مجلة جامعة ام القرى للعلوم التربوية والنفسية

الموقع الإلكتروني: https://uqu.edu.sa/jep

أثر نماذج نظرية الاستجابة للمفردة وطريقة التقدير في دقة تقدير معالم المفردات وقدرات الأفراد لاختبار كاتل للذكاء

Effect Item Response Theory Models and Estimation Method on Accuracy Estimating Item, Ability Parameters of Cattell Intelligence Test

د. سعود بن شايش بشير العنزي*

أستاذ مشارك في القياس والتقويم التربوي، كلية التربية والآداب، جامعة الحدود الشمالية، عرعر، المملكة العربية السعودية.

Dr. Saud Bin Shaish Basher Alenezi*

Associate Professor of Educational Measurement and Evaluation, College of Education and Arts, Northern Border University, Arar, Kingdom of Saudi Arabia.

تاريخ التقديم: 2023/1/25 تاريخ القبول:2023/3/15

Received:25/1/2023 Accepted: 15/3/2023

الملخص: هدف البحث الحالي 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, بلغ المتوسط الحسابي لأعمارهم الزمنية (20.80) 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. وجود فروق ذات دلالة إحصائية في دقة تقدير بارامتر التمييز، ترجع لاختلاف النموذج المستخدم لصالح النموذج الثنائي. عدم وجود فروق ذات دلالة إحصائية في دقة تقدير بارامتر التمييز، ترجع لاختلاف طريقة التقدير، أو التفاعل بين النموذج المستخدم وطريقة التقدير. وجود 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.

الكلمات المفتاحية: نظرية الاستجابة للمفردة، طريقة التقدير، دقة التقدير، اختبار كاتل للذكاء.

Abstract: The current research aimed to identify the effect of selecting item response theory models and estimation methods and the interaction between these two variables on the accuracy of estimating item parameters of the Cattell intelligence test and the abilities of individuals. The research instruments used were Cattell's third-scale intelligence test (form A). The participants were (1104) male and female participants from Northern Border University in the Kingdom of Saudi Arabia; the mean for their ages was (20.80) with a standard deviation of (1.31). Using the statistical programs (SPSS- BILOG-MG) the statistical techniques (Two-way ANOVA, independent samples t-test, Scheffe test), the results revealed that there were statistically significant differences in the accuracy of difficulty parameter estimation due to the difference in the model used in favor of one model. There were no statistically significant differences in the accuracy of difficulty parameter estimation due to the difference in the estimation method or the interaction between the model and the estimation method. There were statistically significant differences in the accuracy of discrimination parameter estimation due to the difference in the model used in favor of the two models. There were no statistically significant differences in the accuracy of discrimination parameter estimation due to the difference in the estimation method or the interaction between the model and the estimation method. There were statistically significant differences in the abilities of individual estimation accuracy due to the difference in the model, estimation method, the interaction between the model used and the estimation method.

Keywords: Item response theory, estimation method, estimation accuracy, Cattell intelligence test.

Doi: https://doi.org/10.54940/ep24224002

***معلومات التواصل** : سعود بن شايش بشير العنزي

البريد الالكتروني الرسمي : dr.saud.shaish@gmail.com

مقدمة

اعتمد العديد من المهتمين بمجال القياس والتقويم التربوي على مبادئ وأسس النظرية الكلاسيكية Classical Theory في بناء الاختبارات والمقاييس وتفسير درجاتها, على الرغم من تعدد أوجه القصور في هذه النظرية, حيث إن إحصاءات المفردات الاختبارية وخصائص الاختبار، تعتمد بدرجة كبيرة على خصائص الأفراد الذين طُبقت عليهم هذه المفردات, ومن ثَمَّ فإن مفردات الاختبار ستكون سهلة جدًّا إذا طُبقت على مجموعة من الطلاب مرتفعي القدرة, وستكون صعبة إذا طُبقت على مجموعة من الطلاب محدودي القدرة, وبسبب تأثر خصائص المفردة بعينة الأفراد المفحوصين؛ فإنه يكون من الصعب تقييم جودة المفردة الاختبارية (Mayer & Zhu, 2013).

ونتيجة لأوجه القصور الموجودة في النظرية الكلاسيكية؛ فقد اتجه علماء القياس والتقويم إلى البحث عن طرائق جديدة في القياس تحقق الموضوعية، وتملك القدرة على التعامل مع متطلبات القياس الجديدة, وقد أثمر هذا البحث عن ظهور نظرية الاستجابة للمفردة Item Response Theory (IRT) (الحولي, 2013).

وتعتمد نظرية الاستجابة على افتراضات قوية مقارنة بالنظرية الكلاسيكية، حيث إن هناك استقلالية للمفردات الاختبارية في قياس القدرات المختلفة (استقلال خصائص مفردات الاختبار عن قدرات الأفراد, والعكس), وأن لكل مفردة اختبارية من مفردات الاختبار خصائصها السيكومترية الخاصة بها, والممثلة في منحنى مُميز لها، يصف احتمالية الحصول على الإجابة الصحيحة أو الخاطئة، بناء على قدرة الفرد وخصائص المفردات التي اختبر بها (Magno, 2009).

وتُعبر نظرية الاستجابة للمفردة عن مجموعة من النماذج الرياضية التي تحدد العلاقة الاحتمالية بين الاستجابات الملاحظة على الاختبار, والسمات الكامنة المسببة لها, التي تُحدد استجابات الأفراد على مفردات الاختبار, وتمثل السمة ببعد كمي متصل، يمكن أن يحدد عليه مواقع الأفراد المستجيبين على مفردات هذا الاختبار, وتظهر هذه النماذج بيانيًّا من خلال دوال رياضية، تربط احتمال الإجابة الصحيحة على المفردة الاختبارية بالقدرة التي تقيسها هذه المفردات (بركات, 2010). وتكمن السمة الرئيسة لنظرية الاستجابة للمفردة، في النظر إلى كل مفردة على حدة دون الاعتماد على الدرجة الكلية, كما أنها تتميز بمبدأ عدم التغير؛ أي إن بارامترات المفردات تكون مستقلة عن قدرات الأفراد, كما أن قدرات الأفراد تكون مستقلة عن بارامترات المفردات, وتتميز أيضًا بإمكانية مقارنة السمات الكامنة للأفراد من مجموعات مختلفة، عندما تطبق عليهم اختبارات ومقاييس تحتوي على مفردات مشتركة, ومقارنة أفراد من المجموعات نفسها عندما تقدم لهم اختبارات مختلفة (Bortolotti et al. 2013).

وتعتمد نظرية الاستجابة للمفردة على مجموعة من الافتراضات التي يجب تحققها في البيانات؛ والتي تتمثل في: افتراض أحادية البعد, والاستقلال الموضعي, ومنحنى خاصية المفردة, والتحرر من السرعة (Reise & Waller, 2003).

ولنظرية الاستجابة للمفردة العديد من التطبيقات المهمة في مجال القياس والتقويم, منها بناء الاختبارات والمقاييس وتحليل مفرداتها, والكشف عن الأداء التفاضلي لمفردات الاختبار, وبناء بنوك الأسئلة, وتشكيل أسس الاختبارات الموائمة المحوسبة, وتدريج مفردات الاختبار وتقنينها, ومعادلة درجات الصور المختلفة للاختبارات (Embretson&Reise, 2000).

وبالرغم من انتشار استخدم نظرية الاستجابة للمفردة في بناء وتطوير العديد من الاختبارات في مجال القدرات العقلية؛ إلا أن الركيزة الأساسية التي يتوقف عليها استخدام هذه النظرية، هي البحث عن العوامل المؤثرة في استخراج معالم مفردات الاختبار والأفراد, ومن هذه العوامل اختيار النموذج الملائم لتدريج مفردات الاختبار, وتحديد طريقة التقدير المستخدمة في استخراج هذه المعالم.

حيث توجد نماذج متعددة لنظرية الاستجابة للمفردة, وذلك لاختلاف الافتراضات المتعلقة بالبيانات الاختبارية, إذ يُفترض أن الأداء في الاختبار قد ينطوي على سمة أحادية البعد, أو على سمة متعددة الأبعاد, لذا؛ يمكن تقسيم هذه النماذج إلى قسمين: أحدهما يسمى نماذج نظرية الاستجابة للمفردة أحادية البعد, والآخر يسمى نماذج نظرية الاستجابة للمفردة متعددة الأبعاد, وينقسم كل منهما إلى نماذج ثنائية الاستجابة ونماذج متعددة الاستجابة (علام,2005). لذا؛ توجد نماذج تتعلق بالمفردات ثنائية الاستجابة مثل النموذج اللوغاريتمي أحادي البارامتر (نموذج راش)، الذي يفترض تساوي معاملات التمييز لجميع المفردات, وانعدام التخمين من قبل الأفراد المفحوصين, فهو يسمح فقط لبارامتر الصعوبة أن يأخذ قيمًا مختلفة, أي إنه يقوم بتقدير بارامتر الصعوبة فقط (Baker, 2001)، والنموذج اللوغاريتمي ثنائي البارامتر (نموذج لورد)، يتضمن بارامترين لتمثيل خصائص المفردة, هما: بارامتر صعوبة المفردة, وبارامتر تمييز المفردة, ومن ثَمَّ فإن تأثير الفرق بين مستوى القدرة وصعوبة المفردة على احتمال الإجابة الصحيحة، يعتمد بدرجة كبيرة على قدرة المفردة على التمييز (Embretson&Reise, 2000), والنموذج اللوغاريتمي ثلاثي البارامتر (نموذج بيرنبوم)، الذي أضاف بارامترًا ثالثًا إلى خصائص المفردات الاختبارية، التي تتطلب الإجابة عليها الاختيار من بين بدائل معطاة, وأطلق عليه بارامتر التخمين, وهذا البارامتر يحدد احتمال أن يجيب فرد يُفترض أن مستوى قدرته منخفض، ومع ذلك يجيب إجابة صحيحة عن غالبية مفردات الاختبار عن طريق التخمين(Wolf, 2013) .

كما يمكن تقدير القدرة وفق نظرية الاستجابة للمفردة بطرق تعتمد

الأرجحية القصوى Maximum Likelihood وتعتمد هذه الطرق على نمط إجابة الفرد, ولا تسمح بتقدير قدرته إلا بعد أن يجيب على مفردة واحدة على الأقل, وتقوم طريقة الأرجحية القصوى بتسجيل وتقدير مستويات قدرة الفرد بمعلومية بارامترات المفردة حسب النموذج المستخدم, وتأخذ في الاعتبار إجابة الفرد إجابة صحيحة عن المفردة الاختبارية, وتوفر العديد من الفروق بين الأفراد أكثر من مجرد تصحيح المفردات, وذلك بسبب دمج المعلومات المتعلقة بنمط استجاباتهم مع بارامترات كل مفردة, كما أنها تقدر خطأ معياريًا لكل فرد، اعتمادًا على كيفية إجابته على مجموعة معينة من المفردات، التي يمكن تحديدها من أي مجموعة فرعية من المفردات التي استجاب لها، بشرط أن يتم تقدير بارامترات المفردات على الميزان نفسه (Weiss, 2004). وكذلك يمكن تقدير القدرة بطرق تعتمد على نظرية بييز Bayes Theory وتتطلب هذه الطرق افتراض احتمالات قبلية للبارامترات، في ضوء اعتبارات نظرية أو إمبريقية, أي إن هذه الطرق تفترض توزيعًا قبليًا للقدرة، ثم تقوم بتعديل هذا التوزيع بناء على استجابة الأفراد على المفردات الاختبارية, وهذا يعني أنه إذا توافرت معلومات عن توزيع القدرة لدى مجموعة من الأفراد المختبرين, فإن الأساليب التي تعتمد على نظرية بييز (أسلوب التقدير البعدي المتوقع Expected A Posterion, وأسلوب التقدير البعدي الأقصى (Maximum A Posterion)، يمكن أن تساعد في الحصول على تقديرات جيدة لهذه القدرة, إذ يمكن قبل تقدير مستويات قدرة مجموعة معينة، افتراض أن قدرة أي فرد في المجموعة لا يختلف عن قدرة الأفراد الآخرين. (Cella et al., 2007).

وقد اهتمت بعض الدراسات والبحوث السابقة بدراسة بعض العوامل المؤثرة على دقة تقديرات بارامترات مفردات الاختبار وتقدير قدرات الأفراد, حيث هدفت دراسة (Kim, 2001) إلى مقارنة دقة تقديرات معالم المفردات والقدرات التي تم الحصول عليها باستخدام طريقة الأرجيحة القصوى الشرطية والهامشية والمشتركة، وطريقة بييز التي تعتمد على توقع التوزيع البعدي, وقد اعتمدت الدراسة على بيانات مولدة بأحد برامج المحاكاة, وتوصلت الدراسة إلى أن نتائج تقديرات معالم المفردات ومعالم قدرات الأفراد، كانت متشابهة في الطرق الأربع للتقدير.

وهدفت دراسة (Garre&Vermunt, 2006) الكشف عن دقة تقديرات معالم المفردات وقدرات الأفراد، باختلاف الطريقة المستخدمة في تقدير هذه المعالم, حيث استخدمت الدراسة في جزء منها بيانات حقيقية بهدف التحقق من أهدافها, وقد توصلت الدراسة إلى أن تقديرات معالم المفردات وقدرات الأفراد التي تمَّ تقديرها بواسطة طريقة بييز، تتمتع بثبات أفضل من التقديرات التي تم الحصول عليها بطريقة الأرجحية القصوى،خصوصًا عند تقدير القدرات على أطراف متصلة القدرة.

وهدفت دراسة (سرحان, 2018) إلى المقارنة بين طريقتي الأرجحية القصوى والقيمة العظمى للتوزيع البعدي في دقة تقدير القدرة، وفق نظرية الاستجابة للمفردة. ولتحقيق أهداف الدراسة، اعتمدت الدراسة على تحليل نتائج اختبار تفسير الرموز الصادر من مركز التعلم السريع, وتوصلت الدراسة إلى وجود فروق في تقدير القدرة بين الطريقتين، ولصالح طريقة القيمة العظمى للتوزيع البعدي، وذلك من خلال انخفاض قيم الخطأ المعياري لها.

كما هدفت دراسة (ضعضع وآخرون, 2020) إلى الكشف عن أثر بعض طرق التقدير عند حجوم عينات مختلفة على تقدير معالم الأفراد والمفردات ودقتها، في ضوء نظرية الاستجابة للمفردة وفق نموذج راش. ولتحقيق أهداف الدراسة، تم توليد بيانات عينة عشوائية باستخدام برنامج Wingen, وقد أسفرت النتائج عن عدم وجود فروق ذات دلالة إحصائية في تقدير معالم المفردة ومعالم الأفراد، ترجع إلى اختلاف حجم العينة وطريقة التقدير والتفاعل بينهما, ووجود فروق ذات دلالة إحصائية في دقة تقدير معالم المفردة ومعالم الأفراد، ترجع إلى اختلاف حجم العينة وطريقة التقدير والتفاعل بينهما.

وهدفت دراسة (Almaleki&Alomrany, 2021) إلى التحقق من أثر طريقة التقدير على دقة تقدير معالم المفردات والأفراد باستخدام النموذج ثلاثي البارامتر. ولتحقيق أهداف الدراسة، تم تصميم اختبار تحصيلي في الكيمياء, وبعد تحليل البيانات توصلت الدراسة إلى أن طريقة التوزيع البعدي المتوقع أظهرت درجة عالية من الدقة في بارامتر الصعوبة وقدرات الأفراد، مقارنة بطريقة الأرجحية القصوى, وعدم وجود فروق ذات دلالة إحصائية في دقة تقدير بارامتري التمييز والتخمين يرجع إلى اختلاف طريقة التقدير.

وهدفت دراسة (القرني, 2022) إلى التعرف على أثر طرق تقدير القدرة للأفراد على دقة معادلة الاختبارات في نظرية الاستجابة للمفردة. ولتحقيق هدف الدراسة، تم استخدام اختبار أوتيس لينون للقدرات, وتم التقدير للقدرة من خلال أربع طرق للتقدير، هي: طريقة الأرجحية العظمى، والتوقع البعدي، والأرجحية العظمى الموزونة، والطريقة البييزية. وقد تم استخدام محك الخطأ المعياري للحكم على الدقة, وقد أظهرت النتائج أن طريقة التقدير بالأرجحية العظمى الموزونة، أفضل نتائج، وأقل خطأ معياريًا مقارنة بباقي الطرق.

ويتضح مما سبق، اهتمام الدراسات والبحوث السابقة بأثر طرق التقدير في تقديرات بارامترات المفردات والأفراد, وقد تباينت نتائج هذه الدراسات والبحوث حول وجود تأثير دال إحصائيًا بين طرق التقدير وعدم وجود تأثير, ويتفق هدف البحث الحالي مع بعض أهداف هذه الدراسات في اهتمامه بالتعرف على أثر طرق التقدير, مراعيًا في ذلك نموذج نظرية الاستجابة للمفردة المستخدم في تقدير بارامترات المفردات، ومدى

ملاءمتها لهذا النموذج. وفي ضوء نماذج نظرية الاستجابة للمفردة, وطرق تقدير معالم المفردات والأفراد, يسعى البحث الحالي إلى التعرف على أثر اختيار نموذج نظرية الاستجابة للمفردة، وطريقة التقدير في دقة تقدير معالم مفردات الاختبار والأفراد.

مشكلة البحث:

تعتمد النظرية الكلاسيكية على مجموعة قليلة من الافتراضات السطحية التي يسهل تحقيقها في الاختبارات؛ وقد أدى ذلك إلى ظهور كثير من الأخطاء في القياس، كحصول الفرد على نسبتين مختلفتين في الذكاء مثلًا، إذا تم اختباره باختبارين مختلفين يقيسان الذكاء, وذلك لاختلاف درجة الخطأ المعياري في الحالتين (Domino& Domino, 2006).

ونتيجة لقلة عدد الافتراضات التي تقوم عليها النظرية الكلاسيكية في القياس وضعف هذه الافتراضات؛ فقد استخدم معظم الباحثين النظرية الكلاسيكية في التحقق من جودة الاختبارات والمقاييس النفسية, وذلك بسبب سهولة تطبيقها في العديد من المواقف الاختبارية وبساطة مخرجاتها, ففي حين تركز نظرية الاستجابة للمفردة بشكل أساسي على مستوى كل مفردة من مفردات الاختبار، تركز النظرية الكلاسيكية على مستوى الاختبار ككل (Eluwa, 2011).

وفي ظل التقدم التكنولوجي في إمكانيات الحاسوب وعملية البرمجة ووفرة العديد من البرامج الإحصائية, فقد اتضح فاعلية نظرية الاستجابة للمفردة في التغلب على كثير من مشكلات القياس النفسي، والتغلب على بعض القيود التي واجهت استخدام النظرية الكلاسيكية في بناء وتطوير الاختبارات والمقاييس النفسية (Aune et al., 2020).

وقد أشارت العديد من البحوث والدراسات السابقة إلى ضرورة استخدام نماذج نظرية الاستجابة للمفردة في بناء وتطوير الاختبارات والمقاييس: حيث يشير Gnambs and Batinic (2011) إلى ضرورة الاعتماد على نظرية الاستجابة للمفردة في البحوث السيكومترية المتعمقة لتقييم الخصائص السيكومترية لأدوات القياس والتقييم, كما

يوصي Yang and Kao (2014) باستخدام نظرية الاستجابة للمفردة للتحقق من صدق الاختبارات؛ كونها توفر طرقًا متعددة لحل الكثير من تحديات القياس، التي يجب معالجتها عند بناء أو تطوير اختبار أو مقياس.

ويرى Holmefyr et al. (2015) أنه ينبغي عدم الاكتفاء باستخدام أساليب النظرية الكلاسيكية في التحقق من الخصائص السيكومترية للاختبارات والمقاييس النفسية, بل يجب أن يمتد التحليل إلى الاعتماد على نماذج نظرية الاستجابة للمفردة, وبخاصة الاختبارات التي تستخدم على نطاق واسع في البحوث النظرية والجوانب التطبيقية في تشخيص وتصنيف قدرات الأفراد.

وقد تباينت نتائج الدراسات في اختلاف بارمترات مفردات الاختبار باختلاف طرق التقدير, ففي حين أشارت دراسة كل من (Kim, 2001؛ ضعضع وآخرون, 2020) إلى عدم وجود فروق في تقدير بارامتر الصعوبة يرجع إلى اختلاف طريقة التقدير, أشارت نتائج دراسة كل من (Garre & Vermunt, 2006؛ سرحان, 2018؛ Almaleki & Alomrany, 2021) إلى وجود فروق في تقدير بارامتر الصعوبة يرجع إلى اختلاف طريقة التقدير.

وتعد اختبارات الذكاء من أكثر الاختبارات شيوعًا واستخدامًا في مجال تشخيص وتصنيف قدرات الأفراد, ونظرًا لتعدد النظريات المفسرة للذكاء؛ فقد تعددت الاختبارات التي تقيسه، والتي صنفت إلى اختبارات فردية واختبارات جماعية حسب طريقة إجرائها, كما صنفت إلى اختبارات لغوية وأخرى أدائية حسب محتواها, في حين صنفها البعض إلى اختبارات متحيزة ثقافيًا، وأخرى متحررة من أثر الثقافة.

وتسعى الاختبارات المتحررة من أثر الثقافة إلى التحكم في الأبعاد التي تتألف منها الثقافة, والتي تصلح لأغراض المقارنة بين الثقافات Cross Cultural, ويعد اختبار كاتل للذكاء أحد الاختبارات المتحررة من أثر الثقافة، الذي أعده العالم الأمريكي ريموند كاتل Cattell, ولها ثلاثة مستويات (مقاييس) تغطي مراحل عمرية متعددة, لكل مقياس من هذه المقاييس صورتان متكافئتان, وتصلح تلك الاختبارات للتطبيق الفردي والجمعي, وقد ترجم اختبار كاتل للذكاء إلى اللغة العربية، وأجريت عليها العديد من الدراسات والبحوث للتحقق من صدقها وثباتها (أبو حطب وآخرون, 2008).

ولقد اعتمدت معظم هذه الدراسات والبحوث على النظرية الكلاسيكية في التحقق من الخصائص السيكومترية لاختبار كاتل للذكاء, بالرغم من وجود بعض القيود المرتبطة بنظرية القياس الكلاسيكية في استخراج خصائص مفردات الاختبار وقدرات الأفراد, حيث يشير كل من Avcu (2021), Bortolotti et al. (2013) إلى أنه نتيجة لوجود العديد من أوجه القصور في تقدير خصائص مفردات الاختبارات بالاعتماد على النظرية الكلاسيكية؛ فإن استخدام نظرية الاستجابة للمفردة في إعداد وتطوير الاختبارات والمقاييس والتحقق من صدقها وثباتها وتحديد بارامتراتها، يعد اتجاهًا عالميًا, ويستخدم على نطاق واسع في مجال التربية, فهي تقدر قدرة الأفراد بدقة مع مراعاة بارامترات مفردات الاختبار (الصعوبة, والتمييز, التخمين), كما أنها تسمح بتدريج المفردات والأفراد على الميزان نفسه، مما يسمح بإجراء مقارنات موضوعية بين الأفراد, وبين المفردات, وبين المفردات والأفراد معًا.

ويتضح مما سبق أهمية استخدام نماذج نظرية الاستجابة للمفردة في تطوير الاختبارات والمقاييس النفسية، خاصة اختبارات القدرات التي تستخدم في

أغراض التشخيص والتصنيف, ونتيجة لكون اختبار كاتل للذكاء أحد الاختبارات التي تستخدم على نطاق واسع في قياس القدرات العقلية؛ كونه أحد اختبارات الذكاء المتحررة من أثر الثقافة, فإن البحث الحالي يسعى إلى التعرف على أثر اختيار نماذج نظرية الاستجابة للمفردة، وطرق التقدير والتفاعل بينهما في دقة تقدير معالم مفردات اختبار كاتل للذكاء وقدرات الأفراد.

وبناءً على ما سبق، يمكن صياغة مشكلة البحث الحالي بالتساؤلات التالية:

1- ما أثر اختيار نموذج نظرية الاستجابة للمفردة (الأحادي, والثنائي, والثلاثي) وطريقة التقدير (الأرجحية القصوى, التقدير البعدي المتوقع, التقدير البعدي الأقصى)، والتفاعل بينهما في دقة تقدير بارامترات مفردات اختبار كاتل للذكاء؟

2- ما أثر اختيار نموذج نظرية الاستجابة للمفردة (الأحادي, والثنائي, والثلاثي) وطريقة التقدير (الأرجحية القصوى, التقدير البعدي المتوقع, التقدير البعدي الأقصى)، والتفاعل بينهما في دقة تقدير قدرات الأفراد؟

فروض البحث:

يسعى البحث الحالي إلى التحقق من الفروض التالية:

1- لا يوجد فروق ذات دلالة إحصائية في دقة تقدير بارامتر الصعوبة، ترجع إلى أثر اختيار نموذج نظرية الاستجابة للمفردة وطريقة التقدير والتفاعل بينهما.

2- لا يوجد فروق ذات دلالة إحصائية في دقة تقدير بارامتر التمييز، ترجع إلى أثر اختيار نموذج نظرية الاستجابة للمفردة وطريقة التقدير والتفاعل بينهما.

3- لا يوجد فروق ذات دلالة إحصائية في دقة تقدير قدرات الأفراد، ترجع إلى أثر اختيار نموذج نظرية الاستجابة للمفردة وطريقة التقدير والتفاعل بينهما.

هدف البحث:

يهدف البحث الحالي إلى التعرف على أثر اختيار نموذج نظرية الاستجابة للمفردة وطرق التقدير، والتفاعل بينهما في دقة تقدير معالم مفردات اختبار كاتل للذكاء وقدرات الأفراد.

أهمية البحث:

يمكن أن تتضح أهمية البحث الحالي في الأهمية النظرية والأهمية التطبيقية كما يلي:

- الأهمية النظرية:

1- إلقاء الضوء على نظرية الاستجابة للمفردة وتطبيقاتها في عملية بناء وتطوير الاختبارات والمقاييس النفسية؛كونها أحد توجهات القياس النفسي الحديث.

2- يربط البحث الحالي بين متغيرات مهمة تؤثر على دقة تقدير بارامترات مفردات الاختبارات وتقدير قدرات الأفراد، وهما نماذج نظرية الاستجابة للمفردة, وطرق التقدير.

- الأهمية التطبيقية:

1- تطوير مقياس كاتل للذكاء واستخراج بارامترات مفرداته، باستخدام نماذج نظرية الاستجابة للمفردة أحادية البعد؛ كونه أحد الاختبارات المتحررة من أثر الثقافة، الذي يستخدم على نطاق واسع في كثير من الدراسات والبحوث.

2- تقديم دلالات إمبريقية يمكن استخدامها في اختيار النموذج المناسب لتقدير بارامترات مفردات اختبار كاتل وطرق التقدير، من أجل الحصول على بارامترات للمفردات والأفراد أكثر دقة.

حدود البحث:

تتمثل حدود الدراسة الحالية فيما يلي:

- حدود زمانية: تتمثل في فترة تطبيق اختبار كاتل للذكاء خلال العام الجامعي 2022/ 2023م.

- حدود مكانية: تتمثل في كليات جامعة الحدود الشمالية في منطقة الحدود الشمالية بالمملكة العربية السعودية.

- حدود موضوعية: أثر اختيار نموذج نظرية الاستجابة للمفردة (الأحادي, والثنائي, والثلاثي) وطريقة التقدير (الأرجحية القصوى, التقدير البعدي المتوقع, التقدير البعدي الأقصى) في دقة تقدير بارامترات المفردات، وقدرات الأفراد لاختبار كاتل للذكاء العام المقياس الثالث (الصورة "أ").

التعريف الإجرائي لمصطلحات البحث:

1- نماذج نظرية الاستجابة للمفردة Item Response Theory Models:

مجموعة من النماذج الرياضية التي تحدد العلاقة الاحتمالية بين الاستجابات الملاحظة على مفردات الاختبار, والسمات الكامنة المسببة لها, التي تُحدد استجابات الأفراد على مفردات هذا الاختبار, وتظهر هذه النماذج بيانيًا من خلال دوال رياضية، تربط احتمال الإجابة الصحيحة على المفردة الاختبارية بالقدرة التي تقيسها هذه المفردات, وقد اعتمد البحث الحالي على ثلاثة نماذج، هي: (أحادي البارامتر, ثنائي البارامتر, ثلاثي البارامتر).

2- طرق التقدير Estimation Methods:

مجموعة من الأساليب الرياضية الاحتمالية التي تستخدم في تقدير بارامترات مفردات الاختبار وقدرات الأفراد على مقياس مشترك, وقد اعتمد البحث الحالي على ثلاث طرق للتقدير، هي: (الأرجحية القصوى, التقدير البعدي المتوقع, التقدير البعدي الأقصى).

3- دقة التقدير Estimationaccuracy :

تعبير وصفي يشير إلى جودة تقدير بارامترات مفردات الاختبار وتقدير

قدرات الأفراد, الذي يعبر عن مدى اقتراب البارامترات المقدرة من القيمة الحقيقية لهذه البارامترات, ويعبر عنها رياضيًّا بالأخطاء المعيارية للتقدير, فكلما قلت الأخطاء المعيارية للتقدير، كان أكثر دقة.

منهج البحث:

استخدم الباحث المنهج الوصفي المقارن لمناسبته لأهداف البحث الحالي, وذلك لحساب الفروق في تقدير بارامترات مفردات اختبار كاتل للذكاء وتقدير قدرات الأفراد، تبعًا لاختلاف نموذج نظرية الاستجابة للمفردة المستخدم وطريقة التقدير.

مجتمع البحث وعينته:

بلغ عدد المشاركين في البحث الحالي (1104) مشاركين من طلاب وطالبات جامعة الحدود الشمالية بالمملكة العربية السعودية، تم اختيارها بطريقة العينة الطبقية حسب النوع (ذكور, إناث) والتخصص (كليات نظرية, كليات عملية), وقد بلغ المتوسط الحسابي لأعمارهم الزمنية (20.80) بانحراف معياري (1.31), والجدول التالي يوضح وصف عينة البحث في ضوء متغير النوع والتخصص:

جدول1: **وصف عينة البحث حسب متغير النوع والتخصص**

النوع	كليات نظرية	كليات عملية	المجموع
ذكور	233	193	426
إناث	473	205	678
المجموع	706	398	1104

أداة البحث:

تمثلت أداة البحث في اختبار كاتل للذكاء (المتحرر من أثر الثقافة) المقياس الثالث الصورة (أ), حيث يعد المقياس الثالث الصورة (أ) أحد مقاييس الذكاء التي أعدها العالم الأمريكي ريموند كاتل لقياس الذكاء, التي يمكن تقسيمها إلى ثلاثة مقاييس, ولكل منها صورتان متكافئتان (أ, ب), وتصلح هذه الاختبارات للتطبيق الفردي والجماعي, وقد أعد المقياس الثالث للغة العربية فؤاد أبو حطب وآخرون, وقد تم اختيار اختبار كاتل للذكاء في البحث الحالي؛كونه من أشهر المقاييس المستخدمة في قياس الذكاء في الوطن العربي, كما أنه اختبار غير لفظي متحرر من أثر الثقافة، ومن ثَمَّ لا تقف الثقافة عائقًا في تطبيقه على فئات من أقطار متعددة, كما أن للاختبار مؤشرات صدق وثبات عالية، تم التحقق منها في العديد من الدراسات والبحوث (أبو حطب وآخرون, 2008).

ويتكون الاختبار من أربعة اختبارات فرعية: الأول (السلاسل) وعدد أسئلته (12) سؤالًا, والثاني (التصنيف) وعدد أسئلته (14) سؤالًا, والثالث (المصفوفات) وعدد أسئلته (12) سؤالًا, والرابع (الشروط) وعدد أسئلته (8) أسئلة, ويتدرج كل اختبار فرعي في مستوى صعوبته, كما يوجد لكل اختبار فرعي مجموعة من الأمثلة التدريبية لتطبيق الاختبار قبل البدء في الإجابة, ويصحح الاختبار بإعطاء درجة (1) للإجابة التي تتفق مع مفتاح تصحيح الاختبار, والدرج (صفر) في حالة الإجابة الخطأ, وقد حظي اختبار كاتل للذكاء الصورة الثانية بالتقنين على أعداد كبيرة ومتنوعة من الأفراد أكثر من الاختبارات الأخرى، سواء كان الاختبار موقوتًا أو غير موقوت.

وقد اعتمد معدو المقياس على التحليل العاملي الاستكشافي للتحقق من صدق المقياس, وقد تشبعت جميع أسئلة الاختبار وكذلك الاختبارات الفرعية على العام, كما ارتفعت قيم معاملات الارتباط بين درجات الأفراد على الاختبار ودرجاتهم على اختبارات أخرى لفظية وغير لفظية, كما تم حساب ثبات المقياس بطريقة التجزئة النصفية، وقد تراوحت قيم معاملات الثبات ما بين (0.92:0.7), كما تم التحقق أيضًا من ثبات المقياس بطريقتي إعادة التطبيق والصور والمتكافئة, وقد تمتع الاختبار بقيم ثبات تراوحت ما بين(0.87:0.69).(أبو حطب وآخرون, 2006).

– التحقق من افتراضات نظرية الاستجابة للمفردة: يُعد التحقق من افتراضات نظرية الاستجابة للمفردة (أحادية البعد, الاستقلال الموضعي, التحرر من السرعة)، أمرًا ضروريًّا قبل تدريج مفردات الاختبار, وقد قام الباحث بفحص استجابات الأفراد للكشف عن الاستجابات التامة والصفرية قبل التحقق من افتراضات النظرية, ولم يُسفر الفحص الأولي لاستجابات الأفراد عن استبعاد أي فرد أو أي مفردة من مفردات الاختبار, وفيما يلي إجراءات التحقق من افتراضات نظرية الاستجابة للمفردة:

أ– التحقق من أحادية البعد: قام الباحث بالتحقق من افتراض أحادية البعد باستخدام التحليل العاملي الاستكشافي بطريقة المكونات الأساسية لاستجابات الأفراد على أسئلة اختبار كاتل للذكاء, حيث تعد العوامل الناتجة من التحليل العاملي مؤشرًا لأحادية البعد، عندما تكون النسبة بين قيمتي الجذر الكامن للعامل الأول والعامل الثاني أكبر من (2) طبقًا لمحك لورد, أو عندما تكون نسبة التباين المفسر على العامل الأول، مقارنة بنسبة التباين المفسر التراكمي للعوامل المستخلصة لا تقل عن (20%) طبقًا لمحك ريكاس, وقد قام الباحث بإجراء التحليل العاملي الاستكشافي وفق الخطوات التالية:

– حساب مدى ملاءمة حجم العينة لإجراء التحليل العاملي باستخدام اختبار K.M.O Test حيث تتراوح قيمة هذا الاختبار بين الصفر والواحد الصحيح, وقد اقترح Kaiser قبول هذا المؤشر الذي لا تقل قيمته عن (0.50), وقد بلغت قيمة إحصائي اختبار K.M.O في تحليل هذا الاختبار (0.700), ومن ثَمّ فإنه يمكن الحكم بكفاية حجم العينة لإجراء هذا التحليل.

– حساب قيمة اختبار بارتليت, الذي يجب أن يكون دالًّا إحصائيًّا، ليدل على أن مصفوفة الارتباط ليست مصفوفة الوحدة (خالية من

العلاقات), وإنما تحتوى على الحد الأدنى من العلاقات, وقد بلغت قيمة اختبار بارتليت (23957.818), وهي قيمة دالة إحصائيًا عند مستوى دلالة (0.001)

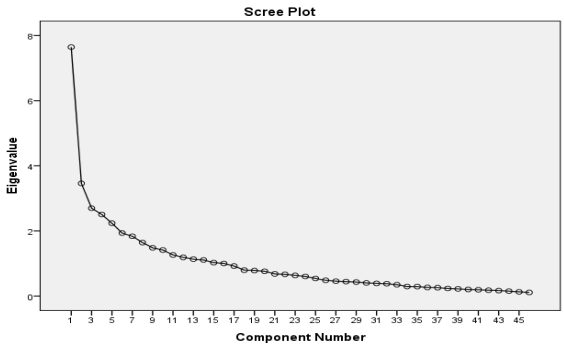
– إجراء التحليل العاملي لمصفوفة الارتباط بطريقة المكونات الأساسية, وقد أخذ الباحث بمحك جيلفورد لمعرفة حد الدلالة الإحصائية للتشبعات الناتجة، وهو اعتبار التشبعات التي تصل إلى (0.30) أو أكثر تشبعات دالة, وفي ضوء نتائج التحليل العاملي، أمكن استخلاص مجموعة من العوامل الرئيسية الجذر الكامن لكل منهما أكبر من الواحد الصحيح, والجدول التالي يوضح الجذور الكامنة والتباين المفسر لبعض العوامل المستخرجة من التحليل العاملي:

جدول2: **قيم الجذور الكامنة ونسبة التباين المفسر والتباين المفسر التراكمي لاختبار كاتل للذكاء**

الاختبار	الجذر الكامن			نسبة التباين المفسر		
	العامل الأول	العامل الثاني	النسبة بينهما	العامل الأول	التراكمي	النسبة بينهما
كاتل للذكاء	7,644	3,459	2,209	16,618	72,988	22.768

يتضح من جدول (2)، أن نتائج التحليل العاملي الاستكشافي أسفرت عن تشبع أسئلة اختبار كاتل للذكاء على عامل واحد فقط, حيث إن النسبة بين الجذر الكامن للعامل الأول والجذر الكامن للعامل الثاني جاءت أعلى من القيمة (2), كما أن العامل الأول يفسر نسبة أكبر من (20%)، مقارنة بنسبة التباين الكلي الذي يفسره جميع العوامل المستخرجة من التحليل العاملي, مما يشير إلى تحقق افتراض أحادية البعد في بيانات اختبار كاتل للذكاء.

كما يمكن أن تتضح أحادية البعد لأبعاد المقياس أيضًا عن طريق فحص التمثيل البيانيScree Plot لقيم الجذور الكامنة لجميع العوامل الممكن استخلاصها من أسئلة الاختبار, ويمكن توضيح ذلك من خلال الشكل التالي:



شكل1: **التمثيل البياني لقيم الجذور الكامنة المستخلصة من التحليل العاملي لاختبار كاتل للذكاء**

يتضح من شكل (1)، أن التغير في قيمتي الجذور الكامنة للعامل الأول والعامل الثاني كبير وسريع، مقارنة بالتغير التدريجي البسيط بين العوامل الأخرى, مما يشير إلى وجود عامل واحد يسيطر على معظم مفرداته, ومن ثَمّ يعد مؤشرًا على أحادية البعد.

ب– التحقق من الاستقلال الموضعي: يعني الاستقلال الموضعي عدم وجود مفردات تعتمد في إجابتها على مفردات أخرى من صورة الاختبار نفسه, وقد قام الباحث بالتحقق من افتراض الاستقلال الموضعي كما يلي:

– الاعتماد على نتائج التحليل العاملي الاستكشافي لمفردات صورتي الاختبار لفحص مصفوفة معاملات الارتباط بين مفردات الاختبار، التي تراوحت ما بين (0.214:0.013) وهي قيم معاملات ارتباط منخفضة, مما يشير إلى عدم وجود مشكلة الاعتماد الخطي بين مفردات المقياس, الذي يعد مؤشرًا على تحقق افتراض الاستقلال الموضعي.

– حساب قيم مؤشر (Q3) لفحص معاملات الارتباط بين البواقي لجميع أزواج المفردات، وذلك باستخدام برنامج (jMetrik. V.4.1.1), وقد تراوحت قيم معاملات الارتباط بين البواقي لجميع أزواج مفردات الاختبار بين (0.1241:0.0010)، وهي قيم أقل من القيمة المطلقة للعدد (0.2), كما بلغ متوسط قيم مؤشر (Q3) لجميع أزواج المفردات (0.0161)، وهي أقل من القيمة المتوقعة لهذا المؤشر، والبالغة (-1" = -0.0222 n"/1), مما يشير إلى تحقق الاستقلال الموضعي.

ج– منحنى خصائص المفردة: يمثل المنحنى المميز للمفردة جوهر نظرية الاستجابة للمفردة, ولكل مفردة منحنى مميز خاص بها حسب النموذج المستخدم, ويقترب فيه احتمال الاستجابة الصحيحة من الصفر عند المستويات المنخفضة من القدرة, ويزيد احتمال الاستجابة الصحيحة عند المستويات المرتفعة حتى يقترب من الواحد الصحيح, وقد تم التحقق من هذا الافتراض من خلال مدى ملاءمة مفردات الاختبار للنموذج المستخدم في التدريج, كما تم استخراج المنحنيات المميزة لمفردات الاختبار باستخدام برنامج بايلوج إم جي (BILOG-MG.3.0)، وفق نماذج نظرية الاستجابة للمفردة أحادية البعد ثنائية الاستجابة.

د– التحرر من السرعة: قام الباحث بإعطاء الوقت الكافي للإجابة على مفردات أسئلة الاختبار, بحيث يكون إخفاق الأفراد في الإجابة على أسئلة الاختبار عائدًا إلى انخفاض قدراتهم، وليس إلى تأثير عامل السرعة في الإجابة, كما قام الباحث بالتحقق من افتراض التحرر من السرعة، عن طريق حساب النسبة المئوية للطلاب الذين تمكنوا من الانتهاء من الإجابة عن جميع أسئلة الاختبار, حيث بلغت النسبة المئوية للطلاب الذين تمكنوا من الانتهاء من الإجابة على جميع الأسئلة (100%), مما يشير إلى تحقق افتراض التحرر من السرعة في الاختبار.

– ملاءمة مفردات الاختبار لنماذج نظرية الاستجابة للمفردة: نظرًا لأن البحث الحالي يهدف إلى الكشف عن أثر اختيار النموذج وطرق التقدير في دقة تقدير بارامترات المفردات والأفراد, فقد تم التحقق من ملاءمة الأفراد لنماذج نظرية الاستجابة للمفردة، وذلك باستخدام بايلوج إم جي (BILOG-MG.3.0), الذي يستبعد الأفراد الذين ترتفع أخطاؤهم المعيارية في أثناء التقدير, من خلال إعطاء القيمة (999.000*) كمؤشر على عدم ملاءمة الأفراد للنموذج, وقد أسفرت نتائج التحليل عن وجود أفراد غير ملائمين, كما قام الباحث بالتحقق من ملاءمة مفردات الاختبار للنماذج أحادية البعد ثنائية الاستجابة (أحادي البارامتر, ثنائي البارامتر, ثلاثي البارامتر)، باستخدام برنامج بايلوج إم جي (BILOG-MG.3.0), الذي يعتمد على مؤشر مربع كاي (CHI.SQ)، التي تتطلب أن تكون قيمتها غير دالة إحصائيًا حتى تكون المفردة ملائمة للنموذج, والجدول التالي يوضح قيم مؤشرات الملاءمة لمفردات الاختبار:

جدول3: مؤشرات كا2 لملاءمة مفردات اختبار كاتل لنماذج نظرية الاستجابة للمفردة

المفردة	أحادي البارامتر		ثنائي البارامتر		ثلاثي البارامتر	
	$CHI.SQ^2$ كا	الدلالة الإحصائية	$CHI.SQ^2$ كا	الدلالة الإحصائية	$CHI.SQ^2$ كا	الدلالة الإحصائية
1	10,3	0,246	12,8	0,172	11,9	0,159
2	14,9	0,061	19,3	0,022	28,5	0,000
3	10,6	0,225	8,9	0,443	11,7	0,163
4	13,5	0,082	7,9	0,544	8,3	0,404
5	12,3	0,138	12,0	0,215	34,7	0,000
6	8,9	0,347	11,3	0,258	13,7	0,091
7	13,6	0,084	11,2	0,261	40,0	0,000
8	12,9	0,116	7,2	0,611	13,9	0,87
9	12,7	0,120	16,1	0,064	12,2	0,138
10	12,2	0,135	5,3	0,809	7,5	0,484
11	4,5	0,813	4,2	0,895	7,5	0,481
12	19,7	0,011	4,6	0,864	3,8	0,877
13	3,0	0,881	4,9	0,843	11,8	0,160
14	7,0	0,536	15,1	0,088	19,9	0,010
15	5,3	0,728	7,5	0,582	11,0	0,201
16	13,9	0,091	14,5	0,105	11,0	0,198
17	13,5	0,083	12,8	0,170	6,1	0,640
18	7,5	0,486	11,0	0,279	12,3	0,136
19	12,7	0,121	16,4	0,058	7,5	0,485
20	10,6	0,155	12,7	0,123	9,8	0,200
21	22,0	0,004	15,1	0,087	7,2	0,515
22	8,4	0,396	4,6	0,868	6,2	0,620
23	29,4	0,003	14,8	0,095	3,4	0,904
24	13,6	0,084	8,2	0,516	11,2	0,190
25	31,9	0,000	20,5	0,015	5,9	0,653
26	11,4	0,165	13,3	0,148	30,9	0,001
27	9,1	0,245	19,0	0,25	12,1	0,140
28	13,8	0,088	15,0	0,090	13,4	0,099
29	4,6	0,797	14,5	0,104	13,7	0,090
30	42,1	0,000	25,7	0,022	49,1	0,000
31	9,6	0,249	14,4	0,072	29,4	0,000
32	37,8	0,000	12,1	0,207	12,2	0,139
33	13,7	0,088	2,7	0,973	3,4	0,904
34	12,8	0,120	9,9	0,361	7,1	0,527
35	13,9	0,086	6,7	0,673	6,3	0,617
36	12,7	0,122	9,7	0,673	12,3	0,137
37	13,7	0,090	7,2	0,617	2,4	0,966
38	11,7	0,164	4,4	0,833	9,2	0,326
39	13,9	0,086	11,7	0,228	10,1	0,258
40	13,8	0,089	11,3	0,258	12,2	0,135
41	2,2	0,975	8,8	0,455	6,7	0,566
42	12,9	0,127	14,1	0,118	22,2	0,004
43	18,6	0,017	3,1	0,962	10,9	0,204
44	4,1	0,847	2,6	0,977	11,2	0,202
45	12,6	0,126	10,0	0,352	10,9	0,206
46	13,6	0,086	14,6	0,102	10,8	0,207

يتضح من جدول (3)، عدم ملاءمة (7) مفردات من مفردات الاختبار للنموذج أحادي البارامتر، حيث جاءت قيم "كا2" لها دالة إحصائيًا، وهي المفردات أرقام (12-21-23-25-30-32-43)، كما يتضح عدم ملاءمة (3) مفردات من مفردات الاختبار للنموذج ثنائي البارامتر، حيث جاءت قيم "كا2" لها دالة إحصائيًا، وهي المفردات أرقام (2-25-30)، كما يتضح أيضًا عدم ملاءمة (8) مفردات من مفردات الاختبار للنموذج ثلاثي البارامتر، حيث جاءت قيم "كا2" لها دالة إحصائيًا، وهي المفردات أرقام (2-5-6-14-26-30-31-42).

- تدريج مفردات أسئلة الاختبار وفق نماذج نظرية الاستجابة للمفردة: قام الباحث بإعادة التحليل لمفردات الاختبار بعد حذف المفردات غير الملائمة من الاختبار، وفقًا لنماذج نظرية الاستجابة للمفردة أحادية البعد؛ وذلك لاستخراج تقدير بارامترات المفردات حسب النموذج المستخدم (أحادي البارامتر، ثنائي البارامتر، ثلاثي البارامتر)، باستخدام برنامج بايلوج إم جي (BILOG-MG.3.0)، وقد جاءت نتائج التحليل كما يلي:

-بالنسبة للنموذج أحادي البارامتر: تراوحت قيم بارامترات الصعوبة ما بين (-0.411) للمفردة رقم (7)، و(1.279) للمفردة (27)، وقد بلغ متوسط قيم بارامتر الصعوبة لمفردات الاختبار (0.382) بانحراف معياري (0.398).

- بالنسبة للنموذج ثنائي البارامتر: تراوحت قيم بارامترات الصعوبة ما بين (-0.402) للمفردة رقم (5)، و(1.441) للمفردة (33)، وقد بلغ متوسط قيم بارامتر الصعوبة لمفردات الاختبار (0.413) بانحراف معياري (0.448)، كما تراوحت قيم بارامترات التمييز ما بين (0.282) للمفردة رقم (33)، و(0.730) للمفردة (32)، وقد بلغ متوسط قيم بارامتر التمييز لمفردات الاختبار (0.491) بانحراف معياري (0.106).

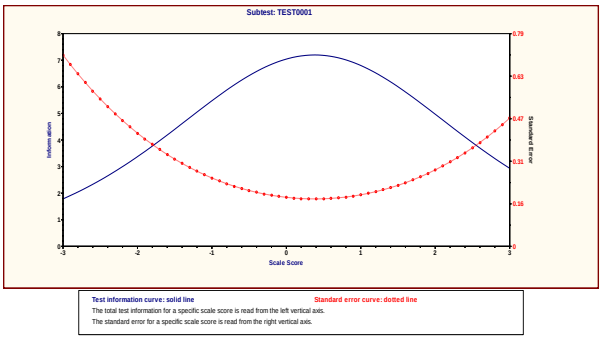
- بالنسبة للنموذج ثلاثي البارامتر: تراوحت قيم بارامترات الصعوبة ما بين (0.368) للمفردة رقم (7)، و(1.812) للمفردة (33)، وقد بلغ متوسط قيم بارامتر الصعوبة لمفردات الاختبار (1.003) بانحراف معياري (0.349)، كما تراوحت قيم بارامترات التمييز ما بين (0.640) للمفردة رقم (3)، و(1.804) للمفردة (21)، وقد بلغ متوسط قيم بارامتر التمييز لمفردات الاختبار (1.038) بانحراف معياري (0.292)، كما تراوحت قيم بارامترات التخمين ما بين (0.132) للمفردة رقم (27)، و(0.305) للمفردة (15)، وقد بلغ متوسط قيم بارامتر التخمين لمفردات الاختبار (0.244) بانحراف معياري (0.044).

- ثبات اختبار كاتل وفق نماذج نظرية الاستجابة للمفردة: قام الباحث باستخراج قيم الثبات الإمبريقي لاختبار كاتل للذكاء، وفق نماذج نظرية الاستجابة للمفردة أحادية البعد ثنائية الاستجابة (أحادي البارامتر، ثنائي البارامتر، ثلاثي البارامتر)، باستخدام برنامج بايلوج إم جي (BILOG-MG.3.0)، كما قام باستخراج الرسوم البيانية لدالة معلومات الاختبار حسب كل نموذج، وقد جاءت نتائج التحليل كما يلي:

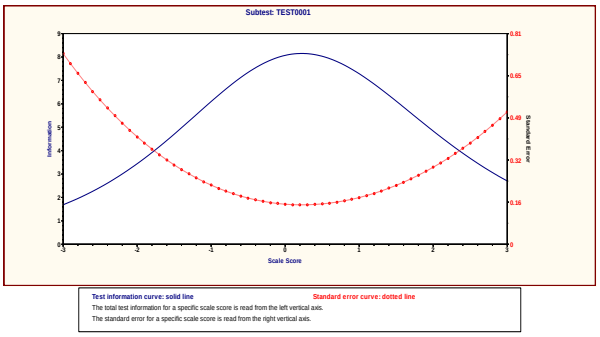
جدول 4: معاملات الثبات الإمبريقي لاختبار كاتل وفق نماذج نظرية الاستجابة للمفردة

النموذج	معاملات الثبات
أحادي البارامتر	0,859
ثنائي البارامتر	0,867
ثلاثي البارامتر	0,791

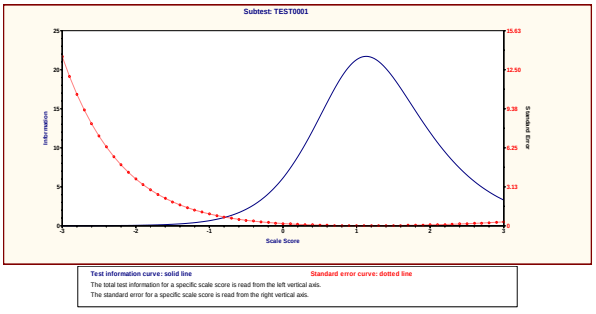
يتضح من جدول (4)، أن معاملات ثبات اختبار كاتل وفق نماذج نظرية الاستجابة للمفردة، جاءت على الترتيب (0.791) للنموذج ثلاثي البارامتر، و(0.859) للنموذج أحادي البارامت، و(0.867) للنموذج ثنائي البارامتر، ويوضح الشكل التالي دالة معلومات الاختبار والخطأ المعياري للتقدير، وفق نماذج نظرية الاستجابة للمفردة أحادية البعد ثنائية الاستجابة:



شكل2: دالة معلومات اختبار كاتل والخطأ المعياري للتقدير عبر مستويات القدرة المختلفة وفق النموذج الأحادي



شكل3: دالة معلومات اختبار كاتل والخطأ المعياري للتقدير عبر مستويات القدرة المختلفة وفق النموذج الثنائي



شكل4: دالة معلومات اختبار كاتل والخطأ المعياري للقياس عبر مستويات القدرة المختلفة وفق النموذج الثلاثي

يتضح من الأشكال السابقة لدوال معلومات الاختبار وفق النماذج الثلاثة, أن قيمة دالة معلومات الاختبار تزداد تدريجيًّا لتصل إلى أعلى قيمة لها عند مستوى القدرة المتوسطة، ثم تبدأ بالانخفاض تدريجيًّا حتى المستويات المرتفعة من القدرة, وهذا يعني أنه باستخدام مفردات الاختبار، يمكن الحصول على معلومات أكثر دقة عن الأفراد ذوي المستوى المتوسط من القدرة, بالرغم من أن قيمة الدالة المعلوماتية للاختبار وفق النموذج الثلاثي، تميل بعض الشيء إلى المستوى المرتفع من القدرة.

إجراءات البحث:

اتبع الباحث الخطوات التالية لتنفيذ البحث الحالي:

1- كتابة الإطار النظري للبحث, والوقوف على بعض الدراسات والبحوث السابقة التي اهتمت بنماذج نظرية الاستجابة للمفردة، وطرق تقدير بارامترات المفردات وقدرات الأفراد.

2- تحديد أداة البحث المتمثلة في اختبار كاتل للذكاء لاستخدمها في البحث الحالي؛ كونه من اختبارات الذكاء المستخدمة على نطاق واسع في البحوث والدراسات السابقة في تحديد قدرات الأفراد وتصنيفهم, كما أنه من الاختبارات المتحررة من أثر الثقافة, بالإضافة إلى تمتعه بمؤشرات صدق وثبات مرتفعة.

3- تطبيق أداة البحث على (1104) مشاركين من طلاب وطالبات جامعة الحدود الشمالية, والتحقق من افتراضات نظرية الاستجابة للمفردة (أحادية البعد, الاستقلال الموضعي, المنحنيات المميزة للمفردة, التحرر من السرعة) في بيانات المشاركين.

4- التحقق من ملاءمة مفردات الاختبار للنماذج أحادية البعد ثنائية الاستجابة (أحادي البارامتر, ثنائي البارامتر, ثلاثي البارامتر)، باستخدام برنامج بايلوج إم جي (BILOG-MG.3.0), واستبعاد المفردات غير الملائمة لكل نموذج.

5- تدريج مفردات الاختبار باستخدام برنامج بايلوج إم جي، وفق نماذج نظرية الاستجابة للمفردة أحادية البعد ثنائية الاستجابة (أحادي البارامتر, ثنائي البارامتر, ثلاثي البارمتر)، وطرق التقدير (الأرجحية القصوى, البعدي المتوقع, البعدي الأقصى), للحصول على بارامترات مفردات الاختبار والأخطاء المعيارية لها, وتقديرات قدرات الأفراد والأخطاء المعيارية لها.

6- رصد الأخطاء المعيارية المستخرجة من تدريج مفردات الاختبار وقدرات الأفراد، حسب نماذج نظرية الاستجابة للمفردة الثلاثة، وطرق التقدير المستخدمة.

7- إجراء المعالجات الإحصائية المناسبة لفروض البحث باستخدام برنامج (SPSS), واستخراج وعرض نتائج الفروض.

8- عرض نتائج فروض البحث والتعليق عليها، وتفسيرها في ضوء الإطار النظري ونتائج البحوث والدراسات السابقة.

9- كتابة استنتاجات وتوصيات البحث في ضوء ما توصل إليه من نتائج, ووضع مجموعة من البحوث والدراسات المقترحة.

الأساليب الإحصائية المستخدمة في البحث:

اعتمد الباحث على الأساليب الإحصائية التالية للتحقق من فروض البحث:

1- مربع كاي (CHI.SQ)

2- تحليل التباين ثنائي الاتجاه Two-Way ANOVA

3- اختبار شيفيه لمعرفة اتجاه الفروق Scheffe test

4- اختبار "ت" لعينتين مستقلتين t-testIndependentsamples

نتائج البحث ومناقشتها:

يعرض الباحث في هذا الجزء نتائج فروض البحث والتعليق عليها، ومناقشتها وتفسيرها في ضوء الإطار النظري والدراسات والبحوث السابقة:

نتائج الفرض الأول:

ينص الفرض الأول على: "لا يوجد فروق ذات دلالة إحصائية في دقة تقدير بارامتر الصعوبة، ترجع إلى أثر اختيار نموذج نظرية الاستجابة للمفردة وطريقة التقدير والتفاعل بينهما".

وللتحقق من هذا الفرض، استخدم الباحث تحليل التباين ثنائي الاتجاه (3×3) لمعرفة أثر اختيار النموذج (أحادي البارامتر, ثنائي البارامتر, ثلاثي البارامتر), وطريقة التقدير (الأرجحية القصوى, التقدير البعدي المتوقع, التقدير البعدي الأقصى), وذلك بعد تدريج مفردات الاختبار، وفق النماذج الثلاثة وطرق التقدير الثلاث، باستخدام برنامج بايلوج إم جي(BILOG-MG.3.0), وذلك للحصول على الأخطاء المعيارية لتقدير بارامتر الصعوبة, والجدول التالي يوضح ذلك:

جدول5: نتائج تحليل التباين ثنائي الاتجاه لمعرفة دلالة الفروق في دقة تقدير بارامتر الصعوبة طبقًا لاختلاف النموذج وطريقة التقدير

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة "F"	الدلالة الإحصائية
النموذج	0,053	2	0,024	33,147***	0,001 دالة
طريقة التقدير	0,000	2	0,000	0,000	1,000 غير دالة
النموذج* طريقة التقدير	0,000	4	0,000	0,000	1,000 غير دالة
الخطأ	0,268	351	0,001		
المجموع الكلي		359			

يتضح من جدول (5) ما يلي:

1- وجود فروق ذات دلالة إحصائية في دقة تقدير بارامتر الصعوبة ترجع لاختلاف النموذج (أحادي البارامتر, ثنائي البارامتر, ثلاثي البارامتر), حيث بلغت قيمة F (33.147), وهي قيمة دالة إحصائيًّا عند مستوى

(0.001). ولمعرفة اتجاه الفروق بين النماذج الثلاثة، استخدم الباحث اختبار شيفيه, والجدول التالي يوضح ذلك:

جدول6: نتائج اختبار شيفيه لمعرفة اتجاه الفروق بين النماذج الثلاثة في دقة تقدير بارامتر الصعوبة

النماذج	أحادي البارامتر	ثنائي البارامتر	ثلاثي البارامتر
المتوسطات	0,0828	0,0981	0,1130
أحادي البارامتر	-		
ثنائي البارامتر	-0,0152**	-	
ثلاثي البارامتر	-0,0301**	-0,0148**	-

يتضح من جدول (6)، وجود فروق ذات دلالة إحصائية عند مستوى (0.01) في دقة تقدير بارامتر الصعوبة بين النموذج (الأحادي والثنائي) لصالح الأحادي, وبين النموذج (الأحادي والثلاثي) لصالح الأحادي, وبين (الثنائي والثلاثي) لصالح الثنائي, حيث انخفضت متوسطات قيم الأخطاء المعيارية لتقدير بارامتر الصعوبة في كل مقارنة.

2- عدم وجود فروق ذات دلالة إحصائية في دقة تقدير بارامتر الصعوبة ترجع لاختلاف طريقة التقدير (الأرجحية القصوى, التقدير البعدي المتوقع, التقدير البعدي الأقصى), حيث بلغت قيمة ف (000,0), وهي قيمة غير دالة إحصائيًا.

3- عدم وجود فروق ذات دلالة إحصائية في دقة تقدير بارامتر الصعوبة ترجع للتفاعل بين النموذج المستخدم وطريقة التقدير, حيث بلغت قيمة ف (000,0), وهي قيمة غير دالة إحصائيًا.

ويفسر الباحث وجود فروق ذات دلالة إحصائية عند مستوى (01,0) في دقة تقدير بارامتر الصعوبة يرجع إلى اختيار نموذج نظرية الاستجابة للمفردة، -حيث أظهرت نتائج التحليل أفضلية النموذج أحادي البارامتر- ، في ضوء خصائص نماذج نظرية الاستجابة للمفردة، حيث يفترض النموذج أحادي البارامتر تساوي معاملات التمييز والتخمين لمفردات الاختبار، مقارنة بالنموذج ثنائي البارامتر وثلاثي البارامتر, حيث يفترض النموذج الثنائي اختلاف معاملات تمييز مفردات الاختبار، ويفترض النموذج الثلاثي اختلاف معاملات تمييز وتخمين مفردات الاختبار.

ويدعم هذه النتيجة ما أشار إليه علام (2005) من أن النموذج أحادي البارامتر بالرغم من استناده إلى افتراضات تتعلق بعدم التخمين، وتساوي جميع المفردات في التمييز وقياسها سمة أحادية البعد؛ إلا أن مخالفة بعض هذه الافتراضات لا يؤثر كثيرًا في دقة التقدير الناتجة عن استخدام هذا النموذج في مواقف تطبيقية معينة، وذلك مقارنة بباقي النماذج.

وتتفق هذه النتيجة جزئيًا مع نتائج دراسة كل من (Kim, 2001؛ ضعضع وآخرون, 2020) في عدم وجود فروق في تقدير بارامتر الصعوبة يرجع إلى اختلاف طريقة التقدير, وتختلف هذه النتيجة مع نتائج دراسة كل من (Garre&Vermunt,2006؛ سرحان, 2018؛ Almaleki&Alomrany, 2021) في وجود فروق في تقدير بارامتر الصعوبة يرجع إلى اختلاف طريقة التقدير.

نتائج الفرض الثاني:

ينص الفرض الثاني على: "لا يوجد فروق ذات دلالة إحصائية في دقة تقدير بارامتر التمييز، ترجع إلى أثر اختيار نموذج نظرية الاستجابة للمفردة وطريقة التقدير والتفاعل بينهما".

وللتحقق من هذا الفرض، استخدم الباحث تحليل التباين ثنائي الاتجاه (2×3) لمعرفة أثر اختيار النموذج (ثنائي البارامتر, ثلاثي البارامتر), وطريقة التقدير (الأرجحية القصوى, التقدير البعدي المتوقع, التقدير البعدي الأقصى), وذلك بعد تدريج مفردات الاختبار وفق النموذجين الثنائي والثلاثي، وطرق التقدير الثلاث باستخدام برنامج بايلوج إم جي, وذلك للحصول على الأخطاء المعيارية لتقدير بارامتر التمييز, والجدول التالي يوضح ذلك:

جدول7: نتائج تحليل التباين ثنائي الاتجاه لمعرفة دلالة الفروق في دقة تقدير بارامتر التمييز طبقًا لاختلاف النموذج وطريقة التقدير

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة "F"	الدلالة الإحصائية
النموذج	1,386	1	1,386	323,318***	0,001 دالة
طريقة التقدير	0,000	2	0,000	0,000	1,000 غير دالة
النموذج* طريقة التقدير	0,000	2	0,000	0,000	1,000 غير دالة
الخطأ	1,016	237	0,004		
المجموع الكلي		242			

يتضح من جدول (7) ما يلي:

1- وجود فروق ذات دلالة إحصائية في دقة تقدير بارامتر التمييز ترجع لاختلاف النموذج (ثنائي البارامتر, ثلاثي البارامتر), حيث بلغت قيمة F (323.318), وهي قيمة دالة إحصائيًا عند مستوى (0.001). ولمعرفة اتجاه الفروق بين النموذجين الثنائي والثلاثي، استخدم الباحث اختبار(ت) لعينتين مستقلتين، والجدول التالي يوضح ذلك:

جدول 8: قيمة "ت" ودلالتها الإحصائية لمعرفة اتجاه الفروق بين النموذجين الثنائي والثلاثي في دقة تقدير بارامتر التمييز

النموذج	العدد	المتوسط الحسابي	الانحراف المعياري	قيمة "ت"	الدلالة الإحصائية
الثنائي	129	0,0559	0,0658	18,132***	0,001 دالة
الثلاثي	114	0,2072	0,0638		

يتضح من جدول (8)، وجود فروق ذات دلالة إحصائية عند مستوى (0.001) في دقة تقدير بارامتر التمييز بين النموذجين (الثنائي والثلاثي) لصالح الثنائي, حيث انخفض متوسط قيم الأخطاء المعيارية لتقدير بارامتر التمييز في النموذج الثنائي.

2- عدم وجود فروق ذات دلالة إحصائية في دقة تقدير بارامتر التمييز ترجع لاختلاف طريقة التقدير (الأرجحية القصوى, التقدير البعدي المتوقع, التقدير البعدي الأقصى), حيث بلغت قيمة ف (0.000), وهي قيمة غير دالة إحصائيًّا.

3- عدم وجود فروق ذات دلالة إحصائية في دقة تقدير بارامتر التمييز ترجع للتفاعل بين النموذج المستخدم وطريقة التقدير, حيث بلغت قيمة ف (0.000), وهي قيمة غير دالة إحصائيًّا.

ويفسر الباحث وجود فروق ذات دلالة إحصائية عند مستوى (0.001) في دقة تقدير بارامتر التمييز بين النموذجين (الثنائي والثلاثي) لصالح الثنائي، في ضوء ملاءمة مفردات الاختبار للنموذج ثنائي البارامتر, وعدم تأثر مفرداته بالتخمين بدرجة كبيرة. فبالنظر إلى قيم بارامتر التخمين، يتبين انخفاض متوسط قيم بارمترات تخمين المفردات, مما قد يشير إلى مناسبة مفردات الاختبار بدرجة أكبر للنموذج ثنائي البارامتر، وحفاظه على عدد أكبر من المفردات مقارنة بالنموذج الثلاثي.

ويدعم هذه النتيجة ما أشار إليه علام (2005)، من أن النموذج ثنائي البارامتر يكون أكثر مناسبة للبيانات التي يقل احتمال التوصل إلى الإجابة الصحيحة عن مفرداتها عن طريق التخمين, وبخاصة إذا كانت المفردات تتباين في تمييزها بين مستويات القدرة المختلفة.

وتتفق هذه النتيجة مع نتائج دراسة كل من (Kim, 2001؛ ضعضع وآخرون, 2020؛2021 ,Almaleki&Alomrany) في عدم وجود فروق في تقدير بارامتر التمييز يرجع إلى اختلاف طريقة التقدير, وتختلف هذه النتيجة مع نتائج دراسة كل من (Garre&Vermunt, 2006؛ سرحان, 2018) في وجود فروق في تقدير بارامتر التمييز يرجع إلى اختلاف طريقة التقدير.

نتائج الفرض الثالث:

ينص الفرض الثالث على: "لا يوجد فروق ذات دلالة إحصائية في دقة تقدير قدرات الأفراد، ترجع إلى أثر اختيار نموذج نظرية الاستجابة للمفردة وطريقة التقدير والتفاعل بينهما".

وللتحقق من هذا الفرض، استخدم الباحث تحليل التباين ثنائي الاتجاه (3×3) لمعرفة أثر اختيار النموذج (أحادي البارامتر, ثنائي البارامتر, ثلاثي البارامتر), وطريقة التقدير (الأرجحية القصوى, التقدير البعدي المتوقع, التقدير البعدي الأقصى), وذلك بعد تدريج مفردات الاختبار، وفق النموذج الأحادي والثنائي والثلاثي وطرق التقدير الثلاث باستخدام برنامج بايلوج إم جي, وذلك للحصول على الأخطاء المعيارية لتقدير قدرات الأفراد, والجدول التالي يوضح ذلك:

جدول9: نتائج تحليل التباين ثنائي الاتجاه لمعرفة دلالة الفروق في دقة تقدير قدرات الأفراد طبقًا لاختلاف النموذج وطريقة التقدير

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة "F"	الدلالة الإحصائية
النموذج	4,016	2	2,008	84,698***	0,001 دالة
طريقة التقدير	0,298	2	0,149	6,284***	0,001 دالة
النموذج* طريقة التقدير	1,220	4	0,305	12,866***	0,001 دالة
الخطأ	235,358	9927	0,024		
المجموع الكلي	240,914	9935			

يتضح من جدول (9) ما يلي:

1- وجود فروق ذات دلالة إحصائية في دقة تقدير قدرات الأفراد ترجع لاختلاف النموذج (أحادي البارامتر, ثنائي البارامتر, ثلاثي البارامتر), حيث بلغت قيمة F (84.698), وهي قيمة دالة إحصائيًّا عند مستوى (0.001), ولمعرفة اتجاه الفروق بين النماذج الثلاثة، استخدم الباحث اختبار شيفيه، والجدول التالي يوضح ذلك:

جدول10: نتائج اختبار شيفيه لمعرفة اتجاه الفروق بين النماذج الثلاثة في دقة تقدير قدرات الأفراد

النماذج	أحادي البارامتر	ثنائي البارامتر	ثلاثي البارامتر
المتوسطات	0,2948	0,3102	0,3435
أحادي البارامتر	-		
ثنائي البارامتر	-0,0154**	-	
ثلاثي البارامتر	-0,0487**	-0,0333**	-

يتضح من جدول (10)، وجود فروق ذات دلالة إحصائية عند مستوى (0.01) في دقة تقدير قدرات الأفراد بين النموذج (الأحادي والثنائي) لصالح الأحادي, وبين النموذج (الأحادي والثلاثي) لصالح الأحادي, وبين (الثنائي والثلاثي) لصالح الثنائي, حيث انخفضت متوسطات قيم الأخطاء المعيارية لتقدير قدرات الأفراد في كل مقارنة.

1- وجود فروق ذات دلالة إحصائية في دقة تقدير قدرات الأفراد ترجع لاختلاف طريقة التقدير (الأرجحية القصوى, التقدير البعدي المتوقع, التقدير البعدي الأقصى), حيث بلغت قيمة ف (6.284), وهي قيمة دالة إحصائيًّا عند مستوى (0.001), ولمعرفة اتجاه الفروق بين طرق التقدير الثلاث استخدم الباحث اختبار شيفيه، والجدول التالي يوضح ذلك:

جدول11: نتائج اختبار شيفيه لمعرفة اتجاه الفروق بين النماذج الثلاثة في دقة تقدير قدرات الأفراد

طرق التقدير	الأرجحية القصوى	التقدير البعدي المتوقع	التقدير البعدي الأقصى
المتوسطات	0,3066	0,3189	0,3215
الأرجحية القصوى	-		
التقدير البعدي المتوقع	-0,0123**	-	
التقدير البعدي الأقصى	-0,0149**	-0,0025	-

يتضح من جدول (11)، وجود فروق ذات دلالة إحصائية عند مستوى (0.01) في دقة تقدير قدرات الأفراد بين طريقة (الأرجحية القصوى والتقدير البعدي المتوقع) لصالح الأرجحية القصوى, وبين طريقة (الأرجحية القصوى والتقدير البعدي الأقصى) لصالح الأرجحية القصوى, حيث انخفضت متوسطات قيم الأخطاء المعيارية لتقدير قدرات الأفراد في كل مقارنة, وعدم وجود فروق ذات دلالة إحصائية في دقة تقدير قدرات الأفراد بين طريقة (التقدير البعدي المتوقع والتقدير البعدي الأقصى).

3- وجود فروق ذات دلالة إحصائية في دقة تقدير قدرات الأفراد ترجع للتفاعل بين النموذج المستخدم وطريقة التقدير, حيث بلغت قيمة ف (12.866), وهي قيمة دالة إحصائيًا عند مستوى دلالة (0.001), ولمعرفة اتجاه الفروق، قام الباحث باستخدام اختبار شيفيه للوقوف على أثر اختلاف طريقة تقدير قدرات الأفراد حسب النموذج المستخدم, والجدول التالي يوضح ذلك:

جدول12: نتائج اختبار شيفيه لمعرفة اتجاه الفروق بين مجموعات التفاعل في دقة تقدير قدرات الأفراد

النموذج	طريقة التقدير		الأرجحية القصوى	البعدي المتوقع	البعدي الأقصى
الأحادي	الأرجحية القصوى	0,2763	-		
	البعدي المتوقع	0,3062	-0,0298**	-	
	البعدي الأقصى	0,3019	-0,0255**	0,0042	-
الثنائي	الأرجحية القصوى	0,3009	-		
	البعدي المتوقع	0,2997	0,0012	-	
	البعدي الأقصى	0,3314	-0,0304**	-0,0316**	-
الثلاثي	الأرجحية القصوى	0,3491	-		
	البعدي المتوقع	0,3510	-0,0018	-	
	البعدي الأقصى	0,3313	0,0178	0,0197	-

يتضح من جدول (12) ما يلي:

- بالنسبة للنموذج الأحادي: وجود فروق ذات دلالة إحصائية عند مستوى (0.01) في دقة تقدير قدرات الأفراد بين طريقة (الأرجحية القصوى وكل من التقدير البعدي المتوقع والتقدير البعدي الأقصى)، لصالح الأرجحية القصوى في كل مقارنة.

- بالنسبة للنموذج الثنائي: وجود فروق ذات دلالة إحصائية عند مستوى (0.01) في دقة تقدير قدرات الأفراد بين طريقة (الأرجحية القصوى والتقدير البعدي الأقصى) لصالح الأرجحية القصوى, وبين طريقة (التقدير البعدي المتوقع والتقدير البعدي الأقصى) لصالح التقدير البعدي المتوقع.

- بالنسبة للنموذج الثلاثي: عدم وجود فروق ذات دلالة إحصائية في دقة تقدير قدرات الأفراد بين طريقة (الأرجحية القصوى والتقدير البعدي المتوقع والتقدير البعدي الأقصى).

ويفسر الباحث وجود فروق ذات دلالة إحصائية عند مستوى (0.01) في دقة تقدير قدرات الأفراد بين نماذج نظرية الاستجابة للمفردة الثلاثة (أحادي البارامتر, ثنائي البارامتر, ثلاثي البارامتر)، لصالح النموذج الأقل بارامترات في كل مقارنة، في ضوء قلة عدد البارمترات المقدرة في كل نموذج، حيث تنخفض الأخطاء المعيارية لتقدير البارامترات بانخفاض عدد البارمترات المقدرة, وذلك قد يكون بسبب أن زيادة عدد البارامترات المقدرة يحتاج إلى كمية أكبر من البيانات، المتمثل في استجابات الأفراد المختبرين بمفردات الاختبار, كما أنه بالرجوع إلى دالة معلومات الاختبار وفق النماذج الثلاثة، فسيتضح أن النماذج الأقل بارامترات تعطي معلومات أكبر، وتغطي مدى من قدرات الأفراد أكبر مقارنة بالنموذج الثلاثي.

ويدعم هذه النتائج ما أشار إليه Lord (1983) في علام (2005)، من أنه كلما زاد عدد بارامترات النموذج، يتطلب ذلك كمية أكبر من البيانات المتمثلة في استجابات المختبرين للمفردات, فإذا لم توجد بيانات كافية كما في غالبية الأحيان، فإن أخطاء التقديرات لنموذج معين سوف تنخفض إذا تم اختيار نموذج عدد بارامتراته أقل، حتى وإن كان أقل ملاءمة.

كما يفسر الباحث وجود فروق ذات دلالة إحصائية عند مستوى (0.01) في دقة تقدير قدرات الأفراد بين طرق التقدير (الأرجحية القصوى والتقدير البعدي المتوقع والتقدير البعدي الأقصى)، لصالح الأرجحية القصوى في كل مقارنة، في ضوء خصائص طريقة الأرجحية القصوى، حيث تعتمد على تقدير مستويات قدرة الفرد بمعلومية بارامترات المفردة حسب النموذج المستخدم, وتأخذ في الاعتبار إجابة الفرد إجابة صحيحة عن المفردة الاختبارية, وتوفر العديد من الفروق بين الأفراد أكثر من مجرد تصحيح المفردات, كما أنها تقدر خطأ معياريًا لكل فرد، اعتمادًا على كيفية إجابته على مجموعة معينة من المفردات، التي يمكن تحديدها من أي مجموعة فرعية من المفردات التي استجاب لها، كما أشار إلى ذلك (Weiss, 2004)

وتتفق هذه النتيجة مع نتائج دراسة كل من (ضعضع وآخرون, 2020؛ القرني, 2002) في وجود فروق في دقة تقدير قدرات الأفراد يرجع إلى اختلاف طريقة التقدير, وتختلف هذه النتيجة مع نتائج دراسة (Kim, 2001)،في عدم وجود فروق في دقة تقدير قدرات الأفراد يرجع إلى اختلاف طريقة التقدير, كما تختلف مع دراسة كل من (Garre&Vermunt, 2006)؛ سرحان, 2018) في تفوق طرق بييز على طريقة الأرجحية القصوى في دقة تقدير القدرة.

توصيات البحث:

في ضوء ما أسفرت عنه نتائج البحث الحالي؛ يقدم الباحث التوصيات التالية:

1- ضرورة مراعاة اختيار نموذج نظرية الاستجابة للمفردة الملائم للبيانات، في ضوء افتراضات كل نموذج للحصول على دقة أعلى في تقدير بارامترات المفردات وقدرات الأفراد.

2-يفضل اختيار نموذج نظرية الاستجابة للمفردة أحادي البارامتر في حالة اقتراب معاملات تمييز المفردات بعضها من بعض، وعدم تأثرها بالتخمين بدرجة كبيرة؛ كونه يعطي أخطاء معيارية أقل, ولا يتأثر بدرجة كبيرة بانتهاك افتراضاته.

3- يفضل استخدام طريقة الأرجحية القصوى في حالة استخدام نماذج نظرية الاستجابة للمفردة أحادية البعد ثنائية الاستجابة، حيث إنها تعطي أخطاء معيارية أقل.

4- ضرورة الأخذ في الاعتبار نموذج نظرية الاستجابة للمفردة المستخدم في تدريج مفردات الاختبار عند اختيار طريقة تقدير قدرات الأفراد, حيث يعطي النموذج أحادي البارامتر دقة أكبر مع طريقة الأرجحية القصوى مقارنة بالطرق البيزية.

مقترحات البحث:

يقترح الباحث الحالي المشكلات والبحوث التالية:

1-أثر اختيار نموذج نظرية الاستجابة للمفردة متعددة الاستجابات وطريقة التقدير في دقة تقدير بارمترات المفردات وقدرات الأفراد.

2-أثر التفاعل بين نماذج نظرية الاستجابة للمفردة وطريقة تقدير القدرة في دقة تقدير قدرات الأفراد باستخدام الاختبارات التواؤمية المحوسبة.

3-أثر اختلاف النموذج المستخدم وطريقة التقدير في دقة معادلة صورتي اختبار وفق نظرية الاستجابة للمفردة.

4-المقارنة بين طريقة الأرجحية القصوى وطريقة بييز في دقة تقدير قدرات الأفراد باستخدام نموذج الاستجابة المتدرجة.

5-الأداء التفاضلي لاختبار كاتل للذكاء وفق نظرية الاستجابة للمفردة في ضوء متغيري الجنس والتخصص.

الإفصاح و التصريحات

تضارب المصالح: ليس لدى المؤلفون أي مصالح مالية أو غير مالية ذات صلة للكشف عنها. المؤلفون يعلنون عن عدم وجود اي تضارب في المصالح.

الوصول المفتوح: هذه المقالة مرخصة بموجب ترخيص اسناد الابداع التشاركي غير تجاري 4.0 الدولي (CC BY- NC 4.0)، الذي يسمح بالاستخدام والمشاركة والتعديل والتوزيع وإعادة الإنتاج بأي وسيلة أو تنسيق، طالما أنك تمنح الاعتماد المناسب للمؤلف (المؤلفين) الأصليين. والمصدر، قم بتوفير رابط لترخيص المشاع الإبداعي، ووضح ما إذا تم إجراء تغييرات. يتم تضمين الصور أو المواد الأخرى التابعة لجهات خارجية في هذه المقالة في ترخيص المشاع الإبداعي الخاص بالمقالة، الا إذا تمت الإشارة الى خلاف ذلك في جزء المواد. إذا لم يتم تضمين المادة في ترخيص المشاع الإبداعي الخاص بالمقال وكان الاستخدام المقصود غير مسموح به بموجب اللوائح القانونية أو يتجاوز الاستخدام المسموح به، فسوف تحتاج إلى الحصول على إذن مباشر من صاحب حقوق الطبع والنشر. لعرض نسخة من هذا الترخيص، قم بزيارة:
https://creativecommons.org/licenses/by-nc/4.0

قائمة المراجع:

أبو حطب، فؤاد، صادق، آمال، وعبدالعزيز, مصطفى. (٢٠٠٦). *كراسة تعليمات اختبار كاتل للذكاء المتحرر من أثر الثقافة, المقياسان الثاني والثالث الصورة أ, ب*. الأنجلو المصرية.

أبو حطب, فؤاد، عثمان, سيد، وصادق, آمال. (٢٠٠٨). *التقويم النفسي*. الأنجلو المصرية.

بركات, مايا. (٢٠١٠). *أثر تصميمات المعادلة ومتوسط صعوبة الاختبارات وتوزيع القدرة على معادلة درجات الاختبارات متعددة الأبعاد باستخدام نظرية الاستجابة للمفردة* [رسالة دكتوراة غير منشورة]. جامعة القاهرة.

الخولي, زياد. (٢٠١٣). اتجاهات معاصرة في القياس الموضوعي (نموذج التقدير الجزئي متعدد التدريج). *مجلة الآداب*, ١٠٦, ٦٥٣-٦٨٠.

سرحان, محمد. (٢٠١٨). *المقارنة بين طريقتي الأرجحية العظمى والقيمة العظمى للتوزيع البعدي في تقدير القدرة وفقًا لنظرية الاستجابة للفقرة* [رسالة دكتوراة غير منشورة]. جامعة بغداد.

ضعضع, هبة، طومان, منار، وطيفور, مصطفى. (٢٠٢٠). أثر حجم العينة وطرائق التقدير في دقة تقدير معالم نموذج راش. *مجلة جرش للبحوث والدراسات*, ٢١ (١), ١٣١-١٧٠.

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

القرني, أحمد. (٢٠٢٢). أثر طرق تقدير القدرة للأفراد على دقة معادلة الاختبارات في نظرية الاستجابة للمفردة. *المجلة العربية للعلوم التربوية والنفسية*, ٦ (٣٠), ١-٢٢.

References

Abu Hatab, F., Sadeq, A., & Abdulaziz, M. (2006). Karasat Taelimat Aikhtibar Katil Lildhaka' Almutaharir Min 'Athar Althaqafati, Almiqyasan Althaani WalthaalithAlsuwrat 'A, Bi. Cairo: Anglo-Egyptian.

Abu Hatab, F., Othman, S., & Sadeq, A. (2008). Altaqwim Alnafsiu., Anglo-Egyptian.

Al-Khouli, Z. (2013). AitijahatMueasirat Fi Alqias Almawdueii (Namudhaj Altaqdir Aljuzyiyi Mutaeadid Altadriji). Journal of Arts, 106, 653-680.

Allam, S. (2005). Namadhij Aliaistijabat Lilmufradat Aliakhtibariat 'Uhadiat Albued Wamutaeadidat Al'abead Watatbiqatiha Fi Alqias Alnafsii Waltarbawii. Dar Al-Fikr Al-Arabi.

Almaleki, D. & Alomrany,A. (2021). The effect of methods of estimating the ability on the accuracy and items parameters according to 3PL model. *IJCSNS International Journal of Computer Science and Network Security, 21*(7), 93-102.

Al-Qarni, Ahmed (2022). Athar Turuq Taqdir Alqudrat Llil'afrad Ealaa Diqat Mueadalat Alaikhtibarat Fi Nazariat Alastijabat Lilmufradati. Almajalat Alearabiat Lileulum Atarbawiat Walnafsia. *The Arab Foundationfor Education, Science and Arts in Egypt, 6* (30), 1-22.

Aune, S., Abal, F., &Attorresi, H. (2020). A Psychometric analysis from the item response theory: step- by- step modeling of a loneliness scale. *Psychological Sciences, 14*(1), 1-15. https://doi.org/10.22235/cp.v14i1.2179

Avcu, A. (2021). Item response theory-based psychometric investigation of SWLS for university students. *International Journal of Psychology and Educational Studies, 8*(2), 27-37. https://doi.org/10.52380/ijpes.2021.8.2.265

Baker, F. (2001). *The Basics of Item Response Theory*. (2nd ed.), The ERIC Clearinghouse on Assessment and Evaluation.

Barakat, M. (2010). *'Athar Tasmimat Almueadalat WamutawasitSueubatalaikhtibaratwatawziealqudratealaa mueadalatdarajatalaikhtibaratmutaeadidatAl'abeadBiaisti khdamNazariatAlaistijabatLilmufrada* [Unpublished PhD thesis]. Cairo University

Bortolotti, S., Tezza, R., Andrade, D., Bornia, A., & Júnior, A. (2013). Relevance and advantages of using the item response theory. *Qual Quant, 47*, 2341-2360.

Cella, D., Gershon, R., Lai, J., & Choi, S. (2007). The future of outcomes measurment: item banking, tailored short-forms, and computerizd adaptive assessment. *Quality of Life Research, 16*, 133-141. DOI: 10.1007/s11136-007-9204-6

Dodouh, H., &Touman, M., & Tayfur, M. (2020). 'Athar Hajm Aleayinat Watarayiq Altaqdir Fi Diqat Taqdir Maealim Namudhaj Rash. *Jerash Journal for Research and Studies, 21*(1), 131-170.

Domino, G., & Domino, M. (2006). *Psychological testing: An introduction*. (2nd ed.). Library of Congress.

Eluwa, O. (2011). Evaluation of Mathematics achievement test: A Comparison between classical test theory (CTT) and item response theory (IRT). *Journal of Educational and Social Research, 1*(4), 99-106.

Embretson, S., & Reise, S. (2000). *Item Response theory for Psychologists*. Lawrence Erlbaum Associates.

Garre, G. & Vermunt, K. (2006). Avoiding Boundary Estimation in Latent Class Analysis by Bayesian Posterior Mode Estimation. *Behaviormentrika,33* (1), 1-17. DOI:10.2333/bhmk.33.43

Gnambs, T. & Batinic, B. (2011). Evaluation of measurement precision with Rasch-type models: The case of the short generalized opinion leadership scale. *Personality and Individual Differences, 50*(1), 53-58. https://doi.org/10.1016/j.paid.2010.08.021

Holmefur, M., Sundberg, K., Wettergren, L., & Langius-Eklof, A..(2015). Measurement properties of the 13-item sense of coherence scale using Rasch analysis. *Quality of Life Research, 24*(6), 1455-1463. DOI: 10.1007/s11136-014-0866-6

Kim, S. (2001). An Evaluation of A Markov Chain Monte Carlo Method for the Rasch Model. *Applied Psychological Measurement, 25*(2), 163-176. https://doi.org/10.1177/01466210122031984

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

Mayer, M., & Zhu, S. (2013). Fair and equitable measurement of student learning in MOOCs: An introduction to item response theory, scale linking, and score equating. *Research& Practice in Assessment, 8*, 26-39.

Reise, S., & Waller, N. (2003). How many IRT parameters does it take to model psychopathology items?. *Psychological Methods, 8*(2), 164- 184. https://doi.org/10.1037/1082-989X.8.2.164

Sarhan, M. (2018). *Almuqaranat Bayn Tariqatay Al'arjahiat Aleuzmaa Walqimat Aleatmaa Liltawzie Albuedii Fi Taqdir Alqudrat Wfqaan Linazariat Alaistijabat Llilfaqra* [Unpublished doctoral thesis]. Baghdad University.

Weiss, D. (2004). Computerized adaptive testing for effective and efficient measurement in counseling and education. *Measurement and Evaluation in Counseling and Development,37*(2), 70-84. DOI:10.1080/07481756.2004.11909751

Wolf, R. (2013). *Assessing the impact of characteristics of the test, common-items, and examinees on the preservation of equity properties in mixed-format test equating* [Unpublished doctoral dissertation]. University of Pittsburg.

Yang, F., & Kao, S. (2014). Item response theory for measurement validity. *Shanghai Archives of Psychiatry, 26*(3), 171-177. doi:10.3969/j.issn.1002-0829.2014.03.010