

2022

## The Differential Performance of the Items of the National Test to Control the Quality of Education in the Subjects of Science, Mathematics, Arabic and English for the Eighth Grade in Jordan

Mifleh Salman  
Mifleh79@yahoo.com

Haider Thatha  
dr.haidarzaza@gmail.com

Follow this and additional works at: <https://digitalcommons.aaru.edu.jo/jpu>



Part of the [Arts and Humanities Commons](#), and the [Social and Behavioral Sciences Commons](#)

### Recommended Citation

Salman, Mifleh and Thatha, Haider (2022) "The Differential Performance of the Items of the National Test to Control the Quality of Education in the Subjects of Science, Mathematics, Arabic and English for the Eighth Grade in Jordan," *Jerash for Research and Studies Journal* *مجلة جرش للبحوث والدراسات*: Vol. 23: Iss. 2, Article 26.

Available at: <https://digitalcommons.aaru.edu.jo/jpu/vol23/iss2/26>

This Article is brought to you for free and open access by Arab Journals Platform. It has been accepted for inclusion in *Jerash for Research and Studies Journal* *مجلة جرش للبحوث والدراسات* by an authorized editor. The journal is hosted on [Digital Commons](#), an Elsevier platform. For more information, please contact [rakan@aarj.edu.jo](mailto:rakan@aarj.edu.jo), [marah@aarj.edu.jo](mailto:marah@aarj.edu.jo), [u.murad@aarj.edu.jo](mailto:u.murad@aarj.edu.jo).

## الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم في مباحث العلوم والرياضيات واللغتين العربية والإنجليزية للصف الثامن الأساسي في الأردن

مفلح حسين عيال سلمان\* وحيدر إبراهيم ظاظا\*\*

### ملخص

هدفت الدراسة الحالية إلى الكشف عن الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم للصف الثامن الأساسي في مدارس المملكة الأردنية الهاشمية الحكومية تبعاً لمتغيري: الجنس. وشكل الصورة الاختبارية. خلال الدورة الاختبارية 2019م باستخدام طريقة مانتل - هانزل المعممة. توفرت بيانات الدراسة من استجابات طلبة الصف الثامن الأساسي والبالغ عددهم (10000) على (158) فقرة اختبارية في مباحث العلوم والرياضيات واللغة العربية واللغة الإنجليزية. وأظهرت النتائج أن فقرات مبحث الرياضيات قد أبدت أعلى نسبة في الأداء التفاضلي. حيث بلغت (77.5%) منها (17) فقرة تبعاً لمتغير شكل الصورة الاختبارية. و(28) فقرة تبعاً لمتغير الجنس ولصالح الإناث. يلي ذلك مبحث اللغة الإنجليزية. حيث بلغت (77.77%) منها (17) فقرة تبعاً لمتغير شكل الصورة الاختبارية. و(24) فقرة تبعاً لمتغير الجنس ولصالح الإناث. أما مبحث العلوم، فقد بلغت (71.42%) منها (12) فقرة تبعاً لمتغير شكل الصورة الاختبارية. و(28) فقرة تبعاً لمتغير الجنس ولصالح الإناث. وأخيراً مبحث اللغة العربية. حيث بلغت (60%) منها (3) فقرة تبعاً لمتغير الجنس. و(11) فقرة تبعاً لمتغير شكل الصورة الاختبارية ولصالح الشكل الإلكتروني. وقد انتهت الدراسة إلى مجموعة من التوصيات ومن ذلك: رفع قدرات العاملين في إعداد الاختبارات الوطنية في مجال صياغة وإعداد فقرات اختبارية ذات مواصفات سيكومترية ذات جودة عالية وتخلو من أي من مصادر الأداء التفاضلي لفقرات. وإجراء دراسة تتناول الأداء التفاضلي لفقرات امتحان الثانوية العامة في مباحث العلوم والرياضيات واللغتين العربية والإنجليزية: لأهمية هذا الامتحان والقرارات المترتبة على نتائجه.

الكلمات المفتاحية: الأداء التفاضلي للفقرة. الاختبار الوطني لضبط نوعية التعليم. مانتل - هانزل المعممة.

## The Differential Performance of the Items of the National Test to Control the Quality of Education in the Subjects of Science, Mathematics, Arabic and English for the Eighth Grade in Jordan.

Mifleh H. Salman, MOE, Jordan.

Haider I. Thatha, The University of Jordan - College of Educational Sciences, Jordan.

© جميع الحقوق محفوظة جامعة جرش 2022.

\* وزارة التربية والتعليم / الأردن. Email: [Mifleh79@yahoo.com](mailto:Mifleh79@yahoo.com)

\*\* الجامعة الأردنية - كلية العلوم التربوية، الأردن. Email: [dr.haidarzaza@gmail.com](mailto:dr.haidarzaza@gmail.com)

## Abstract

The current study aimed to detect the differential item functioning (DIF) of the items of the Jordanian National Test to control educational quality for the 8th grade in the government schools of Jordan according to gender, model of test presentation in 2019 using Generalized Mantel – Haenszel. The data of the study was available from the responses of the 8th grade students, totaling (10,000), to (158) test items in science, mathematics, Arabic and English subjects. The results showed that the mathematics subjects showed the highest percentage of (DIF) Where (77.5%) of them amounted to (17) items according to the model of test presentation, and (28) items according to the gender variable in favor of females. This is followed by the English language subject, where (77.77%) of them amounted to (17) items according to the model of test presentation, and (24) items according to the gender variable in favor of females. As for the science subject, it amounted to (71.42%) of them (12) items according to the model of test presentation, and (28) items according to the gender variable in favor of females. And finally: the Arabic language stage, where (60%) of them amounted to (3) items according to the gender variable, and (11) items according to the test format and in favor of the electronic format. The study concluded with a set of recommendations and proposals, including: raising the capabilities of workers preparing national tests in the field of drafting and preparing test items with psychometric specifications of high quality and devoid of any (DIF) sources, and conducting a study dealing with the (DIF) of high school exam paragraphs in science subjects mathematics, Arabic and English; Because of the importance of this exam and the decisions resulting from its results.

**Keywords:** Differential item functioning, The national educational quality control test, Generalized Mantel – Haenszel.

## خلفية الدراسة

أولى المختصون في مجال القياس النفسي والتربوي قضية عدالة الاختبار، Fairness Test، بالغ الاهتمام والتقدير: نظراً لأهمية القرارات التي يتم اتخاذها استناداً إلى نتائج الاختبارات. سواء في اختيار الأفراد. أو تصنيفهم تبعاً لقدراتهم أو خصيلهم. أو توجيه الطلبة مهنيًا. أو تحديد الصفوف المناسبة لهم. إذ ينصف الاختبار بالعدالة عندما يكون صادقاً لجميع الأفراد والمجموعات: ويُعطي فرصاً متساوية لجميع المفحوصين لإبداء المعرفة والمهارات المكتسبة المتعلقة بغرض الاختبار (Roever, 2005).

ويرى وايبيرج (Wiberg, 2007) أن ضمان عدالة الاختبار من التحديات الرئيسية في استخدامه. حيث تُعد العدالة جزءاً من صدق الاختبار وتدعم المساواة في التعامل مع المتقدمين

له. والاتساق في معاني الدرجات للجميع: وتحقق العدالة عندما يكون هناك دالة تناظر بين الدرجات المتحققة للطلبة وقدراتهم الكامنة.

ولتحقيق عدالة الاختبار يؤكد وايبيرج (Wiberg, 2007) الحاجة إلى مجموعة من الإجراءات المعيارية في بناء الاختبارات. فالاختبارات الجيدة يتم التخطيط لها وتطوير فقراتها وتدقيقها وتحليلها بأساليب مختلفة تسمح بانتقاء أفضل الفقرات وتضمينها في النسخة النهائية من الاختبار لضمان الحصول على فقرات فاعلة لا غموض فيها. فالفقرات ذات الخصائص السيكومترية الضعيفة في الاختبار تؤثر سلباً في خصائصه السيكومترية؛ وعليه فإن تحديد مثل هذه الفقرات الغامضة من خلال تحليل الفقرات له أهمية كبيرة في مراجعة الاختبار وتطويره.

ويُعدُّ فحص الأداء التفاضلي للفقرة (DIF) Differential Item Functioning جزءاً مهماً في تحليل الفقرات. إذ إن خلو الاختبار وفقراته وموهاته من الأداء التفاضلي من أهم الشروط التي ينبغي توافرها في الاختبار قبل نشره. وقد وضعت الجمعية الأمريكية للبحث التربوي (AERA) (American Educational Research Association AERA, 1999) الأداء التفاضلي ضمن قواعد النشر للاختبارات. وأشارت إلى أنه إذا استخدمت الاختبارات في اتخاذ القرارات فلا بد من الحصول على دالة تجريبية واضحة من أجل دعم الادعاءات بأن فقرات الاختبار لا تعكس أداء تفاضلياً للفقرة عبر المجموعات المختلفة من المفحوصين. كما وضعت الجمعية الأمريكية لعلم النفس (APA) (American Psychology Association, 1988) الأداء التفاضلي أيضاً ضمن قواعد النشر. ونصت على أنه يجب على مطوري الاختبارات أن يشيروا إلى طبيعة الأدلة المتعلقة بملاءمة الاختبار للمجموعات ذات الخلفيات العرقية أو اللغوية المختلفة (النوافلة, 2017).

وقد أكد المجلس الوطني للقياس في التربية The National Council on Measurement in Education (NCME, 2014) على أن الأداء التفاضلي للفقرة هو دليل على صدق البنية الداخلية للاختبار.

ولمعرفة الأداء التفاضلي للفقرة، فلا بد من المرور على تعريفه اصطلاحاً. إذ يُشير مصطلح الأداء التفاضلي للفقرة إلى اختلافات في الأداء على الفقرة عبر مجموعات مختلفة من الأفراد بعد مزاجية تلك المجموعات فيما يتعلق بالسمة التي يفترض أن تقيسها الفقرة والاختبار الذي تظهر فيه (Dorans & Holland, 1994). حيث إن الكشف عن الفقرات ذات الأداء التفاضلي لا يكفي من أجل إدعاء عدم العدالة في الاختبار؛ فعدالة الاختبار تعتمد على كيفية استخدامه. وتحيز الاختبار هو فرق منتظم في الدرجة الكلية للاختبار ضد مجموعة معينة (Camilli, 2006).

وقد أشار هيرمان وديبو (Herrmann & DeBoe, 2011) أن مصمم الاختبار لا يريد ظهور الأداء التفاضلي في أي من فقرات اختباره؛ لذلك يحاول عند كتابته لهذه الفقرات أن يميز بين المفحوصين من مستويات القدرة المختلفة. أي يريد الوصول إلى الوضع المثالي. حيث إن احتمال إجابة المفحوص بشكل صحيح عن الفقرة يعتمد على مستواه في السمة التي يتم قياسها. وبذلك يحصل الطلبة ذو القدرة نفسها على الدرجة نفسها.

إن مفهوم أداء الفقرة التفاضلي مرادف للتحيز الإحصائي Statistical Bias المتمثل في إعطاء تقدير أعلى أو أدنى من القيمة الفعلية لمعلمة Parameter أو أكثر في النموذج الإحصائي Camilli (2006) وهو خاصية إحصائية للفقرة تختلف عن تحيز الفقرة؛ فالفقرة المتحيزة تظهر أداءً تفاضلياً لكن الأداء التفاضلي لا يكفي لأن تكون الفقرة متحيزة. فمفهوم تحيز الفقرة أعم ويكمن في التفسير (Camilli & Shepard, 1994).

وكون الأداء التفاضلي للفقرة مفهوم إحصائي، فإنه يُستدل عليه من دالة مشتقة إحصائياً للتعبير عن الفروق في احتمالات الاستجابة الصحيحة للفقرة بين مجموعتين لهما مستوى القدرة نفسه (Grujter & Kamp, 2005). ويتم الكشف عنه من خلال فحص العلاقة بين الاحتمالات المشروطة للاستجابة الصحيحة على الفقرة. وبين السمة الكامنة المراد قياسها. حيث يتم مقارنة منحنيات خصائص الفقرة بين المجموعتين اللتين يتم رسمهما وفق النموذج المستخدم معلّمي أم لا معلّمي (Gybles, 2004).

ونظراً لاحتمالية أن يهدّد وجود الأداء التفاضلي لل فقرات صدق الاختبار، فالكشف عنه وإزالته يُعدّ ضرورياً للقياس الصادق والمتحرر من التحيز في الفقرات (العجلوني والسوالمه، 2016) ومن هنا ظهرت العديد من الطرق الإحصائية للكشف عن الأداء التفاضلي، منها ما يعتمد على نظرية القياس الكلاسيكية (طريقة الصعوبة المحولة للفقرة، وطريقة مؤشر التمييز، وطريقة قليل التباين، وطريقة تحليل الموهبات)، ومنها ما يعتمد على نظرية الاستجابة للفقرة (IRT) Item Response Theory (مقارنة منحنيات خصائص الفقرة، ومقارنة المعالم المقدرة للفقرة، ومقارنة مطابقة نماذج استجابة الفقرة للبيانات في المجموعات المختلفة)، ومنها ما يعتمد على التحليل العاملي وعلى الإحصائي مربع كاي. إلا أن أكثر الطرق شيوعاً ضمن هذه المجموعة هي: طريقة مانتل - هانزل Mantel-Haenszel، وطريقتي مربع كاي chi-square statistic وللإستجابات الصحيحة، ومربع كاي للإستجابات الصحيحة والخطأ، وطريقة الانحدار اللوجستي Logistic regression methods (علام، 2005).

وبالنظر لكل طريقة من الطرق السابقة نجد أن لها إيجابياتها وسلبياتها وأسسها الرياضية في حساب الأداء التفاضلي لل فقرات والاختبار. ولعل الطريقة المستخدمة في هذه الدراسة وهي طريقة مانتل - هانزل المعممة، وهي من الطرق اللامعلمية الأكثر شهرة وتعامل مع فئات الفقرات كبيانات أسمية بينما يتعامل اختبار مانتل مع الطبيعية الرتبة لفئات الفقرات متعددة الاستجابات. ويتم استخدامها عندما تكون متغيرات العامل والفقرة ضمن المقاييس التصنيفية؛ الأسمية أو الرتبة. فالبيانات الناجمة يمكن تلخيصها في الجداول المخصصة (Q) بالأبعاد (R×C).

وتُعدّ طريقة مانتل هانزل من أكثر طرق النظرية التقليدية انتشاراً في الكشف عن الأداء التفاضلي الفقرة وذلك لسهولة حساباتها وإجراءاتها وتقوم على مقارنة متوسطات المجموعات في الأداء على الاختبار ككل أو على فقراته المتكافئة وذلك باستقصاء التحيز بين مجموعتين أحدهما تُسمى: المجموعة المرجعية والأخرى تُسمى: المجموعة المستهدفة وهي المجموعة التي تتأثر بتحيز الفقرات. والفرضية الصفرية هي أنه ليس هناك ارتباط شرطي بين علامة الفقرة وعضوية المجموعة (Fidalgo & Madeira, 2008).

وتتعامل طريقة مانتل - هانزل المعممة (Generalized Mantel - Haenszel) مع فئات الفقرات كبيانات أسمية بينما يتعامل اختبار مانتل مع الطبيعية الترتيبية لفئات الفقرة في الفقرات متعددة الاستجابات. ويتم استخدامها عندما تكون متغيرات العامل والفقرة ضمن المقاييس التصنيفية: الأسمية أو الترتيبية. فالبيانات الناتجة يمكن تلخيصها في الجداول المخصصة (Q) بالأبعاد  $(R \times C)$ . والفرضية الصفرية التي يتم اختبارها هي عدم وجود ارتباط. مما يشير إلى أن توزيع متغير الاستجابة عشوائي على مستويات العامل. وفي الفقرات متعددة الاستجابات تقدم أكثر من استجابتين. ومن أهم أمثلتها فقرات مقياس ليكرت.

وتُعدُّ مراقبة مستوى مخرجات النظام التربوي وتقييمه من المهام الضرورية لدفع عجلة التنمية في المجتمعات. حيث تُوفر نتائج الدراسات التقويمية الكبيرة والشاملة والمستمرة بيانات موضوعية تتعلق بالكثير من المؤشرات التربوية ولا سيما إتقان الطلبة للكفايات التعليمية الرئيسة التي يسعى النظام التربوي إلى تحقيقها لدى الطلبة. وفي ظل المتابعة والتقييم المتعلق بمؤشرات الجودة ومعاييرها أصبح من الضروري تنفيذ اختبارات وطنية تكشف عن مدى جودة التعليم ونوعيته. ليصار إلى توظيف نتائجها في التخطيط وبرامج التطوير والتحديث. ومن هذه المنطلق شرعت إدارة الامتحانات والاختبارات في وزارة التربية والتعليم في الأردن منذ العام 2000م بإعداد الاختبار الوطني لضبط نوعية التعليم وجودته وإعداد تقارير بنتائجه حيث تعقد لجنة التخطيط في وزارة التربية والتعليم اجتماعاً سنوياً لمناقشته والاستفادة من توصياته؛ وتزويد المعنيين بنسخ منه؛ ليصار إلى توظيفها في التخطيط على جميع المستويات (وزارة التربية والتعليم، 2016).

ويُحلُّ الاختبار الوطني لضبط نوعية التعليم مرجعاً مهماً يمكن الاستفادة من نتائجه في تحديد جوانب القوة وجوانب الضعف في أداء الطلبة تمهيداً لوضع الحلول المناسبة ومعالجتها. واتخاذ القرارات التربوية حيثما يلزم ذلك. ويُعدُّ هذا الاختبار وثيقة يستفاد من نتائجها في تحسين تعلم الطلبة. ويتم تطبيق هذا الاختبار في صفوف المرحلة الأساسية: الرابع والثامن والعاشر في مدارس المملكة الأردنية الهاشمية الحكومية والخاصة. والمدارس التابعة لوكالة الغوث الدولية. والمدارس التابعة لمديرية التعليم والثقافة العسكرية (وزارة التربية والتعليم، 2019).

ونظراً لأهمية الاختبار الوطني لضبط نوعية التعليم في تقديم تغذية راجعة لوضعي السياسات التربوية والمنهجية في تطوير وتحسين تعلم الطلبة وتعليمهم. فإن الحاجة تبدو ماسة بأن يتصف هذا الاختبار بمواصفات الاختبار الجيد وأن يخلو من التحيزات المتعددة والأداء التفاضلي سواء في فقراته أو موهاته لا سيما أن وجود الأداء التفاضلي للفقرة مؤشر بأن هناك عوامل مرتبطة تنتمي إلى مجموعة معينة تؤثر في احتمالية الإجابة الصحيحة للفقرة. وبالتالي فإنها تهدد التقييم العادل بين الطلبة.

ومحلياً فقد تمت دراسة الاختبار الوطني لضبط نوعية التعليم وفقراته والأداء عليه من زوايا عدة مثل دراسة الطراونة (2011)، ودراسة البرصان (2013)، ودراسة فريحات (2014)، ودراسة العمري وشطناوي (2016)، ودراسة الحجاج (2019)، وأبو مسرة والسوالمه (2021).

وفي ضوء ما تقدم. يمكن القول أنَّ الاختبار الوطني ما زال بحاجة إلى المزيد من البحث لتمكين العاملين في تقييم أداء طلبة المرحلة الأساسية من ضبط نوعية التعليم؛ فالمتبع للجهود التي تبذلها وزارة التربية والتعليم في تطوير اختبارات ضبط جودة التعليم يجد أنها

جهود دؤبة ومنظمة تهدف إلى مواكبة الاتجاه العالمي في تمكين الطلبة من المهارات والمعارف والمعلومات في موضوعات دراسية معينة والتي تحاكي الاختبارات الدولية مثل اختبار القدرة القرائية PIRLS واختبار PISA. فالاختبارات الدولية والقرارات التي تبني على نتائجها بدأت تفرض على المدرسة أن تغير في أدوارها، وكذلك الحال بالنسبة للمناهج والمعلمين، والتقييمات التي يتبعها المعلمين في تقييم أداء الطلبة. فالاختبارات التي تُستخدم في ضبط جودة التعليم ونوعيته لم تعد تقتصر على تقييم خصيل الطلبة للمعارف المتضمنة بين دفتي الكتاب المدرسي، وإنما بدأت تربط تقييم تلك المهارات بالحياة والمواقف التي يتوقع من الطالب أن يصادفها في حياته العملية. والتأكد من أن الطلبة تعلموا وحققوا النتائج المتوقعة. وأنهم قادرين على الأداء وفق مستويات معرفية وأدائية تفرضها تلك الاختبارات الجديدة (البطوش والقرعان، 2018).

واستكمالاً لجهود وزارة التربية والتعليم في الأردن والباحثين الذين بدلوا جهوداً حثيثة لزيادة جودة التعليم في المرحلة الأساسية من خلال تطوير الاختبار الوطني لضبط نوعية التعليم في أربعة مباحث. كمساهمة فاعلة في متابعة جهود الباحثين السابقين واستكمالاً لجهود رفع سوية الاختبارات الوطنية التي تُعدها وزارة التربية والتعليم في الأردن.

#### مشكلة الدراسة

من الملاحظ أنه عند تطبيق الاختبارات بمختلف أنواعها تظهر نتائج متباينة لكل اختبار على المجموعات المستهدفة بذلك الاختبار؛ الأمر الذي يثير مجموعة من التساؤلات حول ما إذا كان هذا التباين في الدرجات يعود إلى طبيعة السمة المراد قياسها ومستوياتها؟ أم يعود إلى طبيعة الاختبار وخصائص فقراته؟ حيث إن بعض الفقرات قد تتضمن أداءً تفضلياً يجعله متحيزاً وهذا من شأنه أن يهدد صدق الاختبار وموثوقيته.

إن التباين في النتائج المتعلقة بالأداء على الاختبارات الدولية أو الوطنية ومن بينها الاختبارات التي تُعدها وزارة التربية والتعليم لضبط نوعية التعليم، الأمر والذي يُسهم في زيادة الجدل حول العوامل أو المتغيرات التي تُسهم في هذه التباينات. فبعض الدراسات والتقارير تفحص النتائج النهائية وتقرأها من منظور ختامي دون أن توضح الأسباب الكامنة وراء التباينات في النتائج. كما أن بعضها الآخر يولد الحيرة في تصنيفها لعدم القدرة على التحقق من كونها تعكس فروقاً حقيقية أم أنها مجرد تنميط متعلق ببعض المتغيرات. وهذا التناقض بين الدراسات يجعل الباب مفتوحاً لفحص التباين في الأداء على أحد أهم الاختبارات الوطنية التي تُعدها وزارة التربية والتعليم لضبط نوعية التعليم في أربعة مباحث. وفحص تباين الأداء عليه تبعاً لعدد من متغيرات خلال الدورة الاختبارية 2019 باستخدام طريقة مانتل هانزل المعممة في سياق تعزيز ضمانة عدالة الاختبار.

#### أسئلة الدراسة

سعت هذه الدراسة إلى توفير إجابة عن السؤال التالي: ما الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم في مباحث العلوم والرياضيات واللغة العربية واللغة الإنجليزية للصف الثامن الأساسي تبعاً لبعض المتغيرات خلال الدورة الاختبارية 2019م باستخدام طريقة مانتل - هانزل المعممة The Generalized Mantel-Haenszel.

ويتفرع عن هذا السؤال عدد من الأسئلة الفرعية، وهي على النحو الآتي:

1. ما الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم في مبحث اللغة العربية للصف الثامن الأساسي تبعاً لمتغيري: جنس الطالب، وشكل الصورة الاختبارية؟
2. ما الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم في مبحث اللغة الإنجليزية للصف الثامن الأساسي تبعاً لمتغيري: جنس الطالب، وشكل الصورة الاختبارية؟
3. ما الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم في مبحث الرياضيات للصف الثامن الأساسي تبعاً لمتغيري: جنس الطالب، وشكل الصورة الاختبارية؟
4. ما الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم في مبحث العلوم للصف الثامن الأساسي تبعاً لمتغيري: جنس الطالب، وشكل الصورة الاختبارية؟

#### أهداف الدراسة

تهدف هذه الدراسة إلى الكشف عن الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم في أربعة مباحث هي: العلوم، والرياضيات، واللغتين العربية والإنجليزية، للصف الثامن الأساسي، وذلك في ضوء بعض المتغيرات؛ لاعتبارات عدة تم التطرق إليها سابقاً.

#### أهمية الدراسة

تبرز أهمية الدراسة من الناحية النظرية في تسليط المزيد من الضوء على الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم وأسبابه. إذ من المتوقع أن الكشف عن الأداء التفاضلي لفقرات الاختبار الوطني سوف يُسهم في تحسين صدق الاختبار وعدالته، وسوف يساعد واضعي الاختبار على تحسين الخصائص السيكومترية للاختبار من خلال وجود عدد من المؤشرات والأدلة التي تحسّن من عملية صياغة الفقرات الجيدة، ومراجعة الأداء التفاضلي للفقرات؛ وبالتالي صدق الاختبار وعدالة تفسير نتائجه، وازدياد الثقة بها.

أما من الناحية العملية، فيؤمل من هذه الدراسة توفير بيانات إحصائية من خلال استخدام طريقة مانتل - هانزل المعممة في الكشف عن الأداء التفاضلي للاختبار الوطني لضبط نوعية التعليم في المباحث الأربعة وتبعاً لعدد من المتغيرات، والتي من شأنها أن تساعد العاملين في مجال بناء الفقرات على الكشف عن الأداء التفاضلي للفقرات، لاستبعاد الفقرات التي تُظهر أداءً تفاضلياً من فقرات الاختبار أو تعديلها.

#### مصطلحات الدراسة وتعريفاتها الإجرائية:

وردت في هذه الدراسة مجموعة من المفاهيم، وعُرفت إجرائياً على النحو الآتي:

1. الأداء التفاضلي للفقرة (Differential Item Functioning (DIF: مؤشر إحصائي للتعبير عن الفروق في الاستجابة عن الفروق في الاستجابة الصحيحة على الفقرة بين مجموعتين من المفحوصين ممن لهم المستوى نفسه من القدرة (Camilli & Shepard, 1994).
2. طريقة مانتل هانزل المعممة (The Generalized Mantel-Haenszel) إحدى طرق النظرية الكلاسيكية المستخدمة في الكشف عن الأداء التفاضلي للفقرات، حيث تقوم هذه الطريقة بالكشف عن الأداء التفاضلي بين مجموعتين فرعيتين من مجتمع الدراسة



إحدهما تُسمّى: المجموعة المرجعية وتُسمّى الأخرى المجموعة المستهدفة. وهي المجموعة التي يعتقد بوجود فقرات ذات أداء تفاضلي فيها (Fidalgo & Madeira, 2008).

3. الاختبار الوطني لضبط نوعية التعليم: اختبار سنوي تُعده وزارة التربية والتعليم مثله بمديرية الاختبارات في إدارة الامتحانات والاختبارات لقياس مدى امتلاك الطلبة في صفوف المرحلة الأساسية: الرابع، والثامن، والعاشر لمهارات التعلم الأساسية ومستويات أداء الطلبة بدلالة مؤشرات الأداء. ومدى امتلاكهم لمهارات اقتصاد المعرفة في أربعة مباحث: اللغة العربية واللغة الإنجليزية والرياضيات والعلوم؛ ويستهدف هذا الاختبار الصف نفسه بشكل دوري كل ثلاث سنوات (وزارة التربية والتعليم، 2013).

#### حدود الدراسة ومحدداتها

تحدد نتائج الدراسة في ما يأتي:

1. الحدود الزمنية: تقتصر الدراسة على الأداء التفاضلي في الاختبار الوطني لضبط نوعية التعليم الذي جرى خلال الفترة الدورية الاختبارية 2019م.
2. الحدود البشرية: تقتصر الدراسة على نتائج طلبة الصف الثامن الأساسي في مدارس المملكة الأردنية الهاشمية.

3. الحدود الموضوعية: تتناول الدراسة الأداء التفاضلي للاختبار لضبط نوعية التعليم

كما تتحدد نتائج هذه الدراسة بدلالات صدق وثبات الاختبار الوطني لضبط نوعية التعليم، وطبيعة المجتمع.

#### الدراسات السابقة

حظي موضوع الأداء التفاضلي باهتمام العديد من الباحثين. ويتكون هذا الجزء من البحوث والدراسات السابقة ذات الصلة بموضوع الدراسة الحالي والتي يُتوقع أن تبرز ملامح التطور في المسارات البحثية التي تنتمي إليها هذه الدراسة. وقد اقتصر الباحثان على الدراسات السابقة التي استهدفت الاختبار الوطني لضبط نوعية التعليم، حسب التسلسل الزمني بدءاً من الأقدم وانتهاءً بالأحدث؛ فقد أجرى الطراونة (2011) التي هدفت إلى الكشف عن الأداء التفاضلي في فقرات الاختبار الوطني لطلبة الصف الثامن في مادتي الرياضيات والعلوم لعام 2010 باستخدام طريقة مانتل - هانزل. حيث تكون الاختبار من (60) فقرة من نوع الاختبار من متعدد منها: (30) فقرة للرياضيات و(30) فقرة للعلوم طُبّق على (67455) طالباً وطالبة. حيث أظهرت النتائج وجود أداء تفاضلي في (9) فقرات لصالح الإناث في اختبار العلوم، و(4) فقرات لصالح الإناث في اختبار الرياضيات.

وأجرى البرصان (2013) دراسة هدفت إلى الكشف عن الأداء التفاضلي لمتغير الجنس لفقرات الاختبار الوطني لضبط نوعية التعليم لمادة الرياضيات للصف العاشر. وذلك لكل من المستويات الأربعة التي تصنف وزارة التربية والتعليم طلبتها فيها. وهي: المستوى المتقدم، والماهر، والماهر جزئياً، والمبتدئ؛ حيث بلغ عدد أفراد عينة الدراسة (43468) طالباً وطالبة اختبروا من طلبة الصف العاشر الأساسي في الأردن بواقع (21330) طالباً و(22138) طالبة. ولتحقيق

أهداف الدراسة تم الحصول على استجابات الطلبة على الاختبار الوطني لمادة الرياضيات المكون من (30) فقرة. وتحليلها لمستويات الطلبة الأربعة باستخدام طريقة مانتل - هانزل للكشف عن الأداء التفاضلي. وكشفت نتائج الدراسة عن الفقرات التي تظهر أداءً تفاضلياً في كل مستوى. وقد كان عدد الفقرات يزداد كلما أجهنا من المستوى الأعلى للمستوى الأدنى. حيث تراوحت نسبة الفقرات التي تظهر أداءً تفاضلياً بين (30%) للمستوى المتقدم و(67%) للمستوى المبتدئ.

كما قامت فريجات (2014) بدراسة هدفت للكشف عن الأداء التفاضلي في اختبائي العلوم والرياضيات في الاختبار الوطني لضبط نوعية التعليم وفقراتها تبعاً لطريقة تقديم الاختبار (ورقي. محوسب) ولتحقيق غرض الدراسة تم استخدام بيانات عينة مكونة من (2734) طالباً وطالبة من طلبة الصف العاشر على اختبائي العلوم والرياضيات والمكونين من (25) و(24) فقرة على التوالي. حيث تم استخدام مؤشر الأداء التفاضلي غير التعويضي كمؤشر للأداء التفاضلي للفقرة. ومؤشر الأداء التفاضلي التعويضي كمؤشر للأداء التفاضلي للفقرة من خلال ارتباطهما بالدرجة الكلية على الاختبار. وقد أظهرت النتائج وجود أداء تفاضلي على مستوى فقرات اختبائي الرياضيات تبعاً لشكل تقديم الاختبار بنسبة (8.33%) وبواقع فقرتين من مجموع الفقرات التي يتكون منها الاختبار. حيث أظهرت الفقرة (28) أداءً تفاضلياً منتظم لصالح الشكل المحوسب من الاختبار. في حين أظهرت الفقرة (13) أداءً تفاضلياً غير منتظم تارة لصالح الشكل الورقي؛ وتارة أخرى لصالح الشكل المحوسب من الاختبار. كما أظهرت وجود أداء تفاضلي على مستوى فقرات اختبائي العلوم تبعاً لشكل تقديم الاختبار بنسبة (12%) وبواقع ثلاث فقرات من مجموع فقرات الاختبار جميعها أظهرت أداءً تفاضلياً غير منتظم تارة لصالح الشكل الورقي؛ وتارة أخرى لصالح الشكل المحوسب من الاختبار.

وهدف دراسة العمري وشطناوي (2016) إلى الكشف عن الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم في مبحث الرياضيات للصف العاشر تبعاً لمتغير الجنس. وقد بلغت عينة الدراسة (2400) طالباً وطالبة بواقع (1200) طالب و(1200) طالبة ولتحقيق هدف الدراسة تم الحصول على البيانات المتعلقة باستجابات الطلبة على الاختبار المكون من (39) فقرة. وتم استخدام مؤشر الأداء التفاضلي اللاتعويضي (NCDIF) في الكشف عن الأداء التفاضلي للفقرات. حيث بلغ عدد الفقرات في صورتها النهائية بعد مطابقة الفقرات للنماذج المستخدمة (25) فقرة مطابقة للنموذج الأحادي المعلمة و(34) فقرة مطابقة للنموذج الثنائي المعلمة. وقد كشفت النتائج عن وجود أداء تفاضلي منتظم (9) فقرات من أصل (15) فقرة من أصل (26) فقرة هي عدد الفقرات المشتركة بين الطلاب والطالبات.

وهدف دراسة الحجاج (2019) إلى فحص الأداء التفاضلي للفقرات عبر المجموعات المتعددة لاختبار الرياضيات الوطني لضبط نوعية التعليم للعام الدراسي 2016/2017م؛ وفقاً لمتغير السلطة المشرفة على المدارس (حكومي. خاص. وكالة الغوث). وتقاطع متغيري الجنس وموقع المدرسة (مدينة. قرية) باستخدام طريقتين للكشف عن الأداء التفاضلي: مربع كاي للورد المعممة. ومانتل -هانزل المعممة. ولأغراض الدراسة تم استخدام استجابات (3000) طالباً وطالبة من الصف العاشر على (40) فقرة اختيار من متعدد. ولكل منها أربعة بدائل. حيث بلغ عدد الفقرات في صورتها النهائية بعد مطابقة الفقرات للنموذج ثلاثي المعلمة (32) فقرة. ووفقاً لتقاطع

متغيري الجنس وموقع المدرسة. كشفت طريقة مانتل - هانزل المعممة عن وجود فقرة واحدة أبدت أداءً تفاضلياً أولياً. وعند استخدام طريقة مربع كاي للورد المعممة أبدت (11) فقرة أبدت أداءً تفاضلياً أولياً.

وهدفت دراسة أبو مسرة والسوالمه (2021) إلى استخدام طريقة أشجار راش للكشف عن الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم في الرياضيات للصف الرابع الأساسي في الأردن. تبعاً للمتغيرات التي توفرها عملية التقييم. وهي: الجنس وموقع المدرسة والسلطة المشرفة على التعليم وطريقة تقديم الاختبار. وتكون مجتمع الدراسة من استجابات (31136) طالباً وطالبة على (25) فقرة من نوع الاختيار من متعدد تطابقت مع نموذج راش وبتطبيق الطريقة على استجابات عينة عشوائية قوامها (1500) من المفحوصين. بينت النتائج أن هناك أداءً تفاضلياً وفقاً لمتغير الجنس فقط: إذ تم الكشف عن (8) فقرات ذات أداء تفاضلي أي مانسبته (32%) من الفقرات. وكان الأداء التفاضلي لصالح الذكور في (5) فقرات ولصالح الإناث في (3) فقرات.

في ضوء استعراض الدراسات السابقة ذات العلاقة بموضوع الدراسة والمتغيرات التي تناولتها. يلاحظ أن جميع هذه الدراسات تشترك في تناول الأداء التفاضلي كإشكالية بحثية تستهدف الاختبار الوطني لضبط نوعية التعليم. وأهمية الكشف عنه كأحد الإجراءات الضرورية لضمان عدالة الاختبار وصدقه. كما استخدمت طرقاً مختلفة في الكشف عن الأداء التفاضلي. مثل: طريقة مانتل - هانزل وطريقة مانتل - هانزل المعممة. والانحدار اللوجستي وطريقة مربع كاي للورد المعممة (Generalized Lord's chi-square test). علماً بأن دراسة (الحجاج، 2019) استخدمت طريقة مانتل هانزل المعممة.

وقد تنوعت الدراسات من حيث المتغيرات التي تناولتها كمتغيرات تصنع الأداء التفاضلي. مثل: متغير الجنس. والمنطقة. وشكل الصورة الاختبارية. ومتغير درجة مخاطرة المفحوصين. وفيما يتعلق بالمباحث الدراسية المستهدفة. فقد اقتصر على مبحثي الرياضيات والعلوم. ومن الملاحظ أن أيّاً من الدراسات السابقة لم يتناول اللغة العربية واللغة الإنجليزية. وقد أفادت الدراسة الحالية من الدراسات السابقة في تحديد منهجية الدراسة. وفي بناء الإطار النظري. ومناقشة وتفسير النتائج لاحقاً.

#### الطريقة والإجراءات

يتضمن هذا الجزء وصفاً لمنهج الدراسة وبياناتها. وعرضاً لأدوات الدراسة. وإجراءات تطبيقها. وكذلك يتضمن وصفاً للطرق والمعالجات الإحصائية التي استخدمت في تحليل البيانات واستخلاص النتائج.

#### منهج الدراسة

تم استخدام المنهج الوصفي التحليلي للكشف عن الفقرات التي تظهر أداءً تفاضلياً في فقرات الاختبار الوطني لضبط نوعية التعليم في مباحث العلوم والرياضيات واللغتين العربية

والإنجليزية للصف الثامن الأساسي للعام الدراسي 2018/2019م لمتغيرات جنس الطالب وشكل الصورة الاختبارية باستخدام طريقة مانتل - هانزل المعممة وذلك لمناسبتها لطبيعة الدراسة وأهدافها.

#### بيانات الدراسة

توفرت بيانات الدراسة الحالية من استجابات طلبة الصف الثامن الأساسي على فقرات الاختبار الوطني لضبط نوعية التعليم في مباحث العلوم والرياضيات واللغتين العربية والإنجليزية في مدارس المملكة الأردنية الهاشمية. والبالغ عددهم (10000) خلال الدورة الاختبارية 2019م. وبين الجدول (1) توزيع أفراد الدراسة وفقاً لمتغيراتها.

الجدول 1: توزيع بيانات الدراسة وفقاً لمتغيراتها

المبحث	الجنس		شكل الصورة الاختبارية	
	الذكور	الإناث	الورقي	إلكتروني
اللغة العربية	4662	5338	8751	1249
اللغة الإنجليزية	4634	5366	8694	1306
الرياضيات	4725	5275	8789	1211
العلوم	4768	5232	8742	1258

#### أداة الدراسة

تكونت أداة الدراسة من الاختبار الوطني لضبط نوعية التعليم في المباحث الأساسية الأربعة للصف الثامن الأساسي في الدورة الاختبارية 2019م. تضمنت فقرات من نوع الاختيار من متعدد لكل منها أربعة بدائل. كما في الجدول (2).

الجدول 2: يبين أعداد الفقرات في المباحث الأساسية في الدورة الاختبارية 2019

المباحث الأساسية	الدورة الاختبارية 2019
اللغة العربية	40
اللغة الإنجليزية	36
الرياضيات	40
العلوم	42

#### إجراءات الدراسة

سارت الدراسة وفق الخطوات الآتية:

- الحصول على الموافقات الرسمية المتعلقة بتسهيل مهمة الباحثان في استيفاء البيانات الخاصة بالاختبار الوطني لضبط نوعية التعليم للصف الثامن الأساسي.
- الحصول على البيانات اللازمة من خلال إدارة الاختبارات والامتحانات في وزارة التربية والتعليم.

- تنقية البيانات وحذف المتغيرات والبيانات التي ليس لها علاقة بالدراسة. والاحتفاظ باستجابات الطلبة على فقرات الاختيار من متعدد على وفق لمتغيري الجنس وشكل الصورة الاختبارية.
- إجراء المعالجات الإحصائية باستخدام برمجية (GMHDIF) والرمز الإحصائية للعلوم الاجتماعية (SPSS).

#### المعالجة الإحصائية

للإجابة عن أسئلة الدراسة استخدمت المعالجات الإحصائية الآتية:

تم استخدام طريقة مانتل - هانزل المعممة للكشف عن الفقرات التي تظهر أداء تفاضلياً على فقرات الاختبار الوطني لضبط نوعية التعليم في المباحث الأربعة (اللغة العربية، واللغة الإنجليزية، والرياضيات، والعلوم) تبعاً لمتغيري: جنس الطالب، وشكل الصورة الاختبارية، على مرحلتين:

المرحلة الأولى: وهي مرحلة أولية تم فيها الكشف عن الفقرات ذات الأداء التفاضلي بالاعتماد على جميع استجابات الطلبة على الفقرات.

المرحلة الثانية: وهي مرحلة نهائية هدفها الكشف عن الفقرات ذات الأداء التفاضلي بشكل نهائي ومنقح بحيث تم استثناء استجابات الطلبة على الفقرات ذات الأداء التفاضلي، واعتمدت على استجابات الفقرات الخالية من الأداء التفاضلي.

## نتائج الدراسة ومناقشتها

السؤال الأول: والذي ينص على: ما الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم في مبحث اللغة العربية للصف الثامن الأساسي تبعاً لمتغيري: جنس الطالب. وشكل الصورة الاختبارية؟ وللإجابة عن هذا السؤال فقد تم استخدام طريقة مانتل - هانزل المعممة التي يوفرها برنامج (GMHDIF) الموضوع من قبل فادلوكوا (Fidalgo, 2010) للكشف عن الفقرات التي تظهر أداءً تفاضلياً في اختبار اللغة العربية لضبط نوعية التعليم للصف الثامن وفقاً للمتغيرات الآتية:

أولاً- جنس الطالب:

جدول 3: نتائج طريقة مانتل-هانزل المعممة للكشف عن الفقرات التي تُبدي أداءً تفاضلياً وفقاً لمتغير جنس الطالب

رقم الفقرة	QMH	احتمالية الخطأ	المجموعة المرجعية		المجموعة البؤرية	
			المتوسط الحسابي	الانحراف المعياري	المتوسط الحسابي	الانحراف المعياري
13	10.825	0.00	0.58	0.58	0.51	0.5
21	10.975	0.001	0.30	0.30	0.32	0.467
40	16.404	0.00	0.56	0.56	0.56	0.497

يلاحظ من جدول (3) أنّ (3) فقرات من أصل (40) فقرة أبدت أداءً تفاضلياً. باستخدام مستوى دلالة ( $\alpha = 0.01$ ) وهي تشكل ما نسبته (7.5%) وهي الفقرات ذوات الأرقام (13، 21، 40). حيث كانت الفقرة رقم (13) لصالح الإناث، والفقرة (21) لصالح الذكور. في لم يتم تحديد الأداء التفاضلي للفقرة (40) لصالح أي من المجموعتين. وتتفق هذه النتيجة مع دراسة الطراونة (2011) أبو مسرة والسوالملة (2021) العمري وشطناوي (2016) التي أظهرت نتائجها وجود أداء تفاضلي في متغير الجنس لصالح الإناث وقد يعود ذلك إلى متغيرات وظروف بيئية وواقعية متشابهة فيما يتعلق بمحتوى اختبار اللغة العربية.

## ثانيًا - شكل الصورة الاختبارية

جدول 4: نتائج طريقة مانتل-هانزل المعممة للكشف عن الفقرات التي تُبدي أداء تفاضلياً وفقاً لمتغير شكل الصورة الاختبارية

رقم الفقرة	QMH	احتمالية الخطأ	المجموعة المرجعية		المجموعة البؤرية	
			المتوسط الحسابي	الانحراف المعياري	المتوسط الحسابي	الانحراف المعياري
3	7.094	0.00	0.65	0.477	0.68	0.467
6	17.054	0.00	0.63	0.483	0.75	0.435
7	18.537	0.00	0.52	0.500	0.65	0.478
9	13.617	0.00	0.48	0.500	0.49	0.500
10	8.236	0.002	0.21	0.404	0.21	0.406
14	7.543	0.006	0.66	0.475	0.76	0.426
16	16.133	0.00	0.78	0.416	0.88	0.325
23	14.258	0.00	0.47	0.499	0.59	0.491
31	15.314	0.00	0.71	0.453	0.83	0.375
33	87.043	0.00	0.52	0.500	0.71	0.452
39	14.924	0.00	0.36	0.481	0.46	0.498

يلاحظ من جدول (4) أنّ (11) فقرات من أصل (40) فقرة أبدت أداء تفاضلياً. باستخدام مستوى دلالة ( $\alpha = 0.01$ ) وهي تشكل ما نسبته (27.5%) وهي الفقرات ذوات الأرقام (3.6، 7.9، 10، 14، 16، 23، 31، 33، 39)، جميعها لصالح الشكل الإلكتروني للاختبار في حين أن الفقرة (10) لم يتم تحديد لصالح أي مجموعة. وتتفق هذه النتيجة مع دراسة فريجات (2014) والتي أظهرت وجود أداء تفاضلي تبعاً لشكل تقديم الاختبار لصالح الشكل المحوسب من الاختبار كما أنها تختلف وتعزى هذه النتيجة إلى استفادة المفحوصين من الشكل الإلكتروني لاختبار بسبب سلاسة التعامل مع الاختبار والاستفادة من مزايا الاختبارات والمتعلقة بإمكانية تغيير تنسيق الصفحة وتغيير حجم الخط في الاختبار الإلكتروني وعدم إمكانية ذلك في الشكل الورقي. وكذلك سرعة الاستجابة في الشكل الإلكتروني من خلال النقر بواسطة الفأرة وليس التظليل بقلم الرصاص على ورقة القارئ الضوئي في الاختبار الورقي. والسرعة في عمليات الحذف والتعديل في الاختبار الإلكتروني. بينما يحتاج استخدام المحاة وإعادة التظليل مرة أخرى في الاختبار الورقي وقتاً أطول. وهذا ربما يؤثر على مقدار تركيز المفحوص للاستدلال إلى الإجابة الصحيحة الأمر الذي يزيد من القلق والتوتر لدى المفحوص: الأمر الذي يزيد من ميل الطلبة إلى التخمين العشوائي للإجابة الصحيحة.

السؤال الثاني: ما الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم في مبحث اللغة الإنجليزية للصف الثامن الأساسي تبعاً لمتغيري: جنس الطالب. وشكل الصورة الاختبارية؟ وللإجابة عن السؤال الثاني فقد تم استخدام طريقة مانتل - هانزل المعممة التي

يوفرها برنامج (GMHDIF) للكشف عن الفقرات التي تظهر أداء تفضلياً في اختبار اللغة الإنجليزية لضبط نوعية التعليم للصف الثامن الأساسي وفقاً للمتغيرات الآتية:

أولاً- جنس الطالب:

جدول 5: نتائج طريقة مانتل-هانزل المعممة للكشف عن الفقرات التي تُبدي أداء تفضلياً وفقاً لمتغير جنس الطالب

رقم الفقرة	QMH	احتمالية الخطأ	المجموعة المرجعية		المجموعة البؤرية	
			المتوسط الحسابي	الانحراف المعياري	المتوسط الحسابي	الانحراف المعياري
1	6.68	0.01	0.43	0.496	0.39	0.488
2	18.773	0.00	0.4	0.49	0.4	0.49
3	116.416	0.00	0.3	0.457	0.35	0.478
4	162.135	0.00	0.45	0.498	0.5	0.5
6	15.029	0.00	0.61	0.487	0.56	0.496
7	43.676	0.00	0.57	0.495	0.53	0.499
9	23.111	0.00	0.73	0.442	0.6	0.489
10	108.005	0.00	0.35	0.478	0.39	0.487
12	31.021	0.00	0.47	0.499	0.46	0.498
15	141.915	0.00	0.67	0.472	0.47	0.499
18	52.835	0.00	0.65	0.476	0.5	0.5
20	7.976	0.00	0.48	0.499	0.44	0.497
21	64.245	0.00	0.68	0.467	0.54	0.499
22	18.847	0.00	0.66	0.475	0.53	0.499
23	69.449	0.00	0.23	0.423	0.29	0.453
24	35.725	0.00	0.69	0.462	0.55	0.498
27	111.71	0.00	0.33	0.471	0.39	0.487
29	149.222	0.00	0.73	0.446	0.53	0.499
30	57.796	0.00	0.7	0.457	0.55	0.497
31	196.901	0.00	0.33	0.471	0.43	0.495
32	46.71	0.00	0.67	0.469	0.53	0.499
34	19.108	0.00	0.6	0.49	0.48	0.499
35	43.7	0.00	0.69	0.463	0.54	0.498
36	41.641	0.00	0.62	0.486	0.49	0.5

يلاحظ من جدول (5) أنّ (24) فقرات من أصل (36) فقرة أبدت أداء تفضلياً. باستخدام مستوى دلالة ( $\alpha = 0.01$ ) وهي تشكل ما نسبته (66.66%) وهي الفقرات ذوات الأرقام (1, 2, 3, 4, 6, 7, 9, 10, 12, 15, 18, 20, 21, 22, 23, 24, 27, 29, 30, 31, 32, 34, 35, 36) منها



(17) فقرات لصالح الإناث وبنسبة (70.83%)، وهي الفقرات ذوات الأرقام: (1، 6، 7، 9، 12، 15، 18، 20، 21، 22، 24، 29، 30، 32، 34، 35، 36)، و(6) فقرات لصالح الذكور، هي الفقرات ذوات الأرقام (3، 4، 10، 23، 27، 31)، في حين لم يتم تحديد لصالح أي مجموعة كانت الفقرة رقم (2)، وتتفق هذه النتيجة مع دراسة الطراونة (2011) أبو مسرة والسوالمه (2021) العمري وشطناوي (2016) التي أظهرت النتائج وجود أداء تفاضلي لصالح الإناث. وقد يعزى تفوق الإناث على الذكور في القدرات اللفظية والتعبيرية في مبحث اللغة الإنجليزية، وربما - أيضا - تهزى النتيجة إلى أن قدرات الإناث السمعية أعلى من البصرية، وإلى طبيعة الإناث والدور الاجتماعي الذي يفرضه مجتمعنا وعاداته ونمط التنشئة الوالدية المتبع.

## ثانيًا - شكل الصورة الاختبارية

جدول 6: نتائج طريقة مانتل-هانزل المعممة للكشف عن الفقرات التي تُبدي أداء تفاضلياً وفقاً لتغير شكل الصورة الاختبارية

رقم الفقر ة	QMH	احتمالية الخطأ	المجموعة المرجعية		المجموعة البؤرية	
			المتوسط الحسابي	الانحراف المعياري	المتوسط الحسابي	الانحراف المعياري
2	44.583	0.00	0.41	0.492	0.35	0.478
4	29.014	0.00	0.47	0.499	0.46	0.499
6	13.803	0.00	0.59	0.492	0.61	0.487
7	63.366	0.00	0.55	0.498	0.52	0.500
8	39.943	0.00	0.70	0.462	0.71	0.456
9	7.789	0.002	0.67	0.470	0.72	0.449
11	8.457	0.058	0.44	0.496	0.47	0.499
12	11.211	0.037	0.47	0.499	0.50	0.500
17	26.45	0.00	0.45	0.497	0.46	0.498
21	9.35	0.00	0.61	0.488	0.71	0.454
24	54.024	0.00	0.62	0.485	0.79	0.409
29	42.924	0.00	0.63	0.483	0.78	0.413
30	38.658	0.00	0.63	0.484	0.78	0.416
31	24.439	0.00	0.37	0.484	0.35	0.476
32	20.571	0.00	0.60	0.489	0.73	0.442
35	7.734	0.002	0.61	0.487	0.73	0.446
36	65.572	0.00	0.55	0.498	0.50	0.500

يلاحظ من جدول (6) أنّ (17) فقرات من أصل (36) فقرة أبدت أداء تفاضلياً. باستخدام مستوى دلالة ( $\alpha = 0.01$ ) وهي تشكل ما نسبته (47.22 %) وهي الفقرات ذوات الأرقام (2، 4، 6، 7، 9، 10، 11، 17، 21، 24، 29، 30، 31، 32، 35، 36) منها (12) فقرة لصالح الشكل الإلكتروني للاختبار وبنسبة (70.58 %). وهي الفقرات ذوات الأرقام (6، 8، 9، 11، 12، 17، 21، 24، 29، 30، 32، 35) و(5) فقرات لصالح الشكل الورقي وبنسبة (29.41 %). وهي الفقرات ذوات الأرقام (2، 4، 7، 31، 36). وتتفق هذه النتيجة مع دراسة فريجات (2014) التي أظهرت وجود أداء تفاضلي تبعاً لشكل تقديم الاختبار لصالح الشكل الحوسبي من الاختبار. ويمكن أن تعزى هذه النتيجة إلى أن هناك العديد من العوامل التي تتفاعل مع قدرة المفحوصين للإجابة عن فقرات الاختبار. ومنها ما يتعلق بشكل الاختبار مثل وجود فواصل للصفحات والأسطر وخصائص تصميم الفقرات وخصائص الفقرة والعوامل المتعلقة بسهولة الحركة في أثناء الإجابة عن فقرات الاختبار ومراجعتها ومعاينتها وقدرة المفحوص على الحذف والتعديل للإجابات الخاطئة.

السؤال الثالث: ما الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم في مبحث الرياضيات للصف الثامن الأساسي تبعاً لمتغيري: جنس الطالب. وشكل الصورة الاختبارية؟ وللإجابة عن سؤال الدراسة الثالث فقد تم استخدام طريقة مانتل - هانزل المعممة التي يوفرها برنامج (GMHDIF) للكشف عن الفقرات التي تظهر أداءً تفاضلياً في اختبار الرياضيات لضبط نوعية التعليم للصف الثامن وفقاً للمتغيرات الآتية:

أولاً- جنس الطالب:

جدول 7: نتائج طريقة مانتل-هانزل المعممة للكشف عن الفقرات التي تُبدي أداءً تفاضلياً وفقاً لمتغير جنس الطالب

رقم الفقرة	QMH	احتمالية الخطأ	المجموعة المرجعية		المجموعة البؤرية	
			المتوسط الحسابي	الانحراف المعياري	المتوسط الحسابي	الانحراف المعياري
1	266.427	0.00	0.63	0.482	0.78	0.414
2	29.24	0.003	0.43	0.495	0.40	0.491
4	8.096	0.00	0.40	0.491	0.45	0.497
5	16.824	0.00	0.38	0.486	0.43	0.495
8	24.091	0.00	0.53	0.499	0.59	0.492
9	112.003	0.00	0.24	0.429	0.16	0.368
10	14.043	0.00	0.44	0.496	0.49	0.500
11	24.445	0.00	0.30	0.458	0.35	0.476
12	47.503	0.00	0.35	0.478	0.30	0.459
13	273.78	0.00	0.51	0.500	0.37	0.483
14	35.052	0.00	0.56	0.497	0.63	0.483
16	7.192	0.00	0.41	0.493	0.45	0.497
17	235.932	0.00	0.67	0.472	0.80	0.398
18	126.564	0.00	0.25	0.431	0.16	0.368
19	127.431	0.00	0.48	0.500	0.60	0.490
21	56.139	0.00	0.28	0.449	0.21	0.410
22	6.287	0.007	0.27	0.442	0.24	0.427
24	33.746	0.000	0.48	0.500	0.55	0.498
26	26.864	0.00	0.23	0.419	0.18	0.385
28	37.598	0.00	0.27	0.445	0.22	0.415
30	66.789	0.00	0.41	0.492	0.50	0.500
31	9.211	0.00	0.38	0.484	0.41	0.492
33	11.231	0.00	0.42	0.494	0.46	0.498
34	6.953	0.001	0.42	0.493	0.45	0.498
37	28.657	0.00	0.35	0.477	0.30	0.458
38	16.25	0.002	0.36	0.479	0.33	0.469

رقم الفقرة	QMH	احتمالية الخطأ	المجموعة المرجعية		المجموعة البؤرية	
			المتوسط الحسابي	الانحراف المعياري	المتوسط الحسابي	الانحراف المعياري
39	36.43	0.00	0.21	0.407	0.17	0.373
40	15.284	0.00	0.25	0.431	0.21	0.409

يلاحظ من جدول (7) أنّ (28) فقرات من أصل (40) فقرة أبدت أداء تفاضلياً، باستخدام مستوى دلالة ( $\alpha = 0.01$ ) وهي تشكل ما نسبته (70 %) وهي الفقرات ذوات الأرقام (1، 2، 4، 5، 8، 9، 10، 11، 12، 13، 14، 16، 17، 18، 19، 21، 22، 24، 26، 28، 30، 31، 33، 34، 37، 38، 39، 40) منها (13) فقرات لصالح الإناث وبنسبة (46.42%)، وهي الفقرات ذوات الأرقام (2، 9، 12، 13، 18، 21، 22، 26، 28، 37، 38، 39، 40)، و(15) فقرات لصالح الذكور وبنسبة (53.57%) وهي الفقرات ذوات الأرقام (1، 4، 5، 8، 10، 11، 14، 16، 17، 19، 24، 30، 31، 33، 34)، وتتفق هذه النتيجة مع دراسة الطراونة (2011) أبو مسرة والسوالمة (2021) العمري وشطناوي (2016) حيث أظهرت النتائج وجود أداء تفاضلي في اختبار الرياضيات لصالح الإناث، وقد يكون الاختلاف في الأداء بين الذكور والإناث على الاختبار عائداً إلى عوامل جينية أو عوامل ثقافية - اجتماعية كذلك التي كونت اعتقاداً لدى الذكور بأن المواد العلمية - خصوصاً الرياضيات - مناسبة لهم أكثر، وأن اللغات والمواد الإنسانية تناسب الإناث أكثر، ما يشكل دافعية وإتقان إيجابية نحو المواد العلمية والرياضيات لدى الذكور أكثر منها لدى الإناث، ومن المعلوم أن إتقان المهارات يرتبط بتقدير الذات والثقة بالنفس.

#### ثانياً - شكل الصورة الاختبارية

جدول 8: نتائج طريقة مانتل-هانزل المعممة للكشف عن الفقرات التي تُبدي أداء تفاضلياً وفقاً لمتغير شكل الصورة الاختبارية

رقم الفقرة	QMH	احتمالية الخطأ	المجموعة المرجعية		المجموعة البؤرية	
			المتوسط الحسابي	الانحراف المعياري	المتوسط الحسابي	الانحراف المعياري
2	8.981	0.00	0.41	0.491	0.40	0.490
3	10.353	0.00	0.35	0.476	0.33	0.470
4	8.638	0.00	0.43	0.495	0.50	0.500
8	15.518	0.00	0.56	0.496	0.53	0.500
11	9.253	0.00	0.32	0.465	0.38	0.485
14	7.706	0.01	0.59	0.491	0.66	0.473
17	33.645	0.00	0.73	0.442	0.84	0.371

0.378	0.17	0.394	0.19	0.00	10.702	18
0.489	0.61	0.499	0.53	0.00	12.358	19
0.467	0.32	0.475	0.34	0.00	10.685	20
0.418	0.23	0.430	0.24	0.01	7.334	21
0.417	0.22	0.433	0.25	0.01	7.149	22
0.380	0.18	0.404	0.20	0.00	8.821	26
0.455	0.29	0.427	0.24	0.00	10.286	28
0.500	0.48	0.488	0.39	0.00	21.393	29
0.489	0.39	0.495	0.43	0.00	16.404	34
0.353	0.15	0.391	0.19	0.00	14.397	39

يلاحظ من جدول (8) أنَّ (17) فقرات من أصل (40) فقرة أبدت أداء تفاضلياً. باستخدام مستوى دلالة ( $\alpha = 0.01$ ) وهي تشكل ما نسبته (42.5%) وهي الفقرات ذوات الأرقام (2, 3, 4, 8, 14, 17, 18, 19, 20, 21, 22, 26, 28, 29, 34, 39) منها (7) فقرات لصالح الشكل الإلكتروني للاختبار وبنسبة (41.17%). وهي الفقرات ذوات الأرقام (4, 11, 14, 17, 19, 28, 29) و(10) فقرات لصالح الشكل الورقي وبنسبة (58.82%). وهي الفقرات ذوات الأرقام (2, 3, 8, 18, 20, 21, 22, 26, 34, 39) وتختلف هذه النتيجة مع دراسة فريجات (2014) التي أظهرت وجود أداء تفاضلي على مستوى فقرات اختبار الرياضيات تبعاً لشكل تقديم الاختبار لصالح الشكل المحوسب من الاختبار وتعزى هذه النتيجة إلى أنَّ الإجابة عن فقرات الرياضيات وما تتضمنه من رموز وعمليات حسابية تتطلب تطلب تأزراً بصرياً أثناء الإجابة عليها ولا يتأتى إلا من خلال استعانة المفحوص بالقلم والورقة لإجراء بعض المحاولات والتأكد من بعض الأفكار ووضع بعض التدوينات على الورقة البيضاء والتي قد ترفع من شدة التركيز وصدق عملية الاستدلال إلى الإجابة الصحيحة وهذه يتعذر تحقيقه على الشكل الإلكتروني لأن التأزر البصري الحركي في تلك الحالة يكون بين الفأرة والشاشة دون مداخلات جانبية تمكن المفحوص من رفع شدة التركيز في الاستدلال على الرغم من تزويد المفحوص بورقة بيضاء في أثناء إجابه للاختبار الإلكتروني الأمر الذي يؤدي إلى الإرباك ويعيق عملية الاسترسال في الإجابة.

السؤال الرابع: ما الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم في مبحث العلوم للصف الثامن الأساسي تبعاً لمتغيري: جنس الطالب، وشكل الصورة الاختبارية؟ وللإجابة عن سؤال الدراسة الرابع فقد تم استخدام طريقة مانتل - هانزل المعممة التي يوفرها برنامج (GMHDIF) للكشف عن الفقرات التي تظهر أداء تفاضلياً في اختبار العلوم لضبط نوعية التعليم للصف الثامن وفقاً للمتغيرات الآتية:

أولاً- جنس الطالب:

جدول 9: نتائج طريقة مانتل-هانزل المعممة للكشف عن الفقرات التي تُبدي أداء تفاضلياً وفقاً لمتغير جنس الطالب

رقم الفقرة	QMH	احتمالية الخطأ	المجموعة المرجعية الإناث	المجموعة البؤرية الذكور
------------	-----	----------------	--------------------------	-------------------------

الانحراف المعياري	المتوسط الحسابي	الانحراف المعياري	المتوسط الحسابي			
0.494	0.58	0.496	0.56	0.00	31.617	1
0.497	0.44	0.497	0.45	0.00	8.902	2
0.47	0.33	0.444	0.27	0.00	65.778	3
0.399	0.8	0.324	0.88	0.00	63.613	4
0.456	0.7	0.425	0.76	0.00	9.673	5
0.473	0.34	0.449	0.28	0.00	55.548	11
0.457	0.7	0.458	0.7	0.00	12.788	12
0.437	0.26	0.471	0.33	0.00	49.659	13
0.499	0.47	0.497	0.45	0.00	27.83	15
0.498	0.55	0.475	0.66	0.00	60.111	16
0.493	0.59	0.496	0.56	0.00	54.006	18
0.411	0.22	0.371	0.16	0.00	52.42	19
0.486	0.38	0.499	0.48	0.00	49.766	24
0.489	0.4	0.5	0.48	0.00	35.318	25
0.499	0.46	0.5	0.52	0.00	10.844	27
0.497	0.45	0.499	0.53	0.00	28.035	28
0.473	0.34	0.465	0.32	0.00	14.685	29
0.48	0.36	0.474	0.34	0.00	12.798	30
0.467	0.32	0.483	0.37	0.00	12.215	33
0.5	0.48	0.493	0.58	0.00	49.242	34
0.5	0.48	0.496	0.57	0.00	26.935	35
0.48	0.36	0.469	0.33	0.00	21.197	36
0.49	0.4	0.5	0.48	0.00	30.26	37
0.492	0.41	0.5	0.48	0.00	24.186	38
0.496	0.57	0.455	0.71	0.00	136.64	39
0.495	0.57	0.435	0.75	0.00	245.62	40
0.499	0.46	0.49	0.6	0.00	122.01	41
0.485	0.62	0.418	0.77	0.00	177.17	42

يلاحظ من جدول (9) أنّ (28) فقرات من أصل (42) فقرة أبدت أداءً تفاضلياً. باستخدام مستوى دلالة ( $\alpha = 0.01$ ) وهي تشكل ما نسبته (66.66%) وهي الفقرات ذوات الأرقام (1، 2، 3، 4، 5، 11، 12، 13، 15، 16، 18، 19، 24، 25، 27، 28، 29، 30، 33، 34، 35، 36، 37، 38، 39، 40، 41، 42) منها (18) فقرات لصالح الإناث وبنسبة (64.29%). وهي: الفقرات ذوات الأرقام (2، 4، 5، 13، 16، 24، 25، 27، 28، 33، 34، 35، 37، 38، 39، 40، 41، 42) و(10) فقرات لصالح الذكور وبنسبة (35.71%) وهي الفقرات ذوات الأرقام (1، 3، 11، 12، 15، 18، 19، 29، 30، 36). وتتفق هذه النتيجة مع دراسة الطراونة (2011) أبو مسرة والسوالمه (2021) العمري وشطناوي (2016) التي أظهرت وجود أداء تفاضلي في فقرات اختبار العلوم لصالح الإناث وقد

## ثانيًا - شكل الصورة الاختبارية

**جدول 10:** نتائج طريقة مانتل - هانزل المعممة للكشف عن الفقرات التي تُبدي أداء تفاضليًا وفقًا لتغير شكل الصورة الاختبارية

المجموعة البؤرية الشكل الإلكتروني		المجموعة المرجعية الشكل الورقي		احتمالية الخطأ	QMH	رقم الفقرة
الانحراف المعياري	المتوسط الحسابي	الانحراف المعياري	المتوسط الحسابي			
0.499	0.46	0.497	0.45	0.000	14.175	2
0.499	0.53	0.496	0.44	0.007	7.237	25
0.448	0.28	0.47	0.33	0.000	35.41	29
0.5	0.5	0.49	0.4	0.000	16.166	31
0.424	0.23	0.443	0.27	0.000	13.479	32
0.484	0.63	0.5	0.52	0.002	10.009	35
0.496	0.56	0.496	0.43	0.000	27.799	37
0.497	0.55	0.497	0.44	0.000	20.267	38
0.438	0.74	0.481	0.64	0.001	10.53	39
0.396	0.81	0.473	0.66	0.000	36.95	40
0.482	0.63	0.5	0.51	0.000	14.804	41
0.357	0.85	0.458	0.7	0.000	50.403	42

2690

## التوصيات والمقترحات

- انتهت هذه الدراسة إلى مجموعة من التوصيات والمقترحات. ومن ذلك:
1. رفع قدرات العاملين في إعداد الاختبارات الوطنية في مجال صياغة وإعداد فقرات اختبارية ذات مواصفات سيكومترية عالية ويخلو من التحيز والأداء التفاضلي من خلال وتعريفهم وتدريبهم بالأداء التفاضلي لفقرات.
  2. توجيه إدارة الامتحانات والاختبارات في وزارة التربية والتعليم لتجريب الاختبارات ومعالجة الفقرات التي تبدي أداء تفاضلياً قبل اعتمادها.
  3. إجراء دراسة تتناول الأداء التفاضلي لفقرات امتحان الثانوية العامة في مباحث العلوم والرياضيات واللغتين العربية والإنجليزية؛ لأهمية هذا الامتحان والقرارات المترتبة على نتائجه.
  4. إعطاء مزيد من الاهتمام في إعداد فقرات الاختبار الوطني لضبط نوعية التعليم في مبحث الرياضيات واللغة الإنجليزية مراعاة مصادر الأداء التفاضلي في هذه المباحث.
  5. التوسع في تنفيذ الاختبار الوطني لضبط نوعية التعليم إلكترونياً وخاصة في مباحث اللغة العربية واللغة الإنجليزية والعلوم.

## المراجع العربية

- البرصان، إسماعيل (2013). الأداء التفاضلي لمتغير الجنس لفقرات الاختبار الوطني الأردني لضبط نوعية التعليم لمادة الرياضيات للصف العاشر. مجلة كلية التربية بالزقازيق، 28 (79)، 1 - 35.
- البطوش، علي والقرعان، محمود (2018). فحص الأداء التفاضلي لفقرات في أدوات تقييم الطلبة لجودة التعليم في الأردن حسب الكلية الأكاديمية باستخدام طريقة مانتل - هانزل العامة. مجلة جامعة القدس المفتوحة للأبحاث والدراسات التربوية والنفسية، 8 (23)، 158-173.
- الحجاج، عبدالوهاب (2019). الأداء التفاضلي لفقرات لاختبار الرياضيات الوطني لضبط نوعية التعليم للصف العاشر للعام 2016/2017 عبر المجموعات المتعددة. أطروحة دكتوراه، جامعة اليرموك، إربد، الأردن.
- الطراونة، محمود (2011). تحري التحيز في فقرات الاختبار الوطني لطلبة الصف الثامن في مادتي الرياضيات والعلوم (2010)، رسالة ماجستير غير منشورة، جامعة مؤتة، الكرك، الأردن.
- العجلوني، جهاد والسوالمه، يوسف (2016). فحص الأداء التفاضلي لموهبات فقرات اختبار الرياضيات في التقييم الوطني لاقتصاد المعرفة وفقاً لجنس الطالب، وموقع المدرسة. مجلة جامعة النجاح للأبحاث (العلوم الإنسانية)، 30 (11)، 2265-2286.



علام. صلاح الدين (2005). نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي. القاهرة: دار الفكر العربي.

العمري. حسان وعبدالله. شطناوي (2016). الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم في الرياضيات للصف العاشر تبعاً لتغير الجنس. مجلة جامعة النجاح للأبحاث (العلوم الإنسانية). 30 (8). 1530 - 1554.

فريحات. نجود (2014). الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم في العلوم والرياضيات للصف العاشر الأساسي تبعاً طريقة تقديم الاختبار. أطروحة دكتوراه غير منشورة. جامعة اليرموك. إربد. الأردن.

أبو مسرة. آيات والسوالمه. يوسف (2021). استخدام طريقة أشجار راش للكشف عن الأداء التفاضلي لفقرات الاختبار الوطني لضبط نوعية التعليم في الرياضيات للصف الرابع الأساسي في الأردن. المجلة الأردنية في العلوم التربوية 17 (1): 119 - 134.

النوافلة. علي (2017). أثر الفقرات ذات الأداء التفاضلي المنتظم في تقديرات معالم فقرات الاختبار والأفراد في ضوء بعض المتغيرات باستخدام بيانات مولدة وفقاً لنموذج استجابة الفقرة ثلاثي المعلمة. أطروحة دكتوراه. جامعة اليرموك. إربد. الأردن.

وزارة التربية والتعليم (2013). نتائج الاختبار الوطني لضبط نوعية التعليم للعام الدراسي 2013/2012.

وزارة التربية والتعليم (2016). نتائج الاختبار الوطني لضبط نوعية التعليم للعام الدراسي 2016/2015.

وزارة التربية والتعليم (2019). نتائج الاختبار الوطني لضبط نوعية التعليم للعام الدراسي 2019/2018.

المراجع الأجنبية:

Camilli, G. and Shepard, L. (1994). *Methods for identifying biased test items*. Thousand Oaks, CA: Sage Publications.

Camilli, G. (2006). Test fairness. In R. L. Brennan (Ed.) *Educational Measurement* (4<sup>th</sup> ed., Vol. 4, pp. 221- 256). Westport: American Council on Education & Praeger Publishers.

Dorans, J. & Holland, W. (1994). *DIF detection description, Mantel-Haenszel and standardization*. In W. Holland & H. Wainer (Eds.). *Differential item functioning* (pp.35-66). Hillsdale, NJ: Lawrence Erlbaum.

Fidalgo, A. & Madeira, J. (2008). Generalized Mantel-Haenszel methods differential item functioning in polytomous items. *Journal of Educational Measurement*, (33)8, 315-332.

- Gierl, M. J., Jodoin, M. G. & Ackerman, T. A. (2000). *Performance of Mantel-Haenszel, Simultaneous Item Bias Test and Logistic Regression when the proportion of DIF items is large*. In Annual Meeting of the American Educational Research Association (AERA).
- Gruijter, D. & Kamp, L. (2005). *Statistical test theory for education and psychology*. Retrieved December 30, 532017 from: [www.leidenuniv.nl](http://www.leidenuniv.nl).
- Gybles, J. (2004). The Performance of some Observed and Unobserved Conditional Invariance Techniques for the Detection of Differential Item Functioning. *Quality of Quantity*, 38,681-702.
- Herrmann-Abell, C. & DeBoer, G. (2011). *Using distractor-driven standards-based multiple-choice items and Rasch modeling to investigate hierarchies of chemistry misconceptions and detect structural problems with individual items*. Chemistry Education.
- Roever, C. (2005). *That's not fair Fairness, bias, and differential item functioning in language testing*. Retrieved on 23 Feb. 2017 from: <http://www2.hawaii.edu/~roever/brownbag.pdf>.
- Uiterwijk, H. & Vallen, T. (2005). Linguistic sources of item bias for second generation immigrants in Dutch tests. *Language Testing*, 22(2), 211-234.
- Wiberg, M. (2007). *Measuring and detecting differential item functioning in criterion-referenced licensing test: A theoretic comparison of methods*.