الخصائص السيكومترية لاختبار التنمية المعرفية في مادة العلوم وفقا لنموذج راش لدى طلبة الصف السابع بسلطنة عمان

أ. فاطمة بنت حمد البادية

وزارة التربية والتعليم، سلطنة عمان

أ. د. علي مهدي كاظم **د. راشد بن سيف المحرزي**

قسم علم النفس، كلية التربية، جامعة السلطان قابوس

الملخص: هدفت الدراسة الى تفحص الخصائص السيكومترية لاختبار التنمية المعرفية في مادة العلوم للصف السابع وفق نموذج راش. يتألف الاختبار من ٢٥ مفردة من نوع الاختيار من أربع بدائل. طبق الاختبار على أفراد الدراسة البالغ عددهم ٨٤٨٤ طالبا وطالبة. وللإجابة عن أسئلة الدراسة تمت معايرة مفردات الاختبار باستخدام برنامج Multilog7.03 لتقدير معلمة الصعوبة للمفردات وقدرات الطلبة ودالة معلومات الاختبار وفق طريقة الأرجحية القصوى الهامشية Marginal Maximum Likelihood. وأشارت النتائج الى تحقق افتراضات نموذج راش ومطابقة الاستجابات لجميع مفردات الاختبار لنموذج راش. بلغت قيمة معامل ثبات الاختبار ٠,٨٣ وفق نظرية الاستجابة للمفردة، كما تراوحت فيم معالم صعوبة مفرداته بين ٢,٠٢ الى – ٠,٧٥ لوجيت. كما بينت النتائج أن قيمة دالة المعلومات كانت أقصى ما يمكن عندما كانت ($\theta = b$) وذلك كما يتوقع نموذج راش، وأن الاختبار يقدم أكبر كمية من المعلومات عند مستويات القدرة المتوسطة ويقدم أقل كمية من المعلومات عند مستويات القدرة العالية والمتدنية.

الكلمات المفتاحية: الخصائص السيكومترية، اختبار التنمية المعرفية، نموذج راش، عمان.

Psychometric Properties of Cognitive Development Test in Science using Rasch Model for Seventh Grade in Sultanate of Oman

Fatma Hamad Albadiya

Prof. Ali Mahdi Kazem

Dr. Rashid Saif Al-Mehrzi

Abstract: The study aimed to investigate the psychometric properties of the Cognitive Development Test in science subject for seventh grade using Rasch IRT model. The test contains 25 items with four choices. The test is administered to 8484 male and female students. The test was calibrated using Multilog 7.03 software to estimate item difficulty parameters, students' ability parameters as well as test information function with Marginal Maximum Likelihood estimation method. Results supported the assumption for Rasch model and item fit for all test items. The marginal reliability with IRT was 0.83. The difficulty parameters ranged between 0.75 – 2.02 logit. Results also showed that the highest level of test information was at the average difficulty of the test which was close to the medium ability, whereas it was at minimal for low and high ability levels.

Keywords: Psychometric properties, cognitive ability test, Rasch model, Oman.

المقدمة

يعد القياس والتقويم التربوي أحد المكونات الرئيسية للمنظومة التربوية وأكثرها تأثيرا في تقدمها ورفع كفاءتها، حيث يسهم تطوره اسهاما ايجابيا في اصلاح وتطوير تلك المنظومة، وعليه فقد حظي مجال تطوير نظم وأساليب القياس والتقويم باهتمام كبير من قبل القائمين على تطوير النظم التعليمية. وتعد الاختبارات أحد الوسائل الرئيسية لعمليات التقويم التربوي حيث تستخدم على نطاق واسع للتعرف على مقدار التقدم في مستوى التحصيل الدراسي للطالب ومقياس لمدى تحقيقه للأهداف التعليمية ويعتمد عليها في صنع القرارات التي تتعلق بعمليتي التعلم والتعليم.

وتعتمد صحة القرارات المتخذة على نوع ودقة المعلومات والنتائج والتغذية الراجعة التي تزود الاختبارات بها، وعليه فان ضبط الاختبارات يؤدي إلى ضبط عملية التعلم، إلا أن العديد من الدراسات وتقارير الندوات والاجتماعات المتعددة للجان جمعيات التقويم في العالم بعامة والعالم العربي خاصة تدل على ضعف واضح في معرفة المعلمين بمواصفات الاختبار الجيد ووجود قصور في اعدادها، فهي تفتقر إلى توافر الجوانب الفنية عند بنائها وأثناء عمليات التطبيق، كما

أنها تعاني من الكثير من المشكلات مثل: تقيد درجة الفرد على الاختبار بمفرداته، وتقيد الخصائص السيكومترية للمفردات الاختبارية بخصائص أفراد عينة التقنين، وعدم توافر خطية القياس، وتغير تكوين ومعنى مفردات الاختبار بمرور الزمن وما يترتب على ذلك من آثار سلبية على مصداقية النتائج المستمدة من تلك الاختبارات (الحموري، ٢٠١١).

لقد ظلت النظرية الكلاسيكية للاختبار تشكل الأساس النظري والعملي في العلوم السلوكية لسنوات عديدة مضت، وقدمت اطارا مرجعيا لبناء وتقييم الاختبارات والمقاييس النفسية والتربوية وتفسير درجاتها. وقد استندت هذه النظرية على نموذج بسيط للقياس والذي ينص على أن درجة الفرد الملاحظة في الاختبار تساوي مجموع الدرجة الحقيقية للفرد في السمة المقاسة والدرجة الخطأ في الاختبار، فلكي نقيس سمة معينة ينبغي تقدير المصادر المتعددة للخطأ التي تؤثر في تباين درجات الاختبارات من أجل التحقق من ثبات هذه الدرجات، بحيث يمكن التعميم من أحد استخدامات اسلوب قياس معين إلى ظروف أخرى يمكن أن يستخدم فيها هذا الأسلوب (Stone & Yumoto, 2004).

ومما سبق نجد أن القياس في النظرية الكلاسيكية يتأثر تأثرا ملحوظا بخصائص عينة الأفراد المفحوصين، وبنوعية المعالجة التعليمية وبصعوبة المفردات التي يشتمل عليها الاختبار حيث تعتمد قيم معلمة التمييز ومعلمة الصعوبة على كل من عينة الأفراد وعينة المفردات. فالمفردة التي يختبر بها أفراد متمكنون تبدو سهلة بالنسبة لهم بينما تبدو صعبة للأفراد غير المتمكنين أو الذين لم يتقنوها. فاذا كانت عينة الأفراد متجانسة القدرة نسبيا فان قيم معاملات تمييز المفردات تكون أقل من القيم التي نحصل عليها من عينة غير متجانسة. كما أن درجات الأفراد تعتمد على عينة المفردات التي يختبرون بها فالفرد الذي يختبر بمفردة سهلة تكون درجته أعلى مما لو اختبر بمفردة صعبة، أي أن الأساليب الكلاسيكية في تحليل المفردات تكون محكومة بعينة الأفراد وعينة المفردات.

من هنا بدأ الشعور بالحاجة إلى فلسفة جديدة في القياس، كما بدأت الدعوة إلى تحقيق الموضوعية في القياس السلوكي كما هو الحال في العلوم الفيزيقية، وعليه ظهر اتجاه حديث في القياس النفسي والعقلي والتربوي يتمثل في نظرية السمات الكامنة والتي اطلق عليها نظرية الاستجابة للمفردة.

وتعد نظرية الاستجابة للمفردة طريقة مشهورة لنمذجة البيانات، بمعنى أنها تحاول نمذجة العلاقة بين المتغير غير الملاحظ (قدرة الفرد)، واحتمال استجابة الممتحن صوابا على مفردة ما (المتغير الملاحظ)، ويتم فيها التركيز على المفردة الاختبارية بدلا من التركيز على الاختبار ككل، كما يمكن من خلالها فصل خصائص المختبرين عن خصائص الاختبار وهو ما يحقق الصدق والدقة في الحكم على النتائج المستمدة من النظرية، كما أن النظرية تقوم على جمع متغيرين في

علاقة رياضية وتمثل نموذج احتمالي لأن طبيعة البيانات تسمح بالتنبؤ باحتمال النجاح على أي مفردة من المفردات المدرجة، كما توفر تقديرات لقدرات الأفراد بوحدات النموذج المستخدم (كاظم، ١٩٨٨).

وتقوم هذه النظرية على مجموعة من الافتراضات الأساسية هي: أحادية البعد وخاصية الاستقلال الموضعي ووتيرية منحنى خاصية المفردة. ويعتبر افتراض أحادية البعد شيئا مرغوبا فيه عند جميع مطوري الاختبارات؛ وذلك لأجل تطوير عملية التفسير لدرجات الاختبارات وتعزيزها. والمقصود به وجود عامل واحد سائد على العوامل الأخرى، بحيث يكون هذا العامل هو السمة التي يقيسها الاختبار. ويقصد بالاستقلال الموضعي أن تكون استجابات الفرد للمفردات المختلفة مستقلة استقلالا احصائيا عند مستوى قدرة معين، وحتى يكون متحققا يجب ان لا تؤثر استجابة المفحوص عن مفردة ما في الاختبار سلبا او ايجابا في استجابته للمفردات الأخرى. اما ما يتعلق بالافتراض الثالث فهو يوفر احتمالية اجابة المفحوصين عن المفردة في مستويات القدرة المختلفة اجابة صحيحة، وكون المنحنى تراكميا فانه يشير الى ان احتمال اجابة المفردة اجابة صحيحة يزداد بازدياد قدرة المفحوص. أما فيما يتعلق بسرعة الأداء فتفترض نماذج النظرية أن عامل السرعة لا يلعب دورا ف ي الاجابة على مفردات الاختبار (حجازي والشريفين، ٢٠١٢).

ويندرج تحت نظرية الاستجابة للمفردة مجموعة من النماذج الرياضية تعرف باسم نماذج الاستجابة للمفردة Item Response Models وهذه النماذج تحقق القياس الموضوعي، وتعالج كثير من مشكلات القياس النظرية والتطبيقية التي عجزت النظرية التقليدية عن مواجهتها، ولعل أكثر هذه النماذج استخداما وانتشارا في مجال القياس والتقويم النفسي والتربوي نموذج راش أحادي المعلم، والنموذج ثنائي المعالم، والنموذج ثلاثي المعالم (الناغي، ٢٠١١).

ويعد نموذج راش من أبسط نماذج نظرية الاستجابة للمفردة لأنه يشتمل على معلم واحد هو معلم صعوبة المفردة لتمثيل الفروق بين المفردات. ولقد قام راش ببناء نموذجه لتفسير أداء الممتحين على الاختبارات، ولقد استمد نموذجة من نقطة بسيطة وهي ان احتمال نجاح شخص في ادائه على مفردة اختبارية معينة يجب أن يعتمد على الفرق بين محصلة قدرة هذا الفرد وعلى صعوبة هذه المفردة (Hwang, 2002).

ويعبر عنه بالعلاقة الرياضية التالية التي تقيس احتمال اجابة الفرد ذي القدرة θ عن المفردة i.

$$P_i(\theta_v) = e(\theta_v - b_i) \Big/ 1 + e(\theta_v - b_i)$$

فعندما يشرع الفرد (v) في الاستجابة لمفردة (i) فان كلا من مستوى القدرة (θ_v) وصعوبة المفردة (b_i) تعبران عن موضع هذا الفرد على متغير المتصل المقاس وتتحكمان في احتمال حدوث الاستجابة المناسبة على المفردات المتدرجة على متصل هذا المتغير $P_i(\theta_v)$ ، وتكون هذه الاحتمالية محكومة بالفرق بين معلم قدرة الفرد وصعوبة المفرد($\theta_v - b_i$)، فاذا كان مستوى قدرة الفرد أكبر من مستوى صعوبة المفردة، كان الفرق أكبر من الصفر وعندئذ يكون احتمال حدوث الاستجابة الصواب أكبر من ٠,٥ ، أما اذا كان مستوى قدرة الفرد أقل من مستوى صعوبة المفردة فان الفرق يكون أقل من الصفر وعندئذ يكون احتمال حدوث الاستجابة الصواب أقل من ٠,٥، أما في حالة تساوي قدرة الفرد مع صعوبة المفردة فان الفرق يكون مساويا للصفر، وعندئذ يكون احتمال حدوث الاستجابة الصواب مساويا لاحتمال حدوث الاستجابة الخطأ ويساوي كل منهما ٠,٥ (Peter & Stephen, 2004).

وقد قامت العديد من الدراسات باستقصاء الخصائص السيكومترية للاختبارات التحصيلية وفق نموذج راش نعرض منها التالي:

قام الشربيني وحجازي (٢٠٠٢) بدراسة هدفت الى تقويم تحصيل تلاميذ مرحلة التعليم الاساسي في العلوم باستخدام نظرية الاستجابة للمفردة، حيث استخدم الباحثان نموذج راش في اعداد وتدريج مفردات بنك أسئلة يتكون من ٢٢٤ مفردة في وحدة المادة للطلبة من الصف الرابع الابتدائي الى الصف الثالث الاعدادي. تكونت عينة الدراسة من ٧٥٤ طالبا وطالبة من مدارس مدينة القاهرة. واستخدم الباحثان برنامج RUMM 2010 في التأكد من مطابقة البيانات لنموذج راش وأوصيا في نهاية دراستهما باستخدام نظرية الاستجابة للمفردة في اعداد اختبارات تشخيصية تكشف عن أهم الصعوبات التي تؤدي الى انخفاض مستوى الطلبة في المادة.

هدفت دراسة الشريفين (٢٠٠٦) الى تقدير الخصائص السيكومترية لاختبار محكي المرجع في القياس والتقويم التربوي وفق نموذج راش. ولتحقيق هدف الدراسة تم بناء اختبار تحصيلي مؤلف من ٥٠ مفردة من نوع الاختيار من أربعة بدائل. طبق الاختبار على أفراد الدراسة البالغ عددهم٢٢٢ طالبا وطالبة. وأشارت النتائج الى مطابقة الاستجابة عن ٣٧ مفردة لافتراضات النموذج، وحذف ١٣ غير مطابقة، وبلغ معامل ثبات الأفراد٠,٩٠ ومعامل ثبات المفردات ٠,٩٥. وتم التحقق من تمتع الاختبار بمظاهر متعددة من الـ ـ ، ما بينت النتائج أن قيمة دالة معلومات الاختبار كانت اقصى ما يمكن عندما كانت($\theta_v = b_i$) وان الاختبار يقدم أكبر كمية من المعلومات عند مستويات القدرة المتوسطة وأقل كمية من المعلومات عند مستويات القدرة المتدنية والعالية.

كما هدفت دراسة بيتر وآهان وستيفن (Peter, Ahan & Stephen, 2004) الى مناقشة التطورات الحديثة في مجال السيكومترية خصوصا تطبيقات تحليلات راش لعملية تطوير الأداة.

وتؤكد على أهمية الخصائص السيكومترية للمقاييس الأساسية. وتقدم الورقة استخدام نموذج راش في تحليل المقاييس والأدوات وتبين كيفية تطبيقه في تطوير الاحتياجات المعتمدة على القياسات. ويؤكد البحث على قدرة نموذج راش في اعطاء قياسات أحادية وان هذه المقاييس تكون معتمدة على نفس نموذج القياس.

وفي دراسة فرانك جوني وآخرون (Franchignoni et. al, 2011) والتي هدفت إلى تحليل فقرات اختبار داش (DASH) في ايطاليا؛ وفق النظرية الكلاسيكية في القياس ونموذج راش أحادي المعلمة، تكونت عينة الدراسة من٢٨٣ مريضًا. أظهرت نتائج الدراسة مطابقة عشرة مفردات لنموذج راش أحادي المعلمة، ودلت النتائج على فاعلية نموذج راش في مطابقة المفردات بشكل أفضل من التحليل العاملي وفق النظرية الكلاسيكية في القياس.

وفي دراسة حموري وصباح (٢٠١١) التي هدفت لتحليل الاختبار الوطني الاردني للتحكم بنوعية تعليم العلوم باستخدام نموذج راش، فقد تم تطبيق الاختبار على عينة عشوائية من طلبة الصف العاشر بلغ عددها٤١٥٥٦ في العام الدراسي ٢٠٠٥-٢٠٠٦، وتم حفظ النتائج ببنك البيانات، ولتحقيق هدف الدراسة تم اختيار عينة عشوائية من البنك بلغت ١٥٠ مشاركا وتكون الاختبار من ٣٠ مفردة من نوع الاختيار من متعدد بأربعة بدائل، وتم تحليل البيانات باستخدام برنامج WINSTEPS وبرنامج RUMM، وأظهرت النتائج أن هناك مشكلة في تناسب مفردتين للنموذج، وأن هناك ٣ مفردات خارج حدود المطابقة المتوقعة، وتراوحت صعوبة المفردات بين- ١,٢١ الى ١,٤ لوجيت، وقد تم حساب ثبات المفردات٠,٨٩ وثبات الأفراد٠,٧٣ وهذ يعني أن تقدير صعوبة المفردات أكثر دقة من تقدير قدرة الأفراد.

هدفت دراسة الخياط (٢٠١٢) إلى التحقق من فاعلية النموذج اللوغارتمي ذي المعلمة الواحدة "نموذج راش" ومطابقة البيانات للنموذج لاختبارTIMSS والذي يقيس المهارات الرياضية لمستوى الصف الثامن الأساسي، ولتحقيق هدف الدراسة تم تطبيق مفردات الاختبار على عينة من طلبة الصف الثامن الأساسي بلغ عددها ٥٩٩ طالبا وطالبة؛ بعد إيجاد الخصائص السيكومترية (الصدق والثبات) لاختبار TIMSS. وللإجابة على أسئلة الدراسة تم استخدام البرنامج الإحصائي RUMM، وأشارت نتائج الدراسة إلى تمتع الاختبار المقدم بالصدق والثبات، حيث بلغت قيمة معامل الثبات للاختبار باستخدام معامل كرونباخ الفا ٠,٩٠، أما معامل الصدق بدلالة المحك بين درجات الطلبة على الاختبار ودرجاتهم في مادة الرياضيات فقد بلغت قيمته٠,٨٧، وأشارت النتائج أيضا إلى مناسبة ٢٥ مفردة من مفردات الاختبار والبالغ عدده ٣٥ مفردة للبيئة المحلية بحيث تمتعت بخصائص إحصائية مناسبة، من حيث معامل صعوبة المفردة.

من جانب آخر لاحظت الباحثة عدم وجود دراسات محلية استخدمت نظرية الاستجابة للمفردة في تقصي الخصائص السيكومترية للاختبارات بشكل عام ولاختبارات برنامج التنمية

المعرفية بشكل خاص بالرغم من أهميتها، في حين أن هناك بعض الدراسات تناولت تقويم وأثر برنامج التنمية المعرفية باعتباره برنامجا فاعلا يمكن من خلاله الارتقاء بمواد العلوم والرياضيات نعرض منها الدراسات التالية:

حيث هدفت دراسة الهشامية (٢٠١٠) الى التعرف على أثر برنامج التنمية المعرفية على التحصيل الدراسي للطالبات في مادة الرياضيات ودوره في الانماء المهني لهن وذلك من خلال تطبيق برنامج التنمية المعرفية على عينة الدراسة المكونة من ٢٩٦ طالبة، وقياس أثر التغير في المعلومات والمهارات الرياضية قبل وبعد تطبيق البرنامج من خلال الاختبار التحصيلي القبلي والبعدي المعدين للدراسة، وأظهرت النتائج أن هناك ارتفاعا كبيرا في المستوى للطالبات، وأن هناك توسع في معارفهن ومعلوماتهن، كما دلت النتائج حماسا واضحا وتفاعلا مستمرا نحو المشاركة في البرامج المطبقة للبرنامج.

وهدفت دراسة تيمورة (٢٠١٢) الى تقويم برنامج التنمية المعرفية من وجهة نظر الطلاب والمعلمين والمشرفين التربويين بمحافظة مسقط، حيث تم تطبيق استبانة على ٧٤٧ من طلبة الصفوف الثامن والتاسع والعاشر و٤٥ من المعلمين والمشرفين التربويين، وأسفرت النتائج الى وجود فروق دالة احصائيا في أهداف البرنامج بين وجهات نظر الطلبة وفقا لمتغيري النوع والصف الدراسي وأن هناك معوقات وتحديات تواجه البرنامج .

كما هدفت دراسة الهنائي (٢٠١٧) الى تقويم أسئلة اختبار الرياضيات في برنامج التنمية المعرفية خلال ثلاث سنوات متتالية، حيث تكونت عينة الدراسة من عينة عشوائية من استجابات أوراق الطلبة على اسئلة اختبار الرياضيات في البرنامج، وقد تم حساب الصدق المحك وذلك بمقارنة نتائج الطلبة في اختبار البرنامج مع تحصيلهم الدراسي لاختبار الرياضيات من كل عام دراسي، وقد بلغ معامل الارتباط بيرسون على التوالي (٠,٤٩)، (٠,٥٤)، (٠,٥٥) اما بالنسبة لثبات الاتساق الداخلي لدرجات الاختبارات فقد تم حسابه بطريقة الفا كرونباخ حيث بلغ على التوالي(٠,٥٥)، (٠,٦٠)، (٠,٦٥) وهي تدل على درجة ثبات (اتساق داخلي) أقل من المتوقع التي تقيس قدرات معرفية وأظهرت نتائج تحليل التباين الاحادي الى وجود فروق ذات دلالة احصائية بين الطلبة في المستوى التحصيل الدراسي لدرجات اختبارات التنمية المعرفية في مادة الرياضيات باختلاف النوع والعام الدراسي.

مشكلة الدراسة

يغلب على الاختبارات والمقاييس المستخدمة في مجال العلوم الانسانية في البيئة العربية اعتمادها عند بناءها وتقنينها على نظرية القياس التقليدية بالرغم من العيوب المرتبطة بها، ومن جانب آخر فان المتابع لما تحويه المؤسسات التربوية العربية، يلاحظ خلوها من اختبارات تحصيلية

مبنية لغرض معين وفق نظرية الاستجابة للمفردة، على الرغم من توافر البرمجيات الإحصائية المتعلقة بالقياس وفق هذه النظرية.

من جانب آخر أشارت نتائج الدراسة الدولية في الرياضيات والعلوم في الأعوام ٢٠٠٧ و ٢٠١١ و ٢٠١٥ الى انخفاض المستوى التحصيلي، حيث حصلت السلطنة على مراكز متأخرة في المادتين، فجاءت في المركز ٣٦ في مادة العلوم ضمن ٤٢ دولة مشاركة ولم تتجاوز المتوسط الدولي البالغ ٥٠٠ نقطة كما أشار الهنائي(٢٠١٦). لذلك جاء تطبيق برنامج التنمية المعرفية في مواد العلوم والرياضيات والمفاهيم الجغرافية بناء على التوجيهات السامية للنهوض بالمستوى التحصيلي للطلبة في الجانبين العلمي والمعرفي. وقد أشار الصارمي (٢٠١١) والهشامية (٢٠١٠) الى أهمية اختبارات البرنامج كونها تعد المجهر التربوي للمدارس والقائمين على العملية التربوية، غير أن نتائج هذه الاختبارات لم يتم تحليلها والوقوف عليها سواء من خلال النظرية الكلاسيكية أو النظرية الحديثة للقياس، الأمر الذي يفتح أبواب للباحثين التربويين المتخصصين في القياس والتقويم للتطرق إلى تحليل هذه الاختبارات والتعرف على خصائصها السيكومترية.

أسئلة الدراسة

١. ما درجة تحقيق نتائج اختبار التنمية المعرفية في مادة العلوم لافتراضات نظرية الاستجابة للمفردة؟
٢. ما درجة ملائمة نموذج راش لاستجابات الطلبة في اختبار التنمية المعرفية في مادة العلوم؟
٣. ما قيم تقديرات معالم مفردات اختبار التنمية المعرفية في مادة العلوم؟
٤. ما مقدار المعلومات التي يقدمها منحنى معلومات اختبار التنمية المعرفية في مادة العلوم لقدرات الطلبة؟

أهمية الدراسة

على الرغم من أن بعض البحوث قد تناولت استخدام نماذج نظرية الاستجابة للمفردة في تدريج الاختبارات النفسية والتربوية إلا أن الدراسة الحالية هي الدراسة المحلية الأولى في حدود علم الباحثة التي تسلط الضوء على استخدام نظرية الاستجابة للمفردة في تحليل اختبارات وطنية مثل اختبار التنمية المعرفية في مواد العلوم والرياضيات، كما أن أهميتها تظهر من خلال عدة اعتبارات وهي:

- يعتبر التحليل مطلبا لإثراء برنامج التنمية المعرفية لأن بواسطته يتم الحصول على التغذية الراجعة التي تسهم في تحديد جوانب القوة والضعف في أدوات البرنامج وبالتالي تطويره وتحقيقه للأهداف المرجوة منه.

- الكشف عن جودة اختبارات مادة العلوم المطبقة في البرنامج وكذلك ايجاد الخصائص السيكومترية لها مما يساعد في تطويرها وكذلك الثقة في نتائجها.
- تسهم في تجويد اختبارات العلوم والرياضيات في البرنامج، والاستعداد الامثل للمشاركات الاقليمية والدولية كالدراسة الدولية TIMSS.
- تساعد الجهات المختصة في اتخاذ قرارات حول كيفية اعداد وتطبيق اسئلة اختبارات مادة العلوم ضمن المواصفات المطبقة حاليا وحسب الخصائص السيكومترية التي تتمتع بها هذه الاختبارات.

حدود الدراسة

تقتصر الدراسة الحالية على بيانات طلبة الصف السابع بمدارس التعليم العام والأساسي بمحافظة شمال الباطنة بسلطنة عمان الذين طبق عليهم اختبار برنامج التنمية المعرفية في مادة العلوم للعام الدراسي ٢٠١٦/٢٠١٧م. كما تقتصر الدراسة الحالية على استخدام أحد نماذج الاستجابة للمفردة أحادية البعد وهو نموذج راش، واستخدام برنامج Multilog7.03 لتقدير قدرة الفرد ومعلمة صعوبة المفردة وفق طريقة الأرجحية القصوى الهامشية Marginal Maximum likelihood.

مصطلحات الدراسة

- نموذج راش Rasch Model: "أحد نماذج نظرية الاستجابة للمفردة الذي يفترض اختلاف مفردات الاختبار في صعوبتها وتتساوى في تمييزها، وتخمينها مساوي للصفر" (علام، ٢٠٠٥).
- صعوبة المفردة Item Difficulty: "أحد معالم دالة الاستجابة للمفردة وتعادل قيمة القدرة التي تناظر احتمال ٥٠% للإجابة الصحيحة عندما تكون قيمة التخمين تساوي صفراً"(Hambelton & Swaminathan, 1985).
- الخصائص السيكومترية: ويقصد بها معالم الصعوبة والتمييز والتخمين للمفردة. والصدق والثبات للاختبار. وفي هذه الدراسة اقتصرت الخصائص السيكومترية على معلم الصعوبة للمفردة والصدق والثبات للاختبار (علام، ٢٠٠٥).
- دالة معلومات المفردة Item Information Function: "هو اقتران رياضي يعبر عن كمية المعلومات التي تمدها المفردة عند كل مستوى من مستويات القدرة " (عبد الوهاب، ٢٠١٣، ص١٣٢٥).
- دالة معلومات الاختبار Test Information Function: "هو اقتران رياضي يعبر عن مجموع دوال المعلومات لمفردات الاختبار" (الزبون، ٢٠١٣، ص١٣٢٥).

- برنامج التنمية المعرفية: "هو برنامج تربوي يتم تنفيذه في مدارس السلطنة (٥-١٠) ابتداء من العام الدراسي ٢٠٠٧/٢٠٠٨م، ويهدف إلى تطوير تعليم وتعلم مواد العلوم الشفهية، والاختبارات التحريرية، والمشاريع والبرامج الداعمة" (وزارة التربية والتعليم، ٢٠١٣، ص٤).

مجتمع الدراسة

يتكون مجتمع الدراسة من جميع طلبة الصف السابع والبالغ عددهم الكلي (٤٨٥٢١) في المدارس الحكومية التي يطبق فيها برنامج التنمية المعرفية في سلطنة عمان خلال العام الدراسي(٢٠١٦/٢٠١٧).

عينة الدراسة

تكونت عينة الدراسة من جميع طلبة الصف السابع بمحافظة شمال الباطنة والذين خضعوا لاختبار التنمية المعرفية في العام الدراسي (٢٠١٦/٢٠١٧) والبالغ عددهم (٨٤٨٤) طالبا وطالبة. تتنوع المناطق في محافظة شمال الباطنة بين الريف والحضر والمستوى المعيشي والاقتصادي والاجتماعي، بحيث يمكن اعتبار محافظة شمال الباطنة أحد المحافظات التي تمثل المجتمع الكلي في سلطنة عمان.

أداة الدراسة

استخدمت الدراسة اختبار التنمية المعرفية في مادة العلوم لطلبة الصف السابع الأساسي للعام الدراسي ٢٠١٦/٢٠١٧م المعد من قبل المديرية العامة للتقويم التربوي بوزارة التربية والتعليم، حيث تم إعداد الاختبار من قبل فرق عمل تضم مختصين من التقويم التربوي والمناهج والإشراف التربوي بالوزارة، وتم تطبيقه خلال الفصل الدراسي الثاني على جميع طلبة الصف السابع بالمدارس الحكومية بالسلطنة، ويتضمن الاختبار ٢٥ مفردة من نمط الاختيار من أربعة بدائل لكل مفردة ٤ درجات. يركز الاختبار على أربعة مجالات (الأحياء، الكيمياء، الفيزياء، المفاهيم الجغرافيا البيئية)، ويشير الجدول (١) إلى مواصفات بناء الورقة الاختبارية (المديرية العامة للتقويم التربوي، ٢٠١٣).

جدول (١)

مواصفات الورقة الاختبارية لمادة العلوم

المجال	معرفة٣٢%	تطبيق٤٠%	استدلال٢٨%	عدد المفردات
الأحياء٣٢%	٣	٣	٢	٨
كيمياء٢٠%	١	٢	٢	٥
فيزياء٢٨%	٢	٣	٢	٧
المفاهيم البيئية٢٠%	٢	٢	١	٥
عدد المفردات	٨	١٠	٧	٢٥

ويتم استخدام الشفرة الالكترونية (الباركود) في دفتر إجابات الطلبة واستمارة الرصد الآلي لضمان سرية البيانات أثناء عملية التصحيح، ودقة الرصد وسرعته، ويتم قراءة إجابات الطلبة ورصدها آليا، لاستخراج النتائج واجراء التحليلات الإحصائية المناسبة.

صدق وثبات أداة الدراسة

بلغ معامل ثبات الاختبار باستخدام معامل ثبات ألفا كرونباخ ٠,٧٥٠ وهو معامل ثبات جيد، كما تم حساب ثبات الاختبار وفق نظرية الاستجابة للمفردة وكات قيمته ٠,٨١٠ مما يؤكد أنه يتمتع بدرجة عالية من الثبات. كما تم حساب معامل الارتباط بين كل مكون من مكونات الاختبار والمجموع الكلي، ويبين جدول (٢) معامل الارتباط بين كل مكون مع الدرجة الكلية للمقياس.

جدول (٢)

معاملات الارتباط بين الأبعاد المختلفة لاختبار التنمية المعرفية مع الدرجة الكلية للمقياس

البعد	معامل الارتباط
الأحياء	٠,٧٨٤*
الكيمياء	٠,٥١٣*
الفيزياء	٠,٦٤٦*
المفاهيم البيئية	٠,٤٨٤*

*دال عند المستوى٠,٠٥

كما تم التحقق من الصدق العاملي للأداة من خلال استخدام التحليل العاملي للمكونات الرئيسية على اجابات الأفراد باستخدام برنامج SPSS وتبين من ذلك سيادة عامل رئيسي واحد في الاختبار. إضافة الى ذلك فإن الإجراءات التي اعتمدت في تحديد مجالات الاختبار وتحليل محتوى كتابة الاهداف السلوكية وكتابة مفردات الاختبار والتأكد من صحة الاجراءات من خلال الاحكام

والتقديرات والتوافق الذي برز من خلال لجان وفرق اعداد الاختبار يعتبر مؤشر آخر على صدق الاختبار (صدق المحتوى).

إجراءات الدراسة

- تم التواصل مع المديرية العامة للتقويم التربوي للحصول على نتائج اختبار التنمية المعرفية لمادة العلوم بعد عملية تصحيح وتحليل الاختبار الكترونيا على مستوى الوزارة.
- تم حذف استجابة ٣٥٥ من الطلبة الذين حصلوا على درجة صفر في جميع المفردات، كما تم التحقق من عدم وجود مفردة أجاب عليها جميع الطلبة اجابة صحيحة أو خطأ وعليه لم يتم استبعاد أي مفردة من مفردات الاختبار.
- تم فحص افتراض أحادية البعد (Unidimensionality) للاستجابات باستخدام التحليل العاملي اعتمادا على طريقة المكونات الأساسية.
- تم التحقق من جودة مطابقة البيانات للنموذج.
- تمت معايرة مفردات الاختبار باستخدام برنامج Multilog7.03 وتقدير معلمة الصعوبة للمفردات وقدرات الطلبة ودالة معلومات الاختبار وفق طريقة الأرجحية القصوى الهامشية Marginal Maximum likelihood.

المعالجة الإحصائية

للإجابة عن أسئلة الدراسة قامت الباحثة باستخدام الاساليب الاحصائية والبرمجيات التالية:

- للإجابة على السؤال الأول تم استخدام برنامج SPSS لإجراء التحليل العاملي الاستكشافي للمفردات بطريقة المكونات الأساسية Principal Component للتحقق من أحادية البعد، كما تم استخدام مؤشر Q3 للتحقق من افتراض الاستقلال الموضعي.
- للإجابة على السؤال الثاني تم استخدام اختبار حسن المطابقة مربع كاي(χ^2) للتحقق من مدى ملائمة البيانات للنموذج.
- للإجابة على السؤال الثالث تم تقدير معلمة صعوبة المفردات باستخدام برنامج Multilog7.03.
- للإجابة على السؤال الرابع تم تقدير دالة المعلومات للاختبار باستخدام برنامج Multilog7.03 عند كل مستوى من مستويات القدرة.

نتائج الدراسة

أولا: الاجابة عن السؤال الأول "ما درجة تحقيق نتائج اختبار التنمية المعرفية في مادة العلوم لافتراضات نظرية الاستجابة للمفردة؟"

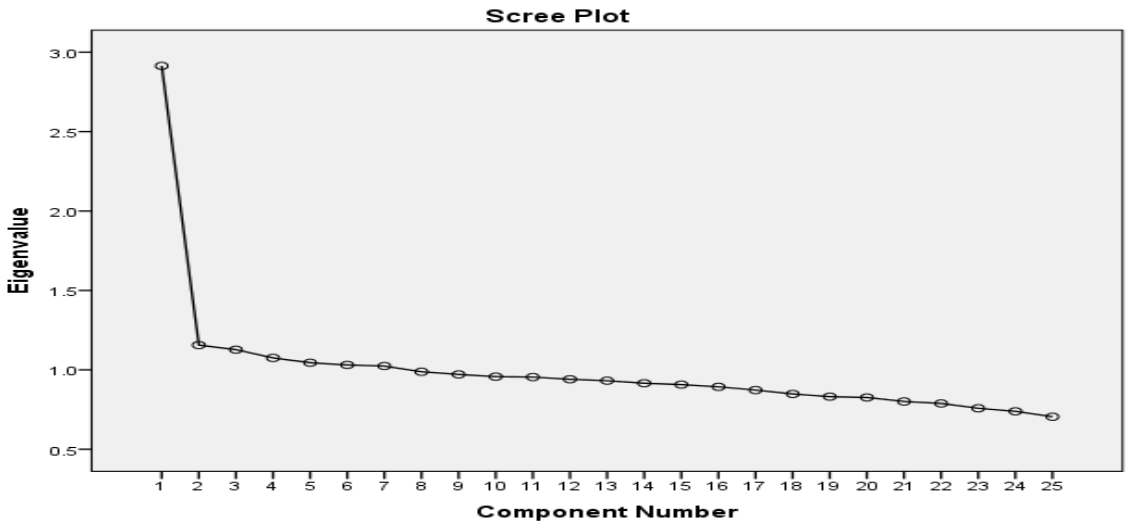
تم التحقق من افتراض أحادية البعد باستخدام التحليل العاملي للمكونات الرئيسية على اجابات أفراد العينة الكلية باستخدام برنامج SPSS. ويبين الجدول (٣) أن هناك ٧ عوامل الجذر الكامن لكل منها أكبر من ١ ويفسرون معا ٣٧,٤٩% من التباين؛ وأن العامل الأول له جذر كامن قيمته٢,٩١ وتفسر ١١,٦٥% وهي قيمة كبيرة مقارنة ببقية الجذور وبالتالي يمكن اعتباره عاملا سائدا، كما أن نسبة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني أكبر من ٢، وتعد هذه القيم كما يشير (Hambleton & Swaminathan, 1985)، ولورد (Lord, 1980) مؤشرات على أحادية البعد.

جدول (٣).

نتائج التحليل العاملي الاستكشافي لبيانات اختبار التنمية المعرفية في مادة العلوم

رقم العامل	الجذر الكامن	نسبة التباين المفسر	نسبة التباين المفسر التراكمي
١	٢,٩١٤	١١,٦٥٤	١١,٦٥٤
٢	١,١٥٦	٤,٦٢٤	١٦,٢٧٨
٣	١,١٢٧	٤,٥٠٨	٢٠,٧٨٦
٤	١,٠٧٥	٤,٣٠١	٢٥,٠٨٧
٥	١,٠٤٥	٤,١٧٩	٢٩,٢٦٦
٦	١,٠٣١	٤,١٢٤	٣٣,٣٨٩
٧	١,٠٢٤	٤,٠٩٦	٣٧,٤٨٥

واذا تم تمثيل الجذور الكامنة للعوامل المكونة للاختبار بيانيا (scree plot) كما هو موضح في الشكل (١) يتضح ارتفاع قيمة الجذر الكامن للعامل الاول مقارنة ببقية العوامل، وهو يؤكد على وجود عامل سائد يعبر عن السمة المقاسة، حيث نلاحظ انحناء الرسم بسرعة من منحنى يعامد تقريبا المحور السيني إلى منحنى أفقي تقريبا بعد العامل الأول.



شكل (١) الجذور الكامنة لاختبار التنمية المعرفية في مادة العلوم

ينظر الكثير من الباحثين الى أن تحقق افتراض الاستقلال الموضعي هو نتيجة محصلة لتحقق افتراض أحادية البعد، وأنه افتراض مكافيء له كما أشار وهامبلتون وجونز (Hambleton & Jones, 1994) وهامبيلتون وسوامينثان (Hambleton & Swaminthan, 1985) والحموري (٢٠١١)، في حين يرى البعض أن الاستقلال الموضعي وأحادية البعد ليسا مفهومًا واحدًا فالاختبار قد يقيس سمتان كامنتان، ولكن تكون المفردات مستقلة بالنسبة للأفراد المتجانسين في كل من هذين البعدين (علام، ٢٠٠٥).

وللتحقق بشكل خاص من افتراض الاستقلال الموضعي لمفردات الاختبار تم استخدام مؤشر (Q3)، حيث اقترحت ين (Yen, 1984) المؤشر Q3 لقياس الارتباط الموضعي بين مفردات الاختبار، وهو مؤشر يعبر عن العلاقة بين البواقي لزوج من المفردات بعد ضبط السمة المقاسة، ويتبع توزيع المؤشر توزيعاً عينياً بمتوسط حسابي (-1/ n-1) وانحراف معياري قدره (1/n-3)، حيث (n) عدد المفردات. وتم استخدام المعادلة التالية (البواقي= درجة الطالب في المفردة– الدرجة المتوقعة) لجميع مفردات الامتحان في برنامج الأكسل، ومن ثم تم حساب معاملات الارتباط للبواقي بمؤشر(Q3) لكل زوج من مفردات بواسطة برنامج الرزم الاحصائية SPSS، وجاء متوسط هذه الارتباطات مساوية (−٠,٠٠١)، كما أن معاملات الارتباط للبواقي كانت أقل من (٠,٠٥)، مما يدل على تحقق افتراض الاستقلال الموضعي، وبالتالي يمكننا القول أن افتراضية الاستقلال الموضعي قد تحقق لمفردات اختبار مادة العلوم باستخدام مؤشر (Q3).

كذلك تم التحقق من افتراض السرعة في الأداء، حيث أنهى جميع الطلبة الاجابة على مفردات الاختبار ضمن الوقت المحدد، بمعنى ان اخفاق الطلبة في اجابة مفردات الاختبار عائد إلى انخفاض قدرتهم وليس تأثير عامل السرعة في اجابتهم.

ثانيا: الاجابة عن السؤال الثاني للدراسة "ما درجة ملائمة نموذج راش لبيانات اختبار التنمية المعرفية في مادة العلوم؟"

للإجابة على السؤال الثاني تم فحص جودة مطابقة المفردات لنموذج راش، وقد تم استخدام اختبار كاي تربيع لنسبة الترجيح likelihood-ratio chi-squared test لاختبار الاختلاف بين الأنماط المتوقعة لاستجابة المفردة والاستجابة الفعلية لها. وتبين من خلال النتائج أن قيمة مؤشر النموذج ككل والبالغة ١,٠٤ أصغر من قيمة مربع كاي الحرجة والبالغة ٣,٨٤١ عند درجة حرية ١ ومستوى دلالة ٠,٠٥، وهي قيمة غير دالة احصائيا الأمر الذي يعني أن هذا النموذج مطابق للبيانات. كما أظهرت النتائج مطابقة جميع المفردات للنموذج، والجدول (٤) يوضح قيمة مؤشر النموذج ككل وكل مفردة على حدة باستخدام نموذج راش للعينة الكلية.

جدول (٤)

قيم مؤشر مربع كاي لمطابقة مفردات اختبار التنمية المعرفية وللنموذج ككل وفق نموذج راش للعينة الكلية

المفردة	المؤشر	المفردة	المؤشر	المفردة	المؤشر	المفردة	المؤشر	المفردة	المؤشر
١	١,٢٠٩	٦	١,٢٠٧	١١	١,٠٥٦	١٦	١,٣٨٦	٢١	١,١٨٧
٢	١,٢١٤	٧	٠,٩٩١	١٢	١,٠٠١	١٧	٠,٨٥٥	٢٢	٠,٩٨٥
٣	١,٤١٦	٨	٠,٩١٤	١٣	٠,٧٨٠	١٨	١,٠١٩	٢٣	١,١٤٦
٤	١,٥٩٧	٩	٠,٨٧٢	١٤	٠,٩١٦	١٩	١,٠٤٩	٢٤	٠,٧٩٨
٥	٠,٩٥٣	١٠	٠,٧٨٢	١٥	١,٠٨٧	٢٠	٠,٧٢٠	٢٥	٠,٩١٧
مؤشر النموذج ككل									١,٠٤

ثالثا: ما قيم تقديرات معالم مفردات اختبار التنمية المعرفية في مادة العلوم؟

للإجابة على السؤال تم استخدام برنامج Multilog7.03 لتقدير معالم نموذج راش المتمثلة في صعوبة المفردات الاختبارية مقدرة بوحدة اللوجيت والخطأ المعياري لتقديرها، والجدول (٥) يوضح هذه التقديرات.

جدول (٥)

تقديرات معلمة الصعوبة والخطأ المعياري لمفردات اختبار التنمية المعرفية في مادة العلوم

المفردة	معلمة الصعوبة	الخطأ المعياري	المفردة	معلمة الصعوبة	الخطأ المعياري
١	٠,١٣	٠,٠٣	١٤	١,٠٨	٠,٠٣
٢	٠,١٥	٠,٠٣	١٥	٠,٣٥	٠,٠٣
٣	-٠,٣٩	٠,٠٣	١٦	-٠,٣٧	٠,٠٣
٤	-٠,٧٥	٠,٠٣	١٧	١,٣٤	٠,٠٣
٥	٠,٥٧	٠,٠٣	١٨	٠,٧٨	٠,٠٣
٦	٠,٠٧	٠,٠٣	١٩	٠,٥٠	٠,٠٣
٧	٠,٧٤	٠,٠٣	٢٠	٢,٠٢	٠,٠٤
٨	٠,٨٨	٠,٠٣	٢١	٠,٣٤	٠,٠٣
٩	١,٠٩	٠,٠٣	٢٢	١,٠٢	٠,٠٣
١٠	١,٤٦	٠,٠٤	٢٣	٠,٣٦	٠,٠٣
١١	٠,٧٥	٠,٠٣	٢٤	١,٤٦	٠,٠٤
١٢	٠,٧٦	٠,٠٣	٢٥	١,٠٤	٠,٠٣
١٣	١,٤٢	٠,٠٤			

يلاحظ من الجدول أن قيم التقديرات لصعوبة المفردات قد تراوحت بين (٢,٠٢) للمفردة رقم ٢٠ الى (-٠,٧٥) لوجيت للمفردة رقم ٤، مما يعني أن أسهل مفردة كانت المفردة ٤ وأن أصعب مفردة كانت المفردة ٢٠، كما بلغ متوسط الاخطاء المعيارية للمفردات (٠,٠٣) وهي قيمة متدنية تشير الى دقة وثبات تقديرات صعوبة المفردات. كما توزعت معاملات الصعوبة بمتوسط قدره (٠,٦٧) لوجيت وانحراف معياري قدره (٠,٦٥) لوجيت، مما يشير الى صعوبة أغلب المفردات على الطلبة.

رابعا: ما مقدار المعلومات التي يقدمها منحنى معلومات اختبار التنمية المعرفية في مادة العلوم لقدرات الطلبة؟

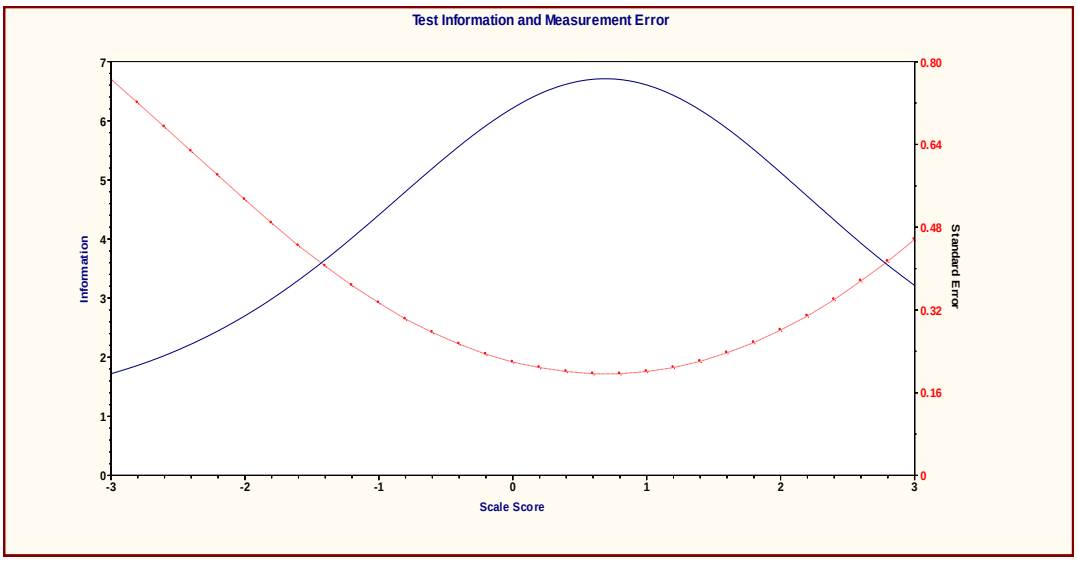
للإجابة على السؤال تم الاعتماد على قيم معلومات الاختبار والتي يقدمها برنامج Multilog7.03 ضمن نتائج عملية التقدير عند ٣١ قيمة مختلفة على متصل القدرة، والجدول (٦) يشير الى تقديرات قيم دالة معلومات الاختبار عند المستويات المختلفة من القدرة.

جدول (٦)

تقديرات قيم دالة معلومات التي يقدمها الاختبار عند مستويات القدرة المختلفة

القدرة	دالة المعلومات	القدرة	دالة المعلومات
٣,٠-	١,٧١٨	٠,٢	٦,٣٧٠
٢,٨-	١,٨٦٠	٠,٤	٦,٥٢٢
٢,٦-	٢,٠٢٧	٠,٦	٦,٥٩٥
٢,٤-	٢,٢٢١	٠,٨	٦,٥٨٦
٢,٢-	٢,٤٤٤	١,٠	٦,٤٩٥
٢,٠-	٢,٦٩٨	١,٢	٦,٣٢٥
١,٨-	٢,٩٨٤	١,٤	٦,٠٨٥
١,٦-	٣,٣٠١	١,٦	٥,٧٨٧
١,٤-	٣,٦٤٦	١,٨	٥,٤٤٥
١,٢-	٤,٠١٦	٢,٠	٥,٠٧٣
١,٠-	٤,٤٠٤	٢,٢	٤,٦٨٦
٠,٨-	٤,٨٠٠	٢,٤	٤,٢٩٧
٠,٦-	٥,١٩٢	٢,٦	٣,٩١٧
٠,٤-	٥,٥٦٩	٢,٨	٣,٥٥٦
٠,٢-	٥,٩١٤	٣,٠	٣,٢١٩
٠,٠	٦,٢٣١		

كما يلخص الشكل (٢) العلاقة بين قيم دالة المعلومات للاختبار والقدرة حيث يمثل الخط المتصل منحنى دالة معلومات الاختبار، والخط المتقطع يمثل الخطأ المعياري للاختبار، ويمثل المحور العمودي الأيسر قيم المعلومات، بينما يمثل المحور العمودي الأيمن قيم الخطأ المعياري.



شكل (٢) دالة معلومات اختبار التنمية المعرفية

ويلاحظ أن قيمة دالة المعلومات للاختبار تتزايد تدريجيا حتى تصل الى أقصى قيمة ممكنة لها (٦,٩) عندما تكون قيمة القدرة (٠,٦) لوجيت وهذه القيمة تقابل متوسط الصعوبة للمفردات (٠,٦) لوجيت. كذلك يلاحظ ان الاختبار يقدم أكبر كمية من المعلومات عند مستويات القدرة فوق المتوسطة ويقدم أقل معلومات عند مستويات القدرة العالية والمنخفضة، كما يظهر من الشكل ارتفاع قيمة الخطأ المعياري للتقدير عند القدرات العالية والمنخفضة، اذ أن الخطأ المعياري في التقدير يرتبط عكسيا مع الجذر التربيعي لدالة معلومات الاختبار وهذا يعني أن الخطأ المعياري في القدرة يكون أقل ما يمكن عند مستويات القدرة التي تناظر أقصى معلومات حسب العلاقة:

$$\mathbf{I(\theta)} = \frac{\mathbf{1}}{\sqrt{\boldsymbol{S.E(\theta)}}}$$

حيث أن $I(\theta)$ تمثل كمية المعلومات عند مستوى القدرة (θ) و $S.E(\theta)$ تعبر عن الخطأ المعياري في تقدير القدرة.

المناقشة

أولا: مناقشة نتيجة السؤال الأول "ما مدى تحقق افتراضات نظرية الاستجابة للمفردة لبيانات اختبار التنمية المعرفية في مادة العلوم وفق نموذج راش؟"

بينت النتائج المتعلقة بالتحليل العاملي للعينة تحقق افتراض أحادية البعد وفق المؤشرات المعتمدة على المكونات الأساسية Principal component. كما تم التحقق من الاستقلال الموضعي اعتمادا على مؤشرين. وقد اتفقت نتيجة السؤال الأول مع الإطار النظري الذي تستند

اليه الدراسة الحالية، كما أشرنا في الإطار النظري فيما يخص فروض نموذج راش، حيث تم التأكد من أن اختبار التنمية المعرفية في مادة العلوم يحقق خاصية أحادية البعد، وهذا بدوره يضيف دليل صدق لتفسير الدرجة وهو صدق البناء الداخلي للاختبار، ويمكن تفسير هذه النتائج بوجود البنية الصحيحة للاختبار وقياسه لمهارات العلوم والتي تمثل مكون واحد وبنية واحدة، مما ساهم في تحقق فرضيات نموذج راش.

ثانيا: مناقشة نتيجة السؤال الثاني "ما درجة ملائمة نموذج راش لاستجابات الطلبة في اختبار التنمية المعرفية؟"

تم التحقق من مدى مطابقة نموذج راش للبيانات المستمدة من اختبار التنمية المعرفية المستخدم في هذه الدراسة باستخدام اختبار مربع كاي، حيث أظهرت نتائج التحليل مطابقة جميع المفردات لنموذج راش، وكذلك أظهرت تطابق توقعات النموذج مع جميع استجابات الأفراد. ويمكن تفسير مطابقة جميع المفردات لنموذج راش في عملية الحذف للبيانات التامة والصفرية. كما ساعد في عملية المطابقة كما أشارت كاظم (١٩٨٨) استقرار صعوبة المفردة بالنسبة لباقي المفردات وذلك عبر المستويات المختلفة لقدرة الأفراد، وانتماء المفردة فعلا إلى مجموعة مفردات الاختبار التي يجب أن تقيس صفة واحدة معينة فقط دون غيرها من الصفات كما يفترض النموذج.

ثالثا: مناقشة نتيجة السؤال الثالث "ما قيم تقديرات معالم مفردات اختبار التنمية المعرفية في مادة العلوم؟

بينت النتائج أن قيم التقديرات لصعوبة المفردات قد تراوحت بين (٢,٠٢) الى (-٠,٧٥) لوجيت، وهذا يدل على سعة مدى القدرات التي يغطيها الاختبار وتمتع الاختبار بمستوى صعوبة أعلى نوعا ما لعينة الدراسة، كما تتوافق أيضا مع ما ذكره هامبلتون (Hambelton, 1989) وعلام (٢٠٠٥) من أن قيم معلم الصعوبة عادة ما يكون بين (-٢ و ٢).

رابعا: مناقشة نتيجة السؤال الرابع " ما مقدار المعلومات التي يقدمها منحنى معلومات اختبار التنمية المعرفية في مادة العلوم لقدرات الطلبة؟"

أظهرت النتائج أن الاختبار يقدم أقل قدر من المعلومات عند مستويات القدرة العالية والمتدنية بينما يقدم أكبر قدر عند مستويات القدرة المتوسطة وفوق المتوسطة، وهذا يتوافق مع ما ذكره الشريفين (٢٠٠٦) وهوانج (Hwang, 2011)، وربما يعود السبب الى عدم تناسب صعوبة المفردات مع المستويات العالية والمتدنية للطلبة.

كما أشارت النتائج الى أن قيمة دالة المعلومات للاختبار تتزايد تدريجيا حتى تصل الى أقصى قيمة ممكنة لها (٦,٩) عندما تكون قيمة القدرة (٠,٦) لوجيت وهذه القيمة تقابل متوسط الصعوبة للمفردات (٠,٦) لوجيت وهذا ينسجم أيضا مع ما يتوقعه نموذج راش اذ أن دالة معلومات تكون أقصى ما يمكن عندما تكون قيمة ($\theta = b$).

التوصيات

توصي الدراسة في ضوء النتائج الآتي:

١. توعية المتخصصين في مجال القياس التربوي بأهمية استخدام نظرية الاستجابة للمفردة في تحليل مفردات الاختبارات التحصيلية ومحاولة تطويرها من خلال المحاضرات وحضور المناقشات العلمية.
٢. استخدام نموذج راش كأحد نماذج نظرية الاستجابة للمفردة في بناء الاختبارات التحصيلية.
٣. عقد دورات تدريبية للمعلمين وكل المشاركين في عمليات التقويم يتم خلالها تدريبهم على كيفية استخدام نموذج راش في تطوير أساليب التقويم وتحليل الاختبارات التحصيلية.

المقترحات

تقدم الدراسة في ضوء النتائج عددا من المقترحات البحثية الآتية:

١. مقارنة الخصائص السيكومترية لاختبار التنمية المعرفية في مادة العلوم وفق نماذج نظرية الاستجابة للمفردة المختلفة.
٢. تقصي الخصائص السيكومترية لاختبار التنمية المعرفية في مادة الرياضيات وفقا لنظرية الاستجابة للمفردة.

المراجع

تيغزة، أمحمد بوزيان (٢٠١٢). **التحليل العاملي الاستكشافي والتوكيدي: مفاهيمهما ومنهجيتهما بتوظيف حزمة SPSS وليزرل LISREL**. عمان: دار المسيرة.

حجازي، أحمد؛ والشريفين، أحمد (٢٠١٤). تقصي الخصائص السيكومترية لاختبار ذكاء جمعي حسب نظرية السمات الكامنة. **مجلة الدراسات التربوية والنفسية، ٨(١)**، ١-١٤.

الحموري، هند (٢٠١١). دراسة استكشافية لملائمة نماذج نظرية الاستجابة للمفردة في بناء اختبار تحصيلي من اعداد المعلم. **مجلة العلوم التربوية والانسانية، ١٢(٢)**، ٤٧-٨١.

الربيعية، تيمورة (٢٠١٢). **تقويم برنامج التنمية المعرفية من وجهة نظر الطلاب والمعلمين والمشرفين التربويين بمحافظة مسقط** (رسالة ماجستير غير منشورة). جامعة السلطان قابوس، سلطنة عمان.

الشربيني، أحلام؛ وحجازي، رضا (٢٠٠٢). تقويم تحصيل تلاميذ مرحلة التعليم الاساسي في العلوم باستخدام نظرية السمات الكامنة. المؤتمر العلمي السادس، **الجمعية المصرية للتربية العملية**.

الشريفين، نضال كمال محمد (٢٠٠٦). الخصائص السيكومترية لاختبار محكي المرجع في القياس والتقويم التربوي وفق النظرية الحديثة في القياس التربوي والنفسي. **مجلة العلوم التربوية والنفسية- البحرين، ٧(٤)**، ٧٩-١٠٩.

الصارمي، يحيي (٢٠٠٨). جوانب تفعيل نتائج الاختبارات المطبقة ضمن برنامج التنمية المعرفية. **مجلة التطوير التربوي-سلطنة عمان، (٤١)**، ٤١-٤٣.

علام، صلاح الدين (٢٠٠٥). **نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي**. القاهرة: دار الفكر العربي.

كاظم، أمينة (١٩٨٨). **دراسة نظرية نقدية حول القياس الموضوعي للسلوك: نموذج راش**. القاهرة: مكتبة الانجلو المصرية.

الزبون، حابس (٢٠١٣). أثر حجم العينة على تقدير دالة المعلومات للاختبار والخطأ المعياري في تقديرها باستخدام النظرية الحديثة في القياس. **مجلة جامعة النجاح للعلوم الانسانية- فلسطين، ٢٧(٦)**، ١٣١٣-١٣٣٤.

الناغي، هبه (٢٠١١). أثر عدد مفردات الاختبار على تقدير قدرات الأفراد ودالة المعلومات لاختبار تحصيلي مرجعي المحك في الرياضيات باستخدام نماذج نظرية الاستجابة للمفردة. ***مجلة كلية التربية* ببورسعيد- مصر، (١٠)**، ٦٧٦-٧٢١.

الهنائي، معن (٢٠١٧). **تقويم أسئلة اختبار الرياضيات في برنامج التنمية المعرفية للصف بسلطنة عمان** (رسالة ماجستير غير منشورة). جامعة السلطان قابوس، سلطنة عمان.

وزارة التربية والتعليم (٢٠١٣). **دليل برنامج التنمية المعرفية للطلبة في العلوم والرياضيات والمفاهيم البيئية**. سلطنة عمان.

وزارة التربية والتعليم (٢٠١٦). **الكتاب السنوي للإحصاءات التعليمية**. سلطنة عمان.

Franchignoni, F., Ferriero, G., Giordano, A., Sartorio, F., Vercelli, S., & Brigatti, E. (2011). Psychometric properties of QuickDASH - a classical test theory and Rasch analysis study. *Manual Therapy, 16*(2), 177-182.

Hambleton, R. K., & Jones, R. W. (1994). Item parameter estimation errors and their influence on test information functions. *Applied Measurement in Education, 7*(3), 171-186.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Springer Science & Business Media.

Hammouri, H. & Sabah, S. (2011). Analysis and assessment of the Jordan National Test for Controlling the Quality of Science Instruction (NTCQSI): A Rasch measurement perspective. *Educational Research and Evaluation, 16*, 451-470.

Harwell, M., Stone, C. A., Hsu, T. C., & Kirisci, L. (1996). Monte Carlo studies in item response theory. *Applied Psychological Measurement, 20*(2), 101-125.

Hwang, D. Y. (2002). *Classical test theory and item response theory: Analytical and empirical comparison*. Paper presented at the Annual Meeting of the Southwest Educational Research Association, Austin, TX. (ERIC Document Reproduction Service No. ED 466779).

Peter, H., Ahan, T., & Stephen, P. M. (2004). *Application of Rasch analysis in the development and application quality of life instruments, peer reviewed journal, 7*(6), 522 – 526.

Stone, M., & Yumoto, F. (2004). The effect of sample size for estimating rasch/IRT parameters with dichotomous items. *Journal of applied Measurement, 5*(1), 48-61.

Yen, W. (1984). Effects of local item dependence on the fit and equating performance of the three- parameter logistic model. *Applied psychological Measurement, 8*, 125-145.