

الأداء التفاضلي لمفردات اختبار القدرة اللفظية في مقياس الخليج للقدرات العقلية المتعددة لطلبة دول مجلس التعاون الخليجي تبعا لمتغيري الجنس والدولة

هلال أحمد المسكري*، راشد سيف المحرزي، عبد الحميد سعيد حسن
وزارة التربية والتعليم، سلطنة عمان جامعة السلطان قابوس، سلطنة عمان

قبل بتاريخ: ٢٠٢٠/٧/٣٠

استلم بتاريخ: ٢٠٢٠/٤/٩

ملخص: هدفت الدراسة الكشف عن الأداء التفاضلي لمفردات اختبار القدرة اللفظية بالمستوى الثاني في مقياس جماس نحو الطالبات الإناث وطلبة سلطنة عمان، وذلك باستخدام طريقتي مانتل هانزل، والصعوبة المحوثة. تكون الاختبار من ٣٠ مفردة من نوع الاختيار من متعدد، بأربعة بدائل، وتكونت عينة الدراسة من بيانات أرييفية لطلبة الصفين الثالث والرابع الابتدائيين في دول مجلس التعاون الخليجي، والبالغ عددهم ٤٢٨٠ طالباً وطالبة. أشارت نتائج الدراسة أن طريقة مانتل هانزل كشفت أن ٦٠% من مفردات الاختبار أظهرت أداءً تفاضلياً تبعاً لمتغير النوع الاجتماعي. بالمثل فإن ٦٠% من مفردات هذا الاختبار أظهرت أداءً تفاضلياً تبعاً لمتغير الدولة. ويُعد هذا الأداء التفاضلي ضعيفاً في معظم المفردات حسب مؤشر D. وقد أظهرت نتائج طريقة الصعوبة المحوثة وجود أداء تفاضلي بنسبة ٣٠% تبعاً لمتغير النوع الاجتماعي و٣٣,٣٣% تبعاً لمتغير الدولة. أما فيما يتعلق بفحص الاتفاق بين طريقة مانتل هانزل وطريقة الصعوبة المحوثة فكانت قيمة معامل ثبات التصنيف كابا ٠,٥٢٤، وبنسبة اتفاق ٧٠% تبعاً لمتغير النوع الاجتماعي ويعتبر توافقاً متوسطاً؛ بينما كانت قيمة معامل ثبات التصنيف كابا ٠,١٥٨، وبنسبة اتفاق ٤٦,٦٧% تبعاً لمتغير الدولة ويعتبر توافقاً ضعيفاً. وأوصت بدراسة الأسباب التي أدت إلى ظهور الأداء التفاضلي في مفردات اختبار القدرة اللفظية بالمستوى الثاني لمقياس جماس التي تم الكشف عنها لتجنبها ومعالجتها.

كلمات مفتاحية: الأداء التفاضلي للمفردة، اختبار القدرة اللفظية، مقياس جماس، طريقة مانتل هانزل، طريقة الصعوبة المحوثة.

Differential Item Functioning of Verbal Ability Test in Gulf Multiple Mental Ability Scale for GCC Students According to Gender and Country

Hilal A. Almaskari*, Rashid S. Almehrzi, & Abdulhameed S. Hassan
Ministry of Education, Sultanate of Oman Sultan Qaboos University, Sultanate of Oman

Abstract: The study aimed to investigate the differential item functioning (DIF) for verbal ability test items in the Gulf Multiple Mental Ability Scale for female students in general and Omani Female students in particular in using Mantel-Haenszel (MH) and the Transformed Item Difficulty (TID) methods. The test consisted of 30 multiple-choice items with four distractors. The study sample consisted of the archive data for 4280 students of the third and fourth grades in GCC countries. The results revealed that Sixty Percent of the items showed DIF related to gender using MH. Similarly, Sixty Percent of the items showed DIF related to country using MH. DIF values were small indicating weak DIF in most items. Results also indicated that DIF using TID was found for Thirty percent of the items related to gender, and 33.33% related to country. Furthermore, Kappa coefficient 0.524 was moderate between MH method and TID for gender and the agreement ratio was 70%. Kappa coefficient 0.158 was weak between MH method and TID for country and the agreement ratio was 46.67%. Based on the study results the researchers recommend investigating the reasons behind the detected differential functioning of some items in the verbal ability test at the second level of the GMMAS scale to avoid and address it.

Keywords: Differential item function, verbal ability test, GMMAS, mantel-haenszel, transformed item difficulty.

*H.almaskari@moe.om

(Functioning, DIF). وقد عرفه البرصان (٢٠١٣) بأنه دالة مشتقة إحصائياً للتعبير عن الفروق في الاستجابة على المفردة بين مجموعتين من المفحوصين ممن هم في نفس مستوى القدرة. إذ إن الهدف من دراسة الأداء التفاضلي هو العمل على جعل الاختبارات خالية من التحيز ما أمكن؛ وذلك لجعل القرارات التي يتم اتخاذها بناءً على نتائج دقيقة وعادلة، وإتاحة المجال لمساواة الأفراد في فرص التعليم، وغيرها من الفرص التي تتطلب الدقة، والمساواة بين المستهدفين.

وقد بين كل من هامبلتون وسوامينا ثان وروجرز (Hambleton, Swaminathan & Rogers, ١٩٩١) سبب استخدام مصطلح الأداء التفاضلي بدلاً من التحيز، وقد استند في ذلك على أن الاستقصاء عن التحيز يتضمن إيجاد أدلة تجريبية لها علاقة بالأداء النسبي لأفراد المجموعة المستهدفة، والمجموعة المرجعية، على مفردات الاختبار. فيعد الدليل التجريبي على الأداء التفاضلي ضرورياً، ولكنه غير كافٍ لإثبات وجود التحيز. لذا يُستخدم مصطلح الأداء التفاضلي بدلاً من التحيز للإشارة إلى الأدلة التجريبية التي يتم الحصول عليها عند الاستقصاء عن التحيز.

بدأ استخدام مصطلح الأداء التفاضلي للمفردة منذ بداية ثمانينات القرن العشرين للدلالة على الطرق، والمعالجات الإحصائية التي تُستخدم للكشف عن الأداء التفاضلي لمفردات الاختبارات، وهناك من يستخدم مصطلح الأداء التفاضلي للمفردة كمرادف لمصطلح التحيز، وهذا الاستخدام ليس صحيحاً؛ حيث إن مصطلح الأداء التفاضلي للمفردة يُستخدم لتعيين المفردات التي يكون احتمال الإجابة الصحيحة مختلفاً بين مجموعتين لهما نفس مستوى القدرة (Camilli & Shepered, 1994). لذا يُعد الأداء التفاضلي للمفردة شرطاً أساسياً لاعتبار المفردة متحيزة، لكنه شرط غير كافٍ، أي أنه إذا أظهرت مفردة ما أداءً تفاضلياً لمجموعة ما، في هذه الحالة لا بد من اتخاذ إجراءات إضافية للحكم على تلك المفردة بأنها متحيزة، على سبيل المثال قد تتضمن تلك الإجراءات تحكيمياً لتلك المفردة

يسعى مُعدّو أدوات القياس إلى تحقيق درجات عالية من صدق النتائج التي يحصلون عليها من تطبيق هذه الأدوات، بحيث تقيس ما وضعت لقياسه. ومن بين الطرق التي يمكن أن تساعد واضعي أدوات القياس في ذلك فحص وتحري وجود أداء تفاضلي للمفردات وتحيزها. فتواجه المختصين في القياس النفسي قضية ارتباط مفردات المقاييس، والاختبارات مع أنماط ثقافية خاصة بمجموعات الأفراد الذين يتقدمون لهذه الاختبارات؛ نتيجة تباين المجموعات بين بعضها وفقاً للنوع الاجتماعي، أو المستوى الاقتصادي، أو غيرها. وقد ساهمت نظريات القياس والتقييم التقليدية والحديثة في تسليط الضوء على موضوع الأداء التفاضلي لمفردات الاختبار (كروكر وألجينا، ٢٠٠٩).

وقد أثبتت العديد من الدراسات أن عدم وجود أداء تفاضلي في مختلف أدوات القياس يُعد ضرورة ملحة (البشاشة، ٢٠١٦؛ Wedman, 2018). إذ قد يظهر الأداء التفاضلي لمفردة ما؛ والذي قد يُفسّر على أنه تحيز بسبب عدة أمور، منها اختلاف المستويات الاقتصادية، والاجتماعية التي تؤدي إلى فروق ثقافية بين المجموعات الفرعية المختلفة في المجتمع، وهو ما يسمى بالتحيز الثقافي، أو بسبب المحتوى أو اللغة؛ حيث لا تكون اللغة مألوفة لبعض المجموعات الفرعية، أو من خلال استخدام الاختبارات الدولية والاختبارات الوطنية التي تدرج تحت التقييم واسع النطاق، أو حسب تصنيفات الطلاب المختلفة على مستوى الطالب، أو المدرسة، أو المنطقة التعليمية، أو الدولة. كما أنه قد يكون هناك أداء تفاضلي بسبب النوع الاجتماعي؛ إذ إن فروقاً في متوسط الأداء بين النوعين قد تظهر لصالح الذكور في الاختبارات ذات الطبيعة الذكورية، ولصالح الإناث في الاختبارات ذات الطبيعة الأنثوية (البرصان، ٢٠١٣).

وتتعدد مساهمات الأداء التفاضلي لمفردات الاختبار، كالدالة التفاضلية للمفردة، ودالة المفردة التفاضلية، والأداء الفارق للمفردة. واستخدمت هذه الدراسة مصطلح "الأداء التفاضلي للمفردة" (Differential Item)

في اختلاف أداء الأفراد تبعاً للمستويات المعيشية والاجتماعية لهم، مثال: (غني/ فقير).

٤. تحييز السكن: التحييز المتمثل في اختلاف أداء الأفراد تبعاً لأماكن سكنهم، مثال: (بادية، قرية، مدينة).

٥. تحييز لنوع الدراسة: تحييز المفردات المتمثل في اختلاف أداء الأفراد من حقول دراسية (تخصصات دراسية) مختلفة للطلبة، مثال (علمي/ أدبي).

للكشف عن المفردات المتحييزة لأشكال التحييز المختلفة، وضع علماء القياس النفسي عدداً من الطرق التي تساعد على كشف الأداء التفاضلي في مفردات الاختبار أو المقياس (حمادنة، ٢٠٠٧). ومن أبرز الطرق التي تستخدم للكشف عن الأداء التفاضلي للمفردة: طريقة مانتل هانزل، وطريقة الصعوبة المحولة، وطريقة تحليل التباين، وطريقة الانحدار اللوجستي، وطريقة معامل التمييز، وطريقة كاي تربيع، وطريقة تحليل إجابة المموه، وطريقة منحى خصائص الفقرة، وطريقة الفرق في معلمة الصعوبة، وطريقة نسبة الأرجحية العظمى.

ظهرت طريقة مانتل هانزل (MH) عام ١٩٥٩م، على يد كل من مانتل وهانزل Haenszel & Mantel، وتم تعديلها عام ١٩٨٨م على يد كل من هولاند وثاير Holland & Thayer، حتى يمكن استخدامها في دراسة الأداء التفاضلي للمفردة، حيث تقوم باستقصاء الأداء التفاضلي بين مجموعتين فرعيتين من مجتمع الدراسة، أحدهما تسمى المجموعة المرجعية وتسمى الأخرى المجموعة المستهدفة، وهي المجموعة التي يُعتقد وجود مفردات ذات أداء تفاضلي ضدها (Nandakumer, 1993).

أعدّ طريقة الصعوبة المحولة كل من أنجوف وفورد في عام ١٩٧١م، إذ تعتمد هذه الطريقة على معامل الصعوبة التقليدي للمفردة والذي يمثل نسبة الأفراد الذين أجابوا إجابة صحيحة عن المفردة، ويرتبط الأداء التفاضلي للمفردة في هذه الطريقة بالاختلاف الدال إحصائياً لمعامل

من قبل مختصين أو تقييماً تجريبياً لها (البرصان، ٢٠١٣). وفرق دورنز وهولاند (Dorans and Holland, 1994) بين مفهوم الأداء التفاضلي، ومفهوم التحييز للمفردة، فالأداء التفاضلي يدرس الخصائص السيكومترية للمفردة من حيث الكيفية التي تعمل بها في المجموعات المختلفة، بينما التحييز يحمل المعنى الاجتماعي، وعدم العدالة والمساواة.

يرى كاميلي وشيبارد (Camilli & Shepard, 1994) أن المفردة غير المتحييزة هي التي تتساوى فيها احتمالات النجاح لجميع المبحوثين في المجموعتين من ذوي القدرة نفسها. والمفردة غير المتحييزة، هي التي تتأثر بنفس مصادر التباين في جميع المجموعات الفرعية (كروكر والجينا، ٢٠٠٩). إن التحييز في المفردة يحدث عندما تؤدي خصائص المفردة إلى اختلاف في أداء الأفراد الذين يمتلكون القدرات نفسها، ولكنهم من مجموعات مختلفة (Hambleton & Rogers, 1995). ذكر كيم (Kim, 2000) أن المفردة المتحييزة هي التي تقيس سمة أخرى غير السمة الأصلية للأفراد في مجموعتين مختلفتين. وبلا شك، فإنه عندما يحتوي المقياس على مفردات متحييزة، فهذا يدل على وجود آثار سلبية في دقة القرارات التي ستؤخذ، وبالتالي عدم تحقق مفهوم العدالة المرجوة منه.

كما وضّح عدد من الباحثين ومنهم أرشيدات (٢٠٠٧)، والطراونة (٢٠١١)، والبشاشة (٢٠١٦)، بأن هناك مصادر مختلفة للتحيز، ومن أبرزها:

١. التحييز العرقي Racial Bias: تحييز المفردات الناتج من اختلاف أداء الأفراد من أعراق مختلفة في المجتمع الواحد، مثل: السود والبيض، أو مواطن وغير مواطن.
٢. تحييز النوع الاجتماعي Gender Bias: تحييز المفردات المتمثل في اختلاف أداء الأفراد حسب النوع الاجتماعي، ذكر/ أنثى.
٣. التحييز الاقتصادي الاجتماعي Socio-Economic Bias: تحييز المفردات المتمثل

حيث يتم من خلالها الحكم على الأداء التفاضلي للمفردة. A: الجزء المقطوع من المحور الرأسي. B: ميل خط الانحدار. x & y : قيم Δ لمعاملات الصعوبة في كل من المجموعتين. r_{xy} : معامل الارتباط بين قيم دلتا للمجموعتين. s_x & s_y : الانحراف المعياري لقيم دلتا في المجموعتين. M_x & M_y : المتوسطات الحسابية لقيم دلتا في المجموعتين. d_i : المسافة بين النقطة ومحور الانحدار. X_i, Y_i : قيم Δ_i في المجموعتين على الترتيب. ويرى أوسترلند (Osterlind, 1983) أنه لا يوجد اختبار للدلالة الإحصائية لبعد المفردات عن خط الانحدار، ولكن يمكن اعتبار المفردات ذات أداء تفاضلي إذا كان بعدها عن هذا الخط بأكثر من واحد انحراف معياري.

أجرى راجو ودراسجو وسليند (Raju, Drasgow & Slind, 1993) دراسة تهدف إلى المقارنة بين ثلاث طرق للكشف عن الأداء التفاضلي للمفردة، وهذه الطرق هي: طريقة كاي تربيع، وطريقة فرق المساحة بين منحنيات خصائص المفردة، وطريقة مانتل هانزل، وتم الحصول على البيانات اللازمة لذلك من تطبيق اختبار لفظي يشتمل على ٤٥ مفردة على عينة بلغ حجمها ٨٣٩ طالباً وطالبة، أشارت النتائج إلى اتفاق طريقة كاي تربيع، وطريقة فرق المساحة بين منحنيات خصائص المفردة في الكشف عن المفردات ذات الأداء التفاضلي، في حين أظهرت طريقة مانتل هانزل عدداً أكبر من المفردات ذات الأداء التفاضلي مقارنة بالطريقتين السابقتين.

وأجرت ستيج (Stage, 2000) دراسة هدفت إلى مقارنة ثلاث طرق للكشف عن الأداء التفاضلي للمفردة، وهي: طريقة الصعوبة المحوّلة، وطريقة منحني خصائص المفردة، وطريقة مانتل هانزل، تبعاً لمتغير النوع الاجتماعي، استخدمت الدراسة الجزء اللفظي من اختبار الاستعداد المدرسي السويدي المكون من ٢٠ مفردة من نوع الاختيار من متعدد، طُبّق الاختبار على عينة من ١٥٠٠ طالباً وطالبة في السويد، أظهرت النتائج درجة اتفاق عالية بين طريقة مانتل هانزل، وطريقة الصعوبة المحوّلة في

الصعوبة بين المجموعتين المرجعية والمستهدفة، وتعتمد على تعديل معامل الصعوبة p بتطبيق عدة خطوات، وتعتبر المفردة ليست ذات أداء تفاضلي إذا كانت تقع على ما يسمى بخط الموازنة، والذي يُرسم باستخدام قيم المعاملات المعدلة للصعوبة في المجموعتين المرجعية والمستهدفة، وإذا وقعت المفردة خارج خط الموازنة، فإنها تعتبر ذات أداء تفاضلي (عمارة، ٢٠٠٨)، وقد لخص ظاظا (٢٠٠٦) إجراءات تطبيق هذه الطريقة وفقاً للآتي: حساب معاملات الصعوبة لكل مفردة (P) داخل المجموعتين المرجعية والمستهدفة، إيجاد القيم المعيارية Z تحت المنحنى الطبيعي المقابل لـ $Q = 1 - P$ للمجموعتين المرجعية والمستهدفة، من أجل أن تكون العلاقة بين مجموعتي صعوبات المفردات خطية، ثم تحويل القيم المعيارية Z إلى تدرج دلتا بمتوسط حسابي ١٣، وانحراف معياري ٤، حسب المعادلة $\Delta = 4Z + 13$.

وأشار أنجوف وفورد (Angoff & Ford, 1971) أنه إذا كان هناك تشابه بين الصعوبة النسبية للمفردات في المجموعتين المرجعية والمستهدفة، فإن التمثيل البياني لقيم دلتا سيكون على شكل خط مستقيم قطري يمر بنقطة الأصل ويُصدّف الزاوية بين المحورين السيني والصادي، ويعدّ هذا الخط المستقيم بمثابة المحك الذي سيستخدم للحكم على الأداء التفاضلي للمفردات؛ حيث سيتم تحديد المفردات ذات الأداء التفاضلي في ضوء بعدها عن هذا الخط القطري، ويكون تحديد هذا الخط بمعرفه ميله B ، والجزء المقطوع من المحور الرأسي A ، حسب المعادلات الآتية:

$$B = \frac{(s_y^2 - s_x^2) + \sqrt{(s_y^2 - s_x^2) + 4r_{xy}s_y s_x}}{2r_{xy}s_x s_y} \quad \text{and} \\ A = M_y - B M_x$$

ثم يتم إيجاد بُعد المفردة d_i عن الخط القطري المستقيم باستخدام المعادلة الآتية:

$$d_i = \frac{B X_i + A - Y_i}{\sqrt{B^2 + 1}}$$

طريقة مانتل هانزل للكشف عن الأداء التفاضلي المنتظم للمفردات، بينما أُستخدِم اختبار بريسلو- داي للتجانس (Breslow-Day test of homogeneity) للكشف عن الأداء التفاضلي غير المنتظم. أشارت النتائج إلى وجود أربع مفردات ذات أداءٍ تفاضليٍّ مرتبطٍ بالنوع الاجتماعي، بينما لم تظهر أي من المفردات أداءً تفاضلياً مرتبطاً بالمنطقة.

وأما دراسة مبارك (٢٠١٠) فهدفت إلى الكشف عن الأداء التفاضلي لمفردات اختبار العلوم في الدراسة الدولية بيزا ٢٠٠٦، تبعاً لمتغيرات الدولة واللغة والانتماء لمنظمة التنمية والتعاون الاقتصادي، باستخدام طريقة مانتل هانزل، وطريقة نسبة الأرجحية العظمى. بلغ عدد مفردات اختبار العلوم ٩٧ مفردة، وتقدم للاختبار ٤٠٠٠٠ طالباً وطالبة، وكانت النتائج كالتالي: وجود ٦٣ مفردة ذات أداءٍ تفاضليٍّ تبعاً لمتغير الدولة، الأردن مقابل غيرها، حسب طريقة مانتل هانزل، والمفردات التي أظهرت أداءً تفاضلياً كان ٩٠% منها لصالح مجموعة الدول الأخرى، ووجود ٢٣ مفردة ذات أداءٍ تفاضليٍّ تبعاً لمتغير الدولة من أصل ٤٦ مفردة، حسب طريقة نسبة الأرجحية، والمفردات التي أظهرت أداءً تفاضلياً كان ٧٠% منها لصالح مجموعة الدول الأخرى، وقد تم تفسير النتائج المتعلقة بارتفاع نسبة المفردات التي أظهرت أداءً تفاضلياً حسب متغيرات الدراسة، إلى أن الاختبار قد لا يكون متحرراً من المناهج، وأنه تأثر بمناهج الدول المنتهية للمنظمة، كما أن مفردات الاختبار الأصلية مكتوبة باللغة الإنجليزية، وترجمة تلك المفردات قد يغير معناها مما يؤدي إلى حذف معلومات مهمة تجعل الاستجابة لها صعبة.

وأجرى مبارك ومبارك (٢٠١٣) دراسة هدفت إلى الكشف عن الأداء التفاضلي في اختبار الرياضيات في الدراسة الدولية بيزا ٢٠٠٦ تبعاً لمتغير اللغة، باستخدام طريقة مانتل هانزل. وبلغ عدد مفردات اختبار الرياضيات ٤٨ مفردة، وتقدم للاختبار ٤٠٠٠٠٠ مخصص، وقد بينت الدراسة وجود ٢٧ مفردة بها أداءً تفاضلياً تبعاً لمتغير اللغة

النظرية التقليدية، حيث بلغت درجة الاتفاق ٩٨%، بينما بلغت درجة الاتفاق بين طريقة منحنى خصائص المفردة وطريقة مانتل هانزل ٩٢%، وبلغت درجة الاتفاق بين طريقة الصعوبة المحوِّلة في النظرية التقليدية، وطريقة منحنى خصائص المفردة حسب النظرية الحديثة ٨٩%.

وفي دراسة أجراها أوتيرويجك وقانن (Uiterwijk & Vallen, 2005) هدفت إلى الكشف عن الأسباب اللغوية المحتملة لظهور الأداء التفاضلي في الاختبار النهائي للتعليم الابتدائي في هولندا، وهذا الاختبار يقيس ثلاثة موضوعات فرعية هي: اللغة، والرياضيات، ومعالجة المعلومات، وقد تكون الاختبار من ٥٤٠ مفردة من نوع الاختيار من متعدد، حيث طُبِّق الاختبار على عينة من الطلاب تبعاً لمتغير العرق، أظهرت نتائج الدراسة وجود ٤٨ مفردة بها أداء تفاضلي لصالح الطلبة المهاجرين و٣٠ مفردة ليس بها أداء تفاضلي، وذلك حسب طريقة مانتل هانزل.

وأجرى حمادنة (٢٠٠٧) دراسة هدفت إلى مقارنة الأداء التفاضلي لمفردات اختبار اللغة الإنجليزية المُطبَّق في ثلاث جامعات أردنية، هي: اليرموك، العلوم والتكنولوجيا، آل البيت، تبعاً لمتغيري النوع الاجتماعي (ذكر، أنثى)، وفرع الثانوية العامة (أدبي، علمي) بحسب طريقة مانتل هانزل، وطريقة فرق المساحة بين منحنيات خصائص المفردة في النموذج اللوجستي ثلاثي المعلمة، والتوصل إلى درجة التوافق بين طريقتي الكشف المستخدمتين. إذ طُبِّق الاختبار على عينة مكونة من ١٩٣٥ طالباً وطالبة. وأشارت النتائج أن أكثر المهارات التي أظهرت مفرداتها أداءً تفاضلياً لمتغير النوع الاجتماعي، هي القواعد والمفردات، وكانت في معظمها لصالح الإناث.

فيما أجرى دريانا (Driana, 2007) دراسة هدفت إلى الكشف عن المفردات ذات الأداء التفاضلي المرتبط بمتغيري النوع الاجتماعي، والمنطقة في اختبار التفوق الرياضي للصف التاسع، والمكون من ٤٠ مفردة من نوع الاختيار من متعدد، تم تطبيق الاختبار على عينة من الطلبة بلغت ١٩١٩٨ من الذكور والإناث، تم استخدام

وأجرى ودمان (Wedman, 2018) دراسة هدفت إلى تعرف أسباب الأداء التفاضلي بين الجنسين على اختبار القبول بالكلية في جامعة سيديني والمكون من ٨٠٠ مضرده، والمطبّق على ٢٥٠٠٠٠ من المتقدمين للاختبار، وللكشف عن الأداء التفاضلي للمفردات تم استخدام طريقة مانتل هانزل، وطريقة الانحدار اللوجستي. وأشارت النتائج إلى أن عدد المفردات اللفظية التي أظهرت أداءً تفاضلياً تبعاً لمتغير النوع الاجتماعي كان أكبر بكثير من تلك المفردات الكمية، ولصالح الإناث، وأوصت الدراسة بإحداث تغييرات في كيفية أخذ عينات المفردات واستخدام صيغة إكمال الجملة في فهم القراءة باللغة الإنجليزية، وبالتالي زيادة الإنصاف بين الجنسين في الاختبار المدرس.

مشكلة الدراسة

ركزت الجهود التي بذلها الباحثون في مجال بناء وتطوير المقاييس والاختبارات بأنواعها كاختبارات القدرات العقلية، والتحصيل، والاتجاهات، والشخصية، وغيرها على إيجاد دلالات الصدق، والثبات، وفاعلية المفردات من حيث مستوى صعوبتها وتمييزها؛ وذلك بهدف جعلها أدوات يمكن الوثوق بها في وصف السمات المراد قياسها، وفي عمليات تقنين هذه الاختبارات، ومعايرتها، واستخدامها في المجالات الخاصة التي صممت لها (كروكر والجينا، ٢٠٠٩).

يظهر الأداء التفاضلي للمفردات في الاختبارات التي تستخدم الألفاظ واللغة أكثر من الاختبارات التي تستخدم الأرقام أو الرموز (حمادنة، ٢٠٠٧). فاختلاف معاني الكلمات والألفاظ بين البيئات الثقافية المختلفة نظراً لاختلاف اللغات واللهجات التي يتحدث بها الأفراد، واختلاف الخبرات التي يعيشها الأفراد كالبينة الزراعية والبيئة البحرية والبيئة الصحراوية وغيرها يؤدي إلى فهم مختلف للمقصود من الألفاظ والذي ينتج منه اختلاف الإجابة على المفردات باختلاف المجموعات الثقافية. وتتعدد الدراسات التي تناولت فحص الأداء التفاضلي والتحيز للمفردات في الاختبارات اللفظية (Mikyung, 2001;

العربية مقابل الإنجليزية) حسب طريقة مانتل هانزل، أي ما نسبته ٥٦%، والمفردات التي أظهرت أداءً تفاضلياً كان ٧٤% منها لصالح مجموعة اللغة الإنجليزية، وكما بيّنت نتائج الدراسة أيضاً أن متوسط الأداء في الأردن وترتيبه قد تحسّن بعد حذف المفردات التي أظهرت أداءً تفاضلياً، والسبب قد يعود إلى أن مفردات الاختبار الأصلية كانت مكتوبة باللغة الإنجليزية، وترجمة تلك المفردات ربما يُغيّر معناها، أو تؤدي إلى حذف معلومات ضرورية تجعل الاستجابة لها صعبة.

كما هدفت دراسة البشاشة (٢٠١٦) إلى الكشف عن الأداء التفاضلي لمتغير النوع الاجتماعي في اختبار PISA لعام ٢٠١٢ في مبحثي العلوم والقراءة، وذلك باستخدام طريقة مانتل هانزل، وطريقة نسبة الأرجحية العظمى، حيث طُبّق الاختبار على ٧٠٣٨ طالباً وطالبة من الصف العاشر بالأردن، وكشفت الدراسة عن وجود أداء تفاضلي لصالح الإناث على ٣٩ مضرده من مفردات اختبار القراءة، وعلى ٤٠ مضرده من مفردات اختبار العلوم، ومضرده واحدة لصالح الذكور في اختبار العلوم فقط وباتفاق جميع الاختبارات المستخدمة.

وقد أجرى العجلوني والسوالمه (٢٠١٦) دراسة، هدفت إلى فحص الأداء التفاضلي لمموهات مفردات اختبار الرياضيات في التقييم الوطني الأردني لاقتصاد المعرفة تبعاً لمتغير النوع الاجتماعي للطالب وموقع المدرسة، وقد تم توظيف ١٥٥٨ طالباً وطالبة في الصف الخامس على ٢٦ مضرده من نوع اختيار من متعدد، تقيس المهارات الأساسية في الرياضيات، ولكل منها أربعة بدائل، وذلك باستخدام طريقة نسبية الأرجحية لمانتل هانزل، أظهرت النتائج أن ١٥ مموهاً أظهر أداءً تفاضلياً تبعاً لمتغير النوع الاجتماعي، بعضها كان جذاباً للإناث والبعض الآخر للذكور، وبحجم تأثير متفاوت، في حين أظهر ١٥ مموهاً أداءً تفاضلياً بالنسبة لمتغير موقع المدرسة، وبحجم تأثير متفاوت، بعضها كان جذاباً لطلبة مدارس المدن وبعضها الآخر لطلبة مدارس الريف.

المقياس صالحاً للتطبيق خليجياً، يرى الباحثون أنه من الضرورة بمكان إجراء هذه الدراسة عبر النوع الاجتماعي وعبر الدولة، مما يُعظّم الفائدة من تطوير واستخدام هذا المقياس في البيئة الخليجية لكشف عن القدرات العقلية والمعرفية لطلبة المراحل الأولى في التعليم العام. وأجابت الدراسة الحالية عن الأسئلة التالية:

١. ما مفردات اختبار القدرة اللفظية بالمستوى الثاني لمقياس جماس، التي تُظهر أداءً تفضلياً لطلبات الصفين الثالث والرابع وطلبة سلطنة عمان باستخدام طريقة مانتل هانزل؟
٢. ما مفردات اختبار القدرة اللفظية بالمستوى الثاني لمقياس جماس، التي تُظهر أداءً تفضلياً لطلبات الصفين الثالث والرابع وطلبة سلطنة عمان باستخدام طريقة الصعوبة المحوّلة؟
٣. ما درجة الاتفاق بين طريقتي مانتل هانزل والصعوبة المحوّلة، في الكشف عن الأداء التفاضلي في مفردات اختبار القدرة اللفظية بالمستوى الثاني لمقياس جماس لطلبات الصفين الثالث والرابع وطلبة سلطنة عمان؟

أهمية الدراسة

تسعى الدراسة الحالية إلى المساهمة في إثراء جانب مهم من مجالات الدراسات التربوية، يتعلق بدراسة طرق الكشف عن الأداء التفاضلي لمفردات اختبار القدرة اللفظية، إذ تُعد هذه أول دراسة عن الأداء التفاضلي لمقياس جماس. ومن الممكن أن تسفر هذه الدراسة عن نتائج يمكن من خلالها الخروج بتوصيات لتطوير مقياس جماس. كما أن الإهتمام بدراسة الأداء التفاضلي في الاختبارات قليل محلياً، مما يجعل من هذه الدراسة عنصر إسهام في هذه الجهود، ولعلها تشجّع باحثين آخرين لإعطاء مزيد من الاهتمام لدراسة الأداء التفاضلي ودراسة التحيز.

(Uiterwijk & Vallen, 2005). وتركز هذه الدراسة على المجال اللفظي الذي يعتمد أساساً على المعالجات اللفظية للكلمات، وفهمها واستيعابها؛ حيث إن المجال اللفظي معني بقياس القدرة اللفظية، والتي تُعرّف بأدائها معرفة اللغة الأم، وفهمها، وإدراك معانيها، والاستدلال اللغوي واستخدامها استخداماً ملائماً (الزيات والمحرزي، ٢٠١١).

ولقد تعددت الطرق التي تُستخدم للكشف عن الأداء التفاضلي لمفردات الاختبار، ومستوى فاعليتها، فهناك طرق حسب النظرية التقليدية وطرق حسب النظرية الحديثة، ومن الطرق حسب النظرية التقليدية: طريقة مانتل هانزل، وطريقة الصعوبة المحوّلة، وطريقة تحليل التباين، وطريقة الانحدار اللوجستي، وغيرها من الطرق. وتتناول الدراسة الحالية تطبيق طريقة مانتل هانزل، وطريقة الصعوبة المحوّلة وفق النظرية الكلاسيكية للقياس في فحص الأداء التفاضلي لمفردات اختبار القدرة اللفظية لكون عدد من الدراسات أثبتت فاعليتها في الكشف عن الأداء التفاضلي للمفردات (Stage, 2000; الجرادات، ٢٠٠٣؛ الهريشات، ٢٠٠٤). ولقد استخدمت هاتان الطريقتان من قبل العديد من الباحثين منهم الجرادات (٢٠٠٣)، والبستنجي (٢٠٠٤)، والهريشات (٢٠٠٤)، والبشاشة (٢٠١٦)، وذلك للكشف عن المفردات ذات الأداء التفاضلي تبعاً لمتغيري النوع الاجتماعي ومنطقة السكن. وقد أظهرت النتائج وجود عدد من المفردات بها أداءً تفضلياً تبعاً لمتغير النوع الاجتماعي، منها لصالح الذكور، ومنها لصالح الإناث، كما أظهرت النتائج وجود أداءً تفضلياً لبعض المفردات حسب منطقة السكن، ومن خلال هذه الدراسات يتبين لنا أن هاتين الطريقتين ذات فاعلية في الكشف عن المفردات ذات الأداء التفاضلي.

قامت الدراسة الحالية بدراسة الأداء التفاضلي لمفردات اختبار القدرة اللفظية بالمستوى الثاني لمقياس جماس لطلبات الإناث (متغير النوع الاجتماعي)، وطلبة سلطنة عمان (متغير الدولة) بالصفين الثالث والرابع الابتدائيين. وكون هذا

مصطلحات الدراسة

الأداء التفاضلي للمفردة: هو دالة مشتقة إحصائياً للتعبير عن الفروق في الإستجابة على المفردة بين مجموعتين من المفحوصين ممن هم في نفس مستوى القدرة (البرصان، ٢٠١٣).

مقياس جِماس " Gulf Multiple Mental Ability " Scale ويكتب اختصاراً (GMMAS): هو مقياس الخليج للقدرات العقلية المتعددة، ويتكوّن من ثلاثة اختبارات تقيس ثلاث قدرات عقلية، وهي: اللفظية، والعددية، والمكانية، ويغطي مقياساً متصلاً للمدى العمري من ٦ - ١٢ سنة، بحيث إن لكل اختبار قدرة ثلاثة مستويات (الزيات والمحززي، ٢٠١١).

حدود الدراسة

تحدد هذه الدراسة في ضوء المجتمع والعينة والأداة التي شملتها، والتي تتمثل فيما يلي:

١. حدود بشرية: طلبة دول مجلس التعاون الخليجي، بالصفين الثالث والرابع الابتدائيين.
٢. حدود مكانية: المدارس الحكومية في دول الخليج العربي.
٣. حدود موضوعية: استخدام طريقتين للكشف عن الأداء التفاضلي: مانتل هانزل، والصعوبة المحوّلة، تبعاً لمتغيري النوع الاجتماعي والدولة.

المنهجية والإجراءات

منهج الدراسة

استخدمت الدراسة المنهج الوصفي، كأحد أشكال التحليل والتفسير العلمي المنظم؛ لوصف ظاهرة أو مشكلة محددة، وتصويرها كمياً عن طريق جمع بيانات، ومعلومات مقننة عن الظاهرة أو المشكلة، وتصنيفها، وتحليلها، وإخضاعها للدراسة الدقيقة (الكساني، ٢٠١٢).

مجتمع الدراسة وعينتها

تكوّن مجتمع الدراسة من جميع الطلاب والطالبات المسجلين بمدارس التعليم العام الحكومية في الصفين الثالث والرابع الابتدائيين

للعام الدراسي ٢٠٠٩ / ٢٠١٠م وقت تقنين المقياس في دول مجلس التعاون الخليجي الست: دولة الإمارات العربية المتحدة، مملكة البحرين، المملكة العربية السعودية، سلطنة عُمان، دولة قطر، دولة الكويت، بحيث تراوحت أعمار الطلبة من عمر سبع سنوات وتسعة أشهر إلى عمر تسع سنوات وشهرين. واستخدمت الدراسة بيانات أرشيفية من العينة التي استخدمت في تقنين مقياس جِماس على طلبة الصفين الثالث والرابع الابتدائيين في دول مجلس التعاون الخليجي عام ٢٠١١، حيث بلغ عددهم ٤٢٨٠ طالباً وطالبة ٢١٥١ ذكور، ٢١٢٩ إناث. وقد تم اختيار الطلبة بحجم متساوٍ من مجتمع الدراسة في جميع دول مجلس التعاون الخليجي بطريقة عشوائية طبقية داخل المجموعات المختلفة للطلبة حسب المنطقة التعليمية، والصفوف الدراسية، والنوع الاجتماعي، وتم اختيار خمس مناطق تعليمية عشوائياً في كل دولة من الدول المشتركة، حيث مثلت المناطق التعليمية الاختلافات بين جميع المناطق التعليمية بكل دولة، بعد ذلك تم اختيار عدد متقارب من الطلبة بطريقة عشوائية طبقية حسب الصفوف الدراسية بالمستوى الثاني للصفين الثالث والرابع الابتدائيين، داخل كل منطقة تعليمية تم اختيارها. وفي داخل كل صف دراسي، تم اختيار عدد متقارب من الطلبة حسب النوع الاجتماعي. ويمتد المدى العمري من سبع سنوات وتسعة أشهر إلى تسع سنوات وشهرين، وتم تطبيق المقياس بشكل شبه فردي، على مجموعة صغيرة تتكون من أربعة طلاب (الزيات والمحززي، ٢٠١١).

أداة الدراسة

استخدمت الدراسة المستوى الثاني من اختبار القدرة اللفظية في مقياس الخليج للقدرات العقلية المتعددة، والذي تم إعداده من قبل فريق المشروع المؤلف من باحثين من جميع دول الخليج العربي، ويتكون الاختبار من ثلاثة اختبارات فرعية، بمجموع ٣٠ مفردة من نوع الاختيار من متعدد: اختبار معاني الكلمات ١٠ مفردات، بحيث يطلب من الطالب اختيار المعنى الصحيح للكلمة، اختبار الأضداد ١٠ مفردات،

قام الزيات والمحرزي (٢٠١١) بحساب ثبات الأداة، وذلك بقياس معامل ثبات إعادة التطبيق لاختبار القدرة اللفظية عن طريق إعادة التطبيق على عينات المستوى الثاني بكل من دولة الكويت، ودولة الإمارات العربية المتحدة، ومملكة البحرين بفارق زمني تراوح بين ثلاثة أسابيع وشهر، وقد جاءت معاملات الثبات ٠.٨٧ لدولة الكويت، و٠.٦٩ لدولة الإمارات، و٠.٧٤ لمملكة البحرين، و٠.٨٥ للدول الثلاث مجتمعة. وأظهرت النتائج معامل ثبات اتساق داخلي عالي لاختبار القدرة اللفظية في المستوى الثاني، إذ إن قيمة معامل ثبات ألفا كرونباخ ٠.٨٨٢ للصف الثالث، و٠.٨٥٦ للصف الرابع. كما أظهرت النتائج أن الخطأ المعياري للقياس ٢.٣ للصف الثالث و٢.٣ وللصف الرابع. وبخصوص معامل ثبات التجزئة النصفية، تم حسابه بتقسيم المفردات ذات الأرقام الفردية مقابل المفردات ذات الأرقام الزوجية، وبلغ ٠.٨٩٢ بالصف الثالث، و٠.٨٧٠ بالصف الرابع. وتُشير هذه المعاملات إلى ثبات اتساق داخلي مرتفع للاختبار. وتم حساب معامل ثبات الاتساق الداخلي لعينة الدراسة وفقاً لمتغيرات الدراسة بواسطة ألفا كرونباخ، وكانت قيمته مرتفعة، حيث بلغت قيمته لعمان ٠.٩٢٥ مقابل ٠.٨٦٨ لباقي دول الخليج، وكانت قيمته للإناث ٠.٨٨٤ مقابل ٠.٨٧٦ للذكور، وهذا مؤشر على ثبات نتائج الاختبار للفئات المستهدفة من الدراسة.

إجراءات الدراسة

بعد تجهيز البيانات للتحليل الإحصائي تم استخدام الحزمة الإحصائية في العلوم الاجتماعية SPSS V.22 لتطبيق طريقة مانتل هانزل وبرنامج Excel لتطبيق طريقة الصعوبة المحولة للكشف عن الأداء التفاضلي للمفردات وفقاً لنوع الاجتماعي والدولة. ولتطبيق طريقة الصعوبة المحولة تم تنفيذ الإجراءات التي لخصها ظاظا (٢٠٠٦)، والتي تتمثل في الآتي: استخراج معامل صعوبة المفردات (P) لكل مفردة في كل مجموعة من المجموعتين المرجعية والمستهدفة، ثم حساب القيم المعيارية Z تحت المنحنى الطبيعي المقابلة ل $Q = 1 - P$ ، لكل مجموعة على

بحيث يطلب من الطالب اختيار الضد الصحيح للكلمة، اختبار التناظر اللفظي ١٠ مفردات، بحيث يطلب من الطالب إدراك العلاقة بين زوج من الكلمات في مطلع السؤال، وقياسها على نظائر تماثلها معطاة في الاختيارات. وقد روعي في بناء واختيار مفردات كل اختبار توظيف عمليتي الإدراك والاستدلال. ويتكون كل اختبار من كراسة أسئلة وورقة إجابة. وكان التصحيح بإعطاء درجة واحدة للإجابة الصحيحة، وصفر للإجابة الخاطئة (الزيات والمحرزي، ٢٠١١).

صدق الأداة وثباتها

درس الزيات والمحرزي (٢٠١١) الإحصاءات الوصفية لاختبار القدرة اللفظية بالمستوى الثاني في الصفين الثالث والرابع الابتدائيين في دول الخليج مجتمعة، وتبين بأن المتوسط الحسابي للدرجة الخام للصف الثالث ١٨.٣ وللصف الرابع ٢٠.٨، في حين أن الانحراف المعياري ٦.٨ و٦.٠ للصفين الثالث والرابع على التوالي؛ وتشير الإحصاءات الوصفية إلى ارتفاع مستوى القدرة اللفظية عند الانتقال من الصف الثالث إلى الصف الرابع بجميع دول الخليج؛ وتشير هذه النتيجة أن القدرة اللفظية تزداد بالخبرات التي يكسبها الطلبة من خلال المناهج الدراسية وتلك التي يمرون بها مع نموهم العمري. وكانت معاملات الارتباط البينية لاختبار القدرة اللفظية موجبة ومتوسطة القوة بين جميع الاختبارات الفرعية للقدرة اللفظية، حيث تراوحت بين ٠.٥٩٣ و٠.٥٩٧. ويشير ذلك إلى ترابط الاختبارات الفرعية للقدرة اللفظية، إذ أن هذه الاختبارات الفرعية تقيس المهارات المرتبطة بالقدرة اللفظية. وأشار الزيات والمحرزي (٢٠١١) أن معاملات ارتباط القدرة اللفظية مع اختبار المصفوفات المتتابعة لرافن موجبة ودالة إحصائياً كمؤشر على الصدق التلازمي للاختبار، ووجود معاملات ارتباط موجبة ودالة إحصائياً بين القدرة اللفظية والتحصيل المدرسي في مادتي اللغة العربية والرياضيات للصفين الثالث والرابع كمؤشر على الصدق التنبؤي لاختبار القدرة اللفظية.

ثلاثة أنواع، وهي: مفردة ذات أداء تفاضلي ضعيف ($0 < D < 1$)، ومفردة ذات أداء تفاضلي متوسط ($1 < D < 1.5$)، ومفردة ذات أداء تفاضلي مرتفع ($D > 1.5$).

متغير النوع الاجتماعي: يُظهر جدول ١ قيم اختبار كاي تربيع لمانتل هانزل، وقيمة الاحتمال، ونسبة الأرجحية، وقيمة D ، لفحص الأداء التفاضلي لمفردات اختبار القدرة اللفظية حسب النوع الاجتماعي (الإناث/ الذكور). يتبين أن قيمة اختبار كاي تربيع لمانتل هانزل تراوحت بين ٠ و ٤٧.٦٤٠، وقيمة D تراوحت بين ٠.٠٠٢ و ١.١٦٩، وتشير النتائج إلى أن ١٨ مفردة في اختبار القدرة اللفظية بالمستوى الثاني لمقياس جماس أظهرت أداءً تفاضلياً، وهي تمثل نسبة ٦٠٪. أما المفردات ١٢ المتبقية في اختبار القدرة اللفظية بالمستوى الثاني لمقياس جماس لم تُظهر أداءً تفاضلياً لأي نوع اجتماعي.

ومن بين ١٨ مفردة، تشير النتائج أن ٩ مفردات من اختبار القدرة اللفظية أظهرت أداءً تفاضلياً نحو المجموعة المستهدفة (الإناث). ومن حيث قوة الأداء التفاضلي في هذه المفردات، كانت ضعيفة لعدد ٨ مفردات حسب مؤشر D ، بينما كانت متوسطة لمفردة واحدة فقط. في المقابل أظهرت ٩ مفردات أداءً تفاضلياً نحو المجموعة المرجعية (الذكور). ومن حيث قوة الأداء التفاضلي في هذه المفردات كانت ضعيفة أيضاً لعدد ٨ مفردات حسب مؤشر D ، بينما كانت متوسطة لمفردة واحدة فقط.

متغير الدولة: يوضح جدول ٢ قيم اختبار كاي تربيع لمانتل هانزل، وقيمة الاحتمال، ونسبة الأرجحية، وقيمة D لفحص الأداء التفاضلي لمفردات اختبار القدرة اللفظية حسب الدولة (سلطنة عُمان/ باقي دول الخليج)، وقد تم ترتيب النتائج وفقاً لقوة الأداء التفاضلي. يتبين أن قيمة كاي تربيع لمانتل هانزل تراوحت بين ٠.٠٠٦ و ٦٥.٨٢٠، وقيمة D تراوحت بين ٠.٠٤٧ و ١.٨٤٠، وتشير النتائج إلى أن ١٨ مفردة من أصل ٣٠ مفردة في اختبار القدرة اللفظية لمقياس جماس أظهرت أداءً تفاضلياً وفقاً لمتغير الدولة (سلطنة عُمان/ باقي دول الخليج)، وهي تمثل

حده، وذلك لجعل العلاقة بين مجموعتي صعوبة المفردات خطية، بعدها تم تحويل القيم المعيارية الزائفة (Z) إلى القيمة المعيارية دلتا (Δ) ليكون متوسطها الحسابي ١٣ وانحرافها المعياري ٤، ثم حساب ميل الخط المستقيم (B)، ومقطعه مع محور الصادات (A)، وتم استخدامها لتمثيل قيم دلتا على رسم بياني، وتم وضع قيم دلتا للمجموعة المرجعية على المحور السيني، وعلى المحور الصادي قيم دلتا للمجموعة المستهدفة، ثم رسم ما يعرف خط الموازنة الأفضل (The best fit line) ثم حساب بُعد النقطة عن المحور الرئيسي (d_i) كمؤشر على وجود الأداء التفاضلي.

النتائج

ما مفردات اختبار القدرة اللفظية بالمستوى الثاني لمقياس جماس، التي تُظهر أداءً تفاضلياً لطالبات الصفين الثالث والرابع وطلبة سلطنة عمان باستخدام طريقة مانتل هانزل؟

تم استخراج قيم اختبار كاي تربيع لمانتل هانزل، وقيمة الاحتمال، ونسبة الأرجحية، وقيمة الدلتا D لكل مفردة من مفردات اختبار القدرة اللفظية. ولتفسير النتائج تم تقسيم المفردات إلى ثلاثة أنواع: المفردات التي لا تظهر أداءً تفاضلياً ($\alpha_{MH} = 1$) وغير دالة احصائياً)، المفردات ذات الأداء التفاضلي لصالح المجموعة المرجعية ($\alpha_{MH} > 1$) ودالة احصائياً)، المفردات ذات الأداء التفاضلي لصالح المجموعة المستهدفة ($\alpha_{MH} < 1$) ودالة احصائياً) (البرصان، ٢٠١٣).

وللحكم على قوة الأداء التفاضلي للمفردات في حالة وجود أداء تفاضلي للمفردة حسب قيمة نسبة الأرجحية، فقد اعتمدت الدراسة المحك الذي يستخدمه مركز الاختبارات والخدمات التربوية (ETS) من خلال حساب قيمة الدلتا D وفقاً للمعادلة الآتية $D = \beta_{MH} = -2.35 \ln(\alpha_{MH})$ (عمارة، ٢٠٠٨). ويطلق على β_{MH} مسمى المؤشر المميز بالإشارة Signed Index، حيث يُستدل من إشارتها على اتجاه الأداء التفاضلي للمفردة. وتُصنّف المفردة في ضوء هذا المحك من حيث قوة الأداء التفاضلي لها إلى

مفردات ضعيفة، وكانت متوسطة لمفردة واحدة فقط، وكانت قوية لمفردتين. في المقابل أظهرت ٩ مفردات أداءً تفضلياً نحو المجموعة المستهدفة (سلطنة عمان). وحسب قوة الأداء التفاضلي وفقاً لمؤشر D، كانت قوة الأداء التفاضلي لعدد ٥ مفردات ضعيفة، وكانت متوسطة لعدد ٣ مفردات، وكانت قوية لمفردة واحدة فقط.

نسبة ٦٠٪. أما المفردات الأخرى المتبقية فلم تظهر أداءً تفضلياً لأي مجموعة وعددها ١٢ مفردة.

وتشير النتائج أن ٩ مفردات من اختبار القدرة اللفظية بالمستوى الثاني لمقياس جماس أظهرت أداءً تفضلياً نحو المجموعة المرجعية (باقي دول الخليج). وحسب قوة الأداء التفاضلي وفقاً لمؤشر D، كانت قوة الأداء التفاضلي لعدد ٦

جدول ١

قيم اختبار كاي تربيع لمانتل هانزل، وقيمة الاحتمال، ونسبة الأرجحية، وقيمة D لمفردات اختبار القدرة اللفظية تبعاً لمتغير النوع الاجتماعي مرتبة حسب اتجاه الأداء التفاضلي

رقم المفردة	كاي تربيع لمانتل هانزل	قيمة الاحتمال	نسبة الأرجحية	D	قوة واتجاه الأداء التفاضلي
٩	٤٧.٦٤٠	*.٠٠٠٠	٠.٦٠٨	١.١٦٩	متوسط نحو الإناث
٥	٣٨.٨١٩	*.٠٠٠٠	١.٥٩٢	١.٠٩٣	متوسط نحو الذكور
٤	١٣.٠٤٨	*.٠٠٠٠	٠.٧٧٨	٠.٥٩٠	ضعيف نحو الإناث
٦	١٥.٩٩١	*.٠٠٠٠	٠.٧٣٢	٠.٧٣٣	ضعيف نحو الإناث
١٣	٥.٥٤٨	*.٠٠١٨	٠.٨٣٣	٠.٤٢٩	ضعيف نحو الإناث
١٦	٦.٤٧٣	*.٠٠١١	٠.٧٩٨	٠.٥٣٠	ضعيف نحو الإناث
٢٤	٩.٨٨٥	*.٠٠٠٢	٠.٧٥٧	٠.٦٥٤	ضعيف نحو الإناث
٢٧	٦.٣٠٢	*.٠٠١٢	٠.٨٣٨	٠.٤١٥	ضعيف نحو الإناث
٢٨	١٤.٧٠١	*.٠٠٠٠	٠.٧٢٨	٠.٧٤٦	ضعيف نحو الإناث
٢٩	١١.٠٣٩	*.٠٠٠١	٠.٧٥٧	٠.٦٥٤	ضعيف نحو الإناث
٢	٤.١٦٠	*.٠٠٤١	١.١٨٦	٠.٤٠١	ضعيف نحو الذكور
٧	٢١.٩٧٧	*.٠٠٠٠	١.٤١٠	٠.٨٠٧	ضعيف نحو الذكور
١٠	٥.١٢٨	*.٠٠٢٤	١.١٧٢	٠.٣٧٣	ضعيف نحو الذكور
١٤	١١.٠٤١	*.٠٠٠١	١.٢٧٤	٠.٥٦٩	ضعيف نحو الذكور
١٥	٩.٢٦٠	*.٠٠٠٢	١.٢٦١	٠.٥٤٥	ضعيف نحو الذكور
٢١	٣٣.٦٥٦	*.٠٠٠٠	١.٤٦٤	٠.٨٩٦	ضعيف نحو الذكور
٢٣	٤.٤٥٧	*.٠٠٣٥	١.١٧٦	٠.٣٨١	ضعيف نحو الذكور
٢٦	١١.٢٥٩	*.٠٠٠١	١.٢٨٩	٠.٥٩٧	ضعيف نحو الذكور
١	٠.٠٤١	٠.٨٣٩	١.٠١٧	٠.٠٤٠	لا يوجد
٣	٠.٠٠٠	٠.٩٨٥	١.٠٠٢	٠.٠٠٥	لا يوجد
٨	٠.٢٠٩	٠.٦٤٧	٠.٩٦٤	٠.٠٨٦	لا يوجد
١١	٠.٧٠٣	٠.٤٠٢	٠.٩١٧	٠.٢٠٤	لا يوجد
١٢	٠.٠٠٩	٠.٩٢٢	١.٠١٠	٠.٠٢٣	لا يوجد
١٧	٠.٠٢٩	٠.٨٦٦	١.٠١٥	٠.٠٣٥	لا يوجد
١٨	٢.٩٥٦	٠.٠٨٦	٠.٨٦١	٠.٣٥٢	لا يوجد
١٩	٠.٤٣٠	٠.٥١٢	٠.٩٤٧	٠.١٢٨	لا يوجد
٢٠	٠.٢١٣	٠.٦٤٥	٠.٩٦٦	٠.٠٨١	لا يوجد
٢٢	٠.٠٠١	٠.٩٧٩	٠.٩٩٩	٠.٠٠٢	لا يوجد
٢٥	١.٢٢٩	٠.٢٦٨	١.٠٨٣	٠.١٨٧	لا يوجد
٣٠	٠.٢١٥	٠.٦٤٢	٠.٩٦٣	٠.٠٨٩	لا يوجد

* دالة احصائياً عند $\alpha \geq 0.05$

جدول ٢

قيم اختبار كاي تربيع لمantel هانزل، وقيمة الاحتمال، ونسبة الأرجحية، وقيمة D لمفردات اختبار القدرة اللفظية تبعاً لمتغير الدولة مرتبة حسب اتجاه الأداء التفاضلي

رقم المفردة كاي تربيع لمantel هانزل	قيمة الاحتمال	نسبة الأرجحية	D	قوة واتجاه الأداء التفاضلي
٥	*٠.٠٠٧	١.٣٦١	٠.٧٢٤	ضعيف نحو باقي الدول
٧	*٠.٠٢٠	١.٢٩٩	٠.٦١٥	ضعيف نحو باقي الدول
١٦	*٠.٠٠٨	١.٣٧٠	٠.٧٤٠	ضعيف نحو باقي الدول
١٩	*٠.٠٠٠	١.٥٢٥	٠.٩٩٢	ضعيف نحو باقي الدول
٢٧	*٠.٠٠٠	١.٤٦٣	٠.٨٩٤	ضعيف نحو باقي الدول
٢٩	*٠.٠٤٥	١.٢٩٥	٠.٦٠٨	ضعيف نحو باقي الدول
٨	*٠.٠٠٣	٠.٧٣٦	٠.٧٢٠	ضعيف نحو عُمان
١٣	*٠.٠٠١	٠.٦٦٢	٠.٩٦٩	ضعيف نحو عُمان
١٥	*٠.٠١٤	٠.٧٦٣	٠.٦٣٦	ضعيف نحو عُمان
٢٦	*٠.٠٠١	٠.٦٨١	٠.٩٠٣	ضعيف نحو عُمان
٢٨	*٠.٠٠٢	٠.٦٩١	٠.٨٦٩	ضعيف نحو عُمان
٢	*٠.٠٠٠	١.٥٨٦	١.٠٨٤	متوسط نحو باقي الدول
١٤	*٠.٠٠٠	٠.٥٨٧	١.٢٥٢	متوسط نحو عُمان
٢٠	*٠.٠٠٠	٠.٥٩٤	١.٢٢٤	متوسط نحو عُمان
٣٠	*٠.٠٠٠	٠.٥٦٦	١.٣٣٨	متوسط نحو عُمان
١١	*٠.٠٠٠	١.٩٨٢	١.٦٠٨	قوي نحو باقي الدول
١٢	*٠.٠٠٠	٢.٠٣٧	١.٦٧٢	قوي نحو باقي الدول
١٠	*٠.٠٠٠	٠.٤٥٧	١.٨٤٠	قوي نحو عُمان
١	٠.١٣٦	١.١٧٩	٠.٣٨٧	لا يوجد
٣	٠.٢٣٩	٠.٨٥٨	٠.٣٦٠	لا يوجد
٤	٠.١٨٨	٠.٨٦٧	٠.٣٣٥	لا يوجد
٦	٠.٠٦٠	١.٢٢٦	٠.٤٧٩	لا يوجد
٩	٠.٨٤٩	١.٠٢٧	٠.٠٦٣	لا يوجد
١٧	٠.٠٦١	١.٢٢٣	٠.٤٧٣	لا يوجد
١٨	٠.٩٣٨	٠.٩٨٠	٠.٠٤٧	لا يوجد
٢١	٠.١٥٨	١.١٥٠	٠.٣٨٢	لا يوجد
٢٢	٠.١٣٠	١.١٧٩	٠.٣٨٧	لا يوجد
٢٣	٠.٩٠٩	١.٠٢٠	٠.٠٤٧	لا يوجد
٢٤	٠.٧٩٣	٠.٩٦٣	٠.٠٨٩	لا يوجد
٢٥	٠.٤٨٠	١.٠٨٤	٠.١٩٠	لا يوجد

للكشف عن المفردات ذات الأداء التفاضلي بطريقة الصعوبة المحوِّلة، تم تحليل البيانات باستخدام برنامج Excel، وتم إيجاد معامل الصعوبة للمجموعتين المرجعية والمستهدفة،

ما مفردات اختبار القدرة اللفظية بالمستوى الثاني لمقياس جماس، التي تُظهر أداءً تفاضلياً لطالبات الصفين الثالث والرابع وطلبة سلطنة عمان باستخدام طريقة الصعوبة المحوِّلة؟

للمجموعتين المرجعية (الذكور) والمستهدفة (الإناث)، ومقدار بُعد النقطة عن المحور الرئيسي، باستخدام طريقة الصعوبة المحولة، وقد تم ترتيب النتائج وفقاً لاتجاه الأداء التفاضلي. يتبين أن معامل صعوبة المفردات متقارب لدى الإناث والذكور فقد تراوح بين ٠.٤٠١ و٠.٨٤٤ للإناث بمتوسط حسابي وقدره ٠.٦٣٨ أما معامل صعوبة المفردات للذكور بين ٠.٤٢١ و٠.٨٣١ بمتوسط حسابي وقدره ٠.٦٤٩،

ثم حساب بُعد النقاط عن خط الاتجاه، والذي يرمز له بالرمز d_j ، حيث يُستخدم هذا البُعد في المحكم على الأداء التفاضلي للمفردات. وسيكون معيار الحكم على وجود الأداء التفاضلي للمفردات - في هذه الدراسة - وفقاً لعدد الانحرافات المعيارية التي تبعتها النقطة (d_j) عن خط الموائمة، وهو ما يزيد عن واحد انحراف معياري.

متغير النوع الاجتماعي: يظهر جدول ٣ معامل الصعوبة لمفردات اختبار القدرة اللفظية

جدول ٣

معامل الصعوبة وبُعد النقطة عن المحور الرئيسي وفق طريقة الصعوبة المحولة لمفردات اختبار القدرة اللفظية تبعا لمتغير النوع الاجتماعي مرتبة حسب اتجاه الأداء التفاضلي

رقم المفردة	معامل الصعوبة		بُعد النقطة	الاتجاه
	بقية الدول	عُمان		
٤	٠.٦٥٧	٠.٥٩٥	٠.٣٨٦	نحو الإناث
٦	٠.٧٢٥	٠.٦٦٧	٠.٣٣٢	نحو الإناث
٩	٠.٦٠٥	٠.٥٠١	٠.٧٣٤	نحو الإناث
٢٧	٠.٥٧١	٠.٥٢٤	٠.٣٣٣	نحو الإناث
٢	٠.٧٧٧	٠.٧٩٤	٠.٣٩٧-	نحو الذكور
٥	٠.٦٢٥	٠.٦٩٤	٠.٦٤٠-	نحو الذكور
٧	٠.٥٦١	٠.٦٠٧	٠.٣٦٦-	نحو الذكور
٢١	٠.٥١٧	٠.٥٨٩	٠.٥٢٠-	نحو الذكور
٢٦	٠.٦٣٢	٠.٦٦١	٠.٣١١-	نحو الذكور
١	٠.٦٣٢	٠.٦٢٣	٠.٠٠٨-	لا يوجد
٣	٠.٧٠٨	٠.٦٩٥	٠.٠٤٣-	لا يوجد
٨	٠.٦٦١	٠.٦٤٦	٠.٠٢٠	لا يوجد
١٠	٠.٤٠١	٠.٤٢١	٠.٠٢٠-	لا يوجد
١١	٠.٨٤٤	٠.٨٣١	٠.١٥٧-	لا يوجد
١٢	٠.٦٨٥	٠.٦٨٠	٠.٠٨٧-	لا يوجد
١٣	٠.٧٠١	٠.٦٦٢	٠.١٨٤	لا يوجد
١٤	٠.٥٩٧	٠.٦٢٨	٠.٢٨٩-	لا يوجد
١٥	٠.٥٣٤	٠.٥٥٩	٠.١٧٦-	لا يوجد
١٦	٠.٧٦٦	٠.٧٣٠	٠.١٢٠	لا يوجد
١٧	٠.٦٢١	٠.٦١٣	٠.٠٠٣-	لا يوجد
١٨	٠.٧١٤	٠.٦٨٢	٠.١١٥	لا يوجد
١٩	٠.٧١٩	٠.٧٠٢	٠.٠١٥-	لا يوجد
٢٠	٠.٥٠٢	٠.٤٨٢	٠.١٩١	لا يوجد
٢٢	٠.٥٦٦	٠.٥٥٥	٠.٠٦٩	لا يوجد
٢٣	٠.٦٥١	٠.٦٦٤	٠.٢٠٥-	لا يوجد
٢٤	٠.٨١٢	٠.٧٧٠	٠.١٦٤	لا يوجد
٢٥	٠.٥٨٢	٠.٥٨٥	٠.٠٥١-	لا يوجد
٢٨	٠.٧٥٢	٠.٦٩٧	٠.٢٨٧	لا يوجد
٢٩	٠.٧٣٢	٠.٦٨٣	٠.٢٥٢	لا يوجد
٣٠	٠.٦١٠	٠.٥٩٠	٠.٠٩٩	لا يوجد

وتُشير نتائج طريقة الصعوبة المحوَّلة إلى أن ٩ مفردات من أصل ٣٠ مفردة في اختبار القدرة اللفظية أظهرت أداءً تفاضلياً تبعاً لمتغير النوع الاجتماعي، وهي تمثل نسبة ٣٠٪؛ حيث كان الأداء التفاضلي لعدد ٥ مفردات نحو المجموعة المرجعية الذكور، بينما كانت ٤ مفردات نحو المجموعة المستهدفة الإناث. أما المفردات المتبقية في اختبار القدرة اللفظية لم تُظهر أداءً تفاضلياً لأي نوع اجتماعي.

متغير الدولة: يظهر جدول ٤ معامل الصعوبة للمجموعتين المرجعية والمستهدفة، ومقدار بُعد النقطة عن المحور الرئيسي، باستخدام طريقة الصعوبة المحوَّلة، تبعاً لمتغير الدولة، وقد تم ترتيب النتائج وفقاً لاتجاه الأداء التفاضلي. يتبين أن معامل صعوبة المفردات للطلبة في سلطنة عُمان تراوح بين ٠.٤٧٨ و ٠.٧٧٨ بمتوسط حسابي قدره ٠.٦٢٩، في المقابل فإن معامل صعوبة المفردات للطلبة بباقي دول الخليج قد تراوح بين ٠.٣٨٢ و ٠.٨٥٥ بمتوسط حسابي قدره ٠.٦٤٦، وتشير النتائج إلى أن طريقة الصعوبة المحوَّلة كشفت عن ١٠ مفردات من أصل ٣٠ مفردة في اختبار القدرة اللفظية بها أداءً تفاضلياً تبعاً لمتغير الدولة، وهي تمثل نسبة ٣٣.٣٣٪. وكان الأداء التفاضلي لعدد ٤ مفردات نحو المجموعة المستهدفة (سلطنة عُمان)، بينما أظهرت ٦ مفردات أداءً تفاضلياً نحو المجموعة المرجعية (باقي دول الخليج). أما المفردات العشرون الأخرى المتبقية في اختبار القدرة اللفظية لم تُظهر أداءً تفاضلياً لأي من المجموعتين.

ما درجة الاتفاق بين طريقتي مانتل هانزل والصعوبة المحوَّلة، في الكشف عن الأداء التفاضلي في مفردات اختبار القدرة اللفظية بالمستوى الثاني لمقياس جماس لطالبات الصفين الثالث والرابع وطلبة سلطنة عمان؟

تم ذلك عن طريق برنامج SPSS باستخدام معامل ثبات التصنيف كابا Kappa، ونسبة الاتفاق بين نتائج طريقتي مانتل هانزل والصعوبة المحوَّلة. وتلخيص الاتفاق بين

وقد ترجع هذه النتيجة إلى الدقة في إعداد مفردات هذا الاختبار ومراجعتها بناءً على معايير دقيقة، منها: مطابقتها للمنهج ومستويات الطلبة، الدقة في صياغة المموهات؛ بحيث لا يكون لها دور واضح في إظهار الأداء التفاضلي لنوع اجتماعي على حساب نوع اجتماعي آخر، ومراعاة الاختلافات الثقافية والاختلاف في الكلمات المستخدمة في هذه المفردات وملاءمتها لكلا الجنسين. وقد ترجع هذه النتيجة إلى أن الاختبار قد كان متحرراً من المناهج الدراسية المستخدمة في دول الخليج العربي، وأنه لم يتأثر بالمعلومات والأهداف التي درسها الطالب في دولته. وقد بين ميكيونغ (Mikyung, 2001) أن ظهور أداء تفاضلي لمفردة ما من حيث اللغة

مع دراسة ودمان (Wedman, 2018) والتي أشارت إلى أن عدد المفردات اللفظية التي أظهرت أداءً تفاضلياً تبعاً لمتغير النوع الاجتماعي كان أكبر بكثير من تلك المفردات الكمية، ولصالح الإناث. وقد اتفقت هذه النتيجة مع دراسة العجلوني والسوالمه (٢٠١٦) التي هدفت إلى فحص الأداء التفاضلي لموهات مفردات اختبار الرياضيات في التقييم الوطني الأردني لاقتصاد المعرفة تبعاً لموقع المدرسة، وذلك باستخدام طريقة نسبة الأرجحية لماتزل هانزل، والتي

قد يكون بسبب استخدام مفردات غير مألوفة لدى أفراد إحدى المجموعات في مضمون المفردة، وهذه المفردات لها معاني مختلفة لدى بعض المناطق.

واتفقت هذه النتيجة مع دراسة البستنجي (٢٠٠٤) عند مقارنتها بين أربع طرق للكشف عن الأداء التفاضلي لمتغير النوع الاجتماعي في مفردات اختبار القدرة اللفظية، والتي كانت نتائجها أن الأداء التفاضلي بشكل عام كان في معظمه لصالح الإناث في اختبار القدرة اللفظية، وكذلك اتفقت

جدول ٤

معامل الصعوبة ويُعد النقطة عن المحور الرئيسي وفق طريقة الصعوبة المحولة لمفردات اختبار القدرة اللفظية تبعاً لمتغير الدولة مرتبة حسب اتجاه الأداء التفاضلي

رقم المفردة	معامل الصعوبة		الاتجاه
	عمان	بقية الدول	
٧	٠.٥٤٥	٠.٥٩١	نحو باقي الدول
١٢	٠.٥٦٤	٠.٧٠٤	نحو باقي الدول
١٧	٠.٥٥٥	٠.٦٢٨	نحو باقي الدول
٢١	٠.٥٢٢	٠.٥٥٩	نحو باقي الدول
٢٢	٠.٥٢٨	٠.٥٦٦	نحو باقي الدول
٢٧	٠.٤٧٨	٠.٥٥٩	نحو باقي الدول
٨	٠.٦٩٤	٠.٦٤٦	نحو عُمان
١٣	٠.٧٢٥	٠.٦٧٣	نحو عُمان
٢٤	٠.٧٧٨	٠.٧٩٣	نحو عُمان
٢٨	٠.٧٥٣	٠.٧١٩	نحو عُمان
١	٠.٥٩٢	٠.٦٣٤	لا يوجد
٢	٠.٧٠٠	٠.٨٠١	لا يوجد
٣	٠.٦٩٧	٠.٧٠٣	لا يوجد
٤	٠.٦٤٤	٠.٦٢٣	لا يوجد
٥	٠.٦٠٠	٠.٦٧٠	لا يوجد
٦	٠.٦٤٧	٠.٧٠٤	لا يوجد
٩	٠.٥٤٤	٠.٥٥٤	لا يوجد
١٠	٠.٥٧٧	٠.٣٨٢	لا يوجد
١١	٠.٧٣٦	٠.٨٥٥	لا يوجد
١٤	٠.٦٥٦	٠.٦٠٤	لا يوجد
١٥	٠.٦٠٢	٠.٥٣٧	لا يوجد
١٦	٠.٦٨٣	٠.٧٥٩	لا يوجد
١٨	٠.٦٦١	٠.٧٠٥	لا يوجد
١٩	٠.٦٢٥	٠.٧٢٦	لا يوجد
٢٠	٠.٥٧٢	٠.٤٧٨	لا يوجد
٢٣	٠.٦٢٣	٠.٦٦٤	لا يوجد
٢٥	٠.٥٦٧	٠.٥٨٧	لا يوجد
٢٦	٠.٦٨٠	٠.٦٤٠	لا يوجد
٢٩	٠.٦٤٢	٠.٧١٩	لا يوجد
٣٠	٠.٦٦٦	٠.٥٨٨	لا يوجد

٢. دراسة الأداء التفاضلي لمفردات الاختبارات التي تطبق على مستوى السلطنة، تبعاً لمتغير النوع الاجتماعي، ومتغيرات أخرى كمتغير المحافظة مثلاً، للوقوف على المفردات ذات الأداء التفاضلي، ومن ثمّ تحليل هذه المفردات بغرض الكشف عن تحيّزها.
٣. مقارنة طريقتي مانتل هانزل، والصعوبة المحوّلة للاختبارات الأخرى في مقياس جماس، ولمتغيرات أخرى.

المراجع

References

- أرشيدات، بسام كامل (٢٠٠٧). تحري التحيّز في اختبار مادة التربية الوطنية لدى طلبة جامعة مؤتة. رسالة ماجستير غير منشورة. جامعة مؤتة، الأردن.
- البرصان، إسماعيل سلامه (٢٠١٣). الأداء التفاضلي لمتغير الجنس لفقرات الاختبار الوطني الأردني لضبط نوعية التعليم لمادة الرياضيات للصف العاشر. مجلة كلية التربية بالزقازيق، ٧٩، ٢٢٩-٢٧٠.
- البستنجي، محمود محمد (٢٠٠٤). مقارنة أربع طرق للكشف عن الأداء التفاضلي لمتغير الجنس في فقرات اختبار عقلية خاصة للفئة العمرية ١٥-١٦ سنة في الأردن. رسالة دكتوراه غير منشورة. جامعة عمان العربية، الأردن.
- البشباشة، خالد محمود (٢٠١٦). الكشف عن الأداء التفاضلي لمتغير الجنس في اختبار PISA الدولي لعام ٢٠١٢. رسالة ماجستير غير منشورة. جامعة مؤتة، الأردن.
- الجرادات، عبدالله بخيت (٢٠٠٣). مقارنة بين طريقة مانتل هانزل وطريقة الصعوبة المحوّلة في الكشف عن تحيّز الفقرات.

أظهرت نتائجها أن ١٥ مفردة أظهرت أداءً تفضلياً بالنسبة لمتغير موقع المدرسة، وبحجم تأثير متفاوت. وقد اختلفت هذه النتيجة مع دراسة مبارك (٢٠١٠) التي هدفت إلى الكشف عن الأداء التفاضلي لمفردات اختبار العلوم في الدراسة الدولية بيزا ٢٠٠٦ تبعاً لمتغيرات الدولة واللغة والانتماء لمنظمة التنمية والتعاون الاقتصادي، باستخدام طريقتي مانتل هانزل ونسبة الأرجحية، والتي فسرت النتائج المتعلقة بارتفاع نسبة المفردات التي أظهرت أداءً تفضلياً تبعاً لمتغير الدولة، إلى أن الاختبار قد لا يكون متحرراً من المناهج.

التوصيات

١. ضرورة التأكد من خلو اختبار القدرة اللفظية بالمستوى الثاني لمقياس جماس من المفردات المتحيّزة.
٢. تحليل محتوى مفردات اختبار القدرة اللفظية بالمستوى الثاني لمقياس جماس التي أظهرت أداءً تفضلياً لطالبات الصفين الثالث والرابع وطلبة سلطنة عمان، بغرض الكشف عن مدى تحيّزها.
٣. دراسة الأسباب التي أدت إلى ظهور الأداء التفاضلي بمفردات اختبار القدرة اللفظية بالمستوى الثاني لمقياس جماس، التي تم الكشف عنها لتجنبها ومعالجتها.
٤. فحص الأداء التفاضلي لمفردات الاختبارات كمؤشر أساسي على صدق الاختبار.

المقترحات

١. إجراء دراسات للكشف عن الأداء التفاضلي لمفردات اختبار القدرة اللفظية بالمستوى الثاني لمقياس جماس باستخدام طرق النظرية الحديثة في القياس.

- رسالة ماجستير غير منشورة. جامعة مؤته، الأردن.
- حمادنة، إياد محمد (٢٠٠٧). الأداء التفاضلي لمفردات اختبار تحديد الكفاءة اللفوية في اللغة الإنجليزية في الجامعات الأردنية: دراسة مقارنة. رسالة دكتوراه غير منشورة. جامعة اليرموك، الأردن.
- الزيات، فتحي؛ والمحرزي، راشد (٢٠١١). مقياس الخليج للقدرة العقلية المتعدد (جماس): التقرير الفني. الرياض: مكتب التربية العربي لدول الخليج.
- الطراونة، محمود زعل (٢٠١١). تحري التحيز في فقرات الاختبار الوطني لطلبة الصف الثامن الأساسي في مادتي العلوم والرياضيات لعام ٢٠١٠. رسالة ماجستير غير منشورة. جامعة مؤته، الأردن.
- ظاظا، حيدر إبراهيم (٢٠٠٦). فاعلية طريقة الكشف عن التحيز في الفقرة المبنية على الإجراءات التحكيمية مقارنة بالطرق التجريبية المبنية على إجراءات النظرية الكلاسيكية في القياس ونظرية الاستجابة للفقرة. رسالة دكتوراه غير منشورة. الجامعة الأردنية، الأردن.
- العجلوني، جهاد؛ والسوالمية، يوسف (٢٠١٦). فحص الأداء التفاضلي لمموهات فقرات اختبار الرياضيات في التقييم الوطني الأردني لاقتصاد المعرفة وفقاً لجنس الطالب وموقع المدرسة. مجلة العلوم الإنسانية، جامعة النجاح للأبحاث، ٣٠ (١١)، ٢٢٦٦-٢٢٨٦.
- عمارة، إيهاب محمد (٢٠٠٨). طرق الكشف عن تحيز مفردات الاختبار: دراسة سيكومترية مقارنة. رسالة ماجستير غير منشورة. جامعة عين شمس، مصر.
- كروكر، ليندا؛ والجينا، جيمز (٢٠٠٩). مدخل إلى نظرية القياس التقليدية
- والمعاصرة. ترجمة زينات يوسف دعنا. عمان: دار الفكر ناشرون وموزعون.
- الكسباني، محمد (٢٠١٢). البحث التربوي بين النظرية والتطبيق. القاهرة: دار الفكر العربي.
- مبارك، مأمون؛ ومبارك، وائل (٢٠١٣). الأداء التفاضلي لاختبار الرياضيات في الدراسة الدولية بيزا ٢٠٠٦. مجلة العلوم التربوية، جامعة الملك سعود، ٢٥ (٢)، ٢٤١-٢٦١.
- مبارك، وائل محمد (٢٠١٠). الأداء التفاضلي لفقرات اختبار العلوم في الدراسة الدولية بيزا ٢٠٠٦. رسالة دكتوراه غير منشورة. جامعة اليرموك. إربد، الأردن.
- الهريشات، محمد (٢٠٠٤). تحري التحيز في فقرات ACER للتفكير الاستدلالي. رسالة ماجستير غير منشورة. جامعة مؤته. الكرك، الأردن.
- Angoff, W. H. & Ford, S. H. (1971). Item-race interaction on a test of scholastic aptitude. *Journal of Educational Measurement*, 10 (2), 95-106.
- Camilli, G., & Shepard, L. (1994). *Methods for identifying bias test item*. Sage publication, USA.
- Dorans, N. J., & Holland, P.W. (1994). DIF detection and description: Mantel-Haenszel and standardization. In P. Holland and H. Wainer (Eds.), *Differential item functioning*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Driana, E. (2007). *Gender differential item functioning on an 11th - grade Mathematics proficiency test in application OHIO*. Unpublished Doctoral Thesis, Ohio University, Proquest information and Learning Company, UM: 3280034.
- Hambleton, R., & Rogers, J. (1995). *Item bias review practical assessment, research and evaluation*. Retrieved July 31, 2003 from <http://edresearch.org/pare/getvn.ASP=4&n=6>.

- Hambleton, R., Swaminithan, H., & Rogers, J. (1991). *Fundamentals of item response theory*. Sage publications, Newbury perk.
- Kim, S. H. (2000). *An investigation of the likelihood ratio test, the mantel test, and the generalized Mantel-Haensel test of DIF*. Paper presented at the annual meeting of the American educational research association, New Orleans, Louisiana.
- Mikyung, K. (2001). Detecting DIF across the different language groups in a speaking test. *Language Testing*, 18 (1), 89-114.
- Nandakumer, R. (1993). A fortran 77 program for detecting DIF through the Mantel-Haensel statistics. *Educational and Psychological Measurement*, 53 (3), 679-685.
- Osterlind, S. (1983). *Test item bias*. Beverly Hills: Sage publications.
- Raju, N. S. Drasgow, F. & Slind, J. (1993). An empirical comparison of the area methods, Lord's chi-square test, and the Mental-Haenszel technique for assessing differential item function. *Educational and Psychological Measurement*, 53 (1), 301-314.
- Stage, C. (2000). Predicting gender differences in word items. A comparison of item response theory and classical test theory. Swedish Council for Research in the Humanities and Social Sciences (HSFR). Retrieved from: <https://bit.ly/3eqxJhe>.
- Uiterwijk, H., & Vallen, T. (2005). Linguistic sources of item bias for second generation immigrants in Dutch tests. *Language Testing*, 22 (2), 211-234.
- Wedman, J. (2018). Reasons for gender-related differential item functioning in a college admissions test. *Scandinavian Journal of Educational Research*, 62 (6), 959-970.