

## معايرة اختبار الدراسة الدولية لقياس مدى تقدم القراءة (PIRLS) بسلطنة عمان باستخدام نظرية الاستجابة للمفردة

أمل بنت خميس الزعابية\*  
و عبد الحميد وراشد بن سيف المحرزي  
وزارة التربية والتعليم، سلطنة عمان | جامعة السلطان قابوس، سلطنة عمان

قُبِل بتاريخ: ٢٠١٩/٥/٩

اُسْتُلم بتاريخ: ٢٠١٩/١/٣

ملخص: هدفت الدراسة إلى معايرة بيانات سلطنة عمان في كتيبات اختبار قياس مهارات التقدم في القراءة (PIRLS) باستخدام نموذج الاستجابة المتدرجة، وفحص الخصائص السيكومترية لها وتحديد المفردات الملائمة وغير الملائمة لافتراضات نموذج الاستجابة المتدرجة المستخدم في تدرّج مفردات كل كتيب. تم استخدام اختبار (PIRLS) لعام ٢٠١١ كأداة للدراسة؛ حيث يتكون من ١٤٦ مفردة اختبارية (منها ٧٤ مفردة ثنائية الاستجابة، و ٧٢ مفردة متعددة الاستجابة). تم تقسيمها إلى ١٣ كتيب بحيث يشمل كل كتيب على نص أدبي ونص علمي، ثم تم توزيعها على ١٣ مجموعة من طلبة الصف الرابع الأساسي بسلطنة عمان بعينة إجمالية بلغت ١٠٣٩٤ طالباً وطالبة. تم التحقق من افتراضات نظرية الاستجابة للمفردة (افتراض أحادية البعد والاستقلال الموضوعي للمفردات) باستخدام برنامج SPSS ثم تم استخدام نموذج الاستجابة المتدرجة لتدرّج مفردات كل كتيب، وذلك باستخدام برنامج Multilog7.03 لتقدير معالم المفردات الاختبارية وقدرات الأفراد ووضعها على تدرّج مشترك. أظهرت النتائج تحقق افتراضات نموذج الاستجابة المتدرجة في بيانات الدراسة، كذلك ملائمة جميع مفردات الاختبار لتوقعات نموذج الاستجابة المتدرجة ما عدا ٨ مفردات، وتمثل ٥% تقريبا من مفردات الاختبار، مما يؤكد تمتع الاختبار بخصائص سيكومترية جيدة (معالم التمييز والصعوبة ودالة معلومات الاختبار) في ضوء نظرية الاستجابة للمفردة.

كلمات مفتاحية: اختبار PIRLS، الخصائص السيكومترية، الصف الرابع، سلطنة عمان.

### Calibrating PIRLS Test in Sultanate of Oman Using Item Response Theory

Amal k. Al-zaabi,\*  
The Ministry Of Education, Sultanate of Oman | Abdulhameed Hassan & Rashid S. Al-mehrzi  
Sultan Qaboos University, Sultanate of Oman

**Abstract:** The study aimed to calibrate Oman data of the PIRLS test using the graded response model and to examine the psychometric properties of it, as well as identify the fit and unfit of its items. PIRLS2011 test booklets were used, which consisted of 146 test items (74 dichotomous and 72 polytomous). Items were divided into 13 booklets; each with two blocks (one literary and one informational). PIRLS test booklets were administered to 13 groups of fourth grade students in Sultanate of Oman with a total sample of 10394 students. Assumptions of IRT (unidimensionality and local independence) were examined and supported. Also, item fit was examined and supported using Samejima's graded response model. The data was analyzed by Multilog7.03 program to estimate both item and ability parameters. Results indicated that the assumptions of IRT were proved. Also, IRT analysis revealed that 8 items showed unfit which represents only 5% of the test items. So, this result confirms that the test has good psychometric properties under the IRT.

**Keywords:** PIRLS test, psychometric properties, fourth grade, Oman.

\*[amal.zaabe@moe.om](mailto:amal.zaabe@moe.om)

والحضارية لممثلي الحكومات، لذلك فإن اختبار بيرلز يقدم معلومات لكل دولة عن منهجية القراءة، والمنهج التنظيمية للمدارس وكذلك الممارسات التعليمية (Mullis, et. al, 2012). حيث أظهرت نتائج الدورات السابقة لهذه الدراسة في ترتيب الدول المشاركة في "PIRLS" ٢٠٠١ و ٢٠٠٦ أن القوة الاقتصادية والكثافة السكانية ليست شرطاً لاحتلال المراتب الأولى في الدراسات العالمية. ونلاحظ ذلك من خلال ترتيب الدول المشاركة مثل السويد أو هولندا، وهي دول صغيرة في حجمها وتعداد سكانها، ولا تملك قوة اقتصاد الدول العظمى قد احتلت مراكز متقدمة جداً في الترتيب العالمي، وتفوقت على الولايات المتحدة الأمريكية وألمانيا وفرنسا وإيطاليا. و يدفعنا هذا إلى الاعتقاد بأن باستطاعة سلطنة عمان منافسة الدول الأخرى، وهي قادرة على تحقيق نتائج جيدة يمكن أن تتحسن وتتطور مع مرور الزمن متى ما تضافرت جهود المعنيين بالعملية التربوية للرقى بالقراءة بشكل خاص والتعليم بشكل عام. كذلك فإن الكشف عن نتائج دورات PIRLS دفع المسؤولين في مجال التربية وقطاع التعليم بدول عديدة احتلت مراتب متدنية، إلى الإسراع باتخاذ الإجراءات العاجلة لتدارس أسباب ذلك، وتحليلها، والبحث عن الحلول المناسبة لجعل الطلبة يلمون بمهارات القراءة والكتابة. وبما أن اختبارات الدراسات الدولية مثل (PIRLS) و (TIMSS) تستخدم المفردات المختلطة والتي تشمل على مفردات ثنائية الاستجابة والمفردات التي تحتاج إلى إجابة إنشائية من الطالب. فإنه يمكن استعمال الأسئلة متعددة الإجابات لتقييم أي واحدة من عمليات الفهم، وبما أنها لا تسمح بكتابة جمل تدعم الإجابة فإنها لا تكون مناسبة تماماً لتقييم قدرة الطالب على تقديم شروح أو تقديرات مفصلة. أما الأسئلة ذات الإجابة الإنشائية تتطلب إنشاء إجابات مكتوبة، ويستعمل هذا النوع لتقييم

حرصت وزارة التربية والتعليم في سلطنة عمان على امتلاك مؤشرات تساعد في تطوير برامج التعليم والتعرف على مستوى التحصيل، وخصوصاً مادة اللغة العربية والرياضيات؛ فقد كان جل تركيزها على تحقيق أعلى مستويات التقدم في مهارتي القراءة والكتابة والعمليات الحسابية لدى طلبة الحلقة الأولى؛ لذلك شرعت في بناء اختبارات تحصيلية ذات جودة عالية من خلال الاستعانة بخبراء مختصين في هذا المجال، كما شاركت في الدراسات الدولية لقياس وتقويم مستويات التحصيل الدراسي مع باقي دول العالم، بهدف الوقوف على مستويات التحصيل ومحاولة تطويرها وتحسينها للوصول بالطلاب العماني لما وصل إليه أقرانه من الدول الأخرى.

ومن الدراسات الدولية التي شاركت فيها السلطنة، هي الدراسة الدولية (PIRLS) لقياس مدى التقدم في القراءة لدى طلبة الصف الرابع الأساسي، والتي كان من أهدافها معرفة كيف يتم تعليم القراءة، ووسائلها، وموادها في مدارسنا، وأيضاً التعرف على كفاءات طلبة الصف الرابع وإنجازاتهم في تحقيق الأهداف المرسومة حسب معايير القراءة. حيث تخضع نتائج هذه الدراسات للتحليل النوعي والكمي من قبل الجمعية الدولية المنظمة لها ومن قبل الدول المشاركة فيها، ويصدر بشأنها تقرير يحتوي على التفاصيل الخاصة بكل دورة من دوراتها (Mullis, Martin, Minnich, Drucker, & Ragan, 2012). كما أن الإستيبيانات المصاحبة للإمتحانات تقدم معلومات إضافية حول تعليم وتعلم المواد المعنية بحيث يستفيد منها المختصون بالمنهج والتقييم التربوي وطرق التدريس (وزارة التربية والتعليم، ٢٠١٤).

من المهم جداً معرفة أن الدول التي خضعت لاختبار PIRLS2011 مختلفة جداً في كثير من الأساسيات منها أنظمة التعليم التي تعكس هذا الاختلاف، وأيضاً التقاليد التاريخية

إلى الإجابة الصحيحة عن طريق التخمين العشوائي للإجابة، وهذا الإرتفاع في نسبة التخمين لبعض مفردات الإختبار يؤدي إلى نوع من التضليل، لأنه لا يعبر عن درجة الطالب الحقيقية في الإختبار، وذلك يؤدي إلى أفضلية النظرية الحديثة في التعامل مع هذه الإشكالية لوجود نموذج يتعامل مع معلم التخمين وهو النموذج ثلاثي المعالم، وأيضا وجود نماذج الإستجابة المتدرجة التي تؤدي إلى التعامل مع الإختلاف في استجابات الطلبة على المفردات متدرجة الإستجابة وتعطي نتائج تعبر عن قدرات الأفراد بدقة مع مراعاة الخصائص السيكومترية لهذه المفردات.

وفي ظل مزايا نظرية الإستجابة للمفردة وتعدد عيوب النظرية الكلاسيكية اتجه أغلب الباحثين إلى استخدام مبادئ نظرية الإستجابة للمفردة في بناء اختبارات جديدة أو التأكد من مطابقة مقاييس معدة مسبقاً لافتراضات نماذج نظرية الإستجابة للمفردة؛ وذلك من أجل تحقيق شروط القياس الموضوعي للسمة أو القدرة المراد قياسها، تلك الشروط المتمثلة في استقلال تقدير خصائص المفردات عن عينة الأفراد المطبق عليهم المقياس، واستقلال تقديرات الأفراد عن مجموعة المفردات المكونة للمقياس (عبد الوهاب، ٢٠١٠).

ومما يؤكد اتجاه الباحثين نحو استخدام مبادئ النظرية الحديثة دراسة الزبيري وقاسم (Zubairi & Kassim, 2016) عندما استخدموا كل من النظرية الكلاسيكية ونموذج راش وهو أحد نماذج الإستجابة للمفردة في تقييم كفاءة مفردات اختبار مكون من ٣٥ مفردة لقياس القراءة في اللغة الإنجليزية، والذي يستخدم من ضمن إجراءات قبول الطلبة في الجامعة الدولية الإسلامية بماليزيا، وقد أثبتت نتائج الدراسة أنه لا يمكن الإكتفاء بالنظرية الكلاسيكية فقط في تحليل مفردات هذا الإختبار، وضرورة الاستفادة من نموذج راش لعدم

أي عملية من العمليات الأربع التركيز على المعلومات المنصوص عليها بوضوح واسترجاعها بنسبة ٢٠% من الإختبار، وتكوين استدلالات مباشرة بنسبة ٣٠%، وتفسير الأفكار والمعلومات ودمجها بنسبة ٣٠%، وفحص وتقييم المحتوى واللغة والعناصر النصية بنسبة ٢٠%، ومع ذلك فهي أكثر ملائمة لتقييم جوانب من الفهم الذي يستدعي الطلبة تقديم دعم لإجاباتهم أو تستدعي بتقديم شرح بالإعتماد على معلوماتهم وتجاربهم (Mullis, Martin, Kennedy, Tong & Sainsbury, 2009).

ونظراً لتوجه الباحثين في مجال القياس والتقويم إلى قياس وتشخيص الجوانب المعرفية على أساس علمي دقيق للإستفادة منها في إرشاد وتوجيه المعلمين للطرق الحديثة في كيفية توصيل المعرفة للطلبة بما يتناسب مع قدراتهم ويؤهلهم للمنافسة مع طلبة العالم في المهارات القرائية، وتحديث المناهج الدراسية بما يساير توجهات المناهج العالمية، ونظراً لما ينتاب النظرية التقليدية من عيوب وقصور، وما تمتاز به النظرية الحديثة في القياس؛ فإن الدراسة تسعى في هذا البحث إلى تدريج مفردات كتيبات اختبار الدراسة الدولية (PIRLS) التي تم تنفيذها على طلبة الصف الرابع الأساسي في العام الدراسي ٢٠١١، وذلك لإمدادهم بنموذج من هذه الاختبارات المدرجة مفرداتها في ضوء نظرية الاستجابة للمفردة والذي قد يفيد في بناء وتحليل اختبارات مماثلة لإختبار قياس مهارات القراءة (PIRLS) والتي كانت تعتمد على النظرية الكلاسيكية في إعدادها وتحليلها، مما يساعد الطلبة على تنمية مهارات القراءة المتقدمة في ضوء نظرية الاستجابة للمفردة.

وليس إعداد الاختبارات المشابهة لنماذج (PIRLS) هو السبب فقط وراء الدراسة للقيام بتدريج مفردات نماذج منها، ولكن هناك سبب لا يقل عن سابقه وهو ارتفاع نسبة الوصول

الحديثة منها دراسة النضيجي (٢٠١٢) التي هدفت إلى إيجاد الخصائص السيكومترية لاختبار المصفوفات المتتابعة المتقدم، والذي تم تطبيقه على عينة بلغت (١٨٠٨) من طلبة المرحلتين المتوسطة والثانوية، وتم إجراء مقارنة بين النماذج اللوغاريتمية لتحديد أيها أكثر ملاءمة لبيانات اختبار المصفوفات المتتابعة المتقدم، وأظهرت النتائج أن النموذج ثنائي المعالم هو الأكثر ملاءمة لبيانات هذا الاختبار.

و دراسة أحمد (٢٠١٠) التي هدفت إلى اعداد بطارية لقياس الإستعداد الأكاديمي، حيث تم استخدام نموذج راش في بناء وتدريب بنك مفردات لقياس الإستعداد الأكاديمي. وتكون المقياس في صورته النهائية من ٢٧٦ مفردة من نوع الاختيار من متعدد، وأظهرت نتائج الدراسة أن بيانات المقياس اتسقت إلى حد كبير وحقت درجة عالية من حسن المطابقة لأسس القياس الموضوعي.

كما هدفت دراسة عبد الوهاب (٢٠١٠) إلى التحقق من افتراضات النظرية الحديثة في بيانات ثلاثة اختبارات (اختبار التفكير الإستدلالي واختبار التفكير الناقد واختبار النماذج المختفية)، التي تم تطبيقها على طلبة مرحلتي البكالوريوس والليسانس بكلية التربية بجامعة المنيا. تم استخدام النموذج ثلاثي المعالم في تدريب اختبار التفكير الناقد واختبار النماذج المختفية، بينما تم استخدام النموذج ثنائي المعالم في تدريب مفردات اختبار التفكير الإستدلالي، وأظهرت نتائج الدراسة تحقق افتراضات النظرية الحديثة في بيانات الإختبارات الثلاثة، وكذلك تم تحديد المفردات الملاءمة لافتراضات النماذج المستخدمة في التدريب وحذف المفردات غير الملائمة.

ويتبين من عرض الدراسات السابقة التفوق الذي أحرزته النظرية الحديثة في القياس، حيث أظهرت نتائجها دقة وموضوعية نماذج السمات الكامنة في إعداد الإختبارات وقياس

تقديم النظرية الكلاسيكية لمعلومات ضرورية عن أداء الطلبة ذوي المستويات المختلفة في القدرة على كل مفردة في الإختبار.

وكذلك دراسة شيربوم وكوهن وكيرن (Scherbaum, Cohen & Kern, 2006) التي هدفت إلى التحقق من الخصائص السيكومترية لمقياس الكفاءة الذاتية العام الذي بناه Sherer et al. (1982) بالمقارنة مع مقياسين آخرين وهي مقياس الكفاءة الذاتية العام المحسوس الذي بناه Schwarzer and Jerusalem (1995) ومقياس الكفاءة الذاتية العام الجديد الذي بناه Chen et al. (2001) وذلك باستخدام نظرية الإستجابة للمفردة، وأوضح النتائج إلى أن المقاييس الثلاثة متشابهة في قياس السمة الكامنة. وعلى الرغم من وجود اختلافات بسيطة بينها إلا أن مقياس Chen et al. (2001) تفوق عليها في معالم تمييز المفردات، ودالة معلومات الإختبار، مما يشير إلى تمتع المقاييس الثلاثة بخصائص سيكومترية مقبولة وقدرتها على التمييز بين الأفراد وقدرة المفردات.

أما دراسة أويشي (Oishi, 2005) فقد هدفت إلى التحقق من الخصائص السيكومترية لمقياس الرضا عن الحياة Satisfaction with Life Scale (SWLS) بين مجموعتين من الأمريكيين والصينيين، باستخدام نموذج المعادلة البنائية ونموذج ثنائي المعالم التابع لنظرية الإستجابة للمفردة، أظهرت النتائج أن استخدام نموذج ثنائي المعالم أعطى معلومات عالية ونتائج موضوعية في الاختلاف بين المتوسطات للعينيتين وأن هذا الإختلاف قد لا يرجع إلى تحيزات المفردات، مما يشير إلى تفوق النظرية الحديثة في الكشف عن الخصائص السيكومترية للمقياس.

وهناك العديد من الدراسات العربية التي توجهت إلى الاستفادة من مزايا النظرية

مستوى أداء الأفراد، على الرغم من اختلاف نوعية الإختبارات المستخدمة وأهدافها. كما أن هناك العديد من الدراسات التي استفادت من بيانات الدراسة الدولية PIRLS منها دراسة راتري (Ratry, 2015) التي استخدمت بيانات اختبار PIRLS 2011 لطلاب أندونيسيا للتحقق من تأثير العوامل المحيطة بالطالب (المنزل والوالدين واتجاهات الطالب نحو القراءة وكذلك الممارسات القرائية خارج المدرسة وتأثير استخدام التقنية والتكنولوجيا) على المعرفة القرائية، وكيف تتفاعل مع معالجة المعلومات والفهم القرائي، حيث تم استخراج مؤشرات هذه العوامل من الاستبيانات المرافقة لاختبار PIRLS 2011. واستخدم الباحث نموذج راش وتحليل الانحدار الخطي الهرمي لتحليل بيانات الدراسة. وكشفت نتائج الدراسة عن وجود دلالة إحصائية لتأثير العوامل المحيطة بالطالب السابق ذكرها على المعرفة القرائية وأدائهم في اختبار PIRLS 2011.

استخدمت دراسة زيمرمان (Zimmerman, 2014) نتائج اختبار PIRLS 2006 لطلاب جنوب إفريقيا في الصف الرابع للوقوف على أسباب تدني مستويات هؤلاء الطلاب وربطها بكيفية تدريس الفهم القرائي، من خلال ملاحظة الإستراتيجيات التي يستخدمها المعلمون في تدريس القراءة بشكل عام وتدريب الفهم القرائي بشكل خاص. وأظهرت نتائج الدراسة التي اعتمدت على المقارنة بين أداء المدارس ذات الإنجاز المرتفع في PIRLS 2006 وذات الإنجاز المنخفض وجود تباين بين الطرق المستخدمة في تدريس الفهم القرائي بين هذه المدارس، كما كشفت عن وجود فجوة بين أداء المعلمين في الحصة الدراسية وبين المستوى المطلوب للفهم القرائي.

استخدمت دراسة زيمرمان (Zimmerman, 2014) نتائج اختبار PIRLS 2006 لطلاب جنوب إفريقيا في الصف الرابع للوقوف على أسباب تدني مستويات هؤلاء الطلاب وربطها بكيفية تدريس الفهم القرائي، من خلال ملاحظة الإستراتيجيات التي يستخدمها المعلمون في تدريس القراءة بشكل عام وتدريب الفهم القرائي بشكل خاص. وأظهرت نتائج الدراسة التي اعتمدت على المقارنة بين أداء المدارس ذات الإنجاز المرتفع في PIRLS 2006 وذات الإنجاز المنخفض وجود تباين بين الطرق المستخدمة في تدريس الفهم القرائي بين هذه المدارس، كما كشفت عن وجود فجوة بين أداء المعلمين في الحصة الدراسية وبين المستوى المطلوب للفهم القرائي.

معرفة أسباب تدني مستوى طلبة نيوزيلاندا في الفهم القرائي وتحليلها وتحديد الإستراتيجيات المناسبة لرفع هذا المستوى، وذلك من خلال المقارنة بين نتائج طلبة نيوزيلاندا في اختبارات PIRLS 2001, 2006, 2011، والتي أشارت إلى تناقص مرتبة نيوزيلاندا في السلم العالمي في كل مرحلة من المرحلة السابقة لها. وكشفت نتائج الدراسة أن المعرفة القرائية للطالب تتأثر بعوامل عدة من أهمها البيئة الثقافية والاقتصادية، لذلك أوصت بضرورة استبدال المناهج المستخدمة في القراءة مثل برنامج القراءة العلاجية Reading Recovery المستخدم منذ ١٥ عاما في نيوزيلندا بمناهج حديثة أثبتت نظريات التعلم المعاصرة فعاليتها مثل نظريات التعلم البنائية Cues Constructivist model.

استخدمت دراسة الصمدي (2011) المرتبة التي حازتها المغرب في التصنيف الدولي في اختبار PIRLS 2011 والتي من خلالها تم دراسة وتحليل أسباب حصول المغرب على المرتبة الأخيرة في هذا الإختبار (ومنها الأدوار المحورية للأسرة والمنزل والمدرسة والبيئة المحيطة بالطالب وتطبيقات المناهج والطرائق البيداغوجية) مما يستدعي من المسؤولين عن القطاع التربوي والتعليمي استثمار هذه النتائج في تقييم المنظومة التعليمية ومعالجة العوامل التي ساهمت في انحدار مؤشرات جودة التعليم في المغرب.

وكما هدفت دراسة فوي ومارتين وميوليس وستانكو (Foy, Martin, Mullis & Stanco, 2011) إلى دراسة الخصائص السيكمترية لاختبار الدراسة الدولية PIRLS 2011 للدول الأعضاء المشاركة في الدراسة الدولية، ومنها سلطنة عمان، ومقارنتها بالمتوسط العالمي، وذلك من خلال استجابات الطلاب على المفردات (الثنائية والمتعددة الإستجابة). ووجد من خلال عرض بيانات سلطنة عمان أن متوسط صعوبة المفردات

وسعت دراسة تونمر وتشابمن وجريني وبرجنو وأرو (Tunmer, Chapman, )

المعرفية لدى الطلاب وبالتالي أثرت على مستويات القراءة لديهم، حيث أن اختلاف مستويات الخبرات لدى الطلاب أثرت على مستوى صعوبة المفردات.

#### مشكلة الدراسة

إن العمل في المجال التربوي يتطلب تحديد درجة معرفة الطلبة وإتقانهم للمهارات المختلفة المراد إكسابهم إياها، وكافة الخبرات التعليمية المقدمة لهم، وما يمتلكون من أنماط التفكير، وما يستطيعون أداءه من أنشطة ومهام وتكليفات، وكذلك معرفة ميولهم واستعداداتهم واتجاهاتهم نحو القضايا المختلفة؛ وذلك كله من أجل اتخاذ قرارات بشأن تقدمهم في البرامج التعليمية والتدريبية والعلاجية المقدمة لهم، ومن ثم انتقالهم إلى مراحل تعليمية متقدمة، وتوجيههم إلى الدراسة والعمل المناسبين لإمكاناتهم وقدراتهم وميولهم واستعداداتهم، وكذلك تصنيفهم والمقارنة بينهم. وهذا يتطلب مقاييس في غاية الدقة والموضوعية والصدق والحساسية للفروق الطفيفة بينهم (عبد الوهاب، ٢٠١٠).

ونظراً لأن النظرية الكلاسيكية تعتمد على مجموعة من الفرضيات البسيطة وأدت إلى كثير من الأخطاء والمشكلات في القياس كحصول الفرد على نسبتي مختلفتين للذكاء إذا تم اختباره باختبارين مختلفين بقياسان الذكاء؛ وذلك لاختلاف درجة الخطأ المعياري في الحالتين وتأثر أداء الفرد على نفس الاختبار تحت نفس الظروف الاختبارية إذا أعيد تطبيقه عليه مرة أخرى؛ وذلك بسبب وجود هذا الخطأ العشوائي في الدرجة التي يحصل عليها (Sijtsma & Junker, 2006). كما أن اعتماد النظرية الكلاسيكية في تحليل المفردات على عينة الأفراد المطبق عليهم المقياس سواء عند حساب الثبات أو معاملات التمييز أو معاملات الصعوبة يؤدي إلى عدم توفر القياس الموضوعي لهذه السمة أو القدرة (Stage, 2003; Gleason, 2008).

الثنائية باستخدام نموذج راش تساوي (٠.٣٧) وهي نسبة أعلى من المتوسط العالمي (٠.١٧) مما يدل على صعوبة المفردات بالنسبة لطلبة السلطنة، ومتوسط صعوبة المفردات المتعددة الإستجابة باستخدام نموذج راش (١.١٤) وهي أيضاً أعلى من المتوسط العالمي البالغ (٠.٩٠). كما تم استخراج معامل التمييز للمفردات ووجد في بيانات سلطنة عمان أن معاملات التمييز معاملات تمييز موجبة جيدة وأقل من المتوسط العالمي لمعاملات التمييز البالغ (٠.٤٨) للمفردات ثنائية الإستجابة و(٠.٦١) للمفردات متعددة الإستجابة.

وهدفت دراسة ميوليس ومارتين وفوي (Mullis, Martin & Foy, 2011) إلى الاستفادة من بيانات ٣٤ دولة في اختبارات الدراسة الدولية TIMSS & PIRLS 2011 للتحقق من العلاقة بين القدرة القرائية التي تم قياسها بواسطة اختبارات PIRLS 2011 وكيفية تأثيرها على أداء طلاب الصف الرابع في فهم مسائل الرياضيات والعلوم في اختبارات TIMSS. طرحت الدراسة فرضيتين هما: الطلاب الذين لديهم قدرة قرائية عالية قادرون على الإجابة على أسئلة اختبار TIMSS والطلاب الذين لديهم قدرة قرائية متدنية سوف تكون استجاباتهم أفضل للمفردات التي لا تتطلب تمكن في القراءة.

أوضحت النتائج صحة فرضية الدراسة أن الطلبة ذوي القدرة القرائية العالية في اختبار PIRLS2011 هم الأفضل في الإستجابة على جميع مفردات اختبار TIMSS بمختلف مستويات صعوبتها، وفي المقابل الطلبة الذين لا يستطيعون القراءة استطاعوا فقط الإستجابة على المفردات ذات الصعوبة المتدنية والتي تحتوي على عدد كلمات ورموز ورسوم توضيحية أقل، وتوضح الدراسة أن السبب في اختلاف نتائج استجابات الطلاب على المفردات يعود إلى الإختلاف في المناهج الدراسية وطرق التدريس، والتي لها تأثير على المجالات

٢. هل تتلائم البيانات المستمدة من استجابات الطلبة على مفردات اختبار PIRLS2011 مع افتراضات نموذج الإستجابة المتدرجة المستخدم في تدرّيج مفردات الإختبار بسلطنة عمان؟

٣. ما قيم تقديرات معالم مفردات اختبار PIRLS2011 وتدرّيجها باستخدام نموذج الإستجابة المتدرجة بسلطنة عمان؟

٤. ما مقدار المعلومات التي يقدمها اختبار PIRLS2011 لقدرات الطلبة بسلطنة عمان؟

#### أهمية الدراسة

- توجيه أنظار العاملين في وزارة التربية والتعليم إلى استخدام أساليب القياس الموضوعي، والمتمثلة في نظرية الاستجابة للمفردة، والإستفادة من خصائص هذه النظرية ومزاياها في التوصل إلى دقة القياس وموضوعيته.
- قد تشجع هذه الدراسة الباحثين على استخدام جميع نماذج نظرية الإستجابة للمفردة في تدرّيج اختبار قياس مهارات القراءة (بيرلز) والمقارنة بينها.
- الدراسة الحالية خطوة للتفاعل مع التطورات الحديثة في مجال القياس النفسي والتربوي، ومواكبتها من خلال إيجاد الخصائص السيكومترية لاختبار قياس مهارات القراءة (بيرلز) ومعايير تفسير الأداء في ضوء نظرية الإستجابة للمفردة.
- قد يفيد تدرّيج هذا الإختبار في الإكتشاف المبكر لل صعوبات التي يعاني منها التلاميذ، مما يؤدي إلى التركيز على الخطوات التي تعالج

وفي الوقت الذي تتعدد فيه عيوب النظرية الكلاسيكية في القياس تزداد مزايا نظرية الإستجابة للمفردة بنماذجها المختلفة، وتمثل هذه المزايا في استقلال خصائص المفردات (الصعوبة والتمييز) عن عينة الأفراد المستخدمة في التحليل، وكذلك استقلال تقدير قدرات الأفراد عن عينة المفردات المكونة للمقياس. كما أن تفسير درجات الأفراد يتم في ضوء المفردات وليس في ضوء الجماعة المرجعية كما في النظرية الكلاسيكية، وتحقق خصائص الميزان الفكري وربما القياس النسبي دون ضرورة أن يكون توزيع مستويات القدرة في المجتمع المستهدف اعتدالياً (علام، ٢٠٠٥).

لذلك من المهم تدرّيج مفردات اختبارات الدراسة الدولية (PIRLS) التي تم تنفيذها على طلبة الصف الرابع الأساسي، وذلك للحصول على تدرّيج يعطي نتائج تعبر عن قدرات الأفراد بدقة مع مراعاة الخصائص السيكومترية لهذه المفردات والذي قد يحتاجه المختصون لبناء وتحليل اختبارات مماثلة لاختبار قياس مهارات القراءة (PIRLS)، بحيث يمكننا المقارنة بين تلك النماذج التي تقيس السمة ذاتها وفقاً لمعيار مشترك تنسب إليه درجات الطلبة في النماذج المختلفة مما يؤدي إلى تقديرات ثابتة عن مستوى تحصيل كل طالب، بحيث تتحقق الدراسة من أن استجابات الطلبة على نماذج الإختبار ماهي إلا انعكاس عن قدراتهم ولا تتدخل فيها صياغة مفردات الإختبار ولا التخمين أو التسرع في الإجابة، وبالتالي فإن مشكلة البحث الحالي تشير مجموعة من التساؤلات يمكن صياغتها على النحو التالي:

١. هل تتحقق افتراضات نظرية الإستجابة للمفردة (أحادية البعد والإستقلال الموضوعي) للبيانات المستمدة من استجابات الطلبة على مفردات اختبار PIRLS2011 بسلطنة عمان؟

PIRLS2011 من سلطنة عمان، والتي تم اختيارها بناء على مواصفات اختيار عينة الدراسة الدولية في جميع الدول المشاركة وفقا لمعايير التصنيف العالمية للتربية التابعة للتصنيف الدولي الموحد للتعليم إسكد ٢٠١١ (ISCED)، وهو التصنيف المرجعي لتنظيم برامج التعليم والمؤهلات ذات الصلة حسب مستويات ومجالات التعليم، حيث تكونت العينة من طلبة الصف الرابع الأساسي (الذكور والإناث) في جميع محافظات السلطنة من مواليد (٢٠٠٠-٢٠١١) البالغ عددهم ١٠٣٩٤ والذين تم إجراء اختبار PIRLS2011 عليهم (IEA, 2011)؛ والتقويم التربوي، (٢٠١١)، ويوضح جدول ١ توزيع عينة الدراسة على كتيبات اختبار الدراسة PIRLS2011.

جدول ١

توزيع عينة الدراسة من طلبة سلطنة عمان على كتيبات اختبار PIRLS2011		
رقم الكتيب	إناث	ذكور
١	٣٤٥	٣٥٥
٢	٣٣٧	٣٤٧
٣	٣٢٠	٣٥٦
٤	٣١٨	٣٦٥
٥	٣٤٠	٣٥٢
٦	٣٤١	٣٥٣
٧	٣٥٠	٣٤١
٨	٣٤٣	٣٥٠
٩	٣٤٠	٣٥١
١٠	٣٦١	٣٢٧
١١	٣٥٧	٣٣٧
١٢	٣١٨	٣٧٢
١٣	١٠٣٦	١٠٨٢
المجموع	٥١٠٦	٥٢٨٨

#### أداة الدراسة

اختبار دولي يتكون من ٥ نصوص أدبية، تتضمن حكاية أو قصة واقعية من التراث أو من حياتنا اليومية وه نصوص علمية تتضمن معلومات أو مواد تثقيفية في مختلف فروع المعرفة، تقترحها الدول المشاركة موزعة على ١٣ كتيب مختلف ثم يطلب من التلميذ

هذه الصعوبات والوقاية من الوقوع فيها مستقبلاً.

#### مصطلحات الدراسة

**نظرية الإستجابة للمفردة:** هي "مجموعة من النماذج الرياضية التي تفترض أنه يمكن التنبؤ بسلوك الأفراد، أو يمكن تفسير أدائهم في اختبار نفسي أو تربوي معين في ضوء خصائص مميزة لهذا الأداء تسمى سمات Traits أو بمعنى آخر وجود واحدة أو أكثر من السمات الأساسية التي تحدد استجابات الفرد على مفردات الإختبار" (هيبة وعمر، ٢٠١١، ص٩).

**نموذج الإستجابة المتدرجة Graded Response Model:** يعد تعميماً للنموذج ثنائي المعالم (2PL) للمفردة ثنائية الإستجابة، بحيث يمثل العلاقة غير الخطية بين مستوى قدرة الفرد واحتمال استجابته استجابة معينة في المفردات متعددة الإستجابة، ويتطلب ذلك عملية ذات خطوتين من أجل تحديد الإحتمال المشروط لاستجابة فرد استجابة معينة (علام، ٢٠٠٥).

**الدراسة الدولية لقياس مدى تقدم القراءة Progress In International Reading (PIRLS) Literacy Study:** اختبار دولي يقوم على أساس المقارنة لقياس قدرات طلبة الصف الرابع في مهارات القراءة بلغتهم الأم، وذلك لتحديد جوانب القوة والضعف لديهم، ومن ثم تطوير تلك المهارات والإرتقاء بها، بما يحقق أهداف المجلس الأعلى للتعليم، ويلبي متطلبات تطوير التعليم في الدولة، ويسهم في تطوير قدرات وكفايات الطلبة (وزارة التربية والتعليم، ٢٠١٤).

#### الطريقة والاجراءات

##### مجتمع وعينة الدراسة

تكون مجتمع الدراسة من جميع طلبة الصف الرابع بسلطنة عمان في العام الدراسي ٢٠١٢/٢٠١١م. واعتمدت الدراسة على البيانات الأرشيفية للعينة الخاضعة لدراسة



(٠.٧٣) وهو بين نصي الكتيبات (١) و(٣)، بينما أقل معامل ارتباط (٠.٦٣) وهو بين نصي الكتيب (٩)، مما يوفر درجة مقبولة من الصدق البنائي لكل كتيب.

## جدول ٢

معاملات الارتباط بين نماذج كل كتيب من كتيبات اختبار

## PIRLS2011

معامل الارتباط بين النصين	الكتيب
**٠.٧٣	١
**٠.٧١	٢
**٠.٧٣	٣
**٠.٦٨	٤
**٠.٦٤	٥
**٠.٦٧	٦
**٠.٦٤	٧
**٠.٦٨	٨
**٠.٦٣	٩
**٠.٦٩	١٠
**٠.٦٩	١١
**٠.٦٧	١٢
**٠.٦٩	١٣

\*\*دال إحصائياً عند مستوى (٠.٠١)

## إجراءات الدراسة

تم تطبيق كتيبات اختبار PIRLS2011 في يوم واحد، بحيث يجلس الطلبة لأداء جزئي الإمتحان في نفس اليوم مع فترة استراحة قصيرة بينهما بحيث يكون الوقت المخصص لفترتي امتحان الدراسة الدولية لقياس مهارات القراءة ما يقارب ١٠ دقائق تقريباً للتحضير للإمتحان بما في ذلك إعداد الطلبة وقراءة التعليمات عليهم، وتوزيع كتيبات الإمتحان، ثم ٤٠ دقيقة للإجابة على أسئلة الجزء الأول من كتيب الأسئلة، وبعد ذلك استراحة قصيرة [١٥ دقيقة]، ثم ٤٠ دقيقة للإجابة على أسئلة الجزء الثاني من كتيب الأسئلة. تم جمع البيانات وترتيبها ثم تم بناء تدرج مشترك يربط بين السمة المراد قياسها ومعلمات المضردات باستخدام تصميم المجموعة الواحدة من خلال المعايرة المنفصلة لكل كتيب Separate Calibration. ثم تم التحقق من افتراضات

قراءة النصين والإجابة عن الأسئلة. وتوزع الكتيبات على التلاميذ بطريقة مسبقة التنظيم من قبل اللجنة المنظمة للدراسة (Mullis, & et. al, 2009; Marian & Jay, 2001).

- يتكون اختبار PIRLS2011 بشكل إجمالي من ١٦٤ مضردة (٧٤ مضردة ثنائية الاستجابة يتم تصحيحها بنظام (١، ٠)، و٧٢ مضردة متعددة الاستجابة ويتم تصحيحها (٠، ١، ٢، ٣).
- يستخدم تصميم الكتيبات تقنية العينة المصفوفية الذي يقسم النصوص والمضردات إلى نماذج أو كتل بحيث يتم تكوين كتيبات الطلبة وفقاً لترتيب منهجي، بحيث يحتوي كل كتيب على نصين.
- النصوص الأدبية من (L1 إلى L4) و النصوص العلمية من (I1 إلى I4) تظهر في ١٢ كتيب بحيث تتكرر كل منها في ٣ كتيبات، أما النص الأدبي (L5) والنص العلمي (I5) في الكتيب (١٣) فليست مرتبطة مباشرة مع أي مجموعة.

في هذه الدراسة تم استخراج معامل ثبات كل كتيب باستخدام برنامج SPSS بطريقة معامل ثبات ألفا كرونباخ وكانت قيم معامل ثبات ألفا كرونباخ لجميع الكتيبات والتي تتراوح بين (٠.٨٣ - ٠.٨٩)، وهي معاملات مرتفعة ومناسبة لأغراض الدراسة مما يشير إلى دقة الأداة وثباتها. أيضاً تم التحقق من صدق الإختبار عن طريق الصدق البنائي، وذلك عن طريق التأكد من العلاقة الارتباطية بين النصين لكل كتيب، ولتحقيق هذا الغرض تم حساب معامل ارتباط بيرسون لمعرفة مدى الارتباط بين النصين في كل كتيب ويوضح جدول ٢ أن معاملات الارتباط بين النصين لكل كتيب دالة إحصائياً عند مستوى ٠.٠١ حيث تتراوح بين (٠.٦٣، ٠.٧٣)، وبلغ أعلى معامل ارتباط

## نتائج الدراسة

**إجابة السؤال الأول:** هل تتحقق افتراضات نظرية الإستجابة للمفردة (أحادية البعد والإستقلال الموضوعي) للبيانات المستمدة من استجابات الطلبة على مفردات اختبار قياس مدى التقدم في القراءة PIRLS2011 بسلطنة عمان؟

### ١) التحقق من افتراض أحادية البعد

يوضح جدول ٣ نتائج التحليل العاملي الإستكشافي لكتيبات اختبار PIRLS2011، التي تظهر أن نمط الجذور الكامنة لجميع العوامل متشابهة عبر الكتيبات وأن عدد العوامل التي يمكن استخراجها باستعمال المحك القائم على الجذر الكامن هو عامل واحد، حيث تشير النتائج إلى أن قيمة الجذر الكامن للعامل الأول مرتفعة ولبقية العوامل منخفضة. كما يوضح الجدول أن الجذر الكامن للعامل الواحد في كل الكتيبات يفسر نسبة مئوية من التباين الكلي تتراوح بين (٢١٪ إلى ٢٤٪)، مما يرجح وجود عامل كامن سائد يفسر النسبة الأكبر من التباين، ونستدل منه على تحقق أحادية البعد في بيانات هذا الإختبار وقياس مفردات الإختبار لسمة واحدة، وهي السمة التي هدف لقياسها.

### ٢) التحقق من الإستقلال الموضوعي للمفردات

يقصد به أن احتمال الإجابة الصحيحة على مفردة ما يكون مستقلا عن ناتج إجابته عن أي مفردة أخرى في الإختبار عند مستوى قدرة محددة، وأشار هامبلتون و سوامينتان أن الإستقلال الموضوعي مكافئ لافتراض أحادية البعد ويعتبر متحقق بتحقيقه (الرحيل، ٢٠١٣). ولكن لمزيد من التحقق تم حساب مؤشر Q3 وهو معامل الإرتباط للبوقي لزوج من المفردات بعد ضبط السمة المقدره (النعيمي، ٢٠١١).

نظرية الإستجابة للمفردة (افتراض أحادية البعد من خلال التحليل العاملي الإستكشافي، وافتراض الإستقلال الموضوعي)، ثم استخدام برنامج Multilog 7.03 لتقدير معالم المفردات وتدريجها باستخدام نموذج الإستجابة المتعددة. والتحقق من مطابقة البيانات لافتراضات نموذج الإستجابة المتدرجة واستخراج مؤشرات البواق المعيارية باستخدام برنامج الأكسل و SPSS.

### المعالجة الإحصائية

للإجابة عن التساؤل الأول تم إجراء التحليل العاملي الإستكشافي للتحقق من افتراض أحادية البعد باستخدام برنامج SPSS على إجابات أفراد عينة كل كتيب على المفردات المكونة للكتيبات، ثم تم استخراج العوامل بطريقة التحليل إلى المكونات الأساسية (Principal Components)، للعامل الأكبر من الواحد الصحيح، كما تم حساب معاملات الإرتباط للبوقي Q3 لكل زوج من مفردات كل كتيب للتحقق من الإستقلال الموضوعي للمفردات. كما تمت الإجابة عن السؤال الثاني عن طريق استخدام برنامج Multilog في تدرج مفردات اختبار PIRLS2011 ووضعها على مقياس مشترك وبرنامج الأكسل للتحقق من مدى ملائمة النموذج لبيانات Goodness Of fit، ثم حساب مؤشر البواق المعيارية Standardized Residuals ومقارنتها مع مربع كاي ( $\chi^2$ ) كمؤشر لحسن مطابقة المفردات أو الأفراد للنموذج. للإجابة عن السؤالين الثالث والرابع تم استخدام برنامج Multilog لاستخراج تقدير معالم المفردات (الصعوبة والتمييز)، واستخراج دالة معلومات الإختبار.

## جدول ٣

## نتائج التحليل العاملي لكتيبات PIRLS2011 للتحقق من أحادية البعد

النسبة التباين المفسر	الجذر الكامن	العوامل	الكتيب	النسبة التباين المفسر	الجذر الكامن	العوامل	الكتيب
٢٣.٧٣	٦.١٧	١	٢	٢٤.٤٣	٦.١١		١
٤.٧٠	١.٢٢	٢		٥.٢٩	١.٣٢		
٤.٤٦	١.١٦	٣		٤.٦٢	١.١٥		
٢١.٤٩	٦.٨٨	١	٤	٢٣.٢٩	٧.٩٢		٣
٥.٣٨	١.٧٢	٢		٤.٧٠	١.٥٩		
٤.٦٣	١.٤٨	٣		٤.٣١	١.٤٧		
٢٢.٣٩	٥.٣٨	١	٦	٢١.٥٨	٥.١٨		٥
٥.١٩	١.٢٥	٢		٥.٤٧	١.٣١		
٤.٨٦	١.١٧	٣		٤.٩٥	١.١٩		
٢٤.٣٦	٧.٠٧	١	٨	٢٢.١٨	٦.٢١		٧
٥.٩٥	١.٧٣	٢		٦.٠٥	١.٦٩		
٤.٢٥	١.٢٣	٣		٤.٧٩	١.٣٤		
٢٤.٤٣	٥.٨٦	١	١٠	٢١.٨٧	٥.٤٧		٩
٥.١٩	١.٢٥	٢		٥.٧٧	١.٤٤		
٤.٨٤	١.١٦	٣		٤.٨٥	١.٢١		
٢٢.٩٠	٨.٢٤	١	١٢	٢٢.٣٩	٥.٨٢		١١
٥.٧٥	٢.٠٧	٢		٥.٢٦	١.٣٧		
٤.٢٤	١.٥٣	٣		٤.٨٦	١.٢٧		
				٢٣.٤٦	٨.٢١		١٣
				٥.٦٢	١.٩٧		
				٤.٩٠	١.٧٢		

بمتوسط مقداره (٠.٠١)، ماعدا الكتيب ٧ ارتفعت هذه القيمة إلى متوسط مقداره (٠.٠٢). وعند فحص جميع جميع معاملات الارتباط للبوقي بين كل زوج من المفردات فوجدت أنها تقل عن (٠.٠٥)، مما يدل على تحقق افتراض الاستقلال الموضوعي. لذلك نستطيع القول بأن كتيبات الإختبار حققت افتراض الاستقلال الموضوعي ويدعم هذه النتيجة دراسة Reese (1999) التي اعتبرت أن

ويوضح جدول ٤ متوسط مؤشر Q3 لكل كتيب من كتيبات الإختبار بعد حساب معاملات الارتباط للبوقي Q3 لكل زوج من مفردات كل كتيب.

حيث أشارت النتائج في جدول ٤ إلى أن افتراض الاستقلال الموضوعي قد تحقق في جميع الكتيبات، حيث أن القيم صغيرة جدا فكانت أقل هذه القيم عند جميع الكتيبات

## جدول ٤

## متوسط مؤشر Q3 عند كل كتيب من كتيبات الإختبار

متوسط Q3	عدد المفردات	رقم الكتيب	متوسط Q3	عدد المفردات	رقم الكتيب
٠.٠١	٢٩	٨	٠.٠١	٢٥	١
٠.٠١	٢٥	٩	٠.٠١	٢٦	٢
٠.٠١	٢٤	١٠	٠.٠١	٣٤	٣
٠.٠١	٢٦	١١	٠.٠١	٣٢	٤
٠.٠١	٣٦	١٢	٠.٠١	٢٤	٥
٠.٠١	٣٥	١٣	٠.٠١	٢٤	٦
			٠.٠٢	٢٨	٧

حيث يتضح مطابقة البيانات مع توقعات نموذج الإستجابة المتدرجة، حيث كانت قيمة مؤشر كل كتيب غير دالة مقارنة مع قيمة مربع كاي الحرجة (٣.٨٤١) عند درجة حرية ١ بمستوى دلالة (٠.٠٥). كما تم التحقق من ملاءمة المفردات للنموذج كل على حدة، وأسفر التحليل عن وجود ٨ مفردات غير ملائمة لافتراضات النموذج وهي المفردات (١٩، ٢٠) المشتركة بين الكتيبات ٣، ٤، و١٢، والمفردات (٢٤، ٢٥، ٢٧) المشتركة بين ٧، ٨، و١٢، والمفردة (١١) المشتركة بين الكتيبات ١، ٢، ١٠، والمفردة (٢٠) المشتركة بين الكتيبات ٥، ٦، ١١، ويمكن أن يعود عدم ملاءمة هذه المفردات للنموذج إلى انخفاض قيمة معلمة التمييز، حيث تقل قيمته لهذه المفردات عن (٠.٤) مما يدل على أن هذه المفردات غير قادرة على التمييز بين الأفراد بشكل جيد، وإلى ارتفاع قيم معلمة الصعوبة التي تتراوح بين (٢ - ٤.٥)، وهذا يعني أن المفردات ذات صعوبة عالية، أما بالنسبة للمفردة (١٢) في الكتيب ٦، فقد تجاوزت حدود مؤشر المطابقة (٣.٨٤) والذي يمثل الحد الأقصى للدلالة على ملاءمة النموذج، ويرجع ذلك إلى أنها ذات صياغة ضعيفة أو أنها تقيس سمة أخرى شديدة الارتباط بالسمة المقاسة (عبد الوهاب، ٢٠١٠).

كما تم التحقق من مطابقة الأفراد للنموذج وأوضحت نتائج التحليل عن وجود ٥٧ طالباً غير مطابقين، ٣ منهم في الكتيب (٢)، وطالب واحد في الكتيب (٣)، و٣ طلاب في الكتيب (٤)، و٤٢ طالباً في الكتيب (٦)، و٣ طلاب في الكتيب (٧)، و٢ في الكتيب (٩)، و٤ طلاب في الكتيب (١١)، حيث كانت قيم المؤشرات لهم أعلى من قيمة مربع كاي الحرجة (٣.٨٤١) وبدرجة حرية ١، عند مستوى دلالة (٠.٠٥)، مما يعني عدم ملاءمتهم لنموذج الإستجابة المتدرجة، وأن نمط استجاباتهم غير واقعي، إما بسبب

القيمة الصغيرة لهذا الإحصائي تدل على كفاءته في الكشف عن الإستقلال الموضوعي، وكذلك دراسة (Sireci 2002) التي أوضحت أن هذا الإحصائي يأخذ قيمة صغيرة عندما يتحقق الإستقلال الموضوعي لمفردات الإختبار.

**إجابة السؤال الثاني:** هل تتلائم البيانات المستمدة من استجابات الطلبة على مفردات اختبار PIRLS2011 مع افتراضات نموذج الإستجابة المتدرجة المستخدم في تدرج مفردات الإختبار بسلطنة عمان؟

للإجابة عن هذا السؤال تم استخدام برنامج Multilog لتدرج مفردات اختبار PIRLS2011 ووضعها على مقياس مشترك، وبرنامج الأكسل للتحقق من مدى ملائمة النموذج للبيانات، حيث أن تقدير معالم نموذج معين من نماذج الإستجابة للمفردة استناداً إلى البيانات الإختبارية لا يضمن أن النموذج يطابق هذه البيانات، فلا يمكن الإستفادة من ميزات هذا النموذج إلا إذا كانت هذه المطابقة مرضية (علام، ٢٠٠٥). وتم حساب مؤشر البواقي المعيارية ومقارنتها مع مربع كاي ( $\chi^2$ ) كمؤشر لحسن مطابقة المفردات أو الأفراد للنموذج، ويوضح جدول ٥ المؤشر العام لمطابقة كل كتيب مع نموذج الإستجابة المتدرجة والذي يمثل المتوسط العام لمؤشر البواقي المعيارية للمفردات.

جدول ٥

المؤشر العام لمطابقة النموذج لمفردات اختبار

PIRLS2011

رقم الكتيب	المؤشر العام	رقم الكتيب	المؤشر العام
١	٠.٨٦	٨	٠.٨١
٢	٠.٨٥	٩	٠.٨٦
٣	٠.٨٨	١٠	٠.٨٩
٤	١.٢٢	١١	٠.٨٥
٥	٠.٨٤	١٢	٠.٨٢
٦	٠.٨٦	١٣	٠.٨٢
٧	٠.٧٧		

كثييات اختبار PIRLS2011 حيث يتضح أن قيم صعوبة مفردات كل كتيب تغطي مدى واسعاً من الصعوبة وإن كانت تميل بصفة عامة إلى أن تكون مفردات مرتفعة الصعوبة وفقاً لمعيار (Choi, 1992) الوارد في النضيي (٢٠١٢) والذي يوضح أن المفردات تكون سهلة إذا قلت عن (-٠.٥٠)، بينما تكون متوسطة الصعوبة إذا تراوحت قيم معلمة الصعوبة من (-٠.٥٠ إلى +٠.٥٠)، و المفردات تكون صعبة إذا كانت قيم معلمة الصعوبة أكبر من (+٠.٥٠)، وتوضح النتائج أن مفردات اختبار بيرلز ٢٠١١ ذات مدى واسع من الصعوبة، كما تؤكد قيم متوسط الأخطاء المعيارية لتقدير معلمة الصعوبة، أن قيم معلمة الصعوبة تتميز بدرجة مقبولة من الدقة والثبات في تقديرها.

تم تقدير معلمة تمييز مفردات كتيبات اختبار PIRLS2011 والذي يدل على قدرة مفردات الإختبار على التمييز بوضوح بين الأفراد المختبرين ذوي القدرة المرتفعة والقدرة المنخفضة، ويوضح جدول ٧ أعلى وأقل قيمة لمعلم التمييز في كل كتيب من كتيبات الإختبار.

اللجوء إلى التخمين أو التسرع في الإجابة (أحمد، ٢٠١٠).

**إجابة السؤال الثالث:** ما قيم تقديرات معالم مفردات اختبار مدى التقدم في القراءة PIRLS2011 وتدرجها باستخدام نموذج الإستجابة المتدرجة بسلطنة عمان؟

تم تقدير معالم مفردات اختبار مهارات مدى التقدم في القراءة PIRLS2011 باستخدام نموذج الإستجابة المتدرجة وتم تقدير احتمالات الإستجابات للمفردات التي تتطلب استجابات متدرجة، والتي تتضمن حساب منحنيات بعدد يساوي عدد العتبات الفارقة لكل مفردة.

كما تم تقدير معالم الصعوبة للمفردات ثنائية الإستجابة، وأيضاً تقدير صعوبة المفردات متعددة الإستجابة عن طريق حساب متوسط قيم العتبات، حيث تشير هذه المعالم إلى مدى سهولة وصعوبة المفردة، وقد اختلفت قيم معلمة صعوبة المفردات من كتيب لآخر، وهذا يدل على اتساع مدى القدرات التي يغطيها الإختبار من مستوى القدرة المنخفض إلى المستوى المرتفع، ويوضح جدول ٦ الإحصاءات الوصفية لقيم معلمة الصعوبة لمفردات كل كتيب من

#### جدول ٦

الإحصاءات الوصفية لقيم معالم الصعوبة لمفردات كتيبات الإختبار

الانحراف المعياري	المتوسط الحسابي	أصغر صعوبة	أكبر صعوبة	عدد المفردات	الكتيب
٠.١٣	٠.٧٢	-٠.٦٦	٢.٢٩	٢٥	١
٠.١٧	٠.٦٤	-١.٠٥	٢.٩٧	٢٦	٢
٠.٢	٠.٨٥	-٠.٦٦	٢.٥٢	٣٤	٣
٠.٢٦	٠.٩٩	-٠.٧٨	٣.٨٣	٣٢	٤
٠.٢٢	١.٠٧	-٠.٨٧	٢.٢٥	٢٤	٥
٠.٢	٠.٧٨	-٠.٢٧	٢.٩٧	٢٤	٦
٠.٢٧	٠.٨٧	-٠.٣٨	٢.٩٦	٢٨	٧
٠.١٦	٠.٧٢	-٠.٢٥	٣.٢٥	٢٩	٨
٠.٢١	١.٠٠	-٠.٩٥	٢.٧٤	٢٥	٩
٠.٢٣	٠.٧٩	-١.٠٤	٤.٥٩	٢٤	١٠
٠.١٨	٠.٩٢	-٠.١٩	٢.٦١	٢٦	١١
٠.٢٤	١.٠٦	-٠.٤٥	٣.٤٠	٣٦	١٢
٠.١٢	١.١٤	-٠.٢٩	٢.٩٧	٣٥	١٣

ذلك شكل ١ (ملحق ١) الذي يعرض منحني خصائص أربعة كتيبات للإختبار (تم الاقتصار عليها حيث تم اختيارها بشكل عشوائي بهدف التوضيح واختصار العرض).

**إجابة السؤال الرابع:** ما مقدار المعلومات التي يقدمها اختبار PIRLS2011 لقدرات الطلبة بسلطنة عمان؟

تم استخراج منحني معلومات الإختبار (Test Information Curve) لكل كتيب من كتيبات الإختبار باستخدام برنامج MultiLog. ويعرض شكل ٢ (ملحق ١) منحني معلومات الإختبار لأربعة كتيبات (الكتيبات ١، ٥، ٦، ٩) من كتيبات الإختبار (تم الاقتصار عليها بهدف التوضيح واختصار العرض). تشير النتائج أن نمط قيم دالة معلومات الاختبار متشابهة إلى حد كبير بين كتيبات الإختبار، مما يشير إلى أن الكتيبات تعطي معلومات متقاربة وعن نفس الفئة من الطلاب، مع وجود اختلافات بسيطة بينها، كما أنها تغطي مدى واسع من القدرات، حيث أن أول قيمة لدالة المعلومات التي قدمتها كتيبات الإختبار هي (٠.٢) تقريبا عند مستوى القدرة (-٣)، وتزداد تدريجيا بتزايد مستوى القدرة لتصل إلى أعلى درجاتها وهي من (١٠ - ١٥) عند مستوى القدرة (+١)، ثم أخذت في التناقص إلى أن وصلت قيمتها (٤-٥) عند القدرة (+٣)، بينما تتراوح أكبر قيمة لدالة المعلومات التي يمكن أن تقدمها كتيبات الإختبار بين مستويات القدرة (٠-٢)، وتحديدًا عند مستوى القدرة (+١). بينما تكون المعلومات التي تقدمها كتيبات الإختبار أقل عند المستويات المنخفضة من القدرة، وهذا يعني أن الإختبار يمدنا بمعلومات أفضل عند المستوى المتوسط والمرتفع من القدرة، ويتفق ذلك مع مدى قيم معلمة الصعوبة، ويؤكد ذلك قيمة معامل ثبات كتيبات الإختبار التي تراوحت بين (٠.٨٢ - ٠.٨٨)، وأيضًا قيمة الخطأ المعياري والتي تدل على دقة القياس والتي بلغت أقصى درجاتها

ويتضح أن توزيع قيم معالم التمييز للإختبار أن مفردات الإختبار تميز بشكل كبير بين الأفراد المختبرين وذلك وفقا لمعيار بيكر (Baker, 2001)، الذي يعتبر أن المفردات تكون ذات تمييز عالي، إذا كان معامل تمييزها يزيد عن (١.٧) ويكون تمييزها متوسط إذا تراوحت قيم معالم التمييز بين (٠.٦٥ - ١.٣٤)، وذات تمييز منخفض إذا تراوحت قيم معالم التمييز بين (٠.٦٤-٠.٠١).

جدول ٧

الكتيب	أعلى قيمة للتمييز	أقل قيمة للتمييز
١	٢.٥٢	٠.٥٣
٢	٢.٤١	٠.٥٢
٣	٣.٧٧	٠.٣٦
٤	٤.٣٧	٠.٢٨
٥	١.٩٠	٠.٣٢
٦	٢.١٠	٠.٥١
٧	٢.٨٨	٠.٣٤
٨	٥.٢٦	٠.٥٤
٩	٢.١١	٠.٣٨
١٠	١.٩٣	٠.٢٧
١١	٢.٣٥	٠.٣٢
١٢	٣.٠٨	٠.٣٦
١٣	٣.٨٣	٠.٥٠

ويظهر تأثير معلمة تمييز المفردات على شكل منحني خصائص المفردة، فنجد أن المفردات ذات التمييز العالي يكون شكل منحني خصائص المفردة لها على شكل حرف S، بينما المفردات ذات المعالم التمييزي المنخفض تأخذ خط شبه المستقيم ولكنه مائل وليس أفقي، مما يدل على أن المفردات لها قدرة تمييزية ولكن ليست كبيرة، حيث لا يوجد اختلاف بين احتمالية الإجابة الصحيحة بين الأفراد مرتفعي ومنخفضي القدرة. كما يوضح جدول ٧ وجود بعض المفردات ذات تمييز منخفض وخصوصا بين القدرات المتدنية. كما تشير هذه النتائج إلى تمتع بعض مفردات كتيبات PIRLS2011 بقدرة تمييزية جيدة في ضوء استجابات طلبة سلطنة عمان على الإختبار ويدعم

المفردة بالمفردات التي تتبعها في الإختبار القرائي يؤثر على الإستقلال الموضوعي.

٤. صدق محتوى الإختبار من حيث تكافؤ مفردات النماذج المكونة لكل كتيب، وارتباط المفردات بالسمة المقاسة (القراءة).

وتتفق نتائج السؤال الاول مع نتائج العديد من الدراسات منها دراسة فخرو وآخرون (٢٠٠٩)، ودراسة Oishi (2005) من حيث الحصول على نتائج ذات جودة عالية باستخدام نظرية الاستجابة للمفردة في الاختبارات متعددة التدرج والمستخدمه لأغراض التحليل، ويمكن الإستفادة من هذه النتيجة أثناء بناء اختبارات الفهم القرائي للصف الرابع، بحيث يتم بناؤها وفقا لجدول مواصفات للتأكد من ارتباط محتوى الإختبار بالسمة المراد قياسها، وكذلك معرفة الوزن النسبي الحقيقي لكل مستوى من مستويات الفهم القرائي مما يساعد المعلم على توزيع المفردات حسب أهميتها، ويكسب الطالب ثقة كبيرة بعدالة الإختبار.

كما أشارت النتائج إلى وجود حسن مطابقة للبيانات المستمدة من الإستجابة على مفردات اختبار PIRLS2011 لتوقعات نموذج الإستجابة المتدرجة، حيث أن جميع قيم مؤشر البواقي المعيارية غير دالة إحصائياً، ماعدا مفردة واحدة فقط. كما يشير إبراهيم والغاوي وعبد الهادي (٢٠٠٩) إلى أن النموذج الأفضل والأكثر ملاءمة للبيانات هو الذي يحدد أقل عدد من المفردات غير المطابقة لهذا النموذج ويعطي قيمة عالية للثبات، ومن ثم يمكن القول بأن هذا الإختبار يتميز بالدقة والموضوعية في القياس حيث يتصف بخصائص سيكومترية جيدة لكل مفردة من مفردات الإختبار. كما أن حسن مطابقة البيانات لنموذج الإستجابة للمفردة، يوضح قدرة هذا النموذج على تفسير بيانات هذا الإختبار وقدرته على تحمل مخالفة

(٠.٦٨) تقريبا عند مستوى القدرة (-٣)، وتتناقص تدريجياً لتصل إلى أقل قيمة لها وهي (٠.١٥ - ٠.١٧) عند القدرة (+١)، ثم تتزايد لتصل إلى (٠.٤٩) تقريبا عند القدرة (+٣)، حيث أن الخطأ المعياري يكون صغيراً كلما اقترب مستوى قدرة الأفراد من مستوى صعوبة المفردات (فخرو وعبد الرحيم ومحمد، ٢٠٠٩).

### المناقشة والاستنتاجات

تناولت هذه الدراسة تفحص الخصائص السيكومترية لاختبار PIRLS2011 الصف الرابع في بيانات سلطنة عمان باستخدام نموذج الإستجابة المتدرجة التابع لنظرية الاستجابة للمفردة. أظهرت النتائج قياس الإختبار لسمة واحدة تتمثل في مستوى القراءة في اللغة العربية من خلال نتائج التحليل العاملي الإستكشافي، حيث أكد Ackerman المشار اليه في الحجازين (٢٠١٤)، أن أحادية البعد لا تفترض فقط بل يجب التحقق منها. ويؤكد Kolen & Brennan (2004) وعبد الوهاب (٢٠١٠) على وجود عامل أو مكون واحد مسيطر يكمن وراء الأداء في الإختبار وهذا المكون هو القدرة أو السمة التي يسعى الإختبار إلى قياسها. وما يدعم تحقق أحادية السمة في اختبار PIRLS2011 تميز بيانات هذا الاختبار بالآتي:

١. الصياغة الجيدة للمفردات بحيث لا يلجأ المصححون إلى التخمين أو التسرع في الإجابة، حيث تشير دراسة بركات (٢٠١٠) إلى أن افتراض أحادية البعد يتأثر بالصياغة غير الملائمة لمفردات الإختبار.
٢. إعداد الإختبار وفقاً لجدول مواصفات مما يوفر درجة عالية من الصدق البنائي.
٣. ترتيب المفردات بشكل جيد، حيث يشير عبد الوهاب (٢٠١٤)، أن ارتباط

مهارات القراءة (PIRLS) الذي يسعى إلى تطوير مهارات الطلبة القرائية المتدنية إلى المتوسط، حيث اشتمل على أسئلة من الممكن على طلبة القدرات المتوسطة اجتيازها إذا ما تلقوا مزيداً من التدريب على كيفية الإجابة عنها. وتتفق هذه النتائج مع دراسة (Foy & et. al, 2011) ودراسة (Mullis & et. al, 2011) في تمتع مفردات اختبار PIRLS2011 بخصائص سيكومترية جيدة وقادرة على إعطاء مؤشرات دقيقة عن قدرات الطلبة. ويمكن توظيف هذه النتيجة في الحقل التربوي من خلال الرجوع إلى المفردات التي احتواها اختبار PIRLS2011 والوقوف على أنماطها وشروط صياغتها والأبعاد التي تعالجها في كل مستوى من مستويات الفهم القرائي ثم إيجاد استراتيجية تساعد الطلاب على كيفية التعامل مع النص القرائي، بحيث يستطيعون الإجابة عن المفردات بمختلف مستوياتهم.

#### التوصيات والمقترحات

تقدم الدراسة في ضوء نتائجها التوصيات التالية:

١. استخدام نتائج تدريج اختبار PIRLS2011 للوقوف على مستويات طلبة الصف الرابع في الفهم القرائي والسعي للرقى بها للوصول إلى المستوى المطلوب.
٢. استخدام نتائج تقديرات معالم مفردات PIRLS2011 للمقارنة بينها وبين معالم المفردات التي يتضمنها المنهج الدراسي للصف الرابع، ومحاولة تطويرها لتناسب المستوى الذي تتطلع إليه الدراسة الدولية في وصول طلبة الصف الرابع إليه.
٣. استخدام نتائج الدراسة في بناء بنك لمفردات اختبار PIRLS باعتباره يتصف بخصائص سيكومترية عالية، مما يسهل الرجوع إليها وتوظيفها في مناهج الصف الرابع الأساسي.

بعض افتراضاته بحيث لا يؤثر في دقة تقديرات المعالم.

كما أن مطابقة المفردات للنموذج يدل على تحقق الاستقلال الموضوعي للمفردات كما يشير (Ueno 2002)، وفخرو وآخرون (٢٠٠٩)، وأيضاً يفسر جدية الطلبة، وعدم الإجابة عن الأسئلة بشكل عشوائي؛ حيث ذكر حمادنة (٢٠٠٩) أن عدم مطابقة بعض البيانات للنموذج بسبب ميل بعض الطلبة من ذوي القدرات المتدنية إلى التخمين عن الأسئلة، وكذلك إخفاق الطلبة ذوي القدرات العالية في التعامل مع المفردات السهلة، وذلك لأنهم يفكرون بأبعد ما تعنيه المفردة أو أنهم يميلون إلى تفسير المفردة بأكثر مما تحتمل. كما تفسر الدراسة أيضاً ملاءمة البيانات للنموذج بسبب وقوع معالم المفردات ضمن المدى المقبول حسب توقعات النموذج، وكذلك ارتباط هذه المفردات مع بعضها البعض في قياسها للقراءة مما يوفر نسبة جيدة من الصدق البنائي. كما توضح النتائج أن مفردات اختبار PIRLS2011 تتمتع بخصائص سيكومترية جيدة، في ضوء نموذج الإستجابة المتدرجة، والتي تسهم في قياس السمة لدى الطلبة بدقة. حيث نجد أن جميع قيم معلم التمييز موجبة ولا يوجد قيم سالبة وذلك يدل على حرص معدي الإختبار على صياغة المفردات بشكل جيد، حيث أشار (Baker 2001) إلى أن المفردات ذات التمييز السالب تدل على سوء صياغة المفردات والتي بالتالي تؤثر على استجابات الأفراد ذوي القدرات المرتفعة. كما أن قلة المفردات غير الملاءمة وانخفاض الخطأ المعياري لتقديرات معالم المفردات يدل على ثبات الاختبار.

كما تتوزع معلمة صعوبة أغلب المفردات على المستويات بين المتوسطة والمرتفعة القدرة من (١٠٠) تقريباً مما يجعلها تساعد على التمييز بشكل أكبر وتعطي معلومات أكثر عن هذه الفئة من الطلبة، وهي معلومات تتناسب مع أهداف اختبار قياس



الحجازين، نايل عيد (٢٠١٤). نماذج نظرية الاستجابة للمفردة متعددة التدرج متعددة الأبعاد وبرامجها الحاسوبية. عمان: دار جليس الكتاب.

حمادنة، إياد محمد (٢٠٠٩). استخدام نظرية الاستجابة للمفردة في بناء اختبار محكي المرجع في الرياضيات وفق النموذج اللوجستي ثلاثي المعلم. مجلة العلوم التربوية والنفسية، ١٠ (٢)، ٢١٥-٢٣٨.

الرحيل، راتب صايل (٢٠١٣). أثر وجود اداء تفاضلي في المفردات المرساوية على دقة المعادلة العمودية لاختبار اوتيس لينون للقدرة العقلية. المجلة الدولية التربوية المتخصصة، ٢ (٨)، ٧٧١-٧٥٤.

الصمدي، محمد سعيد (٢٠١٤). مشاركة الأطفال المغاربة في المباراة الدولية لقياس الكفاية القرائية بيرلز PIRLS. البيداغوجي، (١)، ٦٩-٧٤.

النفيعي، عبد الرحمن أحمد (٢٠١٢). الخصائص السيكمترية لاختبار المصفوفات المتتابعة المتقدم في ضوء نظرية الاستجابة للمفردة الاختبارية. مجلة كلية التربية، ٢ (١٤٧)، ٢١٤-١٧٥.

النعيمي، عز الدين (٢٠١١). أثر الزيادة في عدد المفردات المرتبطة على الخصائص السيكمترية للاختبار. مجلة اتحاد الجامعات العربية للتربية وعلم النفس، ٩ (٣)، ٢٢-١.

عبد الوهاب، محمد محمود (٢٠١٠). استخدام نماذج الاستجابة للمفردة الاختبارية في تدرج مفردات بعض الاختبارات المعرفية. (رسالة دكتوراه غير منشورة). جامعة المنيا، جمهورية مصر العربية.

عبد الوهاب، محمد محمود (٢٠١٤). التحقق الامبريقي من تكافؤ افتراضي أحادية البعد والاستقلال الموضوعي للمفردات.

٤. تبني استراتيجيات للتعامل مع النص القرائي توفر مناخاً ملائماً للطالب للوصول إلى تعلم ذي معنى قائم على الفهم يساعده على كيفية التعامل مع المفردات الإختبارية.

كما تقدم الدراسة في ضوء نتائجها المقترحات البحثية التالية:

- إجراء دراسة مقارنة بين النظرية الكلاسيكية ونظرية الاستجابة للمفردة في تدرج اختبارات الدراسة الدولية PIRLS2011.
- إجراء دراسة لمعرفة أثر اختلاف حجم العينة على دقة تدرج مفردات اختبار PIRLS2011 باستخدام النموذج الاستجابة المتدرجة.
- إجراء دراسة حول الخصائص السيكمترية لاختبار PIRLS لعام ٢٠١٦ ومقارنتها بالخصائص السيكمترية لاختبار PIRLS2011 التي تم الحصول عليها في هذه الدراسة.

## المراجع

## References

أحمد، محمد محمد (٢٠١٠). بناء وتدرج بطارية لقياس الاستعداد الأكاديمي للقبول بالجامعات باستخدام الاختبارات مرجعية المحك CRT ونظرية الاستجابة للمفردة IRT. مجلة كلية التربية، ٧ (١٤٤)، ٦٤١-٥٥٧.

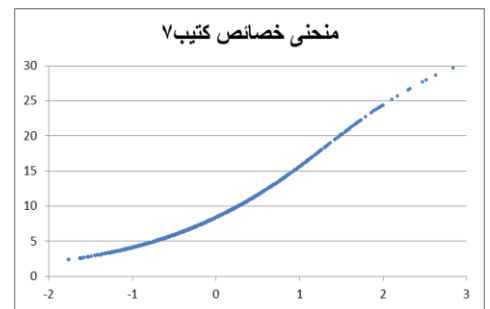
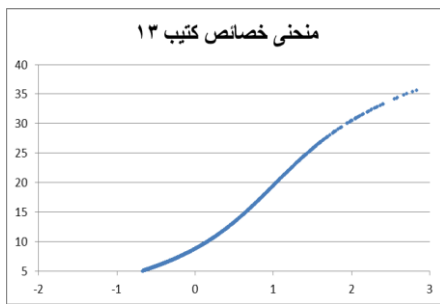
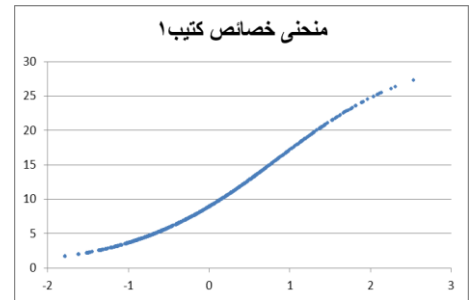
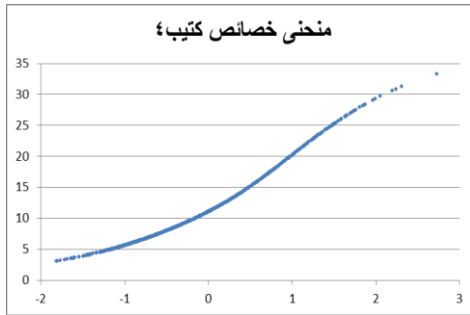
إبراهيم، نسرين اسماعيل؛ الغاوي، هالا عبد الحميد؛ عبد الهادي، سوسن اسماعيل (٢٠٠٩). مقارنة بين النماذج اللوغاريتمية لنظرية الاستجابة للمفردة من حيث ملائمتها للبيانات. مجلة البحث العلمي في التربية- مصر، ١٠ (٤)، ٢٩٥-٣٢٨.

- Baker, F. B. (2001). *The basics of item response theory*. USA: Eric.
- Foy, P., Martin, O., Mullis, V. S., & Stanco, G. (2011). *Methods and procedures- reviewing the TIMSS and PIRLS 2011 achievement item statistics*. Chestnut Hill: Boston college. Retrieved from <http://timssandpirls.bc.edu>. 18/12/2016.
- Gleason, J. (2008). An evaluation of mathematics competitions using item response theory. *Notices of The Ams*, 55 (1), 8-15.
- IEA's Progress in international reading literacy study. (2011). *Methods and procedures- sampling implementation*. Chestnut Hill: Boston college. Retrieved from <http://timssandpirls.bc.edu>. 18/12/2016.
- IEA's Progress in international reading literacy study. (2011). *Population coverage and sample participation rates- Appendix- C*. Chestnut Hill: Boston college.
- Kolen, M., J. & Brennan, R.L. (2004). *Test equating, scaling, and linking: methods and practice*. New York: Springer.
- Marian, S., & Jay, C. (2001). Developing the PIRLS reading assessment. In Campbell, J. R., Kelly, D. L., Mullis, I.V. S., Martin, M.O., & Sainsbury, M. (2001). *Framework and specifications for PIRLS assessment 2001. (2nd ed.)*. Chestnut Hill, MA: Boston college.
- Mullis, V. S., Martin, O., Kennedy, A. M., Tong, L., Sainsbury, M. (2009). *PIRLS 2011 assessment framework*. Chestnut Hill: Boston college.
- Mullis, V., Martin, O., Minnich, A., Drucker, T., & Ragan, A. (2012). *PIRLS 2011 encyclopedia- education policy and curriculum in reading- volume 1:AK*. Chestnut Hill: Boston college.
- Mullis, V., Martin, O., Minnich, A., Drucker, T., & Ragan, A. (2012). *PIRLS 2011 encyclopedia- education policy and curriculum in reading- volume 2:LZ*. Chestnut Hill: Boston college.
- Mullis, V. S., Martin, O., & Foy, P. (2011). The impact of reading ability on
- النشر في المجلة التربوية، الكويت، مجلس البحث العلمي.
- علام، صلاح الدين محمود (٢٠٠٥). نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي. القاهرة: دار الفكر العربي.
- فخرو، حصة عبد الرحمن؛ وعبد الرحيم، أنور رياض؛ محمد، محمد (٢٠٠٩). فاعلية استخدام نموذج سلم التقدير في تحليل مفردات مقياس مداخل الدراسة لدى طلبة جامعتي المنيا وقطر. مجلة البحث في التربية وعلم النفس، ٢٢ (٢)، ١٣٥-١٧٥.
- المديرية العامة للتقويم التربوي (٢٠١١). التقرير الوطني للدراسة الدولية لقياس مهارات القراءة بيرلز ٢٠١١. سلطنة عمان.
- المديرية العامة للتقويم التربوي (٢٠١١). دليل المشرف على إدارة الامتحانات- الصف الرابع- المسح الرئيسي. سلطنة عمان.
- المديرية العامة للتقويم التربوي (٢٠١١). دليل منسقي المدارس TIMSS & PIRLS 2011. سلطنة عمان.
- هيبه، محمد أحمد علي؛ عمر، محمود أحمد (٢٠١١). تكافؤ قياس القائمة المختصرة للعوامل الخمسة للشخصية بين الجنسين في ضوء نظرية الاستجابة للمفردة ونمذجة المعادلة البنائية. مجلة القراءة والمعرفة- مصر، ١١٥، ٩١-١٣١.
- وزارة التربية والتعليم (٢٠١٤). دليل الدراسة الدولية لقياس مهارات القراءة (PIRLS). سلطنة عمان.

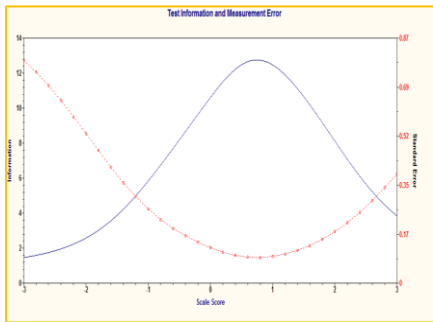
- TIMSS mathematics and science achievement at the fourth grade: An analysis by item reading demands. *TIMSS and PIRLS 2011*, 67-108.
- Oishi, S. (2005). The concept of life satisfaction across cultures: An IRT analysis. *Journal of Research in Personality*, 40(4), 411-423.
- Ratri, S. Y. (2015). Student factor influencing Indonesian student reading literacy based on PIRLS data 2011. *Journal Of Education*, 1(1), 24-32.
- Reese, L. M. (1999). *A classical test theory perspective on LSAT local item dependence*. LSAC research report series. Statistical report.
- Stage, C. (2003). *Classical test theory or item response theory: The swedish experience*. Santiago, Chile.
- Scherbaum, C. A., Cohen-Charash, Y., & Kern, M. J. (2006). Measuring general self-efficacy: A comparison of three measures using item response theory. *Educational and Psychological Measurement*, 66(6), 1047-1063.
- Sijtsma, K. Junker, B.(2006). Item response theory: Past performance, present development, and future expectations. *Behaviormetrika*, 33 (1), 75-102.
- Tunmer, W. E., Chapman, J. W., Greaney, K. T., Prochnow, J. E., & Arrow, A. W. (2013). Why the New Zealand National Literacy Strategy has failed and what can be done about it: Evidence from the Progress in International Reading Literacy Study (PIRLS) 2011 and Reading Recovery monitoring reports. *Australian Journal of Learning Difficulties*, 18(2), 139-180.
- Ueno, M. (2002). An extension of the IRT to a network model. *Behaviormetrika*, 29(1), 59-79.
- Yang, W. L., & Houang, R. T. (1996). *The effects of anchor length and equating method on the accuracy of test equating: Comparisons of linear and IRT-based equating using an anchor-item design*. Paper presented at the annual meeting of the American Educational Research Association, New York.
- Zenisky, A., Hambleton, R., & Sireci, S. (2002). *Effects of local item dependent on the validity of IRT item, test and ability statistics*. Unpublished Phd thesis. University of Massachusetts.
- Zimmerman, L. (2014). Lessons learnt: Observation of Grade 4 reading comprehension teaching in South African schools across the Progress in International Reading Literacy Study (PIRLS) 2006 achievement spectrum. *Reading & Writing-Journal of the Reading Association of South Africa*, 5(1), 1-9.
- Zubairi, A. M., & Kassim, N. L. A. (2016). Classical and Rasch analyses of dichotomously scored reading comprehension test items. *Malaysian Journal of ELT Research*, 2(1), 1-26.

ملحق ١

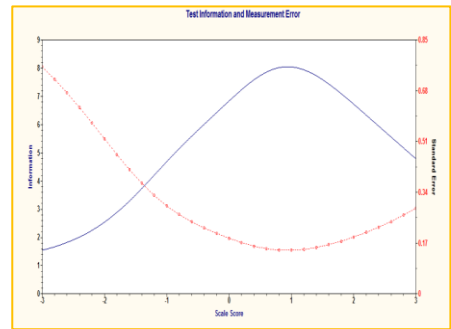
قائمة الأشكال:



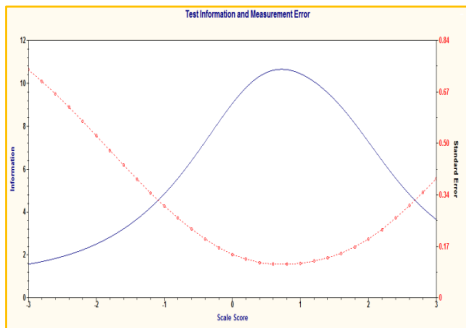
شكل ١: منحني خصائص الكتيبات (١، ٤، ٧، ١٣) باستخدام نموذج الإستجابة المترتبة



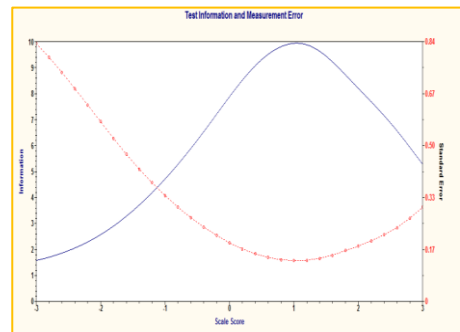
كتيب ٥



كتيب ١



كتيب ٩



كتيب ٦

شكل ٢: منحنيات معلومات الكتيبات (١، ٥، ٦، ٩)