Journal of Educational and Psychological Sciences
Volume (5), Issue (14) : 30 Apr 2021
P: 46 - 62

ISSN: 2522-3399

مجلة العلوم التربوية والنفسية
المجلد (5)، العدد (14) : 30 إبريل 2021 م
ص: 46 - 62

Detection of the effect of the Percentage of Items with DIF on the strength and efficacy of the singular characteristic curve test in the three-parameter model

Alaa Nawwaf Al-Harbi
Hafr Al-Batin University || KSA
Fatima Ramzi Al-Madani
Taibah University || KSA

Abstract: The current research is following the simulation studies in IRT to experiment with the effect and power of the ICC's method on DIF items. To achieve the experiment,data were generated for three tests consisting of 100 items to each one containing (10% -20%-30%) items with DIF. These three tests were applied through the simulation program to 3000 individuals with normal abilities are distributed in average (0) and standard deviation (1).

The results of the study indicated the effectiveness of the ICC's test by using the Rajo equation (1988), which measuring the area below the curve between the Reference groups and the Focal groups in detection items of (10% -20% -30%) amount of DIF and counting them in each of the three tests. The results showed that the strength of the test was (0, 619) for the DIF of 10% and the test strength of the DIF of 20% is amount (0.703) and the strength of the test was (0786) for the DIF of 30%, all of Power Values are higher than the average rating by Cohen, D test to derive the size of the impact. And we recommend taking the results in instruments containing differential items.

Keywords: percent of items, DIF, ICC, Power of test, Effect size, effective.

الكشف عن أثر نسبة المفردات ذات الأداء التفاضلي على قوة وفاعلية اختبار منحنى خصائص المفردة في النموذج ثلاثي المعلم

آلاء نواف الحربي
جامعة حفر الباطن || المملكة العربية السعودية
فاطمة رمزي المدني
جامعة طيبة || المملكة العربية السعودية

الملخص: هدفت الدراسة إلى القيام بتجربة حول موضوع أثر نسبة المفردات ذات الأداء التفاضلي على قوة وفاعلية اختبار منحنى خصائص المفردة. مستخدمةً أسلوب المحاكاة؛ ولتحقيق التجربة تم توليد بيانات لثلاثة اختبارات مكون كل منها من 100 مفردة اختبارية، تحوي على الترتيب نسب تفاضل (10%-20%-30%). طُبقت عبر برنامج المحاكاة على 3000 فرد، تتوزع قدراتهم توزيعاً طبيعياً بمتوسط مقداره (صفر) وانحرافه المعياري (1).

وأشارت نتائج البحث إلى فاعلية اختبار منحنى خصائص المفردة في الكشف عن نسب المفردات التفاضلية (10%-20%-30%)، وعددها في كل اختبار من الاختبارات الثلاث عبر استخدام معادلة راجو (1988)، التي تُعنى بقياس المساحة تحت المنحنى بين المجموعتين المرجعية والمستهدفة. وبعدها تم المقارنة بين المساحات الثلاث المستخرجة من كل اختبار، لاكتشاف قوة اختبار منحنى خصائص المفردة لكل اختبار من الاختبارات الثلاث. وأظهرت النتائج أن قوة الاختبار بلغت قيمتها (0، 619) للاختبار ذو الأداء التفاضلي 10%، وبلغت قيمة

قوة الاختبار للاختبار ذو الأداء التفاضلي 20% مقدار (0، 703)، أما قوة الاختبار للاختبار ذو نسبة الأداء التفاضلي 30% بلغت قيمتها (0، 786)، وجميعها أتت بمقدار قوة أعلى من المتوسط حسب تصنيف اختبار كوهين، ما يعني فعالية اختبار منحنى خصائص الفقرة في استخراج المفردات ذات الأداء التفاضلي، موصين بأخذ نتائجه في الأدوات التي تحوي أداء تفاضلياً.

الكلمات المفتاحية: نسبة المفردات، الأداء التفاضلي، منحنى خصائص المفردة، قوة الاختبار، حجم التأثير، الفاعلية.

المقدمة.

نظراً لأهمية موضوع الاختبارات، وشغوله لحيزٍ كبير من الاهتمام بسبب ما يبنى عليه من القرارات، فقد حظيّ موضوع بناء الاختبارات وتجريبها والتأكد من فعاليتها بالكثير من العمل والجهد. ومع ظهور حركة القياس الأولى -أو ما تعرف بالنظرية التقليدية للقياس CTT- تطورت أدوات القياس التي كان يستخدمها التربويون والمهتمون بعلم الاختبارات كوسائل للكشف عن مواطن الخلل في الاختبارات، بناءً على احصاءاته ومدى ملائمته للعينة، وهل هو صادق لدرجة يمكن الاعتماد فيها على نتائجه؟ وهل من الممكن أن يكون ثابتاً بحيث يعبر مهما جرى تطبيقه عن القدرة الحقيقية للمفحوص رقم 1 أو رقم 2؟ (الطريري، 2014).

ثم ظهرت نظرية جديدة في القياس، والتي تُعرف بنظرية الاستجابة للمفردة IRT. أضافت النظرية الحديثة بُعداً لامعاً وأفقاً واعداً يُمكن من بناء أدوات قوية ومتحررة من أي عينة للمفحوصين، وكذلك يُمكنّ من الحصول على تقدير عالي الدقة لقدرة المفحوصين المقاسة. وهذه الأداءات تتمثل على منحنى خصائص المفردة، والذي يُمثل الواجهة القوية لهذه النظرية، حيث تم تقدير جميع القدرات الممكنة للأفراد على محور القدرة الأفقي، واحتمالات استجاباتهم المقدرة مُسبقاً على المحور الرأسي، مما أدى لتحسين الأدوات وتحريرها من أي عينة يتم تطبيقها عليها (علام، 2013).

وكون الاختبارات هي من أهم محاور دراسات وأبحاث القياس التربوي، فقد وجب التأكد من ملائمة هذه الاختبارات، وصدقها في قياس ما وضعت لقياسه، سواء كانت اختبارات تحصيلية أو اختبارات نفسية أو اختبارات عقلية. ومن ضمن القضايا التي تواجه بنود أو مفردات الاختبار ومدى صدقها في قياس السمة الموضوع لأجلها، هو موضوع التحيز المتعلق بدالة المفردة التفاضلية Differential Item Functionالمعروفة اختصاراً DIF. والتي تعني وجود مفردات تسببت بأداء فارق بين أفراد العينة، نظراً لأن هذه المفردات تتحيز لفئة دون الأخرى، كالبيض والسود مثلاً أو الأغنياء والفقراء أو الإناث والذكور. ومن المهم الكشف عن هذه المفردات وإزالتها، لضمان صلاحية الاختبار وقياسه الصادق للسمة المراد قياسها بشكل عادل بين كل أفراد العينة. ومن هذا المنطلق جاءت دراسات كثيرة واقعية، أو عن طريق أسلوب توليد البيانات (المحاكاة)، تدرس وتختبر العديد من الأوضاع المختلفة لوجود هذه المفردات ذات الأداء التفاضلي وأفضل الطرق للكشف عنها. ويذكر اللبدي (2008): أنه في النظريتين التقليدية والحديثة للقياس التربوي ظهرت عدة اختبارات، الغرض منها الكشف عن دالة المفردة التفاضلية. وفي هذه الدراسة تم تناول إحدى طرق النظرية الحديثة للقياس، وهو اختبار منحنى خصائص المفردة ICC، ومدى فعاليته في الكشف عن نسب التحيز لاختبار طويل نسبياً ومدى قوة هذا الاختبار.

مشكلة الدراسة وأسئلتها:

يعد وجود الأداء التفاضلي لبعض المفردات في الاختبار مهدداً لسلامة ودقة نتائج ذلك الاختبار، ويتعداه الى تهديد درجة الصدق والموضوعية التي تُبنى عليها القرارات والتفسيرات الناتجة عن نتائج الاختبار المقصود، مما يشكك في عدالة تلك القرارات.

ويمكننا تحديد مشكلة الدراسة في السؤال الرئيس التالي: (ما أثر نسبة المفردات ذات الأداء التفاضلي على قوة وفاعليه اختبار منحنى خصائص المفردة)، وتنبثق منه التساؤلات التالية:

1. ما أثر وجود المفردات ذات الأداء التفاضلي بنسبة 10%على طريقة اختبار منحنى خصائص المفردة؟
2. ما أثر وجود المفردات ذات الأداء التفاضلي بنسبة20% على طريقة اختبار منحنى خصائص المفردة؟
3. ما أثر وجود المفردات ذات الأداء التفاضلي بنسبة30% على طريقة اختبار منحنى خصائص المفردة؟
4. هل توجد فروق دالة إحصائياً على طريقة اختبار منحنى خصائص المفردة تبعاً لاختلاف نسب وجود المفردات ذات الأداء التفاضلي مرة 10% ومرة 20% ومرة 30%؟

أهداف الدراسة:

تهدف الدراسة إلى:

1. تحديد أثر وجود المفردات ذات الأداء التفاضلي بنسبة 10%على طريقة اختبار منحنى خصائص المفردة.
2. قياس أثر وجود المفردات ذات الأداء التفاضلي بنسبة20% على طريقة اختبار منحنى خصائص المفردة.
3. معرفة أثر وجود المفردات ذات الأداء التفاضلي بنسبة 30% على طريقة اختبار منحنى خصائص المفردة.
4. معرفة مدى وجود فروق دالة إحصائياً على طريقة اختبار منحنى خصائص المفردة تبعاً لاختلاف نسب وجود المفردات ذات الأداء التفاضلي مرة 10% ومرة 20% ومرة 30%.

أهمية الدراسة:

تنبع الأهمية العلمية للدراسة في إمكانية العمل بنتائجها على أرض الواقع، وذلك عند مواجهة اختبار أو استبيان طويل من 100 مفردة ويحوي مفردات ذات أداء تفاضلي بنسب مختلفة. وبناءً على توصيات دراسة أبوعلام والمرابحة (2015)، نجد أنه من الأهمية بمكان دراسة أثر النسب المختلفة للأداء التفاضلي في الاختبار وأثرها على فاعلية منحنى خصائص المفردة الذي يعد أحد طرق الكشف عن الأداء التفاضلي. واستخدام النموذج ثلاثي المعلم، حيث إن هذا النموذج يتناسب بطبيعته مع نوع الأسئلة الموضوعية التي يتوفر فيها معالم المفردة الثلاثة (الصعوبة-التمييز-التخمين). وعليه فإن المعلمين ومعديّ ومطبقيّ الاختبارات النفسية يمكنهم اللجوء لهذه الدراسة في حال رغبتهم في معرفة الفروق التي قسمت أداء الطلاب إلى مجموعتين، مجموعة لم يظهر ضدها التحيز ومجموعة ظهر ضدها التحيز في الأسئلة. أي أن الباحث أو مصمم الاستبيان أو الاختبار النفسي الطويل يستطيع الوثوق بنتائج هذه الدراسة في معرفة مدى تأثير نسبة المفردات ذات الأداء التفاضلي على قوة الاختبار الإحصائي. كما أنه إضافة في مجال الدراسات التي تناولت طرق الكشف عن الأداء التفاضلي وتحديداً لإجراءات اختبار منحنى خصائص المفردة.

حدود الدراسة:

- استخدام اختبار طريقة خصائص منحنى المفردة.
- استخدام النموذج ثلاثي المعلم.
- الاقتصار على الأداء التفاضلي المنتظم.
- استخدام أسلوب المحاكاة بتوليد بيانات باستخدام برنامج WinGen-3.
- الاقتصار على ثلاث نسب فقط لوجود الأداء التفاضلي في الاختبار 10%و20%و30%.

مصطلحات البحث:

التحيز Bias:

هو وجود مفردات في الاختبار تتسبب في أداء تفاضلي بين أفراد العينة وتكون إما اختلاف جنس أو عرق أو بيئة مما يتسبب بأداء فارق بينهم.

الأداء التفاضلي Differential performance:

هو دلالة إحصائية تعبر عن الفروق في الاستجابة للمفردة بين مجموعتين لديها نفس القدرة. (أبوعلام والمرابحة، 2014) وتُقاس إجرائياً عن طريق قياس الفروق بين منحنيّ الاستجابة للمجموعتين.

الأداء التفاضلي المنتظم Uniform DIF:

ويعني احتمال إجابة المفردة بشكل صحيح في إحدى المجموعتين أكثر من الأخرى وعبر كل مستوى من مستويات القدرة بين نفس المجموعتين.

اختبار مقارنة منحنى المفردة Item Character Curve Test:

أحد الطرق الاختبارية لكشف تحيز المفردات في نظرية الاستجابة للمفردة IRT ويعتمد إما على إيجاد الفروق بصرياً عبر الرسومات البيانية أو إحصائياً عن طريق حساب مساحة المجموعتين المرجعية والمستهدفة تحت المنحنى (اللبدي، 2008) ويعني إجرائياً استخدام الطرق الإحصائية عن طريق معادلة راجو للكشف عن مقدار الفرق التفاضلي.

نسبة المفردات ذات الأداء التفاضلي Percentage of DIF:

أي عدد الأسئلة أو المفردات التي تبدي دالة تفاضلية مقسوم على عدد مفردات الاختبار. (أبوعلام والمرابحة، 2015)

قوة الاختبار الإحصائي Power of Test:

ويعني رفض الفرض الصفري حينما يكون بالفعل خاطئاً (الشاردي، 2015) وتعني في اجرائياً نسبة الكشف عن دالة المفردة التفاضلية بين المجموعتين المرجعية والمستهدفة. (أبوعلام والمرابحة، 2015)

أسلوب المحاكاة Simulation:

هي محاكاة حاسوبية تنتج معلومات دقيقة ذات علاقات ارتباط عالية وهي مقاربة علمية صحيحة لفهم قوانين الطبيعة وإجراء التجارب. (بيندر وهيرمان 1997)

المبحث الأول- الإطار النظري والدراسات السابقة

أولاً- الإطار النظري:

التحيُّز هو مصطلح ضد العدالة الاختبارية، ويطلق على الاختبار صفة التحيُّز حينما تختلف قيمة المتوسط الحسابي في المجموعة المرجعية Reference Group عنها في المجموعة المستهدفة Focal Groupكما يذكر ثورانديك

(أيوب، 2008). إلا أنه توجد فروقات بين مصطلحيّ التحيز والعدالة الاختبارية رغم اتصالهما؛ فالتحيُّز يدل على ممارسات تمييزية غير قانونية في فقرات الاختبار (علام، 2013) ويمكن إحصائياً أن تُقدّر عن طريق مخرجات الاختبار الرقمية. أما العدالة هي قضايا أخلاقية وفلسفية قد لا يوافقها المنطق لعدم وجود دليل إحصائي عليها. ووفقاً لنظرية الاستجابة للمفردة فإننا نستخدم مصطلح دالة المفردة التفاضلية Different Item Function (DIF) حينما يظهر فرق في الاستجابة بشكل صحيح على المفردات الاختبارية بين المجموعات من نفس أصحاب القدرة (اللبدي، 2008).

مفهوم دالة المفردة التفاضلية وأنواعها:

لمفهوم دالة المفردة التفاضلية كما يذكر أبوعلام (2014) استخدامين، أحدهما موضوعي يُستخدم في أدبيات الموضوع والآخر إحصائي يُستخدم في الدراسة أو البحث ويعمل على تحديد المفردات التي تعمل بشكل مختلف بين مجموعات من المختبرين. وبعد تحديد المفردات التفاضلية تتم مناقشتها مناقشة منطقية، والبحث في أسباب صعوبتها لدى مجموعة من المختبرين عن غيرهم ثم إذا ثبت تحيزها نقوم بحذف المفردات المتسببة بالتحيز. ونشير إلى أننا نقارن بين منحنيّ خصائص المفردة في كلٍ من المجموعتين المرجعية R والمستهدفة F حينما نستخدم نظرية الاستجابة للمفردة للكشف عن دالة المفردة التفاضلية DIF، بشرط تساوي القدرة في كلا المجموعتين. أما المفردة التي تكون فيها منحنيات خصائص المفردة متماثلة في كل مجموعة فهي فقرة لا يوجد فيها أداء تفاضلي.

افتراضات النظرية الحديثة ونماذجها:

الافتراضات: -

لنظرية القياس الحديثة، والتي تُعرف بنظرية الاستجابة للمفردة IRT، مجموعة افتراضات يجب أولاً أن تتوفر كشرط يسمح باختيار أحد نماذج النظرية الحديثة لفحص البيانات. وهذه الافتراضات يذكرها عبابنة (2008) كالتالي:

1- أحادية البعد :Unidimensionality

تعني أحادية البعد أن الاختبار يقيس سمة واحدة مشتركة فقط والتي تُعرف بقدرة المفحوص $\boldsymbol{\theta}$ ويمكن التوصل لمعرفة هذه السمة من خلال إجراء عملية التحليل العاملي المتمثلة بإنتاج مصفوفة ارتباط بين جميع المتغيرات.

2- الاستقلال الموضعي Independence Local:

ويعني أن الأداء على فقرتين لنفس المفحوص على هذا الاختبار تكون مستقلة استقلالاً إحصائياً، أي أن استجابة المفحوص على مفردة ما لا تؤثر سلباً أو إيجاباً على إجابته على مفردة أخرى.

3- سرعة الأداء Speediness:

وتعني أن الاختبار لا يتم تحت شرط السرعة، أي أنه عند إجراء الاختبار لا يحدد زمن لإنهاء الاختبار.

نماذج النظرية الحديثة في القياس: -

1- النموذج أحادي المعلم (راش):

النموذج الأحادي يضع كل مفردة اختبارية بموقع خاص بها على ميزان الصعوبة b، وكل المفردات فيه لها نفس التمييز a ، ويتمثل بالمعادلة التالية:

$$P_i(\theta) = \frac{1}{1+e^{-D(\theta-b_i)}},$$

معادلة رقم (1)

حيث P(Θ): هي احتمال الإجابة الصحيحة لصالح ب القدرة i.

E هي الأساس اللوغاريتمي ويساوي 2، 7183 وD قيمة ثابتة وتساوي 1، 7.

2- النموذج الثنائي المعلم:

ويضع هذا النموذج في عين الاعتبار أنه بتغيّر الصعوبة b من مفردة لأخرى فإن قدرة التمييز a لكل مفردة أيضاً تكون متغيرة وفق المعادلة التالية:

$$P_i(\theta) = \frac{1}{1+e^{-Da_i(\theta-b_i)}}.$$

معادلة رقم (2)

3- النموذج الثلاثي المعلم:

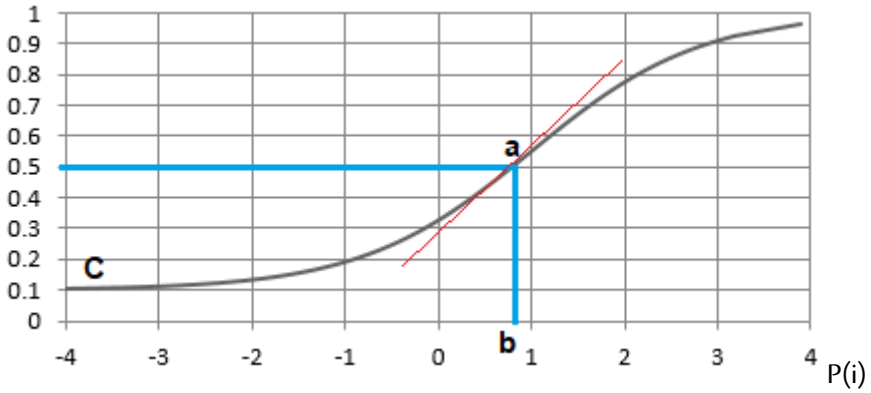
يزيد هذا النموذج بالإضافة إلى معلميّ الصعوبة والتمييز بمعلم ثالث وهو التخمين c ويحسب وفق المعادلة التالية:

$$P_i(\theta) = c_i + (1-c_i)\frac{1}{1+e^{-Da_i(\theta-b_i)}},$$

معادلة رقم (3)

منحنى خصائص الفقرة: Item Characteristics Curve

هو إضافة من أهم إضافات النظرية الحديثة لعملية القياس التربوي. ويُعّرف منحنى خصائص المفردة والمعروف اختصاراً ICC بأنه رسم بياني، تنعكس فيه القدرة أو السمة الكامنة على المحور الأفقي، واحتمال الإجابة الصحيحة على المحور الرأسي. ويذكر علام (2013) أنه لكل مفردة اختبارية منحنى خاص بها. وعن طريق الاشتقاقات الرياضية يتم تقدير قدرات الأفراد الذين اجابوا عن تلك المفردة. ويتضح من خلال المنحنى موضع صعوبة الفقرة وانعكاسه على محور قدرات الأفراد، ومقدار الميل الذي يمثل قيمة التمييز، وكذلك قيمة التخمين الغير ملامس للمحور الأفقي. والرسم التالي هو مثال لمنحنى خصائص مفردة اختبارية، حيث إن a هي التمييز(الميل) وb هي الصعوبة وc هي التخمين:



شكل (3) منحنى خصائص المفردة موضح عليه تدريج القدرة واحتمالات الإجابة

طرق الكشف عن دالة المفردة التفاضلية:

الكشف عن دالة المفردة التفاضلية يتم بالعديد من الطرق تبعاً لنوع نظرية القياس، هل هي النظرية التقليدية CTT أم النظرية الحديثة IRT (نظرية الاستجابة للمفردة)؟ هناك مثلاً طرائق عديدة في نظرية IRT، مثل طريقة نسبة الأرجحية وطريقة لورد/مربع كاي وطريقة راش وطريقة منحنى خصائص المفردة. ونذكر عموماً ما ذكره شبرد Shepard (1981) وإرنسون Ironson (1983)، حين قاما بمراجعة طرق الكشف عن مفردة الدالة التفاضلية DIF التابعة لنظرية الاستجابة للمفردة ووجدوا أنها كلها تشترك بنفس الخطوات التالية:

1- مقارنة خصائص منحنيات المفردة.
2- مقارنة معالم المفردات.
3- مقارنة مطابقة النموذج للبيانات في المجموعات المختلفة. (اللبدي، 2008)

طريقة منحنى خصائص المفردة: -

في طريقة اختبار منحنى خصائص المفردة ICC نقوم بحساب المساحة تحت المنحنيات لكل من المجموعتين المرجعية R والمستهدفة F بشرط أن تكون معالم المفردات على نفس المستوى من القدرة وإلا فإنه ستنتج فروقات في المساحة قد تكون بسبب اختلاف معالم المفردات بين المجموعتين.

وقد اقترح راجو (Raju 1988) معادلة لحساب الفروق بين منحنيّات المفردة بين المجموعتين المرجعية والمستهدفة كالتالي:

$$Area = (1-c)\times\left|\frac{2(a_1-a_2)}{Da_1a_2}\ln\left(1+\exp\left(\frac{Da_1a_2(b_2-b_1)}{a_2-a_1}\right)\right)-(b_2-b_1)\right|$$

معادلة رقم (4)

حيث: a1 تمييز المجموعة المرجعية a2 تمييز المجموعة المستهدفة.

b 1 صعوبة المجموعة المرجعية b2 صعوبة المجموعة المستهدفة.د

c قيمة التخمين وD ثابت ويساوي 1، 7.

قوة الاختبار الإحصائي Power of Test:

هي احتمالية لها قيمة بين (0-1)، وقوة الاختبار هي المعيار الأساسي لاختبار الفروض الإحصائية من ناحية مستوى الدلالة. وتعني كما هو معروف: قرار رفض الفرض الصفري وهو خاطئ بالفعل ما يعني اتخاذ قرار صحيح.

حجم التأثير Effect Size: -

هو توجه جديد في ميدان البحث التربوي ومن خلال حسابه نستطيع إيجاد قوة الاختبار. وقيمة حجم التأثير تشير إلى الحد الذي يمكن فيه التنبؤ بالمتغير التابع عن طريق المتغير المستقل.(الشاردي، 2015)

وتعمل مؤشرات حجم التأثير دور الوسيط في حساب قوة قبول أو رفض الفرض الصفري في العديد من البحوث التجريبية وقد أوجد كوهين طرق عديدة لحساب حجم التأثير منها معادلة Cohen's D:

$$D = \frac{M1-M2}{SD}$$ معادلة رقم (5)

حيث: M1هي متوسط المجموعة الضابطة وM2 متوسط المجموعة التجريبية

وSD هو الانحراف المعياري لإحدى المجموعتين.

إلا أن المعادلة السابقة لا تصح إلا في حال تجانس التباين بين المجموعتين التجريبية والضابطة، ويذكر نصار(2006) والشاردي (2015) أنه في حالة عدم تساوي عدد أفراد المجتمع الأول مع عدد أفراد المجتمع الثاني عندها يتم قسمة فارق المتوسطين على جذر ناتج التباين الممزوج Pooled Variance الذي يدمج تباين المجتمعين في تباين واحد وفق المعادلة:

$$S^2 p = \frac{S^2 1(n1-1)+S^2 2(n2-1)}{n1+n2-2}$$ معادلة رقم (6)

ويتم الحكم على مقدار حجم التأثير وفقاً لكوهين كالتالي:

1- حينما تكون D: 0, 20 فإن حجم التأثير يكون صغير.
2- حينما تكون D: 0, 50 فإن حجم التأثير يكون متوسط.
3- حينما تكون D: 0, 80 فإن حجم التأثير يكون كبير.

ثانياً- الدراسات السابقة

- **دراسة الحجيلي (2012):** استخدمت فيها أسلوب توليد البيانات، وتعرفت من خلاله على قوة تقدير صعوبة المفردة والقدرة وفعالية معادلة الاختبار، بوجود أداء تفاضلي للفقرة عن طريق استخدام نموذج أحادي المعلم (راش)، لعينتين بطول 1600، 800 وطولي اختبار أحدهما 12 فقرة والآخر 20 فقرة بوجود مفردات مشتركة بينهما. وأظهرت النتائج إلى وجود ارتباطات قوية ودالة بين تقديرات الصعوبة والقدرة، وقد أشارت النتائج إلى أن قيم الخطأ المعياري تختلف بشكل دال عند اختلاف طول الاختبار وحجم العينة والتفاعل بينهما.
- **دراسة سيبرت، وستراك Seybert & Strak (2013):** هدفت الدراسة إلى استخدام نماذج النقطة المثالية نظراً لكونها مشهورة في الأبحاث والتدريبات، ولكن حسب قول الباحث فإننا لانعرف عن الطرق التقليدية في كشف التحيز سوى القليل في مجال النقطة المثالية، لذا فحين إجراء دراستهم بأسلوب المحاكاة باستخدام معالم مختلفة وجدوا أداء جيد لمربع كاي وDFIT، وجدوا أن منحنى خصائص المفردة يربط البيانات الغير ممثلة.

- **دراسة أبوعلام والمرابحة (2014):** هدفت الدراسة إلى الكشف عن فعالية طريقتين من طرق كشف التحيز هما طريقة الأرجحية العظمى وطريقة اختبار تحيز المفردة المتزامن. وتم توليد بيانات لأحجام مختلفة من العينات 1200، 800، 400 لاختبارين أحدهما 25 فقرة والأخر 50 فقرة. وخلصت النتائج إلى أن الخطأ من النوع الأول يقل عند زيادة عدد مفردات الاختبار، وعند زيادة عدد أفراد العينة لكل من طريقتي نسبية الأرجحية العظمى وطريقة تحيز المفردة المتزامن.
- **دراسة ساري وهاغينز Sari & Huggins (2015):** هدفت الدراسة إلى مقارنة طريقتين من طرق كشف التحيز، هما طريقة الأزواج وطريقة مقارنة المجموعات المركبة، وذلك بتوليد اختبار من 60 فقرة. وخلصت النتيجة إلى أن كمية المفردات التفاضلية التي تم الكشف عنها عبر الطريقتين لم تكن هي نفسها.
- **دراسة اجباره (2015):** هدفت إلى دراسة نسبة مفردات ذات تحيز مصممة بنسب مختلفة 10%-20%-40% على بيانات مولدة. كانت التجربة عبارة عن اختبار بطول 40 فقرة وعينة مكونة من 1000 فرد لكل من الطريقتين الانحدار اللوجستي والارجحية العظمى. وخلصت النتائج من ناحية قوة الاختبار إلى وجود فروق دالة إحصائياً بين المتوسطات الحسابية تعزي لنسبة المفردات ذات الأداء التفاضلي وكانت لصالح النسبة الأقل. كما كشفت النتائج عن عدم وجود فروق دالة إحصائياً بين المتوسطات الحسابية تعزي للكشف عن الأداء التفاضلي. وأظهرت النتائج عدم وجود فروق دالة إحصائياً بين المتوسطات الحسابية لقوة الاختبار تعزي لتفاعل المتغيرين.
- **دراسة أبوعلام والمرابحة (2015):** درست قوة اختبار تحيز المفردة المتزامن على نسبة المفردات المتحيزة في الاختبار مع حجم العينة باستخدام النموذج ثلاثي المعلم. حيث قامت الدراسة على توليد بيانات 1200، 800، 400 فرد، واختيار نسب ذات أداء تفاضلي مختلفة 10%-20%-30% لاختبار يتكون من 50 مفردة. وخلصت النتائج إلى أن قوة اختبار تحيز المفردة المتزامن تتزايد بزيادة نسب المفردات ذات الأداء التفاضلي. وأشارت النتائج كذلك إلى أن قوة الاختبار تزداد أيضاً بزيادة حجم العينة، وبالنسبة لتفاعل حجم العينة ونسبة المفردات التي تبدي أداء تفاضلياً، فقد تأثرت قوة اختبار تحيز المفردة المتزامن بزيادة حجم العينة وبالنسب ذات التحيز التفاضلي التي تبلغ 20%-30%.
- **دراسة فورتونا وماتورو Fortuna & Maturo (2018):** ركزت على كيفية استخدام طريقة الخوارزميات التصنيفية في تحليل منحنى خصائص المفردة ومنحنى دالة المعلومات بطريقة الخوارزميات التصنيفية. ويتم عن طريق تحليل دالة المكونات الرئيسية لتجنب نسبة الفقد للمعلومات المهمة في المجال، ولتقديم طرق جديدة تساعد المختصين على تقييم الطلاب. ولاختبار منحنى خصائص المفردة عبر طريقة الخوارزميات التصنيفية، تم تجهيز منحنى خصائص المفردة لاختبار رياضيات مكون من 36 فقرة ظهرت 3مفردات فيه سهلة جداً و2صعبة جداً. وبطريقة تحليل المكونات الرئيسية تم تغطية نطاق القدرة كاملاً بتقسيم المفردات ما بين صعب ومتوسط وسهل، لأجل الحفاظ على تغطية جميع جوانب القوة لدى المختبرين. ثم تم اختبار طريقة الخوارزميات التصنيفية، بعد تقسيمها للمفردات على منحنى خصائص المفردة على 3 مناطق ملونة وترتيبية سهل-متوسط-صعب، أمكن بعدها حذف مفردات من المستويات الثلاث. لنخرج باختبار أكثر ملاءمة وأقل في عدد المفردات لأن الأسئلة الكثيرة تعني تشابه سلوك احتمالات الاجابة بشكل صحيح.

3- منهجية الدراسة وإجراءاتها.

منهجية البحث:

المنهج المستخدم هو منهج المحاكاة Simulation study method (ساري وهاغينز، 2015)، حيث إنه هو المنهج المناسب لطبيعة الدراسة الحالية، والذي يتطلب الضبط الدقيق لجميع المتغيرات لاستخلاص أثر المتغير المستقل في تصميم الدراسة الحالية وهو نسبة المفردات ذات الأداء التفاضلي على المتغير التابع وهو طريقة منحنى خصائص المفردة، وذلك بتطبيق3 اختبارات عشوائية كل منها من 100 فقرة بنسب تفاضل على الترتيب (10%-20%-30%) بـ 100 مرة إعادة.

أ- مصادر البيانات:

تم استخدام بيانات مولدة عبر برنامج المحاكاة، حيث إن البرمجيات كانت هي العمود الفقري الأساس في تفاصيل إجراءات الدراسة، ونوضح هذه البرمجيات ومواضع استخدامها في كامل الدراسة عبر التالي:

أ- برنامج WinGen-3:

حيث تم من خلاله توليد البيانات الخاصة بالبحث حسب العينة المدروسة وهي (3000 فرد)، كذلك تم استخراج الاستجابات ل100 فقرة اختبارية افتراضية بالنموذج ثلاثي المعلم، وذلك لكل اختبار من اختبارات النسب التفاضلية الثلاث وإحداث الأداء التفاضلي برمجياً من خلال نافذة DIF في البرنامج.

ب- برنامج BilogMG-3:

تم من خلاله تحليل البيانات بالنظرية الحديثة للقياس، وذلك لإيجاد معالم المفردات (الصعوبة-التمييز-التخمين) لجميع التكرارات الـ100 للمجموعتين المرجعية R والمستهدفة F من كل اختبار من الاختبارات الثلاث.

ت- برنامج Exile 2016:

وتم من خلاله حساب العديد من الخطوات المتعلقة بالدراسة وهي: إيجاد متوسطات الـ100 تكرار لكل اختبار من الاختبارات الثلاث لإيجاد المعالم الحقيقية لكل مفردة في المجموعة المرجعية R وفي المجموعة المستهدفة F. ثم حساب معادلة راجو (Raju, 1988) بين المجموعتين R وF لإيجاد المساحة بين منحنياتهما، وذلك عن طريق وضع عمودين لمعالم المجموعتين RوF وتطبيق المعادلة. وبعدها العمل على إيجاد قيمة القطع Cut-Off Value عبر عدد من عينات المجموعة المرجعية. ثم القيام بتطبيق اختبار (ت) لدلالة الفروق بين المتوسطات، وأخيراً تطبيق معادلة كوهين D لاستخلاص حجم التأثير.

إجراءات البحث:

للإجابة عن أسئلة الدراسة تم إتباع إجراءات متسلسلة أولها كان القيام بمحاكاة 3 اختبارات، مكون كل منها من 100 مفردة اختبارية، تحوي على الترتيب نسب تفاضل (10%-20%-30%). تم توليد البيانات ثم تحليلها لاستخلاص معالم مفردات المجموعتين المرجعية Rوالمستهدفة F لكل اختبار من الاختبارات الثلاثة، ليتسنى بعدها تطبيق طريقة اختبار منحنى خصائص المفردة، وذلك بإيجاد المساحة بين المفردات لكل اختبار من الاختبارات الثلاثة، والحكم على وجود أداء تفاضلي أم لا للمفردة من خلال إيجاد قيمة القطع كما جاء في أيوب(2008) عن طريق اخذ عينات عشوائية من المجموعة المرجعية R لكل نسبة من نسب التفاضل وإيجاد أكبر قيمة بين منحنياتها والتي كانت كالتالي:

جدول (1) يوضح نقاط القطع لكل نسبة من نسب الأداء التفاضلي المدروسة

النسبة	نقطة القطع
%10	0، 539706
%20	0، 683105
%30	0، 677964

ويتضح من الجدول التالي قيمة الدلالة لاختبار (ت) بين المجموعتين من المفردات (التي تبدي أداء تفاضلي والتي لاتُبدي أداء تفاضلي)، لكل نسبة من النسب والتي جاءت جميعها $\alpha \leq 0.05$، مايعني أنه توجد فروق في المساحات مابين المفردات التي تُبدي والتي لاتبدي لصالح المفردات التي تُبدي أداء تفاضلي:

جدول (2) يوضح قيمة الدلالة لكل نسبة من نسب الأداء التفاضلي المدروسة

النسبة	الدلالة α
%10	0، 003525
%20	0، 000142
%30	0، 000017

4- عرض النتائج ومناقشتها.

- **الإجابة على السؤال الأول: ما أثر وجود المفردات ذات الأداء التفاضلي بنسبة 10%على طريقة اختبار منحنى خصائص المفردة؟**

وللإجابة تم استخلاص النتائج بعد استخدام معادلة راجو Raju (1988) في إيجاد الفروق عن طريق تطبيقها ببرنامج Exile2016. والجدول التالي يوضح الفروق بين المساحات بين المجموعتين المرجعية R والمستهدفة F لكل مفردة من المفردات الـ100 للاختبار الأول والذي يحوي نسب تفاضل 10%:

جدول (3) الفروق بين منحنيات المجموعتين المرجعية والمستهدفة للاختبار ذو نسبة الأداء التفاضلي 10%

رقم المفردة	المساحة	رقم المفردة	المساحة	رقم المفردة	المساحة	رقم المفردة	المساحة
1	0.00951	26	0.061409	51	0.387939	76	0.002217
2	0.087049	27	0.165342	52	0.2951	77	0.024679
3	0.054944	28	0.037608	53	0.098605	78	0.269079
4	0.244513	29	0.013349	54	0.227695	79	0.092673
5	0.082188	30	0.178745	55	0.012836	80	0.38794
6	0.352781	31	0.067846	56	0.119079	81	0.084226
7	0.115279	32	0.048346	57	0.098334	82	0.157943
8	0.021294	33	0.044464	58	0.141416	83	0.038211
9	0.438607	34	0.033271	59	0.143722	84	0.130395
10	0.02302	35	0.141528	60	0.084706	85	0.001958
11	0.305807	36	0.249109	61	0.044661	86	0.019282
12	0.265041	37	0.363832	62	0.07712	87	0.04059
13	0.046525	38	0.148088	63	0.216928	88	0.182681

رقم المفردة	المساحة	رقم المفردة	المساحة	رقم المفردة	المساحة	رقم المفردة	المساحة
14	0.117843	39	0.028927	64	0.154643	89	0.064956
15	0.029519	40	0.206188	65	0.385868	90	0.114251
16	0.010637	41	0.200916	66	0.110887	91	*0.786286
17	0.028682	42	0.023746	67	0.114451	92	*2.28241
18	0.331481	43	0.099103	68	0.284553	93	*5.645393
19	0.034536	44	0.40849	69	0.219451	94	*1.642423
20	0.02796	45	0.388458	70	0.157451	95	*8.714048
21	0.314591	46	0.387752	71	0.224203	96	*2.014958
22	0.058036	47	0.368548	72	0.187782	97	*2.964541
23	0.216123	48	0.28295	73	0.231721	98	*11.61608
24	0.275242	49	0.085246	74	0.146114	99	*1.542306
25	0.3459	50	0.25778	75	0.198587	100	*3.293244

(*) تشير إلى تجاوز المساحة لنقطة الحد التي تساوي = 0.539706

- **الإجابة على السؤال الثاني: ما أثر وجود المفردات ذات الأداء التفاضلي بنسبة20% على طريقة اختبار منحنى خصائص المفردة؟**

وللإجابة على السؤال تم اتباع نفس الإجراءات المطبقة في السؤال الأول والجدول التالي يوضح نتائج اختبار منحنى خصائص المفردة للاختبار ذو نسبة التفاضل 20% بواقع 20 مفردة تفاضلية:

جدول (4) الفروق بين منحنيات المجموعتين المرجعية والمستهدفة للاختبار ذو نسبة الأداء التفاضلي 20%

رقم المفردة	المساحة	رقم المفردة	المساحة	رقم المفردة	المساحة	رقم المفردة	المساحة
1	0.116299	26	0.028341	51	0.046704	76	0.086409
2	0.060134	27	0.160464	52	0.019222	77	0.183909
3	0.034512	28	0.033959	53	0.410696	78	0.105624
4	0.067227	29	0.026698	54	0.03075	79	0.301802
5	0.045319	30	0.023662	55	0.186116	80	0.381714
6	0.218789	31	0.077217	56	0.314578	81	*0.863711
7	0.180511	32	0.054455	57	0.194636	82	*8.088365
8	0.416167	33	0.469122	58	0.086179	83	*1.521329
9	0.091734	34	0.041894	59	0.135447	84	*1.812593
10	0.130936	35	0.04618	60	0.191846	85	*2.401522
11	0.233108	36	0.068565	61	0.067265	86	*0.95015
12	0.015459	37	0.050764	62	0.006916	87	*1.131325
13	0.259421	38	0.199116	63	0.230981	88	*0.915307
14	0.261432	39	0.245496	64	0.227867	89	*2.464971
15	0.224324	40	0.104169	65	0.282571	90	*11.79255
16	0.091761	41	0.197635	66	0.210331	91	*2.686007
17	0.09531	42	0.054616	67	0.201967	92	*3.292844

رقم المفردة	المساحة	رقم المفردة	المساحة	رقم المفردة	المساحة	رقم المفردة	المساحة
18	0.169635	43	0.038726	68	0.041455	93	*1.093552
19	0.075501	44	0.035631	69	0.126372	94	*2.773169
20	0.268968	45	0.134452	70	0.115368	95	*1.544392
21	0.28499	46	0.254643	71	0.229466	96	*1.563484
22	0.043613	47	0.003411	72	0.214737	97	*4.17073
23	0.160127	48	0.034642	73	0.118647	98	*0.848746
24	0.074573	49	0.107312	74	0.034505	99	*6.693359
25	0.053294	50	0.139826	75	0.002635	100	*2.313471

*تشير إلى تجاوز المساحة لنقطة الحد التي تساوي = 0.683105

- **الإجابة على السؤال الثالث: ما أثر وجود المفردات ذات الأداء التفاضلي بنسبة30% على طريقة اختبار منحنى خصائص المفردة؟**

وللإجابة تُشير النتائج في الجدول أدناه إلى المساحات تحت المنحنى للمفردات ذات الأداء التفاضلي بنسبة 30% من اختبار بطول 100 مفردة موضحة كالتالي:

جدول (5) الفروق بين منحنيات المجموعتين المرجعية والمستهدفة للاختبار ذو نسبة الأداء التفاضلي 30%

رقم المفردة	المساحة	رقم المفردة	المساحة	رقم المفردة	المساحة	رقم المفردة	المساحة
1	0.008832	26	0.141453	51	0.103268	76	*1.405034
2	0.151999	27	0.228647	52	0.348184	77	*10.09817
3	0.453519	28	0.22991	53	0.259979	78	*1.158799
4	0.094714	29	0.268328	54	0.120892	79	*2.040161
5	0.031445	30	0.269061	55	0.062143	80	*0.837682
6	0.081939	31	0.18529	56	0.19828	81	*1.440185
7	0.102104	32	0.074057	57	0.211071	82	*2.753558
8	0.157091	33	0.509883	58	0.093057	83	*1.588208
9	0.064694	34	0.015485	59	0.13815	84	*2.849756
10	0.163666	35	0.001287	60	0.082842	85	*9.398829
11	0.163501	36	0.083105	61	0.070036	86	*1.228197
12	0.064783	37	0.004721	62	0.00421	87	*1.421082
13	0.061132	38	0.049571	63	0.032227	88	*2.640813
14	0.071816	39	0.133954	64	0.353855	89	*1.024128
15	0.078091	40	0.02307	65	0.16312	90	*0.839025
16	0.030907	41	0.271503	66	0.119459	91	*2.21059
17	0.309656	42	0.282879	67	0.173593	92	*4.203789
18	0.027483	43	0.067264	68	0.007023	93	*0.836416
19	0.400183	44	0.425485	69	0.506788	94	*3.241663
20	0.065874	45	0.105769	70	0.173641	95	*0.732747
21	0.172934	46	0.008345	71	*0.982148	96	*0.737216

رقم المفردة	المساحة	رقم المفردة	المساحة	رقم المفردة	المساحة	رقم المفردة	المساحة
22	0.206097	47	0.023036	72	*1.691767	97	*0.882676
23	0.223664	48	0.063534	73	*0.767044	98	*4.423376
24	0.015467	49	0.285176	74	*0.797816	99	*2.091868
25	0.310819	50	0.095308	75	*6.515877	100	*1.141191

*تشير إلى تجاوز المساحة لنقطة الحد التي تساوي = 0.677964

- **الإجابة على السؤال الرابع: هل توجد فروق دالة إحصائياً على طريقة اختبار منحنى خصائص المفردة تبعاً لاختلاف نسب وجود المفردات ذات الأداء التفاضلي مرة 10% ومرة 20% ومرة 30%؟**

وللإجابة ومن خلال هذا السؤال، ندرس مقارنة قوة اختبار منحنى خصائص المفردة بين نسب التفاضل لثلاثة اختبارات طويلة مكونة من 100 مفردة، تظهر نسبة التفاضل على الترتيب (10%-20%-30%). وفيما يخص موضوع إيجاد قيمة قوة الاختبار فإنه من غير المنطقي المقارنة فقط من خلال الاعتماد على مخرجات قيم المساحات بين المفردات التفاضلية لعدة أمور أولها؛ أن عدد المفردات يختلف حسب النسبة التفاضلية في كل اختبار من اختبارات الثلاث والثاني؛ عدم وجود قاعدة لضبط حجم الفروق في مخرجات المساحة ما يتسنى على أساسه المقارنة بين النسب الثلاث –على حد علم الباحثتين-. وعلى ضوء ذلك تم اعتماد طريقة لإيجاد قوة الاختبار من البيانات المستخرجة للمساحة تحت المنحنيات وفق طريقة اختبار منحنى خصائص المفردة. فبعد تطبيق اختبار (ت) اتضح وجود فروق في المساحات داخل الاختبار الواحد يوضحها الجدول التالي:

جدول (6) قيمة اختبار (ت) لاختبارات نسب التفاضل الثلاث

نسبة التفاضل	%10		%20		%30	
نوع المفردات	غير تفاضلية	تفاضلية	غير تفاضلية	تفاضلية	غير تفاضلية	تفاضلية
المتوسط	156701 ،0	168 ،050 ،4	138636 ،0	079 ،946 ،2	151148 ،0	327 ،399 ،2
التباين	0.014748	12.58862	0.011439	8.007899	0.016062	5.749033
العدد	90	10	80	20	70	30
درجات الحرية	9		19		29	
اختبار T	3.46992-		4.43598-		5.13256-	
الدلالة	0.003525		0.000142		0.000017	

ومما سبق فإنه ومن خلال استخدام برنامج Exile2016، فقد تم استخراج قيمة التباين المشترك لكل اختبار من الاختبارات الثلاث، ومن ثم اجراء معادلة كوهين D على كل واحد منها، من خلال إيجاد الفرق بين المجموعتين الضابطة والتجريبية المحددة لكل اختبار ومن ثم قسمتها على التباين المشترك، لإيجاد حجم التأثير الذي ظهرت نتائجه المُعبرة عن قوة الاختبار لكل نسبة من نسب التفاضل في الجدول التالي:

جدول (7) نتيجة اختبار كوهين D لقوة الاختبار لكل نسبة من نسب التفاضل

النسبة	قوة الاختبار
%10	619 ،0

النسبة	قوة الاختبار
%20	703 ،0
%30	786 ،0

النتائج في هذا المبحث تُظهر أن اختبار منحنى خصائص المفردة فعّال في الكشف عن المفردات ذات الأداء التفاضلي، أياً كانت نسبة وجودها (10%-20%-30%). كما أنه وفق ما عُرض من نتائج تزداد قوته في إيجاد المفردات التفاضلية كلما زاد عدد المفردات التفاضلية ونسبة وجودها في الاختبار.

مناقشة النتائج:

لقد تبيّن من خلال المبحث السابق؛ أن طريقة اختبار منحنى خصائص المفردة والتي تمت في هذه الدراسة، كانت فعّالة في الكشف عن مقدار التفاضل، وذلك حسب نسبة الأداء التفاضلي للمفردات التي تم تصميمها لتُبدي أداءات تفاضلية على الترتيب بنسب (10%-20%-30%) وبمقدار (10-20-30) مفردة تفاضلية في الاختبارات الثلاث. الاختبارات المصممة مكون كل منها من 100 مفردة اختبارية المطبقة عبر أسلوب المحاكاة على 3000 فرد، موزعة قدراتهم توزيعاً طبيعياً بمتوسط (0) وانحراف معياري (1). وتم المقارنة بين مخرجات طريقة منحنى خصائص المفردة في حساب المساحات بين المنحنيات لمفردات الاختبارات الثلاث.

وكما هو مُلاحظ بالجدول رقم (3) في المبحث الثاني –الخاص بالنتائج- فإن العشرة فقرات الأخيرة من الاختبار الخاص بنسبة التفاضل 10% والتي تم اختيارها لتُبدي أداء تفاضلي، قد تجاوزت نقطة القطع. ما يعني أنها بالفعل صُممت لتُبدي أداء تفاضلي وبالفعل أبدت أداء تفاضلي بين الأفراد، ما يعني فاعلية طريقة منحنى خصائص المفردة في كشف التحيز. وعند النظر إلى المساحات في الجدول (3) فإن أكبر مساحة تفاضلية كانت في الفقرة 98 حيث بلغت 11 وأقل مساحة تفاضلية كانت للفقرة 91 حيث بلغت 0، 786286 وقاربت قيمة القطع المحددة.

وتُظهر النتائج في الجدول رقم (4) تجاوز مساحات المفردات العشرون الأخيرة قيمة القطع، ما يعني أن اختبار منحنى خصائص المفردة جاء فعالاً أيضاً في إيجاد المفردات المصممة التي تُبدي أداء تفاضلي بنسبة 20% من أصل 100 مفردة اختبارية. ويُلاحظ أن أكبر قيمة للمساحة هي للمفردة رقم 90 تلها 83 بمساحتين الأولى باتساع 11 والثانية باتساع 8 كما أن أقل مساحة تفاضلية والتي قاربت من نقطة القطع كانت للمفردتين رقم 81 و98.

ويظهر في الجدول رقم (5) تجاوز جميع المفردات التفاضلية نقطة القطع، حتى وإن قارب بعضها قيمة النقطة إلا أنها صُنفت كمفردات تُبدي أداء تفاضلي لتحقيقها الشرط لذلك. ويُلاحظ أن هناك 3 مفردات كانت ذات مساحات كبيرة للأداء بين المجموعتين المرجعية R والمستهدفة F وهي المفردات (75-77-85) حيث بلغ مداها من 6 إلى 10، أما أقربها لنقطة القطع فكانت المفردات رقم (73-74-95-96-97).

وبعد قراءة نصار (2006) والشاردي (2015)، تم تطبيق معادلة كوهين D وهي الوسيط في إيجاد قوة الاختبار لكل نسبة من النسب التفاضلية الثلاث. والتي أظهرت نتائجها أن قوة الاختبار تزداد كلما زاد عدد المفردات التفاضلية ونسبة وجودها في الاختبار، حيث كانت قوة الاختبار لنسبة الأداء التفاضلي 10% تبلغ (0، 619) وهي قيمة تدل على أن قوة الاختبار أعلى من المتوسط، وكذلك قوة الاختبار لنسبة التفاضل 20% التي جاءت قيمتها (0، 703) وهي قيمة تدل على أن قوة الاختبار أعلى من المتوسط أيضاً، وأخيراً قوة الاختبار لنسبة التفاضل 30% حيث بلغت قيمتها (0، 786) وتدل على أن قوة الاختبار أعلى من المتوسط من حيث القوة حسب تصنيف قوة الاختبار لمعادلة حجم التأثير لكوهين D.

ووُفقاً للإطار النظري في تفسير حجم التأثير وفقاً لاختبار كوهين D، فإنه وبما يظهر من الجدول رقم (7) أنه عند وجود مفردات ذات أداء تفاضلي نسبتها 10% و20% فإن قوة الاختبار تكون أعلى من المتوسط، وتُقارب قوة الاختبار للمفردات التفاضلية بنسبة 30% الحد الأعلى للحكم على الاختبار بأنه قوي بمقدار طفيف. وتتفق نتائج هذه الدراسة مع دراسة أبوعلام والمرابحة (2015) في أن قوة الاختبار تزداد بزيادة المفردات ذات الأداء التفاضلي، وتتفق نتائج هذه الدراسة كذلك مع دراسة أيوب (2008) على فعالية طريقة منحنى خصائص المفردة في إيجاد المفردات ذات الأداء التفاضلي.

التوصيات والمقترحات.

بناءً على النتائج التي تم التوصل إليها توصي الباحثتان وتقترحان بالتالي:

1- استعمال طريقة منحنى خصائص المفردة ICC في الكشف عن المفردات التفاضلية عبر تجارب أخرى باستعمال طريقة المحاكاة.
2- تطبيق طريقة منحنى خصائص المفردة على بيانات حقيقة لاختبارات يُشتبه في وجود أداء تفاضلي بين مختبريها كالذكور والإناث على سبيل المثال لا الحصر.
3- تطبيق هذا البحث على بيانات حقيقية عن طريق المقاييس النفسية ومقاييس الاتجاهات والاستبانات الطويلة، والتي يُشتبه في كونها تحوي مفردات تفاضلية ودراسة حجم التأثير ونسبة وجود الأداء التفاضلي فيها ليسهل اكتشاف مقدار الأداء التفاضلي في هذه المقاييس وتعديلها لإزالة أثر التحّيز منها ليضمن المطبقون لها العدالة الاختبارية لمقاييسهم البحثية.

خاتمة:

وأخيراً، فإن هذا النوع من الدراسات يُمثل توجه حديث في دراسة مواقف اختبارية عديدة، مُساهماً في تطوير أدوات القياس، وشاملاً لجميع قدرات المختبرين في منحنى خصائص المفردة الواحدة ومنحنى خصائص مفردات الاختبار مجتمعة. والأصل هو وجود القدرات لدى جميع الأفراد، كونها في المستوى الطبيعي الذي كتبه الله تعالى في أساس خلقه للأرض ومن عليها، فالكل لديه القدرة بشرط توفر المفردة العادلة من غير تفضيل فئة على أخرى، في أدوات يفترض فيها العدل للاستخلاص معلومات أو بيانات تؤدي أغراضاً نبيلة.

قائمة المراجع

أولاً- المراجع بالعربية

- اجباره، محمد عدنان محمد (2015). أثر نسبة المفردات ذات الأداء التفاضلي على الخطأ من النوع الأول وقوة الاختبار في الكشف عنه باستخدام طريقتي الانحدار اللوجستي ونسبة الأرجحية العظمى، (رسالة دكتوراة غير منشورة)، جامعة اليرموك، كلية العلوم التربوية.
- أيوب، عبد الكريم محمود عبد (2008). أثر خصائص توزيع القدرة وطول الاختبار وحجم العينة في الكشف عن الدالة التفاضلية للفقرة: دراسة محاكاة (رسالة ماجستير)، جامعة اليرموك، كلية التربية.
- الحجيلي، خالد هديبان سليمان (2012). أثر طول الاختبار وحجم العينة في دقة تقدير معلمة صعوبة المفردة والقدرة ومعادلة الاختبارات بوجود الأداء التفاضلي للفقرة، (رسالة دكتوراه) جامعة اليرموك، كلية التربية.

- دي ايالا، آر جي (2017). النظرية والتطبيق في نظرية الاستجابة للفقرة (ترجمة عبد الله الكيلاني وإسماعيل البرصان). الرياض: دار جامعة الملك سعود للنشر.
- رينولدز، سيسل آر (2013). إتقان القياس النفسي الحديث النظريات والتطبيق (ترجمة صلاح الدين محمود علام). عمان: دار الفكر.
- الشاردي، محمد (2012). تأثير حجم العينة على قوة الاختبار الإحصائي، (رسالة ماجستير)، جامعة أم القرى، كلية التربية.
- الطريري، عبد الرحمن (2014). القياس النفسي والتربوي نظريته – أسسه – تطبيقاته، الرياض: مكتبة الرشد.
- عبابنه، عماد غصاب (2008). جدوى التحول من النظرية الكلاسيكية في القياس إلى نظرية الاستجابة للفقرة: تحليل الإمكانيات والمحددات، عالم التربية، مصر، 9(27).
- علام، رجاء محمود؛ والمرابحة، عامر (2014). أثر طول الاختبار وحجم العينة على تضخم الخطأ من النوع الأول لإحصائيات طرق الكشف عن دالة المفردة التفاضلية: دراسة محاكاة، عالم التربية، مصر، 15 (47)، 95-124.
- علام، رجاء محمود؛ والمرابحة، عامر (2015). أثر نسبة المفردات ذات الأداء التفاضلي وحجم العينة على قوة وفاعلية اختبار تحيز المفردة المتزامن: دراسة محاكاة، المجلة العربية للعلوم الاجتماعية: المؤسسة العربية للاستشارات العلمية وتنمية الموارد البشرية، مصر، 2(7)، 139-165.
- علام، صلاح الدين محمود (2000). القياس والتقويم التربوي والنفسي أساسياته وتطبيقاته وتوجهاته المعاصرة. القاهرة: دار الفكر العربي.
- اللبدي، نزار راسم محمد (2008). مقارنة أربع طرائق للكشف عن دقة المفردة التفاضلية: دراسة محاكاة، (رسالة دكتوراه غير منشورة)، الجامعة الأردنية، كلية الدراسات العليا.
- نصار، يحيى حياتي (2006). استخدام حجم الأثر لفحص الدلالة العملية للنتائج في الدراسات الكمية. الأردن: مجلة العلوم التربوية والنفسية.

ثانياً- المراجع بالإنجليزية

- Binder, K. & Heermann, W.D. (1997). Monte Carlo Simulation in Statistical Physics. Springer-Verlag Berlin Heidelberg GmbH
- Crocker, L., & Algina J. (2006); Introduction to Classical and Modern Test Theory, Wadsworth Pub Co.-
- Fortuna, F., & Maturo, F. (2018). K-means clustering of item characteristic curves and item information curves via functional principal component analysis, Springer Science+ Business Media B.V., part of Springer Nature 2018, Publish online: 3March2018.
- Frye, Curtis D. (2013). Microsoft Excel 2013 Step by Step. Microsoft Office.
- Han, T. & Hambelton, K. (2014). User's Manual for WinGen3. University of Massachusetts Amherst.
- Sari, H., & Huggins, A. (2015). Differential Item Functioning Detection Across Two Methods of Defining Group Comparison, Educational and Psychological Measurement, 75(4) 648–676.
- Seybert, J., & Strak, S. (2013). Detecting DIF with Ideal Point Models, Applied Psychological Measurement, 38(2) 151–165.