِّل تأثير واحد فقط بحجم متوسط (من 0.06 إلى أقل من 0.14)، وهو التفاعل بين النموذج والطريقة.

وللتعرف على اتجاه الفروق، تم إجراء مقارنات متعددة (Post Hoc) باستخدام اختبار بونفِروني (Bonferroni)، كما هو موضح في الجدول (3).

جدول 3

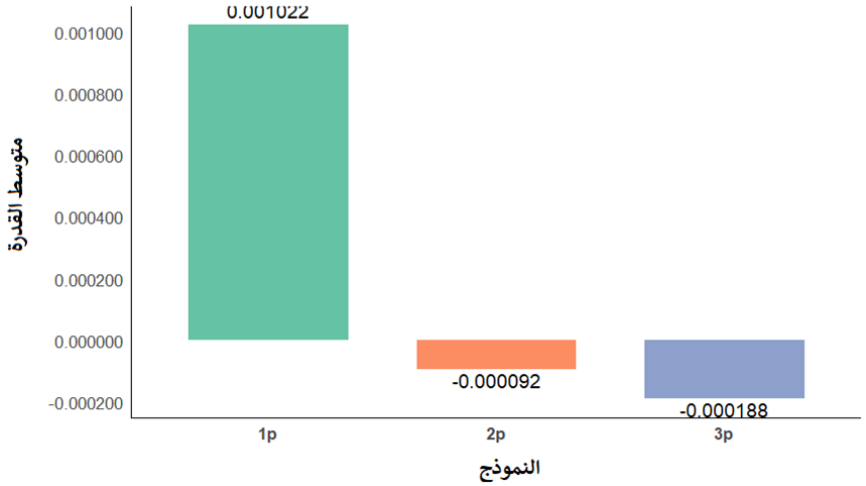
المقارنات البعدية للنماذج في متغير القدرة

الفرق الثنائي	فرق المتوسطين	الخطأ المعياري	مستوى الدلالة	التأثير
النموذج الأحادي – النموذج الثنائي	.001114099	.0005403035	.118	غير مؤثر
النموذج الأحادي- النموذج الثلاثي	.001209835	.0005403035	.075	غير مؤثر
النموذج الثنائي – النموذج الثلاثي	.000095736	.0005403035	1.000	غير مؤثر

ويمكن توضيح متوسطات النماذج الثلاثة في القدرة المقدرة في شكل (1).

شكل 1

تأثير النموذج في القدرة:



يتضح من الجدول (3) عدم وجود فروق ثنائية دالة إحصائيًا بين النماذج أسهمت في فروق القدرة المقدّرة، على الرغم من أن الشكل (1) يُظهر أن متوسط القدرة في النموذج الأول يبدو أعلى ظاهريًا من النموذجين الآخرين، إلا أن هذا الفرق الظاهري لا يمكن الاعتماد عليه، نظرًا لارتفاع قيم الانحرافات المعيارية مقارنة بالمتوسطات، مما يؤدي إلى تقليل الفروق الإحصائية وجعلها غير دالة.

جدول 4:

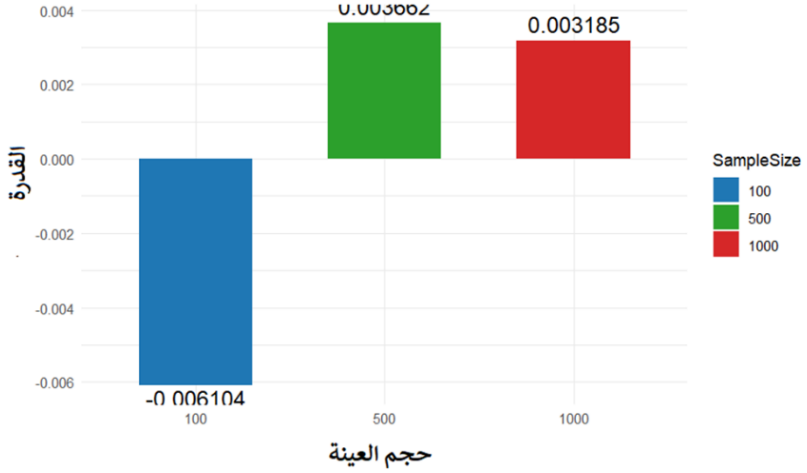
المقارنات البعدية لحجم العينة في متغير القدرة:

الفرق الثنائي	فرق المتوسطين	الخطأ المعياري	مستوى الدلالة	التأثير
100*500	-.009765467*	.0005403035	.000	مؤثر
100*1000	-.009289250*	.0005403035	.000	مؤثر
500*1000	.000476217	.0005403035	1.000	غير مؤثر

ويمكن توضيح متوسطات أحجام العينة الثلاثة في القدرة المقدرة في شكل (2).

شكل 2:

تأثير حجم العينة في القدرة:



يتضح من الجدول (4) وجود فرقين ثنائيين بين أحجام العينات، أسهما في ظهور فروق دالة إحصائيًا في القدرة، ما أدى إلى اختلاف القدرة المقدّرة باختلاف حجم العينة، هذان الفرقان هما بين الحجمين (100-500) و(100-1000)، وذلك لصالح الحجمين الأكبرين 500 و1000، كما يظهر في الجدول (1)، ويتأكد هذا أيضًا من الشكل (2)، الذي يبيّن تفوّق القدرة المقدّرة عند حجمي العينة 500 و1000 مقارنة بحجم العينة 100.

جدول 5:

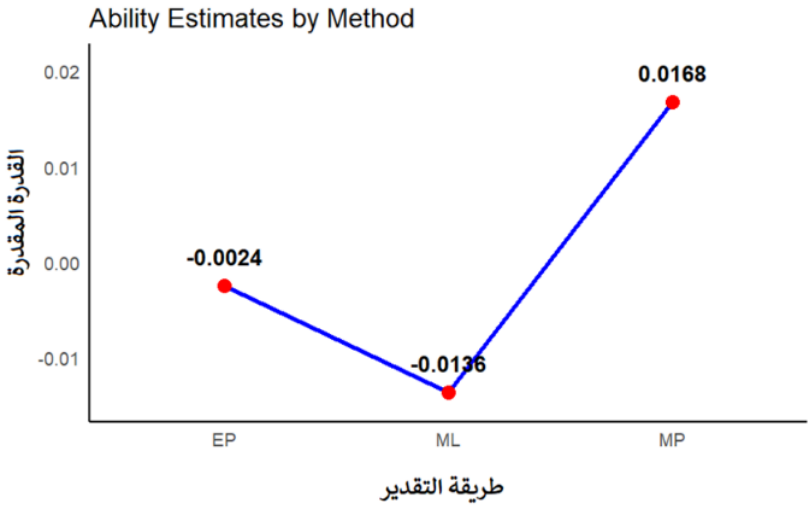
المقارنات البعدية لطريقة التقدير في متغير القدرة:

الفرق الثنائي	فرق المتوسطين	الخطأ المعياري	مستوى الدلالة	التأثير
الاحتمال الأقصى* التقدير البعدي الأقصى	-.030417221*	.0005403035	.000	مؤثر
الاحتمال الأقصى*التقدير البعدي المتوقع	-.011171199*	.0005403035	.000	مؤثر
التقدير البعدي الأقصى*التقدير البعدي المتوقع	.019246022*	.0005403035	.000	مؤثر

ويمكن توضيح ذلك في شكل 3

شكل 3:

تأثير طريقة التقدير في القدرة المقدرة:



يتضح من جدول (5) أن جميع الفروق الثنائية بين طرق التقدير كانت دالة إحصائيًا، مما يشير إلى أن طريقة التقدير تؤثر في القدرة المقدَّرة، على الرغم من أن حجم التأثير كان صغيرًا، ووفقًا لبيانات جدول (1)، فإن اتجاه الفروق الثنائية كان على النحو الآتي: الفرق بين طريقتي التقدير بطريقة الأرجحية العظمى (الاحتمال الأقصى) (ML) وطريقة التقدير البعدي الأقصى (MP) كان لصالح طريقة التقدير البعدي الأقصى (MP)، والفرق بين طريقتي التقدير البعدي الأقصى (MP)، والتقدير البعدي المتوقع (EP) كان أيضًا لصالح طريقة التقدير البعدي الأقصى (MP)، في حين كان الفرق بين طريقتي التقدير البعدي المتوقع EP والأرجحية العظمى (الاحتمال الأقصى) ML لصالح طريقة التقدير البعدي المتوقع EP، كما يُظهر الشكل (3) فروقًا واضحة في القدرة المقدَّرة بين كل زوج من طرق التقدير، مما يعزز النتائج الكمية ويدعم وجود تأثير فعلي لاختلاف طريقة التقدير في نتائج القدرة.

جدول 6:

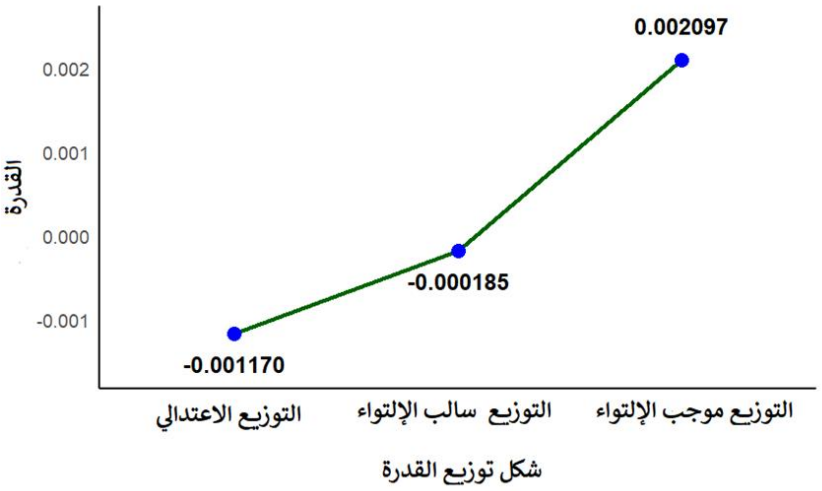
المقارنات البعدية لتوزيع القدرة في متغير القدرة:

الفرق الثنائي	فرق المتوسطين	الخطأ المعياري	مستوى الدلالة	التأثير
توزيع طبيعي * التواء موجب	-.003266839*	.0005403035	.000	مؤثر
توزيع طبيعي * التواء سالب	-.000984776	.0005403035	.205	غير مؤثر
التواء سالب * التواء موجب	-.002282063*	.0005403035	.000	مؤثر

ويمكن توضيح ذلك في شكل 4:

شكل 4:

تأثير شكل التوزيع على القدرة:



يتضح من الجدول (6) وجود فرقين ثنائيين دالين إحصائيًا لمتغير توزيع القدرة، مما يشير إلى أن اختلاف شكل توزيع القدرة يؤدي إلى تغير في القدرة المقدّرة، رغم أن حجم التأثير كان ضئيلًا جدًا (لا يُذكر)، وتظهر اتجاهات هذين الفرقين، كما هو موضح في الجدول (1)، وذلك على النحو الآتي:

الفرق بين التوزيع الاعتدالي والتوزيع الملتوي التواءً موجبًا، لصالح التوزيع الملتوي التواءً موجبًا، والفرق بين التوزيع الملتوي التواءً موجبًا والتوزيع الملتوي التواءً سالبًا، أيضًا لصالح التوزيع الملتوي التواءً موجبًا.

أما الشكل (4)، فيُظهر بوضوح هذا الفارق بين القدرة المرتبطة بالتوزيع موجب الالتواء من جهة، وكل من التوزيع سالب الالتواء والتوزيع الطبيعي (الاعتدالي) من جهة أخرى، ورغم وجود فرق ظاهري بين التوزيع سالب الالتواء والتوزيع الطبيعي (الاعتدالي) في القدرة المقدّرة، إلا أن الانحرافات المعيارية الكبيرة المصاحبة لكل توزيع، كما هو موضح في الجدول (1)، قد قللت من دلالة هذا الفرق إحصائيًا.

تفسير النتائج ومناقشتها

يتضح من جدول (2) أن تأثير النموذج على قدرات الأفراد بالكاد يُعد دالًا إحصائيًا عند مستوى (0.05)، مع حجم تأثير معدوم تقريبًا (0.000)، وبالرجوع إلى جدول (1)، نجد أن متوسطات النماذج الثلاثة في القدرة كانت على الترتيب: 0.001022، 0.000092-، و0.000188-. وتشير هذه النتائج إلى أنه كلما زاد عدد معلمات النموذج، انخفضت قيمة القدرة، إلا أن هذا الانخفاض لا يصل إلى مستوى الدلالة العملية، كما يتضح من نتائج المقارنات البعدية في جدول (3) عدم وجود فروق ثنائية دالة إحصائيًا تفسر وجود تأثير للنموذج على قدرة الأفراد، مما يشير إلى الحاجة لإجراء دراسات إضافية لتأكيد هذه النتائج أو نقضها.

كما يتضح من جدول (2) أن تأثير حجم العينة على قدرات الأفراد دال إحصائيًا عند مستوى (0.05)، إلا أن حجم التأثير يُعد ضعيفًا جدًا (0.005)، وبالرجوع إلى جدول (1)، تظهر متوسطات القدرة لأحجام العينات الثلاثة على النحو الآتي: 0.006104- لعينة حجمها 100، و0.003662 لعينة حجمها 500، و0.003185 لعينة حجمها 1000، وتشير هذه النتائج إلى أن أعلى متوسط قدرة تحقق عند حجم عينة 500، يليه 1000، ثم 100، وتظهر نتائج المقارنات البعدية في جدول (4) وجود فروق دالة إحصائيًا بين حجمَي 500 و100، كذلك بين 1000 و100، في حين لم تُسجَّل فروق دالة بين حجمَي 500 و1000، ما يشير إلى تقاربهما في التأثير، وبذلك يمكن القول إن زيادة حجم العينة حتى 500 تُسهم في تحسين القدرة المقدَّرة، بينما لا تُحدث الزيادة بعد هذا الحد تأثيرًا ملموسًا.

ويتضح أيضاً من جدول (2) أن تأثير طريقة التقدير على قدرات الأفراد دال إحصائيًا عند مستوى (0.05)، مع حجم تأثير صغير (0.039)، ويتضح من جدول (5) أن الفرق بين طريقتي الأرجحية العظمى (الاحتمال الأقصى) (ML) والتقدير البعدي الأقصى (MP) كان لصالح التقدير

البعدي الأقصى (MP)، إذ بلغت قيمة الفرق 0.0192 وهي موجبة ودالة إحصائيًا، مما يشير إلى أن طريقة التقدير البعدي الأقصى MP تنتج تقديرات أعلى نسبيًا للقدرة مقارنة بطريقة الأرجحية العظمى ML، أما المقارنة بين طريقة الأرجحية العظمى (الاحتمال الأقصى) (ML) وطريقة التقدير البعدي المتوقع (EP) فقد أظهرت فرقًا سالبًا مقداره 0.0304-، مما يعني أن طريقة التقدير البعدي الأقصى EP تعطي تقديرات أعلى من طريقة الأرجحية العظمى ML بشكل دال إحصائيًا، وأظهرت المقارنة بين طريقتي التقدير البعدي المتوقع EP وطريقة التقدير البعدي الأقصى MP فرقًا سالبًا مقداره 0.0117-، ما يشير إلى تفوق طريقة التقدير البعدي الأقصى MP في إنتاج تقديرات أعلى مقارنة بطريقة التقدير البعدي المتوقع EP، ويُعزز الشكل (3) هذا التباين بين الطرق الثلاث، إذ يُبيّن اتجاه التقديرات ومتوسطاتها، ويُلاحظ أن طريقة الأرجحية العظمى ML أعطت أدنى متوسط لقدرات الأفراد، بينما كانت طريقة التقدير البعدي الأقصى MP هي الأعلى، تليها التقدير البعدي المتوقع EP.

كما يتضح من جدول (2) أن تأثير توزيع القدرة على قدرات الأفراد دال إحصائيًا عند مستوى (0.05)، مع حجم تأثير معدوم تقريباً (0.000)، وبالرجوع إلى جدول (6)، يظهر وجود تأثير دال إحصائيًا لشكل توزيع القدرة على التقديرات، إذ أظهرت المقارنات البعدية فروقًا في بعض أشكال التوزيع، فنجد الفرق بين التوزيع الطبيعي (الاعتدالي) والتوزيع الموجب الالتواء كان دالًا إحصائيًا عند مستوى دلالة (0.000)، بقيمة فرق سالبة (0.0033–)، مما يشير إلى أن التقديرات الناتجة عن التوزيع الموجب الالتواء كانت أعلى من تلك الناتجة عن التوزيع الطبيعي، والفرق بين التوزيع الطبيعي والتوزيع السالب الالتواء لم يكن دالًا إحصائيًا (p = .205)، ما يشير إلى عدم وجود فرق جوهري بين هذين الشكلين من التوزيع في تقدير القدرة، أما الفرق بين التوزيع السالب الالتواء والتوزيع الموجب الالتواء كان دالًا إحصائيًا، وبقيمة فرق بلغت (0.0023–)، بما يدل على أن التوزيع الموجب الالتواء يؤدي إلى تقديرات قدرة أعلى من التوزيع السالب الالتواء، ويُعزز ذلك ما هو موضح في الشكل (4)، إذ يتبين أن أعلى تقديرات للقدرة ظهرت مع التوزيع الموجب الالتواء (0.002097)، في حين أن أدنى التقديرات نتجت عن التوزيع الطبيعي (0.001170–)، بينما جاء التوزيع السالب الالتواء في المنتصف تقريبًا.

التوصيات

في ضوء نتائج الدراسة التي كشفت عن وجود فروق ذات دلالة إحصائية في دقة تقدير قدرات الأفراد تعزى إلى كل من: نوع النموذج المستخدم (أحادي، ثنائي، ثلاثي الأبعاد)، وحجم العينة، وشكل توزيع القدرة، وطريقة التقدير، بالرغم من أن معظم الفروق لم ترتقِ إلى دلالة عملية، تُوصي الدراسة بما يأتي:

1. ضرورة مراعاة خصائص النموذج المستخدم في التحليل عند تقدير قدرات الأفراد، خاصة في الدراسات التي تستخدم نماذج الاستجابة للمفردة متعددة الأبعاد، لما للنموذج من تأثير ملحوظ على نتائج التقدير.

2. التأكيد على أهمية اختيار حجم عينة مناسب، إذ أظهرت النتائج أن العينات الصغيرة قد تؤدي إلى اختلافات غير دقيقة في تقدير القدرات، وإن كانت ذات دلالة إحصائية دون دلالة عملية.

3. الاهتمام بشكل توزيع القدرة عند إعداد البيانات أو تحليلها، إذ تبين أن التوزيعات الملتوية، وخاصة الملتوية التواءً موجبًا، قد تؤثر على نتائج التقدير بشكل طفيف، مما يتطلب الانتباه لهذا العامل في تصميم الدراسات المستقبلية.

4. اختيار طريقة التقدير الأنسب (ML، MP، EP) بناءً على طبيعة البيانات والغرض من الدراسة، إذ تبين أن طريقة التقدير لها تأثير على دقة التقدير، وإن كان التأثير صغيرًا.

5. تشجيع الباحثين في مجال القياس النفسي والتربوي على إجراء دراسات محاكاة إضافية تأخذ في الحسبان عوامل جديدة أو مستويات إضافية من العوامل الحالية.

6. الاعتماد على مقاييس الدلالة العملية إلى جانب الدلالة الإحصائية، لما لها من أهمية في تقديم نتائج أكثر واقعية وقابلة للتطبيق في البيئات التربوية والنفسية.

البحوث المقترحة:

1- دراسة أثر عدد معلمات النموذج، وحجم العينة، وطريقة التقدير، وشكل توزيع القدرة على دقة تقدير القدرة في النماذج متعددة الأبعاد.

2- دراسة أثر عدد معلمات النموذج، وحجم العينة، وطريقة التقدير، وشكل توزيع القدرة على دقة تقدير القدرة، باستخدام عينات واقعية.

3- تأثير متغيرات حجم العينة وشكل توزيع القدرة وطريقة التقدير على دقة تقدير النموذج رباعي المعلم.

4- مقارنة النموذج رباعي المعلم والنموذج خماسي المعلم في دقة تقدير معالم الأفراد.

قائمة المراجع

اولاً: المراجع العربية

البادية، فاطمة حمد وابن كاظم، علي بن مهدي والمحرزي، راشد بن سيف. (2018). أثر حجم العينة على دقة تقدير خصائص المفردة والقدرة في اختبار التنمية المعرفية في مادة العلوم لطلبة الصف السابع بسلطنة عمان. *مجلة دراسات، 73*، (106) ، ص، 125. مسترجع من http://search.mandumah.com/Record/94773

بيكر، فرانك. (2010). *أسس نظرية الاستجابة للفقرة*. (ترجمة عبد الرحمن الطريري والسيد محمد أبو هاشم وسوسن شلبي). مركز النشر والمطابع جامعة الملك سعود.

الحمدانية، منار مازن عبد الله و النصراوين معين سلمان سليم. (2020). مقارنة بين الطريقة البيزية وطريقة الأرجحية العظمى في دقة تقدير معلمة القدرة ومعلمة الصعوبة وفق نموذج راش باستخدام بيانات مولدة محاكاة. *مجلة جامعة عمان العربية للبحوث - سلسلة البحوث التربوية والنفسية، 4*، (1)، 111 - 144. مسترجع من http://search.mandumah.com/Record/1091790

الحواري، أروى. (2015). أثر طول الاختبار وشكل توزيع القدرة في تقديرات قدرة الأفراد وفق نموذج راش في نظرية استجابة الفقرة. مجلة جامعة النجاح للأبحاث - *العلوم الإنسانية، 29*،(8)، 1463 - 1488. مسترجع http://search.mandumah.com/Record/931120

الخرشة، طه عقلة (2023). تقصي أثر حجم العينة وعدد فقرات الاختبار على دقة تقدير معالم الفقرات وقدرات الأفراد في ضوء عدد نقاط التربيع في برنامج بايلوج. *مجلة كلية التربية 113*،(1)، 39. مسترجع من http://search.mandumah.com/Record/1458389

الخوالدة، نايف عارف والنصراوين، معين سلمان (2020). دقة تقدير معالم الفقرات عند استخدام حالتين من النموذج اللوجستي الثلاثي المعلمة *مجلة جامعة عمان العربية للبحوث - سلسلة البحوث التربوية والنفسية ،4*(1).428 ص 399. مسترجع من http://search.mandumah.com/Record/1091866

شرفاوي، *عبد الكريم وبن نابي، نصيرة (مدير بحث). (2022). طرق تقدير معالم الفقرة والقدرة وأثرها في دقة التقدير باستخدام نظرية الاستجابة لمفردة* (راش أنموذجا) (*Doctoral dissertation, University of algiers2 Abu El Kacem Saad Allah* جامعة الجزائر2 أبو القاسم سعد الله).

شكار، هاجر وبن نابي، نصيرة (2024) طرائق تقدير معالم الفقرة والقدرة في نظرية الاستجابة للمفردة وفق نموذج راش. *مجلة البحوث التربوية والتعليمية 13*، عدد خاص، 491 - 514. مسترجع من http://search.mandumah.com/Record/1469195

طيفور، مصطفى وطومان، منار وضعضع، هبة. (2020). أثر حجم العينة وخطوط التقدير في دقة التقدير معالم نموذج راش، *مجلة جرش للبحوث والدراسات، 21*(1)، 131-170.

عبد الوهاب، محمد محمود. (2019). أثر طريقة تقدير الدرجات ونموذج الاستجابة للمفردة في دقة تقدير بارامترات الأفراد والمفردات لاختبار صواب وخطأ متعدد. *المجلة الدولية للأبحاث التربوية، 43*(2)، 59-85.

العنزي، سعود شايش. (2023). أثر نماذج الاستجابة للمفردة وطريقة التقدير في دقة تقدير معالم المفردات وقدرات الأفراد لاختبار كاتل للذكاء. *مجلة جامعة أم القرى للعلوم التربوية والنفسية*. (4) 15. مسترجع من https://doi.org/10.54940/ep24224002

القضاة، محمد عواد. (2020). أثر حجم العينة في دقة تقدير قدرة الفرد ودالة معلومات الاختبار. *مجلة كلية التربية بالمنصورة، 11* (5)، 1605-1632.

كريشان، أمجد سالم والعناتي، جهاد محمد. (2023). أثر شكل التوزيع لدرجات الطلبة على الاختبارات الوطنية في دقة تقدير دالة معلومات الفقرة وثبات الاختبار. *دراسات: العلوم التربوية. 50*(2)، 102-116.

محمود، سوميه شكري. (2017). أثر شكل توزيع القدرة على ملاءمة المفردات ودقة تقدير معلم الصعوبة في نموذج راش. *دراسات عربية في التربية وعلم النفس*، (86)، 543 - 571. مسترجع من http://search.mandumah.com/Record/941548

النفيعي، عبد الرحمن عبد الله. (2024). مقارنة قوة طرق اختيار نموذج الاستجابة للمفردة عديدة الاستجابة الأكثر مطابقة للبيانات في ضوء اختلاف عدد رتب الاستجابة وشكل توزيع السمة المقاسة وحجم العينة. *مجلة جامعة تبوك للعلوم الإنسانية والاجتماعية، 4* (2)، 564 - 594 مسترجع من http://search.mandumah.com/Record/1501687

Arabic References

Al-Bādiya, Fāṭima Ḥamd, Ibn Kāẓim, ʻAlī bin Mahdī, & Al-Muḥrizī, Rāshid bin Saif. (2018). Athar ḥajm al-ʻaynah ʻalā daqqat taqdīr khaṣāʼiṣ al-mufrada wa al-qudra fī ikhtibār al-tanmiyya al-maʻrifiyya fī mādda al-ʻulūm li-ṭullāb al-ṣaff al-sābiʻ bi-Sulṭanat ʻUmān. *Majallat Dirāsāt*, 73(106), 125. Retrieved from http://search.mandumah.com/Record/94773

Baker, Frank. (2010). *Usūs naẓariyyat al-istiǧāba li-l-faqra* (Trans. ʻAbd al-Raḥmān al-Ṭurayrī, al-Sayyid Muḥammad Abū Hāshim, & Sawsan Shalabī). Riyadh: Markaz al-Nashr wa al-Maṭābiʻ, Jāmiʻat al-Malik Saʻūd.

Al-Ḥamdāniyya, Manār Māzin ʻAbd Allāh, & Al-Naṣrāwīn, Muʻayyin Salmān Salīm. (2020). Muqārana bayna al-ṭarīqa al-Bayziyya wa al-ṭarīqa al-arjaḥiyya al-ʻuẓmā fī daqqat taqdīr muʻallamat al-qudra wa

mu'allamat al-ṣu'ūba wa-fq model Rāsh bi-istikhḍām bayānāt muwallaḍa muḥākā. *Majallat Jāmi'at 'Umān al-'Arabiyya li-l-Buḥūth – Silsilat al-Buḥūth al-Tarbawiyya wa al-Nafsiyya*, 4(1), 111–144. Retrieved from http://search.mandumah.com/Record/1091790

Al-Ḥawwārī, Arwā. (2015). Athar ṭūl al-ikhtibār wa shakl tawzī' al-qudra fī taqdīrāt qudrat al-afrād wa-fq model Rāsh fī naẓariyyat al-istiğāba li-l-faqra. *Majallat Jāmi'at al-Najāḥ li-l-Buḥūth – al-'Ulūm al-Insāniyya*, 29(8), 1463–1488. Retrieved from http://search.mandumah.com/Record/931120

Al-Kharsha, Ṭaha 'Uqla. (2023). Taqṣī athar ḥajm al-'aynah wa 'adad faqarāt al-ikhtibār 'alā daqqat taqdīr mu'allam al-faqrāt wa qudrāt al-afrād fī ḍaw' 'adad nuqāṭ al-tarbī' fī barnāmaj Baylūj. *Majallat Kulliyyat al-Tarbiyya*, *113*(1), 39. Retrieved from http://search.mandumah.com/Record/1458389

Al-Khawālda, Nāyif 'Ārif, & Al-Naṣrāwīn, Mu'ayyin Salmān. (2020). Daqqat taqdīr mu'allam al-faqrāt 'ind istikhdām ḥālatayn min al-model al-lūjistī al-thulāthī al-mu'allama. *Majallat Jāmi'at 'Umān al-'Arabiyya li-l-Buḥūth – Silsilat al-Buḥūth al-Tarbawiyya wa al-Nafsiyya*, *4*(1), 399–428. Retrieved from http://search.mandumah.com/Record/1091866

Sharafāwī, 'Abd al-Karīm, & Bin Nābī, Naṣīra (Dir. of research). (2022). Ṭurūq taqdīr mu'allam al-faqra wa al-qudra wa atharuhā fī daqqat al-taqdīr bi-istikhdām naẓariyyat al-istiğāba li-mufrada (Rāsh Model). Doctoral dissertation, University of Algiers 2 Abu El Kacem Saad Allah.

Shkār, Hājar, & Bin Nābī, Naṣīra. (2024). Ṭarā'iq taqdīr mu'allam al-faqra wa al-qudra fī naẓariyyat al-istiğāba li-l-mufrada wa-fq model Rāsh. *Majallat al-Buḥūth al-Tarbawiyya wa al-Ta'līmiyya*, 13(Special Issue), 491–514. Retrieved from http://search.mandumah.com/Record/1469195

Ṭayfūr, Muṣṭafā, Ṭūmān, Manār, & Ḍa'ḍa', Hiba. (2020). Athar ḥajm al-'aynah wa khuṭūṭ al-taqdīr fī daqqat taqdīr mu'allam model Rāsh. *Majallat Jarash li-l-Buḥūth wa al-Dirāsāt*, *21*(1), 131–170.

'Abd al-Wahhāb, Muḥammad Maḥmūd. (2019). Athar ṭarīqat taqdīr al-darajāt wa model al-istiğāba li-l-mufrada fī daqqat taqdīr barāmitr al-afrād wa al-mufradāt li-ikhtibār ṣawāb wa khaṭa' muta'addid. *Al-Majalla al-Dawliyya li-l-Buḥūth al-Tarbawiyya*, *43*(2), 59–85.

Al-'Anzī, Sa'ūd Shāyish. (2023). Athar namādhij al-istiğāba li-l-mufrada wa ṭarīqat al-taqdīr fī daqqat taqdīr mu'allam al-mufradāt wa qudrāt al-afrād li-ikhtibār Kātell li-l-ḏakā'. *Majallat Jāmi'at Umm al-Qurā li-l-'Ulūm al-Tarbawiyya wa al-Nafsiyya*, 4(15). Retrieved from https://doi.org/10.54940/ep24224002

Al-Quḍāh, Muḥammad 'Awād. (2020). Athar ḥajm al-'aynah fī daqqat taqdīr qudrat al-fard wa dāllat ma'lūmāt al-ikhtibār. *Majallat Kulliyyat al-Tarbiyya bi-l-Manṣūra*, 11(5), 1605–1632.

Krīshān, Amjad Sālim, & Al-'Anātī, Jihād Muḥammad. (2023). Athar shakl al-tawzī' li-darajāt al-ṭullāb 'alā al-ikhtibārāt al-waṭaniyya fī daqqat taqdīr dāllat ma'lūmāt al-faqra wa thabāt al-ikhtibār. *Dirāsāt: al-'Ulūm al-Tarbawiyya*, 50(2), 102–116.

Maḥmūd, Sūmiyya Shukrī. (2017). Athar shakl tawzīʿ al-qudra ʻalā mulāʾamat al-mufradāt wa daqqat taqdīr muʻallam al-ṣuʻūba fī model Rāsh. *Dirāsāt ʻArabiyya fī al-Tarbiyya wa ʻIlm al-Nafs*, 86, 543–571. Retrieved from http://search.mandumah.com/Record/941548

Al-Nafiʻī, ʻAbd al-Raḥmān ʻAbd Allāh. (2024). Muqārana quwwat ṭurūq ikhtiyār model al-istiğāba li-l-mufrada al-ʻadīda al-istiğāba al-akthar muṭābaqa li-l-bayānāt fī ḍawʾ ikhtilāf ʻadad rūtub al-istiğāba wa shakl tawzīʿ al-simma al-maqūsa wa ḥajm al-ʻaynah. *Majallat Jāmiʻat Tabūk li-l-ʻUlūm al-Insāniyya wa al-Ijtimāʻiyya*, 4(2), 564–594. Retrieved from http://search.mandumah.com/Record/1501687

ثانياً: المراجع الانجليزية

Al-Tarawnah, E., & Al-Qahtani, M. (2022). The effect of test length on the accuracy of estimating ability parameter in the two and three parameter logistic models: comparison by using the Bayesian method of expected prior mode and maximum likelihood estimation. *Journal of Educational and Social Research, 12*.(1) DOI: https://doi.org/10.36941/jesr-2022-0015

Birnbaum, A. (1968). *Some latent trait models and their use in inferring an examinee's ability*. In F. M. Lord & M. R. Novick (Eds.), Statistical theories of mental test scores (pp. 397–479). Addison-Wesley.

Choi, S. W., Cook, K. F., & Dodd, B. G. (1997). Parameter recovery for the partial credit model using MULTILOG. *Journal of Outcome Measurement, 1*,(2) 114-142.

Dai, S., Vo, T., Kehinde, O., He, H., Xue, Y., Demir, C., & Wang, X. (2021). *Performance of polytomous IRT models with rating scale data: An investigation over sample size, instrument length, and missing data*. Frontiers in Education.

Embretson, S.E., & Reise, S.P. (2000). *Item Response Theory for Psychologists* (1st ed.). Psychology Press. https://doi.org/10.4324/9781410605269

Fathi, R. Al-alem, M. (2022). Accuracy Of Estimating Item Parameters And Individual' Ability Of Three Parameter Item Response Theory Model Using Joint Maximum Likelihood. *Journal of Positive School Psychology*. 2022,. 6,(8),pp. 1744-1759.

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Sage Publications.

Fisher, R. A. (1922). On the interpretation of χ^2 from contingency tables, and the calculation of P. *Journal of the Royal Statistical Society, 85*(1), 87–94. https://doi.org/10.2307/2340521

Fisher, R. A. (1925). *Statistical methods for research workers*. Oliver and Boyd.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1979). *Fundamentals of item response theory*. Sage Publications.

Law, A. M., & Kelton, W. D. (2000). *Simulation modeling and analysis*. (3rd ed.). McGraw-Hill.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Lord, F. M. (1953). *On the statistical treatment of football numbers. Psychometrika, 18*(1), 31–44. https://doi.org/10.1007/BF02289259

Lord, F. M. (1980). *Applications of item response theory to practical testing problems. Lawrence Erlbaum Associates.* https://www.winsteps.com/a/Linacre-MFRM-book.pdf

Lord, F. M. (1983). *Applications of item response theory to practical testing problems. Hillsdale, NJ: Lawrence Erlbaum Associates.*

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement, 17*(4). https://www.psychometricsociety.org/sites/main/files/file-attachments/mn17.pdf

Spearman, C. (1904). "General intelligence," objectively determined and measured. *The American Journal of Psychology*, *15*(2), 201–292. https://doi.org/10.2307/1412107

Thorndike, E. L. (1910). *Journal of Educational Psychology: 1.*(1) pp,5-12Teachers College, Columbia University. https://psycnet.apa.org/record/1926-01409-001

Thorndike, E. L. (1914). *Educational psychology: Volume III—Mental work and fatigue and individual differences.* Teachers College, Columbia University.

Thurstone, L. L. (1931). Multiple factor analysis. *Psychological Review*, 38(5), pp,406–427. https://doi.org/10.1037/h0071093

Thurstone, L. L. (1935). *The vectors of mind: Multiple-factor analysis for the isolation of primary traits.* University of Chicago Press.

Wright, B. D., & Stone, M. H. (1979). *Best test design.* MESA Press.