

المقارنة بين طريقتي مانتل-هانزل والانحدار اللوجستي في الكشف عن الأداء التفاضلي للبند

د. ريم شهاب حسن*

(الإيداع: 19 آيار 2025، القبول: 23 تموز 2025)

الملخص:

نظراً لأهمية الكشف عن أداء البند التفاضلي الذي يُهدّد وجوده عدالة الاختبار، وفي ظل تعدد الطرائق التي تُستخدم في الكشف عنه، هدف البحث الحالي إلى المقارنة بين طريقتي مانتل هانزل والانحدار اللوجستي في الكشف عن الأداء التفاضلي للبند. ولتحقيق هذا الهدف تم تطبيق مقياس الرهاب الاجتماعي الذي أعده رولين ووي والمكون من (36) بنداً على عينة من طلبة كليتي التربية والاقتصاد في جامعة طرطوس تكونت من (686) طالباً وطالبة، بواقع (410) إناث، و(276) ذكور.

تمت المعالجة الإحصائية باستخدام حزمة (diffR)، وأظهرت النتائج ما يأتي:

- 1- حددت طريقة مانتل هانزل (7) بنود ذات أداء تفاضلي للجنس من أصل (36) بنداً لمقياس الرهاب الاجتماعي، بينما حددت طريقة الانحدار اللوجستي (12) بنداً، وكان هناك اتفاق بين الطريقتين في تحديد (5) بنود.
- 2- تراوحت أحجام تأثير الأداء التفاضلي للبنود التي حددتها طريقة مانتل هانزل بين المعتدلة والكبيرة، بينما كانت جميع أحجام تأثير الأداء التفاضلي للبنود التي حددتها طريقة الانحدار اللوجستي صغيرة.
- 3- توزعت تفضيلات البنود ذات الأداء التفاضلي بين الذكور والإناث وفقاً لطريقتي مانتل هانزل والانحدار اللوجستي، كما توزعت تفضيلات البنود المشتركة بين الطريقتين بشكل متشابه بين الذكور والإناث.

الكلمات المفتاحية: أداء البند التفاضلي، طريقة مانتل-هانزل، طريقة الانحدار اللوجستي.

*مدرس في قسم تربية الطفل في كلية التربية بجامعة طرطوس

Comparison of Mantel–Haenzel and Logistic Regression Methods in Detecting the Differential Item Functioning

Dr. Reem Shihab Hasan*

(Received: 19 April 2025, Accepted: 23 July 2025)

Abstract:

Given the importance of the detecting of the differential item functioning, that the existence of it threatens the fairness of the test, and in light of the multiplicity of methods used to detect it, the current research aimed to compare the Mantel–Haenzel and logistic regression methods in detecting the differential item functioning. To achieve this goal, the Social Phobia Scale prepared by Rollin and Wei, consisting of (36) items, was applied to a sample of students from the Faculties of Education and Economics at Tartous University, consisting of (686) male and female students, (410) females and (276) males.

The statistical processing was done using the (difR) package, and the results showed the following:

- 1- The Mantel–Haenzel method identified (7) items with differential item functioning for gender out of (36) items for the social phobia scale, while the logistic regression method identified (12) items, and there was agreement between the two methods in identifying (5) items.
- 2- The effect sizes of differential functioning of items identified by the Mantel–Haenzel method ranged from moderate to large, while all effect sizes of differential functioning of items identified by the logistic regression method were small.
- 3- The preferences of items with differential functioning were distributed between males and females according to the Mantel–Haenzel and logistic regression methods, and the preferences of items shared between the two methods were distributed similarly between males and females.

Key Words: Differential Item Functioning, Mantel–Haenzel Method, Logistic Regression Method.

*Lecturer in the Department of Child Education– Faculty of Education– Tartous Univer

مقدمة البحث

تُعدّ تحليلات الكشف عن أداء البند التفاضلي (DIF) Differential Item Functioning إجراء أساسياً وضرورياً للتحقق من عدالة الاختبارات والمقاييس، فقد وضعتها الجمعية الأمريكية للبحث التربوي (AERA) American Education Research Association ضمن قواعد النشر للاختبارات، وأشارت إلى أن شركات الاختبار يجب أن تقدّم دليلاً بأن بنود الاختبار ليس لها أداء تفاضلياً وتأثيراً سلبياً عبر المجموعات المختلفة من الممتحنين (Diaz, Brooks & Johanson, 2021, 378)، كما ووضعتها الجمعية الأمريكية لعلم النفس (APA) American Psychology Association ضمن قواعد النشر، ونصّت على أن مطوري الاختبارات يجب أن يشارروا إلى طبيعة الأدلة المتعلقة بملاءمة الاختبار للمجموعات ذات الخلفيات العرقية أو اللغوية المختلفة (سلمان ووظا، 2022، 2671)، كما وأكد إصدار معايير الاختبارات التربوية والنفسية عام (2014) على أن أداء البند التفاضلي هو دليل صدق قائم على البنية الداخلية (Elyan & Aljodeh, 2024)، ويمثّل وجوده حضور أبعاد أخرى مزعجة تتطلّف على السمة التي ضُمت البنود لقياسها (خليل وإسماعيل، 2021، 666).

يعبّر أداء البند التفاضلي عن اختلاف احتمالات الاستجابة الصحيحة للأفراد من مجموعات مختلفة على بنود الاختبار بعد أن تتم مطابقتهم في السمة الكامنة التي يقيسها هذا الاختبار (Asamoah, 2020, 13). وقد تعددت طرائق الكشف عن أداء البند التفاضلي لوجود مدخلين نظريين تنبثق عنهما، هما نظرية الاختبار الكلاسيكية (CTT) Classical Test Theory، ونظرية الاستجابة للبند (IRT) Item Response Theory، ونظراً لما تتمتع به طرائق النظرية الكلاسيكية من ميزات تتمثل في أنها تعتمد على الدرجات الخام الملاحظة على الاختبار، وبأن افتراضاتها أسهل للممارسة العملية (Basman, 2023, 146)، ولا تتطلب تحققاً من أن البيانات تحقق افتراضات النماذج المستخدمة لتقدير بارامترات الصعوبة والسهولة والتخمين كما في طرائق نظرية الاستجابة للبند (الأحمدي، 2020، 290)، وتستخدم عينات صغيرة (Basman, 2023, 146)، مقارنة بطرائق نظرية الاستجابة للبند التي تتطلب استخدام عينات كبيرة الحجم، ووقتاً أطول وتكلفة أكثر (العبد القادر والسوالمه، 2021، 255)، ونظراً لما تتمتع به طريقتا مانتل هانزل (Mantel- (M-H) Haenzel، والانحدار اللوجستي (LR) Logistic Regression اللتين تنتميان إلى طرائق النظرية الكلاسيكية من ميزات تتمثل في سهولتهما من الناحية الحسابية، وسهولة تفسير نتائجهما (العجلوني والسوالمه، 2019، 52؛ Duncan, 2006، 26)، بالإضافة إلى كونهما توفران مقياساً لحجم التأثير، واختباراً للدلالة الإحصائية، ولا تحتاجان إلى عينات كبيرة (خليل وإسماعيل، 2021، 667؛ Macinnes, 2009, 26)، وانطلاقاً مما توصلت إليه نتائج العديد من الدراسات حول فاعلية كل منهما في الكشف عن البنود ذات الأداء التفاضلي، جاء البحث الحالي للمقارنة بين هاتين الطريقتين في الكشف عن الأداء التفاضلي، لمعرفة إن كان لهما نفس الفعالية، أم أن لإحدهما أفضلية على الأخرى في ذلك.

1- مشكلة البحث:

انطلاقاً من كون الكشف عن البنود ذات الأداء التفاضلي لم يعد مطلباً ثانوياً، بل ضرورة لكل الاختبارات، لأن وجوده يكون مؤشراً على عدم تحقق عدالة الاختبار بشكل عام (Salubayba, 2013, 76)، ويهدد صدقه (Kurnaz & Yildiz, 2023, 436)، كما يهدد التفسير الصحيح للفروق بين المجموعات التي تخضع له (Allahyari, 2020, 96)، وانطلاقاً من توصيات الدراسات المتعلقة بالكشف عن الأداء التفاضلي لبنود الاختبار باستخدام طريقتين على الأقل للكشف عن الأداء التفاضلي لتحسين موثوقية وثبات النتائج، حيث أن الطرائق المختلفة توفر نتائج مختلفة في مواقف معينة (Basman, 2019, 155 ; Villas, 2023).

ونظراً لتباين نتائج الدراسات حول مدى دقة أساليب الكشف عن أداء البند التفاضلي، وأفضلية بعضها على البعض الآخر، ولاختلاف نتائج الدراسات التي قارنت بين طريقتي مانتل هانزل والانحدار اللوجستي، فقد توصلت بعض الدراسات إلى أفضلية

طريقة مانتل هانزل على طريقة الانحدار اللوجستي كدراسة (Bastug, 2016)، ودراسة (Villas, 2019)، بينما توصل بعضها إلى أفضلية طريقة الانحدار اللوجستي على طريقة مانتل هانزل كدراسة (حسن، 2020)، ودراسة (Alahmadi, 2019)، ودراسة (Simbolon, Supriyati & Naga, 2019)، ودراسة (Moralez-Oetiz & Pena-Gardenas, 2019)، ودراسة (Abbott, 2021)، في حين توصل البعض الآخر إلى أن للطريقتين نفس الفعالية في الكشف عن البنود ذات الأداء التفاضلي كدراسة (حسن، 2020)، ودراسة (Terzi & Yakar, 2018)، تتمثل ضرورة إجراء البحث الحالي في مجال المقارنة بين طريقتي مانتل هانزل والانحدار اللوجستي في الكشف عن البنود ذات الأداء التفاضلي، للمساهمة في توفير معلومات حول أفضلية أو عدم أفضلية إحداها على الأخرى، وتتلخص مشكلة البحث بما يأتي: " المقارنة بين طريقتي مانتل هانزل والانحدار اللوجستي في الكشف عن الأداء التفاضلي للبنود".

2- أهمية البحث:

تكمن أهمية البحث في النقاط الآتية:

- 1- أهمية دراسة أداء البند التفاضلي، والكشف عنه في بنود الاختبارات بوصفه خطوة رئيسية في بناء اختبارات عادلة غير متحيزة تتمتع بالخصائص السيكومترية اللازمة.
- 2- تطوير مقياس الرهاب الاجتماعي من خلال الكشف عن البنود ذات الأداء التفاضلي، ليُصار إلى محاولة تعديلها أو حذفها واستبدالها بما يضمن تحقق عدالته.
- 3- تأكيد فاعلية طريقتي مانتل هانزل والانحدار اللوجستي كأسلوبين إحصائيين في تحديد البنود ذات الأداء التفاضلي.
- 4- توجيه أنظار الباحثين إلى ضرورة الكشف عن الأداء التفاضلي لبنود المقاييس التي يستخدمونها، واعتبارها من إجراءات التأكد من شروط صلاحية استخدامها.
- 5- توجيه أنظار الباحثين إلى ضرورة استخدام طرائق كشف عن أداء البند التفاضلي أثبتت فعاليتها أكثر من غيرها.
- 6- تنبيه الباحثين لمراجعة نتائج مقارنة الذكور والإناث في استجاباتهم على مقياس الرهاب للكشف عن الفروق الحقيقية بينهم بعد استبعاد الفروق غير الحقيقية الناتجة عن الأداء التفاضلي لبعض البنود.
- 7- إن هذا البحث هو الأول - في حدود علم الباحثة- الذي قارن بين طريقتي مانتل هانزل والانحدار اللوجستي في الكشف عن الأداء التفاضلي لبنود مقياس الرهاب الاجتماعي.
- 3- أهداف البحث: للبحث الحالي هدف رئيس يتمثل في المقارنة بين طريقتي مانتل هانزل والانحدار اللوجستي في الكشف عن الأداء التفاضلي لبنود مقياس الرهاب الاجتماعي وفقاً لمتغير الجنس، وينفّرع عن هذا الهدف ما يأتي:
 - 1- الكشف عن بنود مقياس الرهاب الاجتماعي التي تُظهر أداءً تفاضلياً باستخدام طريقتي مانتل هانزل (MH) والانحدار اللوجستي (LR).
 - 2- الكشف عما إذا كان هناك اتفاق بين طريقتي مانتل هانزل والانحدار اللوجستي في تحديد بنود مقياس الرهاب الاجتماعي ذات الأداء التفاضلي، وتحديد أيهما حددت عدداً أكبر من البنود مُقارنة بالأخرى.
 - 3- تحديد كيفية توزّع أحجام تأثير بنود مقياس الرهاب الاجتماعي التي تمّ تحديدها على أنها ذات أداء تفاضلي باستخدام كل من الطريقتين، وتحديد أيهما الأفضل بناء على أحجام التأثير التي حددتها.
 - 4- تحديد كيفية توزّع تفضيلات البنود ذات الأداء التفاضلي التي تمّ تحديدها في كل من الطريقتين بين المجموعتين المقارنتين (الذكور والإناث)، وتحديد مدى تشابه تلك التفضيلات بالنسبة للبنود المشتركة بينهما.
 - 4- أسئلة البحث: إن تحقيق أهداف البحث تطلب الإجابة عن الأسئلة الآتية:
 - 1- ما بنود مقياس الرهاب الاجتماعي التي تُظهر أداءً تفاضلياً باستخدام طريقتي مانتل هانزل (MH) والانحدار اللوجستي (LR)؟

- 2- هل اتفقت طريقتا مانتل هانزل والانحدار اللوجستي في تحديد بنود مقياس الرهاب الاجتماعي ذات الأداء التفاضلي؟ وهل حددت إحداهما عدداً أكبر من البنود مقارنة بالأخرى؟
- 3- كيف توزعت أحجام تأثير بنود مقياس الرهاب الاجتماعي التي تم تحديدها على أنها ذات أداء تفاضلي باستخدام كل من الطريقتين؟ وأي الطريقتين كانت أفضل بناء على أحجام التأثير التي حددتها؟
- 4- كيف توزعت تفضيلات هذه البنود التي تم تحديدها في كل من الطريقتين بين المجموعتين المقارنتين (الذكور والإناث)؟ وهل كانت تلك التفضيلات متشابهة في البنود المشتركة؟
- 5- مصطلحات البحث وتعريفاته الإجرائية:

- **أداء البند التفاضلي (DIF) Differential Item Functioning**: هو دالة إحصائية لفحص وتقييم عدالة القياس عبر مجموعات مختلفة من المفحوصين، عندما تكون هذه المجموعات متشابهة في القدرة (Khalid, Shafiq & Ahmed, 2021, 71). بمعنى آخر، هو دالة إحصائية تعبر عن الفروق في أداء البند بين مجموعات الأفراد المستجيبين له، والذين يمتلكون المستوى نفسه من السمة الكامنة (القدرة) (Khoeruroh & Retnawati, 2020, 1).

ويُعرف أداء البند التفاضلي (DIF) إجرائياً بأنه: قيمة نسبة الأرجحية المحسوبة وفقاً لطريقتي مانتل هانزل والانحدار اللوجستي التي تُظهر وجود فروق في احتمالية الاستجابة الصحيحة على كل بند من بنود مقياس الرهاب الاجتماعي بين مجموعتي الذكور (المجموعة البؤرية) والإناث (المجموعة المرجعية) في عينة الدراسة (طلبة كليتي الاقتصاد والتربية في جامعة طرطوس) عندما تكون قيمة دلالتها أصغر أو تساوي (0.05).

- **طريقة مانتل هانزل (MH) Mantel-Haenszel**: هي واحدة من أبرز طرائق الكشف عن أداء البند التفاضلي (DIF) في ظل النظرية الكلاسيكية (Liu, 2017, 30)، من خلال مقارنة توزيع الاستجابات على البند بين مجموعتين مدرستين، تُمثل إحداهما المجموعة المرجعية، بينما تُمثل الأخرى المجموعة البؤرية (المُستهدفة) (Khoeruroh & Retnawati, 2020, 3)، وهي مفيدة للكشف عن دالة البند التفاضلي في البنود ثنائية الدرجة، كما يمكن أن تمتد للبنود متعددة الدرجات (Wiberg, 2007, 7).

- **طريقة الانحدار اللوجستي (LR) Logistic Regression**: هو طريقة بارامترية، تُستخدم في الكشف عن أداء البند التفاضلي مع البنود ثنائية الدرجة، ومتعددة الدرجات (Apinyapibal, Lawthong & Kanjanawasee, 2015, 24). وتكشف هذه الطريقة عن الأداء التفاضلي بنوعيه المنتظم وغير المنتظم، ونقارن فيها بين ثلاثة نماذج للانحدار تأخذ بعين الاعتبار قدرة الفرد والمجموعة التي ينتمي إليها والتفاعل بينهما (Allahyari, 2020, 97).

6- الدراسات السابقة:

1- دراسة آبوت (Abbott, 2021): نيجيريا

عنوان الدراسة: حساسية الكشف عن الأداء التفاضلي باستخدام طرائق كاي مربع لورد ومنطقة راجو والانحدار اللوجستي ومانتل هانزل والتوحيد القياسي وصعوبة البند المحولة، بالمقارنة، باستخدام R.

DIF Detection Sensitive of Lord's Chi-square, Raju's Area, Logistic Regression, Mantel-Haenszel, Standardization and Transformed item Difficulties Methods, In Comparison, Using R.

هدفت هذه الدراسة إلى مقارنة ست طرائق للكشف عن الأداء التفاضلي باستخدام (R) هي كاي مربع لورد ومنطقة راجو والانحدار اللوجستي ومانتل هانزل والتوحيد القياسي والصعوبة المحولة للبند في اختبار الاختيار من متعدد لامتحان الثانوية العامة في علوم الكمبيوتر المكون من (60) بنوداً يُجاب عليها بشكل ثنائي، بعد تطبيقه على (400) طالباً وطالبة من طلبة السنة الأولى الذين يدرسون تكنولوجيا معلومات واتصالات، منهم (200) ذكور و(200) إناث.

استُخدم برنامج R من أجل تحليلات أداء البند التفاضلي باستخدام الطرائق الستة، وتوصلت الدراسة إلى أن طريقتي مانتل هانزل والانحدار اللوجستي كشفتتا عن بنود ذات أداء تفاضلي، ولكن كانت طريقة الانحدار اللوجستي أكثر فعالية من طريقة مانتل هانزل، حيث كشفت الأولى عن (16) بنوداً ذات أداء تفاضلي، بينما كشفت الثانية عن (8) بنود.

2- دراسة (حسن، 2020): سورية

عنوان الدراسة: الكشف عن دالة البند التفاضلي (DIF) في بناء الاختبارات النفسية.

هدفت هذه الدراسة إلى الكشف عن الأداء التفاضلي حسب متغير الجنس لبنود المقاييس العيادية العشرة لاختبار مينسوتا متعدد الأوجه للشخصية الخاص بالمرهقين، وبنود اختبار القدرات المعرفية باستخدام أربع طرائق للكشف عن أداء البند التفاضلي منها طريقتي مانتل هانزل، والانحدار اللوجستي. ولتحقيق هذا الهدف تم تطبيق اختبار القدرات المعرفية على عينة مكونة من (2005) طالباً وطالبة من الصفين الثاني والثالث الثانوي في محافظة طرطوس في العام الدراسي (2018-2019)، وتطبيق اختبار مينسوتا أيضاً على عينة مكونة من (2005) طالباً وطالبة من الصفين الثاني والثالث الثانوي في محافظة طرطوس خلال العامين الدراسيين (2015-2016) و(2018-2019)، ومثل الذكور المجموعة المستهدفة، ومثلت الإناث المجموعة الرجعية.

استُخدمت الحزمة (difR) من لغة البرمجة الإحصائية (R) لإجراء تحليلات الأداء التفاضلي باستخدام الطريقتين، وتوصلت الدراسة إلى أن طريقتي مانتل هانزل والانحدار اللوجستي حددتا نفس البنود ذات الأداء التفاضلي في اختبار القدرات المعرفية وحددتا نفس التفضيلات لهذه البنود ونفس أحجام التأثير التي كانت من النوع الصغير، وفيما يتعلق بالمقاييس العيادية العشرة لاختبار مينسوتا، أولاً، اتفقت الطريقتان في تحديد نفس عدد وأرقام البنود ذات الأداء التفاضلي في خمسة مقاييس، بينما لم تتفقان تماماً في تحديد تفضيلات هذه البنود في إلا في أربعة من المقاييس واختلفتا في تحديد التفضيلات للمقياس الخامس في بعض البنود، كما أنهما اختلفتا في تحديد أحجام التأثير، حيث كانت أحجام التأثير بطريقة الانحدار اللوجستي صغيرة للبنود ذات الأداء التفاضلي في أربعة مقاييس، وتراوحت بين الصغيرة والمعتدلة لبنود المقياس الخامس، بينما في طريقة مانتل هانزل تراوحت بين الصغيرة والمتوسطة لبنود مقياسين، وبين بين الصغيرة والكبيرة لبنود مقياسين آخرين، وبين الصغيرة والكبيرة والمتوسطة لبنود المقياس الخامس. ثانياً، تفوقت طريقة الانحدار اللوجستي على طريقة مانتل هانزل وحددت بنوداً إضافياً على البنود نفسها التي حددتها الأخيرة، واتفقت الطريقتان في تحديد تفضيلات البنود المشتركة للمقاييس جميعها، إلا أنهما اختلفتا في تحديد أحجام التأثير لمقياسين، حددتها طريقة الانحدار اللوجستي على أنها صغيرة، بينما حددتها طريقة مانتل هانزل بين الصغيرة والمتوسطة في مقياس، وبين الصغيرة والكبيرة في المقياس الثاني، أما فيما يتعلق بالمقياس الثالث فتراوحت أحجام التأثير بين الصغيرة والمتوسطة بطريقة الانحدار اللوجستي، وبين الصغيرة والمتوسطة والكبيرة بطريقة مانتل هانزل. ثالثاً، اتفقت الطريقتان في تحديد نفس عدد البنود ذات الأداء التفاضلي في مقياس واحد، واتفقتا في تحديد أربعة بنود متشابهة، واختلفتا في تحديد البند الخامس، واتفقتا في تحديد تفضيلات البنود المشتركة، إلا أنهما اختلفتا في تحديد أحجام تأثيرها التي تراوحت بين الصغيرة والكبيرة بطريقة مانتل هانزل، وكانت صغيرة بطريقة الانحدار اللوجستي. رابعاً، تفوقت طريقة الانحدار اللوجستي على طريقة مانتل هانزل حيث حددت نفس البنود التي حددتها طريقة مانتل هانزل في أحد المقاييس مع بنود إضافية عليها، إلا أن الطريقتين اتفقتا في تحديد تفضيلات هذه البنود، وفي تحديد أحجام التأثير لها والتي كانت من النوع الصغير.

3- دراسة الأحمدى (Alahmadi, 2019): السعودية

عنوان الدراسة: تقييم التحيز المحتمل في تقييم المهوبة: مقارنة ميدانية لتقنيات IRT و Non-IRT (البارامترية واللابارامترية) للكشف عن أداء البند التفاضلي.

Evaluating Possible Bias in a Giftedness Assessment: An Empirical Comparison of IRT and Non-IRT (Parametric and Non-Parametric) Techniques for Detecting Differential Item Functioning.

هدفت هذه الدراسة إلى مقارنة مجموعات مختلفة من الأفراد الذكور والإناث باستخدام طرائق عدة للكشف عن أداء البنود التفاضلي منها طريقتي مانتل هانزل، والانحدار اللوجستي، لتحديد أي هذه الطرائق أفضل للكشف عن الأداء التفاضلي، واعتمدت الباحثة على بيانات تم تجميعها من تطبيق بطارية قياس المواهب والإبداع المكونة من (100) بند على طلبة الصف التاسع في المملكة العربية السعودية بلغ عددهم (12572) طالباً وطالبة ممن خضعوا لتقييم القدرات المعرفية المتعددة (تقييم الموهبة) خلال عام (2017)، مثل الذكور فيها المجموعة البؤرية، ومثلت الإناث المجموعة المرجعية.

استُخدمت حزمة (diffR) من برنامج (R) من أجل تحليلات أداء البنود التفاضلي، وتوصلت الدراسة إلى أن طريقة الانحدار اللوجستي كانت الأكثر فاعلية من طريقة مانتل هانزل في الكشف عن البنود ذات الأداء التفاضلي، إذ حددت عدد من البنود ذات الأداء التفاضلي أكبر مما حددته الأخيرة. وكان هناك اتفاق بين الطريقتين في تحديد عدد من البنود، وتروحت أحجام تأثير الأداء التفاضلي للبنود المحددة بين الصغير والكبير والمتوسط باستخدام طريقة مانتل هانزل، بينما كانت أحجام تأثير البنود التي حددتها طريقة الانحدار اللوجستي جميعها من النوع الصغير عدا بنوداً واحداً.

4- دراسة سيمبولون وسوبرياتي وناجا (Simbolon, Supriyati & Naga, 2019): أندونيسيا

عنوان الدراسة: حساسية طرائق الكشف عن أداء البنود التفاضلي على امتحان الرياضيات الوطني في مقاطعة شمال سومطرة، أندونيسيا.

Sensitivity of Differential Item Functioning Detection Method on National Mathematics Examination in North Sumatera Province, Indonesia.

هدفت هذه الدراسة إلى فحص حساسية ثلاث طرائق منها طريقتي مانتل هانزل، والانحدار اللوجستي في الكشف عن أداء البنود التفاضلي بين الجنسين في اختبار الرياضيات المكون من (40) بنوداً، طبق الاختبار على عينة مكونة من (5000) طالب وطالبة، مثل الذكور فيها المجموعة المرجعية، ومثلت الإناث المجموعة البؤرية المستهدفة.

استُخدمت في الدراسة ثلاثة حجوم للعينة من أجل دراسة حساسية الطريقتين في الكشف عن أداء البنود التفاضلي، واستُخدم برنامج (R) من أجل تحليلات الكشف، وتوصلت الدراسة إلى أن طريقتي مانتل هانزل والانحدار اللوجستي كشفتنا عن نسب مختلفة من البنود ذات الأداء التفاضلي مع أحجام العينات المختلفة، وأطوال الاختبار المختلفة، وكانت طريقة الانحدار اللوجستي أكثر حساسية، إذ حددت جميع البنود التي حددتها طريقة مانتل هانزل على أنها ذات أداء تفاضلي بالإضافة إلى بنود أخرى لم تحدها الأخيرة.

5- دراسة موراليز أوتيز وبينا كارديناس (Moralez-Oetiz & Pena-Cardenas , 2019): أمريكا

عنوان الدراسة: أداء البنود التفاضلي في النسخة الإسبانية من اختبار التذكر الانتقائي اللفظي باستخدام عينات من الإسبانين في الولايات المتحدة والمكسيكيين والإسبانين.

DIF in the Spanish Version of the Verabl Selective Reminding Test using Samples from Hispanics in the United State, Maxisans and Spaniards.

هدفت هذه الدراسة إلى استخدام طريقتي مانتل هانزل والانحدار اللوجستي لفحص وجود بنود ذات أداء تفاضلي في النسخة الإسبانية من اختبار التذكر الانتقائي اللفظي (VSRT) المكون من (12) كلمة غير مرتبطة ببعضها بعضاً، بعد

تطبيقه على عينة مكونة من (617) ذكراً وأنثى، عدد الإسمان فيها (211)، وعدد المكسيكيين (201)، وعدد الإسمانيين في الولايات المتحدة (205).

استخدمت الدراسة برنامج (R) وحزمة (diffR) من أجل تحليلات أداء البند التفاضلي باستخدام الطريقتين، وتوصلت إلى أن طريقتي مانتل هانزل والانحدار اللوجستي كانتا فعاليتين في الكشف عن البنود ذات الأداء التفاضلي، ولكن الأخيرة كانت أكثر فاعلية من الأولى لأنها كشفت عن عدد أكبر من البنود، كما توصلت الدراسة إلى وجود اتفاق بين الطريقتين في تحديد البنود ذات الأداء التفاضلي، باستثناء ثلاثة بنود كان هناك تناقض بينهما في تحديدها.

6- دراسة فيلاس (Villas, 2019): الفلبين

عنوان الدراسة: أداء البند التفاضلي في رياضيات الصف الثامن باستخدام الانحدار اللوجستي ومانتل هانزل والتحليل المنطقي للبيانات.

Differential Item Functioning in Grade 8 Math Using Logistic Regression, Mantel-Haenszel and Logical Data Analysis.

هدفت هذه الدراسة إلى الكشف عن البنود ذات الأداء التفاضلي للجنس في اختبار تحصيل الرياضيات المكون من (40) بنداً من نوع الاختيار من متعدد في الاحتمالات والإحصاء، بعد تطبيقه على عينة مكونة من (200) طالب وطالبة من طلبة الصف الثامن تم اختيارهم بطريقة عشوائية. مثل الذكور المجموعة المستهدفة، ومثلت الإناث المجموعة المرجعية. استخدم برنامج SPSS من أجل تحليلات أداء البند التفاضلي باستخدام طريقتي مانتل هانزل والانحدار اللوجستي، وتوصلت الدراسة إلى أن الطريقتين حددتا عدداً من البنود ذات الأداء التفاضلي، وكانت طريقة مانتل هانزل أفضل، إذ حددت بنداً إضافياً على البنود التي حددتها طريقة الانحدار اللوجستي، واختلفت الطريقتان أيضاً في تحديد بنود ذات أداء تفاضلي بتفضيل مختلف لمجموعتي الدراسة (الذكور والإناث)، وفي أحجام تأثير مختلفة للأداء التفاضلية، إذ كانت جميع البنود التي حددتها طريقة مانتل هانزل ذات حجم تأثير كبير وجميعها فضلت الذكور، بينما تنوعت تفضيلات البنود بين الذكور والإناث حسب طريقة الانحدار اللوجستي، وتنوعت أحجام تأثير الأداء التفاضلي لها بين الصغيرة والمتوسطة والكبيرة.

وأوصت الدراسة باستخدام طريقتين على الأقل في الكشف عن أداء البند التفاضلي من أجل تقديم مزيد من التبرير لعرض البند ذي الأداء التفاضلي.

7- دراسة أوكاندا وأوثون وأجاك وأوليش (Ukanda, Othun, Agak & Oleche, 2019): أمريكا

عنوان الدراسة: فعالية إحصاءات مانتل هانزل والانحدار اللوجستي في الكشف عن أداء البند التفاضلي ضمن شروط مختلفة من حجم العينة وتوزيع القدرة وطول الاختبار.

Effectiveness of Mantel-Haenszel and Logistic Regression Statistics in Detecting Differential Item Functioning under Different Conditions of Sample Size, Ability Distribution and Test Length.

هدفت هذه الدراسة إلى التحقق من تأثيرات حجم العينة، وتوزيع القدرة، وطول الاختبار على عدد اكتشافات البنود ذات الأداء التفاضلي باستخدام طريقتي مانتل هانزل والانحدار اللوجستي.

ومن أجل تحقيق هذا الهدف تم توليد (18) مجموعة بيانات ثنائية التفرع باستخدام برنامج WinGen3، وتمت محاكاة عينات لظروف مختلفة باستخدام تقنية أخذ العينات العشوائية التطبيقية من استجابات (2000) فرداً، وكان لكل من المجموعتين المرجعية والمستهدفة ثلاثة أحجام (حجمان صغيران وحجم كبير). واستخدم برنامج SPSS من أجل تحليلات أداء البند التفاضلي باستخدام طريقة الانحدار اللوجستي، وبرنامج Exel للحصول على أحجام تأثير الأداء التفاضلي

لطريقة مانتل هانزل، وتوصلت الدراسة إلى أن طريقة الانحدار اللوجستي كانت أفضل من طريقة مانتل هانزل في الكشف عن عدد أكبر من البنود ذات الأداء التفاضلي، وفي الكشف عن البنود ذات الأداء التفاضلي الصغير والمعتدل بغض النظر عن توزيع القدرة، وحجم العينة، وطول الاختبار، ولكن كانت طريقة مانتل هانزل أفضل في الكشف عن البنود ذات الأداء التفاضلي من النوع الكبير.

8- دراسة ترزي وياكار (Terzi & Yakar, 2018): تركيا

عنوان الدراسة: تحليلات الأداء التفاضلي للبنود والمشتت على امتحان القبول في المدرسة الثانوية التركية.

Differential Item and Differential Distractor Functioning Analyses on Turkish High School Entrance Exam.

هدفت هذه الدراسة إلى إجراء تحليلات الأداء التفاضلي على البنود (DIF) والمشتتات (DDF) لاختبار القبول في المدرسة الثانوية، استخدمت طريقتي مانتل هانزل والانحدار اللوجستي لتحديد البنود ذات الأداء التفاضلي في اختبار القبول المكون من (20) بنوداً في الرياضيات، بعد تطبيقه على (27952) طالباً وطالبة من طلبة الصف الثامن منهم (55%) ذكور، و(45%) إناث، واختيرت ثلاثة أحجام للعينات هي (500-1000-2000)، وتم تنفيذ مئة تكرار لفحص تأثير حجم العينة على الكشف عن الأداء التفاضلي للبنود.

توصلت الدراسة إلى أن طريقتي مانتل هانزل والانحدار اللوجستي كانتا متماثلتين ولم تحدد أي بنود ذات أداء تفاضلي عند حجمي العينة (500-1000)، بينما عند حجم العينة (2000)، حددت طريقة الانحدار اللوجستي (4) بنود على أنها ذات أداء تفاضلي صغير، وحددت طريقة مانتل هانزل (3) بنود ذات أداء تفاضلي معتدل، بندان فضلاً للذكور، وبند فضلاً للإناث، واشتركت الطريقتين في تحديد بندين.

9- دراسة كوكوك وغول وغول (Cokluk, Gul & Gul, 2016): تركيا

عنوان الدراسة: فحص أداء البنود التفاضلي لأشكال اختبار مختلف ترتيب بنوده حسب مستويات الصعوبة.

Examining Differential Item Functioning of Different Item ordered Test forms

According to Item Difficulty Levels.

هدفت هذه الدراسة إلى استخدام طرائق مانتل هانزل والانحدار اللوجستي وكاي مربع لورد ومنطقة راجو في فحص أداء البنود التفاضلي في ثلاث صيغ مختلفة: صيغة عشوائية وصيغة متسلسلة من البنود الأسهل إلى البنود الأصعب، وصيغة متسلسلة من البنود الأصعب إلى البنود الأسهل، لاختبار تحصيل البنود المكون من (20) بنوداً متعددة الخيارات لكل منها أربعة بدائل، بعد تطبيقه على عينة مكونة من (578) طالباً وطالبة من طلبة الصف السابع تم سحبهم بطريقة عشوائية من سبع مدارس في أنقرة.

استُخدمت حزمة (difR) من برنامج (R) من أجل تحليلات أداء البنود التفاضلي، وتوصلت الدراسة إلى أن طريقة الانحدار اللوجستي حددت عدد من البنود ذات الأداء التفاضلي أكبر من العدد الذي حددته طريقة مانتل هانزل في الصيغ الثلاث، واتفقت الطريقتان في تحديد بنود مشتركة على أنها ذات أداء تفاضلي، وحددت طريقة الانحدار اللوجستي في كل صيغة عدد إضافي على البنود التي حددتها طريقة مانتل هانزل على أنها ذات أداء تفاضلي، واتفقت الطريقتان في صيغتين للاختبار بتحديد أحجام تأثير متوسطة للأداء التفاضلي، بينما اختلفتا في الصيغة الثالثة، إذ كانت البنود التي حددتها طريقة مانتل هانزل ذات أحجام تأثير كبيرة، بينما تراوحت أحجام تأثير البنود التي حددتها طريقة الانحدار اللوجستي بين المتوسطة والكبيرة.

10- دراسة باستوغ (Bastug, 2016): تركيا

عنوان الدراسة: مقارنة أربعة إجراءات لأداء البنود التفاضلي في وجود تعددية الأبعاد.

A Comparison of four Differential Item Functioning Procedure in the presence of multidimensionality.

هدفت هذه الدراسة إلى مقارنة أربع طرائق للكشف عن أداء البند التفاضلي منها طريقتي مانتل هانزل، والانحدار اللوجستي. ولتحقيق هذا الهدف تم توليد البيانات وفق النموذج ثنائي البارامتر باستخدام برنامج (UNIGEN)، وتم إنشاء مجموعتين من البيانات تحتوي كل منهما على (1000) فرداً، و(25) بند، وإحدهما لا أداء تفاضلي فيها، بينما الأخرى فيها أداء تفاضلي في خمسة بنود.

استُخدم برنامج (MANTEL) للحصول على إحصاء مانتل هانزل، و برنامج (SAS) للحصول على إحصاء الانحدار اللوجستي، وتوصلت الدراسة إلى أن طريقة مانتل هانزل كشفت عن عدد أكبر من البنود ذات الأداء التفاضلي في حال وجود أداء تفاضلي مقارنة بطريقة الانحدار اللوجستي، وتوزعت تفضيلات البنود التي حددتها الطريقتان بين الذكور والإناث، أما في حال عدم وجود أداء تفاضلي، توصلت الدراسة إلى أن طريقة مانتل هانزل تفوقت على طريقة الانحدار اللوجستي التي لم تحدد أي بند على أنه ذو أداء تفاضلي.

التعليق على الدراسات السابقة وموقع الدراسة الحالية منها:

اتفقت جميع الدراسات السابقة التي تم ذكرها فيما بينها ومع الدراسة الحالية بالهدف التي سعت إليه وهو المقارنة بين طريقتي مانتل هانزل والانحدار اللوجستي في الكشف عن البنود ذات الأداء التفاضلي وفقاً لمتغير الجنس، إلا أن هذه الدراسات اختلفت فيما بينها في تحديدها لمجموعتي الدراسة، فحددت الدراسة رقم (4) الإناث كمجموعة مستهدفة (بؤرية) والذكور كمجموعة مرجعية، وانعكس تحديد هاتين المجموعتين ليكون الذكور هم المجموعة المستهدفة والإناث هنّ المجموعة المرجعية في باقي الدراسات وبما يتفق في هذا التحديد مع الدراسة الحالية، باستثناء الدراستين ذاتا الرقمين (7-10) لم تحددان المتغير الذي دُرس وفقه الأداء التفاضلي للبند، تم توليد البيانات في الدراستين ذاتا الرقمين (7-10)، بينما استُخدمت بيانات حقيقية في باقي الدراسات كما في الدراسة الحالية، وبالنسبة لنوع الاختبار المستخدم، فقد استخدمت الدراسات ذات الأرقام (2-3-5-6) مقاييس واختبارات نفسية متفقة بذلك مع الدراسة الحالية، بينما استخدمت الدراسات ذات الأرقام (1-4-8-9) اختبارات تحصيل، وفيما يتعلق بالبرامج الإحصائية المستخدمة، اتفقت معظم الدراسات فيما بينها ومع الدراسة الحالية في استخدام برنامج (R)، باستثناء الدراسة رقم (10) استخدمت برنامجي (MANTAL & SAS)، والدراستين ذاتا الرقمين (6-7) اتفقتا باستخدام برنامج (SPSS)، واستخدمت الدراسة رقم (7) أيضاً برنامج (Exel). وفيما يتعلق بالنتائج، توصلت جميع الدراسات إلى أفضلية طريقة الانحدار اللوجستي على طريقة مانتل هانزل في الكشف عن البنود ذات الأداء التفاضلي وتحديدها، باستثناء الدراستين ذاتا الرقمين (6-10) توصلنا إلى نتيجة معاكسة.

7- الإطار النظري للبحث:

مفهوم أداء البند التفاضلي (DIF):

إن أداء البند التفاضلي أحد مظاهر التحيز الملحوظ، وهو خطأ منتظم أو غير صادق حول كيف يقيس بند الاختبار الأفراد من مجموعة معينة، ويخلق تشويهاً أو انحرافاً في نتائج الاختبار لأعضاء هذه المجموعة (Camilli & Shepard, 1994, 8)، ويشير إلى اختلاف احتمال الإجابة الصحيحة على بند معين للأفراد الممتحنين من مجموعات سكانية مختلفة (Lee & Suh, 2018, 328)، ويكون هذا الاختلاف بسبب خاصية أو سمة ذات صلة أو غير ذات صلة بالسمة (القدرة) التي يتم قياسها مثل الجنس، أو العرق، بعد أن يتم التحكم بالفروق بينهم في هذه القدرة المقاسة (Raykov, Marcoulides, Lee & Chang, 2013, 898)؛ أي إن أداء البند التفاضلي أداة لقياس الفروق في دالة الاستجابة للبند عبر المجموعات المتطابقة في قدرتها (Zhang, 2015, 1)، بحيث لا تتمتع إحدى المجموعتين بفرصة متساوية لتأييد البند،

والإجابة عليه بشكل صحيح على الرغم من أن أفرادها لهم نفس مستوى قدرة أفراد المجموعة الأخرى (Penfield, 2001, 236، أرشيدات، 2007، 5).

وبعبارة أخرى، ينطوي أداء البند التفاضلي على تقييم العلاقات الشرطية بين الاستجابة للبند، وعضوية المجموعة (Teresi, Ramirez, Lai & Silver, 2008, 540) على القدرة التي يتم قياسها بواسطة هذا البند (Teresi & Fleishman, 2007, 34)، ويكون مؤشراً على أن البند متحيز لإحدى المجموعتين المقارنتين على أساس الجنس، أو العرق، أو الموقع الجغرافي (McCallum, 2003, 22).

والبند ذو الأداء التفاضلي هو البند الذي يكشف عن فروق غير متوقعة في الأداء بين مجموعتين من الأفراد من نفس مستوى القدرة (Sandilands, 2014, 4)، ويتعلق بالصعوبة النسبية غير المفسرة (Orlawence & Queensoap, 2019, 81)، بحيث يكون أكثر صعوبة لمجموعة من مجموعة أخرى لها نفس مستوى القدرة للمُراد قياسها في الاختبار (الصمادي والزهراني، 2014، 5، 261، Embreston & Reise, 2000).

فإذا كان احتمال الإجابة على البند (i) هو (U_i)، وكان (θ) هو مستوى القدرة، و (G) هو عضوية المجموعة، فإن المعادلة العامة لدالة البند التفاضلي تُعطى بالعلاقة:

$$(U_i | \theta, G) \neq (U_i | \theta)$$

وفي حالة البنود الثنائية، تكون العلاقة كالتالي:

$$P(U_i = 1 | \theta, G) \neq P(U_i | \theta)$$

حيث: P : احتمال، U_i : متغير مؤشر البند للإجابة الصحيحة (1)، وللإجابة الخاطئة (0)، θ : القدرة، G: عضوية المجموعة (Paek & Wilso, 2011, 1024).

وهذا يعني أن الفرصة للإجابة على البند بشكل صحيح لا تعتمد فقط على القدرة (θ)، ولكن تعتمد أيضاً على عضوية المجموعة (G).

ولما كانت عضوية المجموعة تحتوي على قيمتين هما: (R) للمجموعة المرجعية، و (F) للمجموعة المستهدفة (البؤرية)، يحدد أداء البند التفاضلي بالعلاقة:

$$P(U_i = 1 | \theta, G = F) \neq P(U_i | \theta, G = R)$$

وهذا يعني أن احتمالية الإجابة الصحيحة على البند للمجموعة المستهدفة يختلف عن احتمال الإجابة الصحيحة على البند للمجموعة المرجعية (Simbolon, Supriyati & Nage, 2019, 1540).

نوعاً أداء البند التفاضلي:

يُصنّف الأداء التفاضلي للبند في نوعين رئيسيين هما: الأداء التفاضلي المنتظم، والأداء التفاضلي غير المنتظم. ينتج الأداء التفاضلي المنتظم عندما لا يحدث تفاعل بين مستوى القدرة، والمجموعة التي ينتمي إليها الفرد (الشريفين، 2018، 606)، وفيه يكون الأداء على البند لصالح مجموعة معينة على حساب المجموعة الأخرى عند جميع مستويات القدرة؛ أي تكون نسبة الذين يجيبون إجابة صحيحة على البند أعلى في كل مستوى من مستويات القدرة (Basman, 2023, 145-146)، وتكون قيمة الفرق بين المجموعتين في احتمال الاستجابة الصحيحة للبند لها نفس الاتجاه والمقدار عبر مستويات القدرة المختلفة (هامبليتوت وميريندا وسبيلبيرغر، 2006، 80)، بينما ينتج الأداء التفاضلي غير المنتظم عندما يكون هناك تفاعل بين مستوى القدرة، وعضوية المجموعة (العمرى وشطناوي، 2016، 1533)؛ أي يكون البند أكثر صعوبة لإحدى المجموعتين مقارنة مع المجموعة الأخرى، بعد أن تتم مطابقة الأفراد على القدرة في أحد طرفي متصل القدرة، وأقل صعوبة لهذه المجموعة من الأفراد في الطرف الآخر من متصل القدرة (Brussow, 2018, 6). وبعبارة أخرى، تكون الفروق في احتمالية الاستجابة على البند مختلفة وغير متماثلة في جميع مستويات القدرة (Almarabeh & Alshammari).

87, 2020, جاد الله، 2024، 14)؛ أي أن احتمالية الإجابة الصحيحة على البند تكون أكبر لإحدى المجموعتين عند مستوى معين من القدرة، وأكبر للمجموعة الأخرى عند مستوى آخر للقدرة (النفيعي، 2023، 546).
طرائق الكشف عن أداء البند التفاضلي:

تُصنّف طرائق الكشف عن أداء البند التفاضلي وفقاً لنظرية القياس الكلاسيكية التي تأخذ بالاعتبار الدرجات الملاحظة، ونظرية الاستجابة للبند التي تأخذ بالاعتبار مجموعة السمات الكامنة (Basman, 2023, 146).
تعد طريقتا مانتل هانزل (MH)، والانحدار اللوجستي (LR) المُستخدَمَتين في البحث الحالي من الطرائق الكلاسيكية، وسنتحدث عنهما بالتفصيل فيما يأتي:

طريقة مانتل هانزل (MH) – Mantel-Haenszel:

تعد طريقة مانتل هانزل (MH) من أشهر الطرائق وأكثرها استخداماً من أجل الكشف عن أداء البند التفاضلي (DIF) (Zhang, 2015, 1)، إذ تقوم بدراسة التحيز بين مجموعتين إحداهما تسمى المجموعة المرجعية، وتسمى الأخرى المجموعة البؤرية أو المُستهدَفة التي تتأثر بتحيز البنود (مبارك ومبارك، 2013، 243).
تختبر طريقة مانتل هانزل (MH) الفرض الصفري الذي ينص على أنه لا توجد فروق في الأداء على البند تُعزى لعضوية المجموعة (مرجعية – مستهدفة) عند كل مستوى من مستويات القدرة (الأحمدي، 2020، 290)، ويتم تطبيقها من خلال تنظيم جدول احتمال أو مصفوفة ثنائية البعد (2*2) لعضوية المجموعة (مرجعية و مستهدفة)، والإجابة على البند (صح وخطأ) وفق ما يأتي:

الجدول رقم (1) : مصفوفة ثنائية البعد (2*2) لتوزع الإجابات (الصحيحة والخاطئة)

وفقاً لنوع المجموعة (مرجعية ومستهدفة) عند مستوى القدرة (j)

المجموع	الإجابة على البند		المجموعة
	خاطئة (0)	صحيحة (1)	
N_{rj}	B_j	A_j	المجموعة المرجعية (R)
N_{fj}	D_j	C_j	المجموعة المستهدفة (F)
T_j	N_{0j}	N_{1j}	المجموع

(Ehis & Judith, 2022, 148–149)

ثم تُستخدَم إحصائية مانتل هانزل كاي مربع (MH_{X^2}) لاختبار الفرضية، وهي:

$$= \frac{(1 \sum_{j=1}^s (A_j - E(A_j)) - 0.5)^2}{\sum_{j=1}^s Var(A_j)} MH_{X^2}$$

حيث:

$E(A_j)$: القيمة المتوقعة لعدد أفراد المجموعة المرجعية الذين أجابوا على البند إجابة صحيحة، وتُعطى بالعلاقة:

$$E(A_j) = \frac{N_{rj} N_{1j}}{T_j}$$

$$Var(A_j): \text{هو تباين } (A_j), \text{ ويُعطى بالعلاقة: } Var(A_j) = \frac{N_{rj} N_{fj} N_{1j} N_{0j}}{T_j}$$

(Liu & Bradley, 2021, 3)

0.5: هو عامل تصحيح الاستمرارية لتحسين التقريب لتوزيع كاي مربع، وهو ضروري من أجل التكرارات الصغيرة (Magis, 850, 2010, Beland, Tuerlinck, & De Boeck), إذ أن قيمة MH_{X^2} تتبع توزيع كاي مربع مع درجة حرية تساوي الواحد.

وتُستخدم نسبة الأرجحية المشتركة (a_{MH}) للدلالة على وجود أداء تفاضلي، وفق العلاقة:

$$= \frac{\sum A_j D_j / T_j}{\sum B_j C_j / T_j} a_{MH}$$

فإذا كانت ($a_{MH} = 1$)، دل ذلك على عدم وجود أداء تفاضلي للبند، لأن احتمالات الإجابة الصحيحة لأفراد المجموعة المستهدفة هي نفس احتمالات الإجابة الصحيحة لأفراد المجموعة المرجعية (De Ayala, 2009, 329–330)، بينما إذا كانت ($a_{MH} > 1$) فإن البند يُظهر أداء تفاضلياً لصالح المجموعة المرجعية، وإذا كانت ($a_{MH} < 1$) فإن البند يُظهر أداء تفاضلياً لصالح المجموعة المستهدفة (Khoeruroh & Retnawati, 2020, 3).

ويتم تحويل نسبة الأرجحية المشتركة (a_{MH}) إلى مقياس دلتا وفق العلاقة:

$$= -2.35 \ln(a_{MH}) \Delta_{MH}$$

وبناء على قيمة (Δ_{MH}) يتم تحديد حجم تأثير الأداء التفاضلي اعتماداً على الفئات المذكورة من قبل مركز خدمة الاختبارات التعليمية (ETS)، وهي: فئة البنود ذات الأداء التفاضلي الصغير عندما تكون ($|\Delta_{MH}| < 1$)، وفئة البنود ذات الأداء التفاضلي المتوسط أو المعتدل عندما تكون ($1 < |\Delta_{MH}| < 1.5$)، وفئة البنود ذات الأداء التفاضلي الكبير عندما تكون ($|\Delta_{MH}| > 1.5$) (Liu & Bradley, 2021, 3)، (الأحمدي، 2020، 291–292).

طريقة الانحدار اللوجستي (Logistic Regression (LR):

يُعدّ الانحدار اللوجستي طريقة بارامترية تكشف عن نوعي الأداء التفاضلي (المنتظم وغير المنتظم) (Zhang, 2009, 1)، وتُستخدم من أجل نمذجة متغير ثنائي من واحد، أو أكثر من المتغيرات التي يمكن أن تكون كمية، أو نوعية (De Ayala, 2009, 331). والمتغير الثنائي التابع هو درجة البند (0، 1)، بينما المتغيرات المستقلة هي: الدرجة الكلية للملاحظة للاختبار، وعضوية المجموعة، والتفاعل بين الدرجة الكلية وعضوية المجموعة (نسبة الاحتمال) (Basman, 2023, 147).

يتم التنبؤ باحتمال الإجابة الصحيحة وفقاً للمعادلة العامة للانحدار اللوجستي التي تُعطى بالعلاقة:

$$(1) \quad P(u=1) = \frac{e^z}{1+e^z}$$

ويُعبر عن (z) بالعلاقة:

$$z = \beta_0 + \beta_1 \theta \quad (2)$$

حيث:

θ : قدرة الفرد (الدرجة الكلية)، و (β_0) بارامتر التقاطع (معامل الانحدار الثابت)، و (β_1) بارامتر الميل (التمييز) (Bastug, 2016, 1255).

تمثل المعادلة (2) التي تتضمن فقط الدرجة الكلية للتنبؤ باحتمال الإجابة الصحيحة النموذج الأول في الانحدار اللوجستي الذي يدل على عدم وجود أداء تفاضلي (Christine, 2014, 20)، وبإدخال متغير عضوية المجموعة (g) للتنبؤ باحتمال الإجابة الصحيحة على البند إلى هذا النموذج، نحصل على النموذج الثاني للانحدار اللوجستي، وتصبح المعادلة كما يأتي:

$$z = \beta_0 + \beta_1 \theta + \beta_2 g \quad (3)$$

حيث:

β_2 : معامل الانحدار المتعلق بعضوية المجموعة، ويتوافق مع متوسط فروق المجموعة في الأداء على البند، ويعبر عنه بالعلاقة ($\beta_2 = \beta_{01} - \beta_{02}$). ويشير (1، 2) إلى المجموعتين المُقارنتين (De Ayala, 2009, 333, Gilmore, 2014, 32)، وبإدخال متغير التفاعل بين الدرجة الكلية وعضوية المجموعة ($g \theta$) إلى النموذج الثاني، نحصل على النموذج الثالث للانحدار اللوجستي، وتصبح المعادلة:

$$z = \beta_0 + \beta_1\theta + \beta_2g + \beta_3(g\theta) \quad (4)$$

حيث:

β_3 : معامل الانحدار المتعلق بالتفاعل بين الدرجة الكلية وعضوية المجموعة، ويُعبّر عنه بالعلاقة $(\beta_3 = \beta_{11} - \beta_{12})$ (Bastug, 2016, 1255).

يتم الكشف عن الأداء التفاضلي بمقارنة النماذج الثلاثة السابقة (1-2-3) مع بعضها البعض، ويتم اختبار ثلاث فرضيات، الأولى للكشف عن الأداء التفاضلي المنتظم، وتتص على أن تأثير متغير عضوية المجموعة على احتمال الإجابة الصحيحة للبند يساوي الصفر؛ أي $(\beta_2 = 0H_0)$ (Atar, 2007, 10)، وتُختبر من خلال مقارنة النموذجين (2) و(3) مع درجة حرية واحدة، فإذا كان $(\neq 0\beta_2)$ و $(= 0\beta_3)$ فهذا يدل على أن البند يُظهر أداءً تفاضلياً منتظماً (Lee, 2016, 2)، بينما الفرضية الثانية التي تتص على أن تأثير متغير التفاعل بين الدرجة الكلية وعضوية المجموعة على احتمال الإجابة الصحيحة يساوي الصفر؛ أي $(\beta_3 = 0H_0)$ (Atar, 2007, 10)، هي للكشف عن الأداء التفاضلي غير المنتظم، وتُختبر من خلال مقارنة النموذجين (3) و(4) مع درجة حرية واحدة، فإذا كانت $(\neq 0\beta_3)$ بغض النظر عما إذا كانت (β_2) تساوي، أو لا تساوي الصفر، فهذا يدل على أن البند يعرض أداءً تفاضلياً غير منتظم (Abedalaziz, 2010, 408)، وتتص الفرضية الثالثة التي تُستخدم للكشف عن نوعي الأداء التفاضلي (المنتظم، وغير المنتظم) معاً، على أن تأثير متغير عضوية المجموعة ومتغير التفاعل بين الدرجة الكلية وعضوية المجموعة على احتمال الإجابة الصحيحة للبند يساوي الصفر؛ أي $(= 0\beta_2)$ و $(\neq 0\beta_3)$ (Jodoin & Gierl, 2001, 332)، تُختبر من خلال مقارنة النموذجين (2) و(4) مع درجتين حرية، فإذا كانت $(\neq 0\beta_2)$ و $(\neq 0\beta_3)$ ، فهذا يدل على وجود أداء تفاضلي للبند (Penfield & Camilli, 2007, 141).

ولتحديد اتجاه تفضيل البند ذي الأداء التفاضلي المنتظم، تتم مقارنة تقدير بارامتر عضوية المجموعة (β_2) بالصفر، فإذا كان $(0 < \beta_2)$ فهذا يدل على أن البند يُفضّل المجموعة المرجعية، وإذا كان $(0 > \beta_2)$ فهذا يدل على أن البند يُفضّل المجموعة البؤرية (Jodoin & Gierl, 2001, 332)، بينما تحديد اتجاه تفضيل البند ذي الأداء التفاضلي غير المنتظم، يكون بمقارنة تقدير بارامتر التفاعل بين الدرجة الكلية وعضوية المجموعة (β_3) بالصفر، فإذا كان $(0 < \beta_3)$ فهذا يدل على أن البند يُفضّل الأفراد ذوي القدرة المرتفعة في المجموعة المرجعية، والأفراد ذوي القدرة المنخفضة في المجموعة البؤرية. وإذا كان $(0 > \beta_3)$ فهذا يدل على أن البند يُفضّل الأفراد ذوي القدرة المنخفضة في المجموعة المرجعية، والأفراد ذوي القدرة المرتفعة في المجموعة البؤرية (Cuevas & Cervantes, 2012, 47, Simbolon & et al, 2019, 1543).

وحجم التأثير الأكثر شيوفاً في طريقة الانحدار اللوجستي هو (ΔR^2) للفرق بين قيم (R^2) (Nagelkerk, 2004)، ويُعبّر عنه بالعلاقة:

$$\Delta R^2 = R^2_1 - R^2_2$$

حيث: (R^2_1) و (R^2_2) تمثل قيم التباين في استجابة البند المرتبطة بالنماذج التي تتم مقارنتها (Finch & French, 2019, 290)، ويتم الحصول على قيمة (R^2) لكل نموذج عن طريق حساب تقدير المربعات الصغرى الموزونة لمتغيرات (الدرجة الكلية، وعضوية المجموعة، والتفاعل بين الدرجة الكلية وعضوية المجموعة) (Ukanda, Othun, Agak & Oleche, 2019, 879).

ويفسر حجم التأثير وفقاً للمبادئ التوجيهية التي حددها جودوين وجيرال (Jodoin & Gierl, 2001)، فعندما تكون $(\Delta R^2 < 0.035)$ تكون البنود ذات أداء تفاضلي صغير (A)، وعندما تكون $(0.035 \leq \Delta R^2 \leq 0.070)$ تكون البنود ذات أداء تفاضلي متوسط (B)، وعندما تكون $(\Delta R^2 > 0.070)$ تكون البنود ذات أداء تفاضلي كبير (C) (Whiteman, Van de Water, 2014, 36, 2011).

8- منهج البحث:

تم اتباع المنهج الوصفي التحليلي الذي يعتمد على تجميع الحقائق والمعلومات، ثم مقارنتها وتفسيرها وتحليلها للوصول إلى تعميمات صحيحة ومقبولة، ويعرّف بأنه المنهج الذي يعتمد على دراسة الظاهرة كما توجد في الواقع، ويهتم بوصفها وصفاً دقيقاً، ويعرّف عنها كفيماً وكمياً (عبيدات، 1993، 219).

9- مجتمع البحث وعينته:

مجتمع البحث: تكون مجتمع البحث من جميع طلبة كليتي الاقتصاد والتربية في محافظة طرطوس والمنتظمين على مقاعد الدراسة في الفصل الدراسي الثاني من العام الدراسي (2018-2019) والبالغ عددهم (2750) طالباً وطالبة وفق إحصائيات كليتي التربية والاقتصاد بطرطوس.

عينة البحث: تم تطبيق المقياس على عينة مؤلفة من (700) طالباً وطالبة تم سحبهم بطريقة عشوائية من طلبة السنوات الدراسية الأربعة في كليتي التربية والاقتصاد في محافظة طرطوس، وقد تم استبعاد (14) ورقة إجابة لأن أصحابها لم يلتزموا بالإجابة على بنود المقياس، وبالتالي تكونت عينة الدراسة من (686) بنسبة سحب بلغت (24.94%) من المجتمع الأصلي، منهم (410) إناث (المجموعة المرجعية) ، و(276) ذكور (المجموعة المستهدفة)، والجدول الآتي يبين توزيع عينة البحث تبعاً لمتغير الجنس.

الجدول رقم (2): توزيع أفراد عينة البحث حسب متغير الجنس

المجموع	الجنس		الكلية
	إناث	ذكور	
321	254	67	التربية
365	156	209	الاقتصاد
686	410	276	المجموع

10- حدود البحث:

- الحدود البشرية: تم تطبيق أداة البحث على عينة من طلبة كليتي التربية والاقتصاد المنتظمين على مقاعد الدراسة للعام الدراسي (2018-2019) في جامعة طرطوس.
- الحدود المكانية: تم تطبيق أداة البحث في كليتي التربية والاقتصاد في جامعة طرطوس.
- الحدود الزمانية: تم تطبيق أداة البحث الفصل الدراسي الثاني من العام الدراسي (2018-2019).
- الحدود الموضوعية: تتمثل الحدود الموضوعية بالمقارنة بين طريقتي مانتل هانزل والانحدار اللوجستي في الكشف عن الأداء التفاضلي للبنود.

11- أداة البحث: مقياس الرهاب الاجتماعي الذي أعده رولين ووي (Roulin & Wee, 1994)، ودرست خصائصه السيكومترية في الجمهورية العربية السورية (حسن، 2017) بعد أن قامت بتطبيقه على عينة من طلبة التعليم الثانوي في محافظة طرطوس.

وصف مقياس الرهاب الاجتماعي إعداد رولين ووي (Roulin & Wee, 1994):

يتكون مقياس الرهاب الاجتماعي من (36) عبارة، ويوجد أمام كل عبارة بديلين هما (نعم، لا) ويطلب من المفحوص أن يختار واحداً منهما بما يتفق مع مشاعره وتصرفاته في مواقف الحياة المختلفة، ويصلح هذا المقياس للتطبيق على الأفراد بدءاً من عمر (12) سنة إلى (90) سنة.

ويصحح المقياس بإعطاء درجة واحدة إذا كانت الإجابة " نعم"، وصفرًا إذا كانت الإجابة " لا"، مع الملاحظة أن بعض العبارات تصحح في الاتجاه العكسي، وتتراوح الدرجة الكلية بين (0-36)، والدرجة المرتفعة تشير إلى أن الفرد يعاني من الرهاب الاجتماعي والعكس صحيح.

توافرت للمقياس دلائل جيدة للصدق من خلال فحص صدق المحتوى، والصدق البنوي بطريقة التحليل العاملي وطريقة الاتساق الداخلي، إذ تراوحت قيم معاملات ارتباط البنود بالدرجة الكلية بين (0.22-0.68)، وكانت جميعها موجبة ودالة إحصائياً عند مستوى دلالة (0.01)، والصدق المحكي بدلالة محك مقياس الخجل، إذ بلغت قيمة معامل الارتباط بين درجات الأفراد على المقياسين (0.823)، ومحك المجموعات الطرفية، كذلك كانت للمقياس معاملات ثبات مرتفعة، حيث بلغت قيمة معامل الثبات بالإعادة (0.91)، وبالتجزئة النصفية باستخدام معادلة سبيرمان براون (0.92)، وبطريقة ألفا كرونباخ (0.85).

12- نتائج البحث ومناقشتها:

للإجابة عن سؤالَي البحث الأول والثاني:

1- ما بنود مقياس الرهاب الاجتماعي التي تُظهر أداء تفاضلياً باستخدام طريقي مانتل هانزل (MH) والانحدار اللوجستي (LR)؟

2- هل اتفقت طريقتا مانتل هانزل والانحدار اللوجستي في تحديد بنود مقياس الرهاب الاجتماعي ذات الأداء التفاضلي؟ وهل حددت إحدهما عدداً أكبر من البنود مقارنة بالأخرى؟

استُخدمت حزمتي (difLogistic – difMH) من حزمة difR، وتم حساب قيم كاي مربع مانتل هانزل وقيم إحصاء نسبة الأرجحية للانحدار اللوجستي، واعتماد مستوى دلالة (0.05) مع درجة القطع (3.8415) لتحديد البنود ذات الأداء التفاضلي، حيث يُعتبر البند الذي تكون قيمة كاي مربع مانتل هانزل وقيمة إحصاء الانحدار اللوجستي له، أعلى من درجة القطع، والقيمة الاحتمالية أقل من مستوى الدلالة، ذا أداء تفاضلي، والجدول الآتي يوضح النتائج.

الجدول رقم (3): نتائج الكشف عن البنود ذات الأداء التفاضلي باستخدام طريقتي مانتل هانزل والانحدار اللوجستي

نتائج طريقة الانحدار اللوجستي			البنود	نتائج طريقة مانتل هانزل			البنود
الدلالة الإحصائية	القيمة الاحتمالية p	قيمة إحصاء نسبة الأرجحية للانحدار اللوجستي		الدلالة	القيمة الاحتمالية p	كاي مربع مانتل هانزل MH_{χ^2}	
دال	**0.0017	9.8605	1	غير دال	0.6807	0.1694	1
دال	**0.0021	9.4529	2	غير دال	0.5927	0.2862	2
غير دال	0.0581	3.5921	3	غير دال	0.4671	0.5288	3
غير دال	0.1147	2.4888	4	غير دال	0.4097	0.6798	4
دال	*0.0213	5.7728	5	دال	**0.0013	10.3244	5
دال	*0.0444	5.2999	6	غير دال	0.6574	0.1967	6
دال	*0.0354	4.4268	7	دال	*0.0407	4.1874	7
غير دال	0.8935	0.0179	8	غير دال	0.6137	0.2549	8
دال	**0.0072	7.2261	9	غير دال	0.3293	0.9516	9
غير دال	0.6818	0.1681	10	غير دال	0.3942	0.7261	10
غير دال	0.3526	0.8642	11	غير دال	0.4298	0.6234	11
دال	*0.0204	5.3777	12	دال	***0.0008	11.3338	12
دال	*0.0257	4.9773	13	غير دال	0.6193	0.2468	13
غير دال	0.9307	0.0076	14	غير دال	0.5534	0.3512	14
دال	***0.0003	13.3854	15	دال	**0.0016	9.9672	15
غير دال	0.7434	0.1072	16	غير دال	0.6177	0.2491	16
غير دال	0.7307	0.1185	17	غير دال	0.4294	0.6244	17
غير دال	0.1646	1.9317	18	غير دال	0.5488	0.3595	18
غير دال	0.8296	0.0463	19	غير دال	0.8998	0.0159	19
دال	**0.0038	8.3883	20	دال	**0.0037	8.4034	20
غير دال	0.1181	2.4423	21	غير دال	0.2291	1.4467	21
غير دال	0.0588	3.5718	22	دال	**0.0058	7.5978	22
غير دال	0.8700	0.0268	23	غير دال	0.9389	0.0059	23
غير دال	0.7021	0.1463	24	غير دال	0.9863	0.0003	24
غير دال	0.2629	1.2535	25	غير دال	0.0877	2.9169	25
غير دال	0.1492	2.0803	26	غير دال	0.5437	0.3687	26
غير دال	0.2163	1.5284	27	غير دال	0.9842	0.0004	27
غير دال	0.3028	1.0620	28	غير دال	0.8786	0.0233	28
غير دال	0.3006	1.0716	29	غير دال	0.4653	0.5331	29
دال	***0.0002	14.2682	30	غير دال	0.0914	2.8495	30
دال	*0.0310	4.6541	31	غير دال	0.7716	0.0843	31
غير دال	0.3741	0.7901	32	دال	*0.0472	3.9382	32
غير دال	0.8636	0.0295	33	غير دال	0.1822	1.7795	33
غير دال	0.2595	1.2714	34	غير دال	0.1681	1.9	34
غير دال	0.8560	0.0329	35	غير دال	0.1654	1.9323	35
غير دال	0.6440	0.2136	36	غير دال	0.3831	0.7607	36

*** دال عند (0.001) " " دال عند (0.01) " " دال عند (0.05)

يظهر الجدول السابق أن طريقة مانتل هانزل حددت البنود (5، 7، 12، 15، 20، 22، 32) على أنها ذات أداء تفاضلي؛ أي كشفت طريقة مانتل هانزل عن (7) بنود ذات أداء تفاضلي من أصل (36) بنوداً لمقياس الرهاب الاجتماعي، أي ما نسبته (19.44%)، بينما حددت طريقة الانحدار اللوجستي البنود (1، 2، 5، 6، 7، 9، 12، 13، 15، 20، 30، 31) على أنها ذات أداء تفاضلي؛ أي كشفت طريقة الانحدار اللوجستي عن (12) بنوداً ذات أداء تفاضلي من أصل (36) بنوداً لمقياس الرهاب الاجتماعي، أي ما نسبته (33.33%).

ويظهر أيضاً من الجدول أن الطريقتين اتفقتا في تحديد خمسة بنود هي (5، 7، 12، 15، 20)، وأن طريقة مانتل هانزل حددت بندين لم تحدهما طريقة الانحدار اللوجستي هما (22، 32)، بينما حددت الأخيرة سبعة بنود لم تحدها الأولى هي (1، 2، 6، 9، 13، 30، 31).

ووفقاً لما سبق تكون طريقة الانحدار اللوجستي قد حددت عدداً أكبر من البنود ذات الأداء التفاضلي مقارنة بطريقة مانتل هانزل، وكانت أفضل منها في ذلك. تتفق هذه النتيجة مع دراسة (Abbott, 2021)، وجزئياً مع دراسة (حسن، 2020)، ومع دراسة (Alahmadi, 2019)، ودراسة (Simbolon, Supriyati & Naga, 2019)، ودراسة (Moralez–Oetiz & Pena–Cardenas, 2019)، ودراسة (Ukanda, Othuon, Aggak & Oleche, 2019)، وجزئياً مع دراسة (Terzi & Yakar, 2018)، ودراسة (Cokluk, Gul & Gul, 2016) والتي توصلت كل منها إلى أن طريقة الانحدار اللوجستي كانت أفضل في الكشف عن البنود ذات الأداء التفاضلي، وحددت عدداً أكبر من هذه البنود مقارنة بطريقة مانتل هانزل، بينما تختلف هذه النتيجة مع دراسة (Villas, 2019)، ودراسة (Bastug, 2016) اللتين توصلتا إلى أن طريقة مانتل هانزل كانت أفضل من طريقة الانحدار اللوجستي.

وبالنسبة لنتيجة اتفاق الطريقتين في تحديد بعض البنود، واختلافهما في تحديد بنود أخرى، فإنها تتفق مع ما توصلت إليه دراسة (حسن، 2020) جزئياً، ودراسة (Alahmadi, 2019)، ودراسة (Moralez–Oetiz & Pena–Cardenas, 2019)، ودراسة (Terzi & Yakar, 2018) وبتختلف مع دراسة (حسن، 2020) جزئياً، ودراسة (Simbolon, 2019)، ودراسة (Supriyati & Naga, 2019)، ودراسة (Cokluk, Gul & Gul, 2016) التي توصلت كل منها إلى أن طريقة الانحدار حددت نفس البنود التي حددتها مانتل هانزل مع بنود أخرى لم تحدها الأخيرة، كما وتختلف مع دراسة (Villas, 2019) التي توصلت إلى أن طريقة مانتل هانزل حددت نفس البنود التي حددتها طريقة الانحدار اللوجستي، مع بند إضافي عليها. الإجابة عن سؤال البحث الثالث: " كيف توزعت أحجام تأثير بنود مقياس الرهاب الاجتماعي التي تم تحديدها على أنها ذات أداء تفاضلي باستخدام كل من الطريقتين؟ وأي الطريقتين كانت أفضل بناء على أحجام التأثير التي حددتها؟"

كانت أيضاً بناء على نتائج الطريقتين باستخدام حزمتي (difMH – difLogistic) من حزمة difR، وبتحديد قيم حجم التأثير وفقاً لطريقة مانتل هانزل (Δ_{MH}) للبنود التي تم تحديدها على أنها ذات أداء تفاضلي، ومقارنتها بالمعايير التي حددتها خدمة الاختبارات التربوية (ETS)، والتي تم ذكرها في الإطار النظري للبحث، وكذلك قيم حجم التأثير وفقاً لطريقة الانحدار اللوجستي (ΔR^2) للبنود التي تم تحديدها على أنها ذات أداء تفاضلي، ومقارنتها بالمعايير التي وضعها جودوين وجيرال (Jodoin & Gierl, 2001)، والتي ذُكرت أيضاً في الإطار النظري للبحث. والجدول الآتي يوضح النتائج.

الجدول رقم (4): قيم أحجام تأثير البنود ذات الأداء التفاضلي ونوعها وفقاً لطريقتي مانتل هانزل والانحدار اللوجستي

البنود	نتائج طريقة مانتل هانزل		نتائج طريقة الانحدار اللوجستي	
	قيمة حجم التأثير (Δ_{MH})	نوع حجم التأثير	قيمة حجم التأثير (ΔR^2)	نوع حجم التأثير
1	ليس له أداء تفاضلي		0.0194	A
2	ليس له أداء تفاضلي		0.0180	A
5	1.4477	B	0.0123	A
6	ليس له أداء تفاضلي		0.0092	A
7	1.1245-	B	0.0117	A
9	ليس له أداء تفاضلي		0.0123	A
12	1.5109	C	0.0092	A
13	ليس له أداء تفاضلي		0.0075	A
15	1.5375	C	0.0295	A
20	1.7156	C	0.0201	A
22	1.3453	B	ليس له أداء تفاضلي	
30	ليس له أداء تفاضلي		0.0244	A
31	ليس له أداء تفاضلي		0.0106	A
32	1.1026-	B	ليس له أداء تفاضلي	

يُظهر الجدول السابق أن أحجام تأثير الأداء التفاضلي للبنود التي حددتها طريقة مانتل هانزل تراوحت بين المتوسطة (المعتدلة) والكبيرة، إذ كانت أربعة بنود وهي (5، 7، 22، 32) ذات أداء تفاضلي معتدل (B)، بينما كانت البنود الثلاث المتبقية وهي (12، 15، 20) ذات أداء تفاضلي كبير (C)، بينما كانت أحجام تأثير الأداء التفاضلي للبنود التي حددتها طريقة الانحدار اللوجستي جميعها من النوع الصغير (A)، ووفقاً لذلك تكون طريقة مانتل هانزل أفضل من طريقة الانحدار اللوجستي لأنها حددت أحجام تأثير معتدلة وكبيرة.

تتفق هذه النتيجة فيما يتعلق بطريقة الانحدار اللوجستي جزئياً مع دراسة (حسن، 2020) لاختبار القدرات المعرفية وعدد من المقاييس العيادية لاختبار مينسوتا، ودراسة (Alahmadi, 2019)، ودراسة (Terzi & Yakar, 2018) والتي توصلت إلى أن أحجام تأثير البنود التي حددتها هذه الطريقة جميعها صغيرة، وتختلف مع نتائج هذه الدراسات فيما يتعلق بطريقة مانتل هانزل، كما وتختلف هذه النتيجة مع نتائج دراسة (Villas, 2019)، ودراسة (Ukanda, Othuon, Aggak & Oleche, 2019)، ودراسة (Cokluk, Gul & Gul, 2016) فيما يتعلق بالطريقتين معاً.

الإجابة عن سؤال البحث الرابع: "كيف توزعت تفضيلات هذه البنود التي تم تحديدها في كل من الطريقتين بين المجموعتين المقارنتين (الذكور والإناث)؟ وهل كانت تلك التفضيلات متشابهة في البنود المشتركة؟"

كانت أيضاً بناء على نتائج الطريقتين باستخدام حزمتي (difLogistic – difMH) من حزمة difR، بتحديد قيم نسب الأرجحية مانتل هانزل (a_{MH}) للبنود التي تم تحديدها على أنها ذات أداء تفاضلي ومقارنتها بالعدد (1)، فإذا كانت تلك القيم أصغر من (1) فإن البند يفضل الذكور (المجموعة البؤرية)، وإذا كانت أكبر من (1) فإن البند يفضل الإناث (المجموعة المرجعية)، وبالنظر إلى قيمة بارامتر نموذج الانحدار (الخاص بالمجموعة) في طريقة الانحدار اللوجستي، فإذا كانت تلك

القيمة موجبة، فإن البند يفصل الإناث (المجموعة المرجعية)، بينما إذا القيمة سالبة، فإن البند يفضل الذكور (المجموعة البؤرية). والجدول الآتي يوضح النتائج.

الجدول رقم (5): تفضيلات البنود ذات الأداء التفاضلي وفقاً لطريقتي مانتل هانزل والانحدار اللوجستي

البنود	نتائج طريقة مانتل هانزل		نتائج طريقة الانحدار اللوجستي	
	نسبة الأرجحية (a_{MH})	المجموعة التي يفضلها البند	قيمة بارامتر نموذج الانحدار (الخاص بالمجموعة)	المجموعة التي يفضلها البند
1	ليس له أداء تفاضلي		موجبة	الإناث
2	ليس له أداء تفاضلي		موجبة	الإناث
5	0.5401	الذكور	سالبة	الذكور
6	ليس له أداء تفاضلي		موجبة	الإناث
7	1.6137	الإناث	موجبة	الإناث
9	ليس له أداء تفاضلي		موجبة	الإناث
12	0.5257	الذكور	سالبة	الذكور
13	ليس له أداء تفاضلي		موجبة	الإناث
15	0.5198	الذكور	سالبة	الذكور
20	0.4819	الذكور	سالبة	الذكور
22	0.5641	الذكور	ليس له أداء تفاضلي	
30	ليس له أداء تفاضلي		موجبة	الإناث
31	ليس له أداء تفاضلي		سالبة	الذكور
32	1.5987	الإناث	ليس له أداء تفاضلي	

يُظهر الجدول أن تفضيلات البنود التي تم تحديدها على أنها ذات أداء تفاضلي وفقاً لطريقتي مانتل هانزل والانحدار اللوجستي، توزعت بين الذكور والإناث. وفقاً لطريقة مانتل هانزل، كانت البنود (5، 12، 15، 20، 22) ذات أداء تفاضلي لصالح الذكور (المجموعة البؤرية)، بينما كان للبندين (7، 32) أداء تفاضلي لصالح الإناث (المجموعة المرجعية)، وبحسب طريقة الانحدار اللوجستي، فضلت البنود (1-2-6-7-9-13-30) الإناث (المجموعة المرجعية)، بينما فضلت البنود (5-12-15-20-31) الذكور (المجموعة البؤرية). وكانت تفضيلات البنود التي تم تحديدها بشكل مشترك بين الطريقتين على أنها ذات أداء تفاضلي لها نفس التفضيلات، حيث فصل البند (7) الإناث، بينما فضلت البنود (5-12-15-20) الذكور. تتفق هذه النتيجة فيما يتعلق بتوزع تفضيلات البنود ذات الأداء التفاضلي بين الذكور والإناث مع دراسة (حسن، 2020)، ودراسة (Bastug, 2016)، ومع دراسة (Villas, 2019) فيما يتعلق بطريقة الانحدار اللوجستي فقط، وفيما يتعلق بنتيجة تشابه توزع البنود التي تم تحديدها على أنها ذات أداء تفاضلي باستخدام الطريقتين معاً، فإنها تتشابه مع دراسة (حسن، 2020) جزئياً، ومع دراسة (Villas, 2019) جزئياً، ودراسة (Terzi & Yakar, 2018).

13- مقترحات البحث: بناء على النتائج التي تم التوصل إليها تم اقتراح ما يأتي:

- 1- دراسة الأداء التفاضلي لمقياس الرهاب الاجتماعي مع طرائق تستند إلى نظرية الاستجابة للبند ومقارنتها مع طريقتي مانتل هانزل والانحدار اللوجستي.
- 2- دراسة أثر حجم العينة على الكشف عن الأداء التفاضلي لمقياس الرهاب الاجتماعي باستخدام طريقتي مانتل هانزل والانحدار اللوجستي.
- 3- دراسة الأداء التفاضلي لمقياس الرهاب الاجتماعي في ضوء متغيرات أخرى غير الجنس باستخدام طريقتي مانتل هانزل والانحدار اللوجستي.
- 4- إجراء دراسة مقارنة لأثر حذف البنود ذات الأداء التفاضلي من مقياس الرهاب الاجتماعي التي حُدّت وفقاً لطريقتي مانتل هانزل والانحدار اللوجستي على خصائصه السيكمترية.

قائمة المراجع

المراجع العربية:

- أرشيدات، بسام كامل. (2007). تحزّي التحيز في اختبار مادة التربية الوطنية لدى طلبة جامعة مؤتة. رسالة ماجستير غير منشورة. جامعة مؤتة، الأردن.
- الأحمدى، شرف. (2020). مقارنة طريقتي مانتل-هانزل ومؤشر قياس المساحة الإشارية لراجو في الكشف عن الأداء التفاضلي لمفردات استبانة مؤشر أساليب التعلم. مجلة البحث العلمي في التربية، (21)، 280-321.
- الشريفين، نضال كمال. (2018). أثر نوع الأداء التفاضلي للفقرات على الخصائص السيكمترية للفقرات والاختبار وفق النماذج المعلمية والنماذج اللامعلمية لنظرية الاستجابة للفقرات. دراسات العلوم التربوية، 45(4)، 605-632.
- الصمادي، عبد الله عبد الغفور ، والزهراني، فهد شرف. (2014). درجاتهم متساوية فهل هم متشابهون: دراسة في ظاهرة تفاضل الأداء الفردي. مجلة كلية الدراسات العليا، جامعة الخليج العربي، 1-17.
- العبد القادر، كوثر والسوالمه، يوسف. (2021). الأداء التفاضلي لفقرات ومموهات اختبار للتفكير الناقد لطلبة المرحلة الجامعية تبعاً لمتغير الجنس. المجلة الأردنية في العلوم التربوية، 17(2)، 251-269.
- العجلوني، جهاد والسوالمه، يوسف. (2019). العلاقة بين الأداء التفاضلي للمموهات والأداء التفاضلي للفقرات في اختبار رياضيات من نوع الاختيار من متعدد. المجلة الأردنية في العلوم التربوية، 15(1)، 49-63.
- العمري، حسان وشطناوي، عبد الله. (2016). الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم في الرياضيات للصف العاشر تبعاً لمتغير الجنس. مجلة جامعة النجاح للأبحاث(العلوم الإنسانية)، 30(8)، 1530-1554.
- النفيعي، عبد الرحمن عبد الله. (2023). الكشف عن القدرة التنبؤية لمقياس تقدير طيف التوحد وعن الأداء التفاضلي لعباراته. مجلة العلوم التربوية والدراسات الإنسانية، 34(3)، 532-566.
- جاد الله، أشرف سعد محمد. (2024). الأداء التفاضلي لمفردات اختبار ستيرنبرج للقدرات الثلاثية في ضوء نظرية الاستجابة للمفردة والتحليل العملي التوكيدي متعدد المجموعات لدى طلاب الجامعة. مجلة التربية، (201)، 1-63.
- حسن، ريم. (2017). دراسة الخصائص السيكمترية لمقياس الرهاب الاجتماعي لدى عينة من طلبة التعليم الثانوي في محافظة طرطوس. مجلة جامعة دمشق للعلوم التربوية والنفسية.
- حسن، ريم. (2020). الكشف عن دالة البند التفاضلي (DIF) في بناء الاختبارات النفسية. رسالة دكتوراه غير منشورة. جامعة دمشق، سورية.

- خليل، محمد المري محمد إسماعيل وإسماعيل، ميمي السيد أحمد. (2021). الأداء التفاضلي لاختبار تحصيلي إلكتروني في مقرر القياس والتقويم لدى طلبة الفرقة الثالثة بكلية التربية بالجامعة باستخدام نموذج التقدير الجزئي. *المجلة التربوية في جامعة سوهاج*، 2(87)، 662-702.
- سلمان، مفلح حسين عيال وظاظا، حيدر إبراهيم. (2022). الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم في مباحث العلوم والرياضيات واللغتين العربية والإنجليزية للصف الثامن الأساسي في الأردن. *مجلة جرش للبحوث والدراسات*، 23(12)، 2669-2696.
- عبيدات، الذوقان وآخرون. (1993). *البحث العلمي، مفهومه، أدواته وأساليبه*. الرياض: دار أسامة للنشر.
- مبارك، مأمون و مبارك، وائل. (2013). الأداء التفاضلي لاختبار الرياضيات في الدراسة الدولية بيزا 2006. *مجلة العلوم التربوية*، 25(2)، 241-261.
- هامليتوت، رونالد وميريندا، بيتر وسبيلبيرغر، تشارلز. (2006). *تكييف الاختبارات التربوية والنفسية للتقييم عبر الثقافات*. (ترجمة: هالة برمدا). المملكة العربية السعودية: مطبعة العبيكان.
- Abbott, W.T. (2021). DIF Detection Sensitive of Lord's Chi-square, Raju's Area, Logistic Regression, Mantel-Haenszel, Standardization and Transformed item Difficulties Methods, In Comparison, Using R. *EPRA International Journal of Multidisciplinary Research (IJMR)*, 7(7), 629-640.
- Abdalaziz, N. (2010). Detecting Gender Related DIF using Logistic Regression and Mantel-Haenszel Approach. *Procedia Social and Behavioral Sciences*, (7), 406-413.
- Al marabheh, A.J., & Alshammari, S.R. (2020). Detection of Sex-Related Differential Item Functioning in Raven's Standard Progressive Matrices Test Using the Mantel-Haenszel Method. *Journal of Education Practice*, 11(29), 86-94.
- Alahmadi, M.T.S. (2019). *Evaluating Possible Bias in a Giftedness Assessment: An Empirical Comparison of IRT and Non-IRT (Parametric and Non-Parametric) Techniques for Detecting Differential Item Functioning*. Doctoral Dissertation. Morgan State University.
- Allahyari, E. (2020). Differential Item Functioning Analysis of the Bar-On Emotional Quotient Inventory (EQ-I) Across Gender and Age Groups: An Iterative Hybrid Ordinal Logistic Regression. *Psychologia*, 14(2), 95-105.
- Apinyapibal, S., Lawthong, N., & Kanjanawasee, S. (2015). A comparison Analysis of the Efficacy of Differential Item Functioning Detection for Dichomously Scored Items among Logistic Regression, SIBTEST, and Rasch Tree Methods. *Procedia-Social and Behavioral Sciences*, 191, 21-25.
- Asamoah, N.A.B. (2020). *Assessing Differential Item Functioning in the Perceived Stress Scale*. Master Thesis, University of Ghana, Arkanasa.
- Atar, B. (2007). *Differential Item Functioning analyses for Mixed Response Data Using IRT Likelihood-Ratio Test, Logistic Regression, and Gllamm Procedures*. Doctoral Dissertation. Florida State University.

- Basman, M. (2023). A Comparison of the efficacies of differential item functioning detection methods. *International Journal of Assessment Tools in Education*, 10(1), 145–159.
- Bastuy, O.Y.O.(2016). A comparison of Four Differential Item Functioning Procedures in the Presence of Multidimensionality. *Academic Journals–Educational Research and Reviews*, 11(13), 1251–1261.
- Brussow, J.A.(2018). *Finding Item–Level Causes of Differential Item Functioning: A Hierarchical Explaining DIF*. Doctoral Dissertation. University of Kansas.
- Camilli, G., & Shepard, L.A.(1994). *Methods of Identifying Biased Test Items*. New Delhi: SAGE.
- Christine, Y.M. (2014). *Differential Item Functioning (DIF) in Personality Testing: Cultural differences and Translation imperfection*. Master Thesis. University of Rotterdam
- Cokluk, O., Gul, E., & Gul, C.D. (2016). Examining Differential Item Functioning of Different Item ordered Test forms According to Item Difficulty Levels. *Educational Sciences in Theory and Practice*, 16(1), 319-330.
- Cuevas, M., & Cervantes, V. (2012). Differential Item Functioning Detection with Logistic Regression. *Mathematics and Science*, 3. 45–59.
- De Ayala, R.J.(2009). *The Theory and Practice of Item Response Theory*. New York: The Guilford PRESS.
- Duncan, S.C. (2006). *Improving The prediction of Differential Item Functioning: A Comparison of the Use of an Effect Size for Logistic Regression DIF and Mantel–Haenszel DIF Methods*. Doctoral Dissertation. The University of Texas A&M.
- Ehis, A.K., & Judith, O.H. (2022). Non–parametric Method of Detecting Differential Item Functioning in Senior School Certificate Examination (SSCE) 2019 Economics Multiple Choice Items. *Kashere Journal of Education*, 3(1), 146–158.
- Elyan, R.M., & Aljodeh, M.M. (2024). The Effectiveness of Mantel Haenzel Log Odds Ratio Method in Detecting Differential Item Functioning Across Different Sample Size and Test Lengths Using Real Data Analysis. *Educational Sciences*, 51(3), 37– 46.
- Embertson, S.E., & Reise, S. P. (2000). *Item Response Theory for Psychologists*. London: Lawrence Erlbaum Associates. Inc.
- Finch, W.H., & French, B.F.(2019). *Educational and Psychological Measurement*. New York: Taylor & Francis.
- Gilmore, W.E.(2014). *A Differential Item Functioning analysis of the New Mexico English Language Proficiency Assessment*. Doctoral Dissertation. The University of New Mexico.

- Jodoin, M.M., & Gierl, M.J. (2001). Evaluating Type I Error and Power Rates using an Effect Size Measure with the Logistic Regression Procedure for DIF Detection. *Applied Measurement in Education*, 14(3), 329-349.
- Khalid, M.N., Shafiq, F., & Ahmed, Sh. (2021). Detection of Differential Item Functioning Using Mantel–Haenszel, Standardization Proportion and BILOG–MG Procedure. *Global Educational Studies Review*, 1(3), 71–78.
- Khoeruroh, U., & Retnawati, H. (2020). Comparison Sensitivity of the Differential Item Function (DIF) detection method. *International Conference on Science Education and Technology*, 1–8.
- Kurnaz, F.B., & Yildiz, H. (2023). Investigating the Sources of differential item functioning: A sample critical thinking motivation scale. *International Journal of Assessment Tools in Education*, 10(3), 434– 453.
- Lee, J.J. (2016). Detecting Differential Item Functioning Using The Logistic Regression Procedure in Small Samples. *Applied Psychological Measurement*, 41(1), 1-14.
- Lee, S., & Suh, Y. (2018). Lord’s Wald Test for Detection DIF in Multidimensional IRT Models: A Comparison of Two Estimation Approaches. *Journal of Educational Measurement*, 55(2), 328-353.
- Liu, M.(2017). *Differential Item Functioning in Large–Scale Mathematics Assessments Comparing the Capabilities of yhe Rasch Trees Model to Traditional Approaches*. Doctoral Dissertation. The University of Toledo.
- Liu, R., & Bradley, K.D. (2021). Differential Item Functioning among English Language Learners on a Larg–Scale Mathematics Assessment. *Studies and Evaluation Faculty Publications, Frontiers in Psychology*, 12, 1–12.
- Macinnes, J.M.W. (2009). *The Mantel–Haenszel Method for Detecting Differential Item Functioning in Dichotomously Scored Items: A Multilevel Approach*. Doctoral Dissertation. University of Florida.
- Magis,D., Beland, S.,Tuerlinck, F., & De Boeck, P.(2010). A general Framework and an R Package for the detection of dicgomous Differential Item Functioning. *Behavior Research Methods Psychonomic*, 42(3), 847–862.
- McCallum, R.S. (2003). *Handbook of Nonverable Assessment(2nd)*. New York: Springer.
- Morales–Ortiz. M .,& Pena–Cardenas,F.(2019). DIF in the Spanish Version of the Verable Selective Reminding Test Using Samples from Hispanics in the United State, Mexicans, and Spaniards. *Frontiers in Psychology*, 10, 1–8.

- Orlawence, G.W., & Queensoap, M. (2019). Using Item Characteristics Curve Approach in Detecting Item Bias in a Chemistry Achievement Test in Nigeria. *Case Studies Journal ISSN*, 8(3), 79-103.
- Paek, I., & Wilso, M. (2011). Formulating the Rasch Differential Item Functioning Model Under the Marginal Maximum Likelihood Estimation Context and its Comparison with Mantel–Haenszel Procedure in short test and small sample conditions. *Educational and Psychological Measurement*, 71(16), 1023-1046.
- Penfield, R.D. (2001). Assessing Differential Item Functioning Among Multiple Groups: A Comparison of Three Mantel–Haenszel Procedures. *Applied Measurement in Education*, 14(3), 235-259.
- Penfield, R.D., & Camilli, G. (2007). Differential Item Functioning and Item Bias. *Handbook of Statistics*, 26, 125-169.
- Raykov, T., Marcoulides, G.A., Lee, Ch., & Chang, Ch. (2013). Studying Differential Item Functioning via Latent Variabl Modeling: A Note on a Multiple–Testing Procedure. *Educational and Psychological Measurement*, 73(5), 898-908.
- Salabayba, T. M. (2013). Differential Item Functioning Detection in Reading Comprehension Test Using Mantel–Haenszel, Item Response Theory, and Logical Data Analysis. *The International Journal of Social Sciences*. 14(1). 76–82.
- Sandilands,D.A.(2014). *Accuracy of Differential Item Functioning Methods in Structurally Missing Data du to BOOKLET Design*. Doctoral Dissertation. University of British, Colombia.
- Simbolon, K., Supriyati, Y., & Nage, D.S.(2019). Sensitivity of Differential Item Functioning Detection Method on National Mathematics Examination in North Sumatera Province, Indonesia. *International Journal of Engineering and advanced Technology*,8(5), 1538–1549.
- Teresi, J.A & Fleishman, J.A. (2007). Differential Item Functioning and health assessment. *Qual Life Res*, 16, 33-42.
- Teresi, J.A., Ramirez, M., Lai, J., & Silver, S. (2008). Occurrences and Sources of Differential Item Functioning (DIF) in patient–reported outcome measures: Description of DIF methods and review of measures of depression, quality of life and general health. *Psychology Science Quarterly*, 50(4). 538-612.
- Terzi, R., & Yakar, L. (2018). Differential Item and Differential Distractor Functioning Analyses on Turkish High School Entrance Exam. *Journal of Measurement and Evaluation and Psychology*, 9(2), 136–149.

- Ukanda, F., Othuon, L., Agak, J., & Oleche, P. (2019). Effectiveness of Mantel–Haenszel and Logistic Regression Statistics in Detecting Differential Item Functioning Under Different Conditions of Sample Size, Ability Distribution and Test Length. *American Journal of Educational Research*, 7(11), 878-887.
- Van De Water, E.(2014). *A meta-analysis of Type I error rates for detecting Differential Item Functioning with Logistic Regression and Mantel–Haenszel in Monte Carlo studies*. Doctoral Dissertation. Georgia State University.
- Villas, J.P. (2019). Differential Item Functioning in Grade 8 Math Using Logistic Regression, Mantel– Haenszel and Logical Data Analysis. *Journal of Education and Practice*,10(12), 138-146.
- Whiteman, H.J.(2011). *Differential Item Functioning and adverse Impact: A comparison of Mantel– Haenszel and Logistic Regression*. Master Thesis. California State University.
- Wiberg, M. (2007). *Measuring and Detecting Differential Item Functioning in Criterion–Referenced Licensing Test: A Theoretical Comparison of Methods*. UMEM University.
- Zhang, M. (2009). *Gender Related Differential Item Functioning in Mathematics Tests: A meta-analysis*. Master Thesis. Washington State University.
- Zhang, Y.(2015). Multiple Ways to detect Differential Item Functioning in SAS. *Education Testing Service*, 1–9.