بناء بنك أسئلة محوسب في ضوء نظرية الاستجابة للمفردة

إعداد

محمود محمد حامد الصفراني*

مقدمة:

تسعى معظم الدول سواء كانت متقدمة أو نامية إلى تطوير نظمها التربوية وإلى تجويد العملية التعليمية من أجل زيادة فاعليتها. وبالرغم من أن مكون التقويم أحد الأركان الرئيسة لهذه النظم، إلا إنه يعد أكثرها تأثيرًا في النظام التربوي بأسره، حيث يسهم في تطويرها تطويرًا فاعلًا وفي تعديل وإصلاح بقية هذه المكونات، وفي توجيه مسار العمل التربوي والنهوض به، وتحديد مدى تحقيق النظام التعليمي لأهدافه المرجوة. ولذلك يحظى مجال تطوير نظم وأساليب التقويم والامتحانات باهتمام كبير، ولقد أولت أغلب تلك الدول اهتمامًا خاصًا بتطوير أساليب التقويم وأدواته، وبذلت في ذلك جهودًا متزايدة؛ تحقيقًا لهذا الهدف (علام، ١٩٩٧ :٨).

وتمثل الاختبارات إحدى أهم الوسائل المستخدمة في عمليات التقويم التربوي، وتعتمد عليها المؤسسات التربوية بصورة أساسية للتعرف على مقدار تحصيل المتعلمين ومدى تحقيقهم للأهداف التعليمية؛ لذلك تعد نتائج الاختبارات أساسًا مهمًا في صنع واتخاذ كثير من القرارات المصيرية لكل من تطبق عليه تلك الاختبارات، فللاختبارات دور أساسي في حياة الطالب في جميع مراحله التعليمية، فمن خلال نتائجها يتحدد مساره التعليمي، ومن ثم فإن دقة وسلامة القرارات التربوية تتوقف على مدى جودة وكفاءة الاختبارات المستخدمة في عملية التقويم، وعلى الأساليب المستخدمة في تفسير وتحليل نتائجها (الوليلي، ٢٠٠٥: ٩٩).

إن التغيرات المرتبطة بـ (ماذا وكيف يتعلم الطلاب؟)، أصبحت تفوق قدرات الاختبارات التقليدية، كما أن فكرة الاستعانة باختبار واحد لتقويم تحصيل المتعلمين لم تعد مناسبة للأهداف

---

* بحث مشتق من رسالة دكتوراه تحت إشراف:

أ.د/ عبد العاطي أحمد الصياد– أستاذ مناهج البحث والإحصاء والقياس والتقويم.

أ.د/ عبد الناصر السيد عامر– أستاذ القياس والإحصاء التربوي– كلية التربية جامعة قناة السويس.

أ.د/ نجلاء عبد الله إبراهيم– أستاذ علم النفس التربوي– كلية التربية جامعة قناة السويس.

التربوية المعاصرة، فاستخدام نظم إدارة التعلم والتقويم البنائي للطالب في مجال دراسي معين، يتطلب توافر نماذج اختبارية متعددة لضمان سرية الاختبارات، بالإضافة إلى تمكين الطالب من إعادة الاختبار أكثر من مرة، وإمكانية تطبيق اختبارات دورية أثناء العام الدراسي، ومقارنة درجاته التي يحصل عليها عبر الزمن. وإن التطور الهائل في تكنولوجيا الاتصالات والمعلومات انعكس على العملية التعليمية، فأصبح التعليم يعتمد على البحث والكتابة والتصور والمحاكاة ...إلخ؛ حيث ظهرت الجامعات الافتراضية، وكذلك اهتمت الجامعات التقليدية بإنشاء مواقع إلكترونية لها على شبكة الإنترنت تقدم الخدمات الإلكترونية المتعددة لطلابها من مكتبات إلكترونية ونظم لإدارة التعلم، وكذلك إتاحة فصول افتراضية للتعلم والتواصل المتزامن بين المحاضر أو المعلم وطلابه؛ ولذلك يشير (Anzaldua,2002) إلى أن عملية التحول من التقويم بمفهومه التقليدي إلى الاعتماد على بنوك الأسئلة، قد أصبحت هدفًا للعديد من الجامعات في مختلف دول العالم ومنها الجامعات الأيرلندية، كما يشير بينيت (Bennett,2002) إلى أن كثيرًا من المدارس والجامعات على مستوى العديد من الولايات بالولايات المتحدة الأمريكية، مثل: ولاية داكوتا الجنوبية South Dakota، وولاية كارولينا الشمالية North Carolina وولاية أوريجون Orego وولاية فرجينيا Virginia وولاية جورجيا Georgia تقدم برامج الاختبارات الرئيسة عن طريق الكمبيوتر داخل المؤسسات التعليمية، وأيضًا عن طريق شبكة الإنترنت، بل تعدى الأمر ذلك؛ حيث قامت ولاية كارولينا الشماليةNorth Carolina بتقديم اختبارات للطلاب المعاقين معتمدة كليةً على الويب. ويشير (الوليلي، ٢٠٠٥) إلى أن خبراء القياس التربوي من أمثال هامبلتون Hambleton، ورايت Wright، وبابام Popham، ولورد Lord وغيرهم يؤكدون ضرورة التحول إلى نظام الاختبارات المتعددة Multi-Tests لقياس التحصيل الدراسي، وهو ما يتطلب الاهتمام ببناء بنوك الأسئلة المحوسبة وبنظريات القياس التي تستند إليها.

وأشار العديد من الدراسات ومنها دراسة كل من (إلياس، ١٩٩٣)، و(Rudner,1998)، و(أبو جلالة،١٩٩٩)، و(بخيت ويركندي،٢٠٠١)، و(الصراف،٢٠٠٢)، و(أبو هاشم،٢٠٠٧)، و(محمود،٢٠١٢)، و(السرحان، ٢٠١٥)، و(محمد، ٢٠١٦)، و(الجمل وآخرون، ٢٠١٧) إلى أهمية بنوك الأسئلة في توفير نماذج اختبارية متعددة وتحقيق الموضوعية في القياس، وخفض القلق والرهبة والتوتر المصاحب لموقف الامتحان وتوفير المفردات والأسئلة المتقنة والمقننة، ومن ثم توفير الوقت والجهد في بناء الاختبارات، حيث يمكن بسهولة ويسر

الحصول على اختبارات جيدة طبقًا لمواصفات وأهداف المقرر الدراسي، وهذا يُمكن المعلم من تطبيق اختبارات تكوينية بنائية لموضوعات ووحدات المقرر الدراسي أولًا بأول.

كما أن الاعتماد على بنوك الأسئلة أثناء تقييم التحصيل الدراسي يتيح إمكانية اتباع اللامركزية في تطبيق الاختبارات، وإمكانية تصميم اختبارات محوسبة خطية أو اختبارات ملائمة لقدرة الطالب، وإمكانية تفريد عملية التقويم، وإمكانية سحب عدة نماذج للاختبار الواحد تسهم في الحد من مشكلة الغش في الاختبارات وتسرب الأسئلة، بالإضافة إلى وجود كلمة السر التي تزيد من السرية، وتمكن واضعي الاختبارات دون غيرهم من الدخول لبنك الأسئلة.

ويشير علام (١٩٩٧) وكاظم (٢٠٠٤) إلى أنه يمكن استخدام أحد نماذج نظرية الاستجابة للمفردة الاختبارية في تدريج المفردات المتعلقة بكل مجموعة متجانسة من الأهداف على حدة، بهدف التوصل إلى ميزان موحد متدرج لكل من صعوبة المفردات وقياسات الأفراد، حيث يتطلب بناء البنك تدريج المفردات Items Calibration ووصلها Items Linking بعد جمع البيانات باستخدام أحد تصميمات جمع البيانات، مثل: تصميم المفردات المشتركة Common Items Design أو تصميم الأفراد المشتركين Common Persons Design أو تصميم المجموعة الواحدة Single Group Design.

وترى كاظم (٢٠٠٤) أن إنشاء بنوك الأسئلة يعد خطوة مهمة في طريق الارتفاع بجودة التعليم، حيث إن ما توفره بنوك الأسئلة من أدوات موضوعية للقياس، يتيح المقارنة العادلة بين أداء الطلاب، وأوصت بضرورة بناء بنوك للأسئلة في المواد المختلفة وبالمراحل المختلفة، كما قام زو وآخرون (Zhu & Others ,2011) ببناء بنك أسئلة محوسب في التربية البدنية في ضوء نظرية الاستجابة للمفردة "نموذج راش"، وأوصى ببناء بنوك يتم تدريج مفرداتها الاختبارية اعتمادًا على أحد نماذج نظرية الاستجابة للمفردة.

مشكلة البحث:

على الرغم من التوجهات الحديثة نحو نظام الاختبارات المتعددة Multi-Tests لقياس التحصيل الدراسي، إلا أن فكرة التحول إلى هذا النوع من الاختبارات والاعتماد عليها في المدارس والجامعات بالدول العربية ما زالت بعيدة بعض الشيء عن التطبيق، وهو ما يضع أمام المهتمين بمجال القياس النفسي والتربوي متطلبًا أساسيًّا وهو زيادة الأبحاث والدراسات المتعلقة ببنوك الأسئلة المحوسبة وبالاختبارات المستمدة منها، وبكيفية تطبيقها ومدى جودتها.

وفي ضوء ما تم ذكره تحددت مشكلة الدراسة الحالية بما يأتي:

بناء بنك أسئلة محوسب في نظرية الاستجابة للمفردة، وتدريج Calibration مفرداته.

أسئلة البحث:

سعى البحث الحالي إلى الإجابة عن الأسئلة الآتية:

١. ما دلالات مطابقة البيانات لنماذج الاستجابة للمفردة (النموذج أحادي البارامتر، النموذج ثنائي البارامتر، النموذج ثلاثي البارامتر)؟

٢. ما قيم بارامتر الصعوبة وبارامتر التمييز المقدرة لمفردات بنك الأسئلة وفقًا للنموذج المناسب من نماذج نظرية الاستجابة للمفردة؟

٣. ما خصائص بنك الأسئلة من حيث مدى القدرة ودالة المعلومات؟

٤. ما خصائص اختبار مسحوب مفرداته من بنك الأسئلة من حيث مدى القدرة ودالة المعلومات؟

أهداف البحث:

هدف البحث الحالي إلى بناء بنك أسئلة محوسب في ضوء نظرية القياس الكلاسيكية ونظرية الاستجابة للمفردة، وتدريج Calibration مفرداته من خلال:

١. التحقق من افتراضات نظرية الاستجابة للمفردة الاختبارية للبيانات المستمدة من استجابات الطلاب على مفردات كل اختبار على حدة وللمفردات المكونة لبنك الأسئلة ككل.

٢. التعرف على دلالات مطابقة المفردات والاختبارات التي تكون منها بنك الأسئلة.

٣. تدريج اختبارات فرعية لقياس تحصيل "مفاهيم تقنية المعلومات" اعتمادًا على تصميم المجموعة الواحدة لتكوين بنك الأسئلة.

أهمية البحث:

يمكن أن يسهم البحث الحالي في:

١. مسايرة التوجهات العالمية والمحلية في بناء بنوك أسئلة وفق الأساليب والنظريات الحديثة لتطبيقها في مراحل التعليم المختلفة.

٢. زيادة اطلاع المهتمين بالقياس التربوي والنفسي في الثقافة العربية على الاتجاهات الحديثة في بناء وتطبيق الاختبارات المستمدة من بنوك الأسئلة.

٣. إعداد بنوك الأسئلة يؤدي إلى إنتاج نماذج اختبارات متنوعة، وهو ما يؤدي إلى تقليل عملية الغش، والتغلب على مشكلات تطبيق الاختبارات، وتحسين دقة القياس، والتقليل من أخطاء القياس، مع السرعة والدقة في التصحيح.

٤. توفير دليل استرشادي وتطبيق عملي لكيفية تدريج بنوك الأسئلة في ضوء نظرية الاستجابة للمفردة للباحثين، وهو ما يحفزهم على إجراء المزيد من الدراسات.

٥. استخدام نظرية الاستجابة للمفردة كأحد الأساليب الحديثة في القياس السلوكي في تدريج بنوك الأسئلة، وهو ما يساعد على نشر وتبسيط فكرة إنشاء بنوك أسئلة.

حدود البحث:

اقتصر البحث الحالي على الحدود التالية:

الحدود الموضوعية: اقتصر البحث على الباب الأول الذي يتناول (تكنولوجيا المعلومات) بمادة الكمبيوتر للصف الأول الإعدادي، وعلى نوعية واحدة من المفردات الاختبارية التي تكون نتائجها ثنائية الدرجة (Dichotomous Items).

الحدود المكانية والبشرية: تم تطبيق البحث على تلاميذ الصف الأول الإعدادي، بمحافظة الإسماعيلية.

الحدود الزمنية: تم تطبيق البحث خلال الفصل الدراسي الأول للعام الدراسي ٢٠١٨-٢٠١٩م.

مصطلحات الدراسة:

بنك الأسئلة المحوسب:

هو "ملف قاعدة بيانات" لنظام إدارة الأسئلة أو المفردات التي تقيس نطاق معرفة معين بعد تدريجها باستخدام أحد نماذج الاستجابة للمفردة وتقدير خصائصها وفقًا لنظرية القياس الكلاسيكية، وهو مؤمن بكلمة سر، وكل مفردة في ملف قاعدة البيانات معرفة برمز وحيد خاص بها يرتبط به معلومات المفردة، بحيث يمكن التعرف على الهدف الذي تقيسه المفردة وفي أي مستوى معرفي، ونص المفردة وطريقة عرضها ونوعها والإجابة الصحيحة ومعاملي الصعوبة والتمييز وفقًا لنظرية القياس الكلاسيكية، وقيمة بارامتر الصعوبة "b" و بارامتر التمييز "a" و بارامتر التخمين "c" وقيمة اختبار حسن المطابقة ودلالته الناتجة من تدريج المفردات وفقًا لأنسب نماذج الاستجابة للمفردة، بالإضافة إلى حقل لتسجيل الملاحظات المتعلقة بالمفردة وتاريخ تطبيقها ومعلومات تعديلها في حالة التعديل، ويوفر نظام إدارة بنك الأسئلة الأدوات التي تسمح للمستخدم أن يبحث عن المفردات وفقًا لخصائصها، وإنشاء اختبار واختيار المفردات التي يتكون منها وعرض خصائص الاختبار الجديد وفقًا لنظرية القياس الكلاسيكية ولنظرية الاستجابة للمفردة قبل تطبيقه، وباستخدام نظام إدارة بنك الأسئلة يمكن تطبيق الاختبارات اعتمادًا على الحاسب الآلي سواء كانت الاختبارات خطية أو موائمة، وأيضًا يمكن طباعة الاختبار وتطبيقه ورقيًا.

نظرية الاستجابة للمفردة:

هي اتجاه معاصر في القياس النفسي يسعى إلى تحقيق الموضوعية وتوفير خطية واستقلالية القياس، كما يسعى إلى تكامل أهداف القياس، وتفترض تلك النظرية وجود واحدة أو أكثر من السمات الأساسية غير الملاحظة التي تحدد استجابات الفرد الملاحظة لمفردات اختبار ما، كما تهتم بالربط بين استجابة الفرد لمفردة اختبارية وخصائص هذه المفردة. ويندرج تحت هذه النظرية عدد من النماذج الرياضية الاحتمالية التي تعين العلاقة المتوقعة بين الاستجابات الملاحظة على الاختبار والسمات أو القدرات غير الملاحظة التي يفترض أنها تحدد هذه الاستجابات(كاظم، ١٩٨٨ : ٤١)

الخلفية النظرية والدراسات السابقة

أولًا: بنوك الأسئلة

يشير ميلمان وأرتر (Millman & Arter,1984) إلى أن مسمى بنك الأسئلة Item Banks قد ورد بتسميات متعددة، مثل: question banks بنك الأسئلة أو item collections مجموعة مفردات أو item reservoirs مخازن الأسئلة أو Item Pools قائمة الأسئلة. ويرى شوبين (Choppin,1990) أن كلمة بنك تعد كلمة غير مناسبة في هذا المجال: نظرًا إلى ما تنطوي عليه من معانٍ ومضامين خارجة عن نطاق المفهوم في إطاره التربوي، وخاصة لمن يستخدم هذا المفهوم لأول مرة؛ ولذلك اقترح الاستعاضة بمفهوم Test Item Libraries مكتبة المفردات الاختبارية (علام،١٩٩٧).

وقد تعددت الآراء التي تناولت مفهوم بنك الأسئلة بالتعريف، ومنها: تعريف سميثهام (Smetherham, 1979: 57) لبنك الأسئلة بأنه مجموعة كبيرة من المفردات المتنوعة في صعوبتها؛ لتلائم جميع مستويات القدرة، وتغطي موضوعات محددة في المنهج، ويمكن منها تكوين اختبارات فرعية.

في حين يعرف فان درلندن وإيجن (Van Der Linden & Eggen,1986) بنك الأسئلة، الذي تم تدريجه ومعايرته Calibrated Item Bank بأنه مجموعة كبيرة من المفردات التي تقيس نفس السمة أو نطاق معرفي معين، والتي تم تخزينها والبارمترات الخاصة بها معًا في الكمبيوتر، حيث تم تقدير البارامترات باستخدام نموذج الاستجابة للمفردة الذي تم مطابقته مع استجابة الممتحنين.

وتشير كاظم (١٩٨٨: ١٠٨) إلى أن بنك الأسئلة يتكون من مفردات اختبارية لعدة اختبارات، تتدرج مفرداتها جميعًا على تدريج واحد مشترك، بحيث تُعرَّف مدى واسعًا من مستويات المتغير موضع القياس، حيث يبدأ بنك الأسئلة بدمج مفردات اختبارين في تدريج واحد؛ ليكون نواة البنك، ثم يتم ربط مفردات نواة بنك الأسئلة بمفردات اختبارات أخرى بحيث يغطي أكبر مدى ممكن للمتغير المقاس.

ويرى برجوف (Burghof, 2001) أن بنك الأسئلة هو مجموعة كبيرة نسبيًّا من المفردات الاختبارية المخزنة في قاعدة بيانات، يسهل الوصول إليها، ومصنفة ومرمزة وفقًا للموضوع، والمستوى الدراسي، ونوع المفردة، ومهارة التفكير، ومستويات بلوم للأهداف المعرفية، ومستوى صعوبة المفردة.

ويشير سكيرز (Squires,2003:1) إلى أن بعض الباحثين يستخدم مصطلح بنوك الأسئلة للإشارة إلى أية مجموعة كبيرة من المفردات، لكن بنك الأسئلة المصمم بشكل صحيح يعني أكثر من ذلك، فهو دليل من المفردات المصنَّفة تبعًا للمحتوى ومعاملات الصعوبة، بحيث يمكن اختيار مجموعة من المفردات في ضوء قواعد محددة، ترتبط بمحتوى المفردة ومستوى صعوبتها.

ويفرق علام (٢٠٠٥: ٢٤٩) بين مفهوم بنك الأسئلة Item Banks ومفهوم قائمة أو ملف الأسئلة Item Pool، في أن أسئلة أو مفردات البنك يتم السحب منها لتجميع اختبارات تكون خصائصها معلومة، وهو ما ييسر تفسير درجات الاختبارات التي تشتمل على مجموعات مختلفة من المفردات المسحوبة من البنك. في حين أن قائمة أو ملف الأسئلة يفتقر إلى تدريج "معايرة" الأسئلة لتحديد خصائصها السيكومترية، وهو ما يجعل المقارنة بين الدرجات على ميزان "مقياس" مشترك غير ممكنة.

وتعرّف شركة أنظمة التقييم[1] (Assessment Systems Company,2020) بنك الأسئلة بأنه عملية إنشاء قاعدة بيانات مقصودة من المفردات الاختبارية المراد قياسها، وذلك بوضع هيكل تنظيمي يتم من خلاله تصنيف العديد من الأسئلة في مستويات القدرة المختلفة للحصول على نماذج اختبارية متعددة؛ حيث تتضمن قاعدة البيانات معرفًا للمفردة ونصها، ومفتاح التصحيح وتصنيفها وخصائصها السيكومترية وفقًا لنظرية الاستجابة للمفردة وأعلى قيمة لدالة المعلومات الخاصة بالمفردة.

[1] تأسست أنظمة التقييم من جامعة مينيسوتا خلال سبعينيات القرن الماضي، وتخصصت في برمجيات القياس وأنظمة بنوك الأسئلة والاختبارات الموائمة CAT.

خصائص ومزايا استخدام بنوك الأسئلة:

أشار العديد من الأدبيات والدراسات، ومنها: (إلياس، ١٩٩٣ : ١٨٢)، (أبو جلالة، ١٩٩٩: ١٤٩)، (بخيت ويركني، ٢٠٠١: ٥)، (الصراف، ٢٠٠٢: ١٨٣)، (Rudner,1998: 2)، (أبو هاشم،٢٠٠٧: ١٦٦-١٦٧)، (كاظم، ٢٠٠٠: ٥٨)، إلى أن من أهم خصائص ومزايا استخدام بنوك الأسئلة، ما يأتي:

١. تحقيق الموضوعية في التقويم، وهذا هو الهدف الأساسي لبنوك الأسئلة، ويعتمد تحقيق هذا الهدف على مجموعة من العوامل التي يمكن أن تؤثر في بناء وتكوين بنك الأسئلة، مثل: فلسفة القياس، ومدى الدقة في صياغة الأسئلة، وتحليلها وتحديد مواصفاتها الإحصائية.

٢. تخفيض عوامل القلق والرهبة والتوتر المصاحبة لموقف الامتحان، وذلك باستخدام عينات من مفردات من البنك للتقويم المستمر أثناء العام الدراسي؛ حتى يألف الطلاب هذا النوع من الأسئلة الموضوعية، ولا يؤثر هذا في سرية الامتحانات.

٣. المفردات والأسئلة المخزنة بها متقنة ومقننة؛ لذا يمكن الاعتماد عليها في بناء اختبارات جيدة لنظم الامتحانات.

٤. مرونة القياس حيث يسهل تشكيل أي اختبار في أي وقت بسرعة وسهولة.

٥. سرعة الحصول على صور متكافئة من الاختبارات تبعًا لمواصفات المفردات المكونة للبنك.

٦. إمكانية المقارنة الموضوعية بين مستويات الطلاب مهما اختلفت الاختبارات المستخدمة، ومهما اختلفت مستويات المجموعات التي ينتمون إليها، طالما أن المفردات قد تم تدريجها ومعايرتها ووصفها في بنك واحد، وهو ما يسمح للمعلمين بمقارنة النتائج ومعرفة مدى تعلم موضوع ما من قبل طلبة مختلفين مع معلمين مختلفين.

٧. إمكانية تصحيح المفردات المعدة في بنك الأسئلة بسهولة وسرعة باستخدام الكمبيوتر وهو ما يوفر الكثير من الوقت والجهد.

٨. توفير وقت وجهد المعلم في بناء الاختبارات، وذلك بحصوله على اختبارات جيدة من بنوك الأسئلة طبقًا لمواصفات وأهداف المقرر الدراسي.

٩. إتاحة الفرصة لقياس العمليات العقلية العليا؛ وذلك نظرًا لاتساع المدى الذي يمكن أن يغطيه البنك من حيث المحتوى أو المستويات المعرفية المختلفة.

١٠. تساعد بنوك الأسئلة على إعداد اختبارات تكوينية بنائية لموضوعات ووحدات المقرر الدراسي أولاً بأول.

١١. إمكانية اتباع لا مركزية تطبيق الاختبارات.

١٢. إمكانية تصميم اختبارات محوسبة خطية أو اختبارات ملائمة لقدرة الطالب.

١٣. إمكانية تفريد عملية التقويم.

١٤. إمكانية سحب عدة نماذج للاختبار الواحد يسهم في الحد من مشكلة الغش في الاختبارات وتسرب الأسئلة، بالإضافة إلى وجود كلمة السر التي تزيد من السرية وتمكن واضعي الاختبارات دون غيرهم من الدخول لبنك الأسئلة.

محددات بنوك الأسئلة:

بمراجعة بعض الأدبيات، والبحوث والدراسات، ومنها: (كاظم، ٢٠٠٠؛ علام، ٢٠٠٥؛ عودة، ٢٠٠٢؛ Nakamura, 2001: 13) تبين أن تطبيق فكرة بنوك الأسئلة تستلزم ما يأتي:

١. الحاجة إلى توافر برمجيات Software متخصصة في إدارة بنوك الأسئلة، بالإضافة إلى برمجيات لتحليل البيانات وفقًا لنظريتي القياس الكلاسيكية والاستجابة للمفردة، وكذلك توافر العنصر البشري المدرب على استخدام هذه البرمجيات.

٢. الحاجة إلى صياغة وإعداد عدد كبير نسبيًا من المفردات أو الأسئلة، وبالتالي الحاجة إلى توافر خبراء في مادة التخصص والمناهج وطرق التدريس لهذا التخصص وفي اللغة وفي القياس والإحصاء والتقويم.

٣. الحاجة إلى تحديث مستمر لبنك الأسئلة من خلال إيداع مفردات جديدة، وتنقيح المفردات المخزنة به باستمرار.

٤. حاجة نظام بنك الأسئلة إلى إدارة تمنع تسرب المفردات والحفاظ عليها.

٥. الحاجة إلى اعتمادات مالية مناسبة لتوفير البرمجيات وأجهزة الكمبيوتر والخبراء والإدارة.

ثانيًا: نظرية الاستجابة للمفردة Item Response Theory

هي نظرية تقوم على افتراض علاقة بين قدرة الفرد واحتمال إجابته الصحيحة (P_i) عن المفردة الاختبارية (i)، حيث تفترض وجود سمة أو مكون فرضي أو متغير كامن أو أكثر تحدد استجابات الفرد الملاحظة لمفردات اختبار ما؛ أي أن كل استجابة مرتبطة بسمات غير ملاحظة أدت إلى هذا النوع من الاستجابة، يطلق عليها سمات Traits أو قدرات كامنة، ويقصد بالسمة الكامنة، بأنها أي مكون نفسي سواء كان متغيرًا شخصيًا أو اجتماعيًا أو معرفيًا، مثل: التحصيل، أو السمات الشخصية، أو القدرات العقلية، أو الاتجاهات (Sijtsma & Molenaar, 2002: 4) (Wilson, 2005: 6)(كاظم، ١٩٨٨: ٤١)، وأن القيمة الاحتمالية لاستجابة فرد لمفردة

اختباريـة تكـون دالـة لكـل مـن السـمة أو القـدرة التـي يفتـرض أن الاختبـار يقيسـها لـدى الفـرد وخصائص المفردة التي يحاول الإجابة عنها.

ويرى ديمارز (Demars,2010:3) أنها نظرية تقوم على نمذجة العلاقات القائمة بين مستوى سمة أو قدرة لدى الفرد، واستجابته لمفردة من مفردات الاختبار.

ويشير محاسنة (٢٠١٣) إلى أن نماذج نظرية الاستجابة للمفردة خرجت إلى حيز الوجود على يد جورج راش Georg Rasch في عقد الستينيات من القرن العشرين؛ وذلك من خلال نموذجه "نموذج راش"، ثم قام رايت "١٩٧٧ – ١٩٨٦" بتطوير هذا النموذج، وقد أشار علام (٢٠٠٠) إلى أنه حدث تقدم سريع في الحركة البحثية والتطبيقية لنظرية الاستجابة للمفردة IRT ابتداءً من عام ١٩٨٦م، وهو العام الذي نشر فيه لورد Lord كتابه "النظرية الإحصائية لدرجات الاختبارات العقلية" الذي جمع بين غلافيه لأول مرة أسس نظرية القياس الكلاسيكية، وأسس نظرية الاستجابة للمفردة الاختبارية، وفي الفترة من عام ١٩٦٩ . ١٩٧٢م قام ساميجيما Samjima بجهود أدت إلى الانتقال من تحليل البيانات الثنائية(صفر،١) إلى البيانات متعددة التدريج، ثم ظهرت محاولات للعديد من خبراء القياس المعاصر لتطوير نماذج الاستجابة للمفردة متعددة الأبعاد MIRT، التي تفسر بدرجة أفضل أداء الأفراد؛ نظرًا للتفاعل بين الأفراد والمفردات، ويعتبر كل من ريكاس Reckase، وماكنلي Mckinley، وفريزر Fraser، ومولييك Mulaik، وسيمبسون Sympson، وأيمبرتسون Embreston وغيرهم من رواد هذا المجال (علام، ٢٠٠٥).

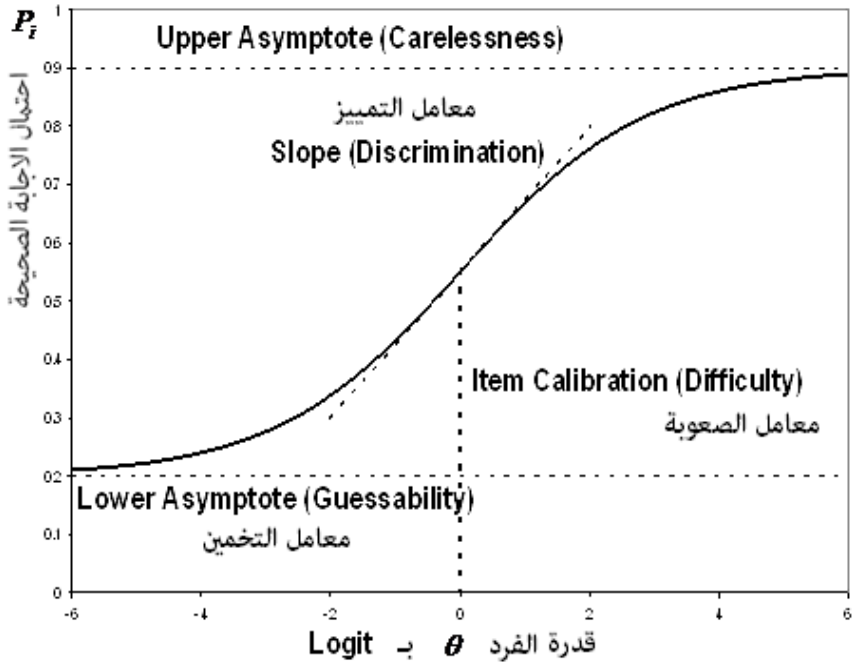
افتراضات نظرية الاستجابة للمفردة:

١. أنه إذا كان الممتحن يعرف الإجابة الصحيحة لمفردة اختبارية، وسيجيب عنها بشكل صحيح؛ يمكن الإشارة إلى ذلك باسم "المعرفة الصحيحة"، وإذا أجاب أحد الممتحنين عن إحدى المفردات الاختبارية بطريقة غير صحيحة، فإنه لم يكن يعرف الإجابة الصحيحة (Weiss & Yoes, 1991: 72-73)؛ أي أن الدالة التي تصف العلاقة بين قدرة الفرد واحتمال نجاحه في الإجابة عن المفردة التي تقيس تلك القدرة هي دالة تزايدية Monotonically Increasing Function، وتسمى دالة خصائص المفردة (Linden, 2016: 4). ويمكن التعبير عن هذه العلاقة برسم بياني يأخذ شكل منحنى، يعرف بالمنحنى المميز للمفردة (ICC) Item Characteristic Curve، الذي يوضح كيفية تغير مستوى السمة في علاقتها بالتغيرات في احتمال الإجابة الصحيحة ويُمثَّل احتمال الإجابة الصحيحة على المحور الرأسي، وتُمَثَّل القدرة

المقاسة على المحور الأفقي، وتقاس بوحدة الترجيح اللوغارتمي (اللوجيت Logit) ويتخذ هذا المنحنى شكل المنحنى اللوغاريتمي التراكمي Logistic Curve؛ أي أن نماذج الاستجابة للمفردة أحادية البعد ثنائية التصحيح، تفترض وجود مُتَّصل للسمة المراد قياسها، وأنه يمكن تقدير احتمال إجابة فرد إجابة صحيحة عن مفردة اختبارية، إذا علمنا موقعه على هذا المتصل، وأنه يزداد احتمال توصل الفرد إلى الإجابة الصحيحة عن مفردة اختبارية بازدياد مقدار السمة لديه، وتُمثَّل دالة خصائص المفردة بالصيغة الرياضية العامة:

$$P_i(\theta) = \frac{e^x}{1+e^x}$$

حيث إن $P_i(\theta)$ احتمال أن يجيب فرد مستوى قدرته (θ) إجابة صحيحة عن المفردة (i)، و(e) أساس اللوغاريتم الطبيعي وهو مقدار ثابت = ٢,٧١٨٣ تقريبًا، و(x) رمز اختياري يُعبَّر عنه بالبارامترات المناسبة لأي من هذه النماذج، والاحتمالية في هذه الدوال تضمن أن الصيغة الرياضية للنموذج تشير إلى وصف أحداث غير مؤكدة (علام، ٢٠٠٥ : ٦٥).



شكل (١): المنحنى المميز للمفردة Item Characteristic Curve (ICC)

يُعرّف المنحنى المميز للمفردة Item Characteristic Curve بأنه انحدار درجات المفردة بالنسبة إلى مجتمع ما من الأفراد على متغير القدرة (Osterlind, 2012: 53, Bandalos,2018:405)، ويوصف هذا المنحنى بالبارامترات التالية:

أ. بارامتر صعوبة المفردة (b_i)Difficulty Parameter: الذي يشير إلى موقع المنحنى على متصل القدرة (موقع المفرد على تدريج الصعوبة لمفردات الاختبار) (Demars,2010:4).

ب. بارامتر تمييز المفردة (a_i) Discrimination Parameter: الذي يشير إلى درجة تمييز المفردة بين مستويات القدرة المختلفة؛ أي انحدار درجة المفردة على القدرة، ويقاس بظل الزاوية التي يميل بها المماس عند منتصف المنحنى المميز للمفردة على محور القدرة (Demars,2010:5).

ج. بارامتر التخمين (c_i): هو احتمال الإجابة الصحيحة لذوي القدرة المنخفضة، الذي يشير إلى ارتفاع المنحنى عند أدنى مستوى للقدرة (c_i)(علام، ٢٠٠٥: ٧٣).

د. بارامتر القدرة Ability Parameter (θ): هو مستوى القدرة لدى الفرد الذي يجيب عن مفردات الاختبار(Demars,2010:3).

٢. أن دالة الاستجابة للمفردة تمثل العلاقة الحقيقية بين المتغيرين الملاحَظ والكامن؛ أي العلاقة بين استجابة الفرد للمفردة وقدرته الكامنة (Hambleton et al., 1991: 9).

٣. أحادية البعد Undimensionality: أن أداء الفرد يُعزَى إلى قدرة أو سمة واحدة، وأن مفردات الاختبار تقيس قدرة واحدة أو سمة واحدة، ويرى ريكاس (Reckase,1979) أنه يتم التحقق منه بإجراء التحليل العاملي الاستكشافي للمفردات بطريقة تحليل المكونات الرئيسة (Principal Components Analysis) وفحص نتائج قيمة الجذر الكامن للعامل الأول إلى قيمة الجذر الكامن للعامل الثاني، فإذا كانت النسبة تساوي أو أكبر من (٢) فقد تحقق شرط أحادية البعد.

٤. استقلال الموضع Local Independence: ويقصد به أن احتمال الإجابة الصحيحة للفرد على مفردة ما يكون مستقلًّا عن (لا يتأثر بـ) استجاباته للمفردات الأخرى المكونة للاختبار، ويعد استقلال الموضع افتراضًا جوهريًّا؛ حيث يتم في ضوئه انتقاء المفردات(Weiss & Yoes, 1991: 73)، (Irwing & Others,2018:63)، (Hambleton & Swaminathan, 1990:23).

ويعني استقلال الموضع أن المُمتَحن يتناول كل مفردة في الاختبار على حِدَة، دون أن يكتسب معرفة من استجاباته للمفردات الأخرى في الاختبار (Osterlind, 2012: 51)؛ أي أن استجابة المُمتَحن للمفردة لا تتأثر إيجابًا أو سلبًا باستجابته لبقية مفردات الاختبار، فإذا كان هناك مفردتان مستقلتان موضعيًّا، فإن احتمال النجاح أو الفشل في الإجابة عن المفردة الأولى لا يؤثر في احتمال النجاح أو الفشل في الإجابة عن المفردة الثانية (Wainer & Mislevy, 2000: 91). ويُعَبر عن استقلال الموضع إحصائيًّا بعدم وجود ارتباط إحصائي بين مفردات الاختبار لدى المُمتَحنين ذوي نفس مستوى القدرة، ويرتبط افتراض استقلال الموضع بافتراض أحادية البعد؛ إذ يعتبر هامبلتون وآخرون (Hambleton & Others,1991: 11) أن استقلال الموضع ينتج تلقائيًّا عن أحادية البعد. كما بيَّن كروكر وألجينا (Crocker and Algina, 1986: 343) أن الافتراضين غير متكافئين، ويرى دي بويك وويلسون (De Boeck and Wilson, 2004: 52) أن استقلال الموضع يعني أن القدرة الكامنة لها أثر جوهري في استقلال المفردات أو ارتباطها، فتكون الاستجابات على المفردات مستقلة عند مستوى محدد من القدرة. ومن أبرز المؤشرات التي استخدمت للكشف عن استقلال الموضع بين المفردات، مؤشر Q3 وهو معامل ارتباط البواقي بين أزواج المفردات عند نفس المستوى من القدرة. وتعتمد قيم البواقي على الفرق بين الدرجة المقاسة والدرجة المتوقعة للنموذج، وتتراوح قيمة المؤشرQ3 بين (−١) و(١)، وتشير القيمة المطلقة الأكبر إلى درجة أكبر من ارتباط الموضع؛ أي درجة أكبر لانتهاك افتراض استقلال الموضع، فالقيمة المطلقة للمؤشرQ3 تساوي أو أكبر من ٠,٢٢٣٦ ($|Q3| \geq \sqrt{0.05} = 0.2236$) تستخدم كحد فاصل يعبر عن وجود تباين مشترك يساوي أو يزيد على (٥%)، ويعكس درجة انتهاك افتراض استقلال الموضع (de Ayala, 2009:134)، (Paek & Cole, 2019:41).

٥. التحرر من السرعة في الإجابة Speediness: تفترض نظرية الاستجابة للمفردة أن عامل السرعة في الإجابة ليس له تأثير في الإجابة عن مفردات الاختبار، وأن الإخفاق في الإجابة عن مفردة اختبارية ما يرجع إلى انخفاض القدرة، ويمكن تقدير ما إذا كان عامل السرعة قد أدى دورًا في الإجابة من خلال تحديد عدد الأفراد الذين لم يتمكنوا من الانتهاء من الإجابة عن جميع مفردات الاختبار(علام، ١٩٨٦).

نماذج نظرية الاستجابة للمفردة ثنائية التصحيح Dichotomous IRT Models

١. النموذج أحادي المَعْلَم (البارامتر) One – Parameter Logistic Model

ويطلق عليه نموذج راش Rash Model؛ حيث وضعه العالم الرياضي الدنماركي جورج راش (George Rash) عام ١٩٦٦م، وطوره العالم الأمريكي رايت (Wright)، ويعتبر حالة خاصة من النموذج ثلاثي البارامتر، والفكرة الأساسية لنموذج راش هي أن احتمالية إجابة الفرد إجابة صحيحة عن مفردة ما تتوقف على ما يملكه الفرد من السمة الكامنة (θ_j)، التي تقيسها المفردة التي صعوبتها (b_i)، فإذا كانت قيمة قدرة الفرد أكبر من صعوبة المفردة، فإن احتمال الإجابة عن السؤال أكبر من ٠,٥ والفرق بين ($\theta_j - b_i$) يكون مقدارًا موجبًا، وإذا كانت قدرة الفرد أقل من صعوبة المفردة فإن احتمال الإجابة عن السؤال أقل من ٠,٥ والفرق بين ($\theta_j - b_i$) يكون مقدارًا سالبًا، وإذا كانت قدرة الفرد تساوي صعوبة المفردة فإن احتمال الإجابة عن السؤال تساوي ٠,٥ والفرق بين ($\theta_j - b_i$) يساوي صفرًا.

والصيغة الرياضية للنموذج أحادي البارامتر، هي:

$$P_i(\theta) = \frac{e^{\theta_j - b_i}}{1 + e^{\theta_j - b_i}}$$

حيث إن $P_i(\theta)$ احتمال أن يجيب فرد (j) إجابة صحيحة عن المفردة (i).

و(θ_j) هي درجة قدرة الفرد (j). و(b_i) هي درجة صعوبة المفردة (i).

٢. النموذج ثنائي المَعْلَم (البارامتر) Two – Parameter Logistic Model

يقوم هذا النموذج على كون مَعْلَمي الصعوبة والتمييز متغيرين، وأن التخمين لجميع المفردات يساوي صفرًا، وتمتاز عملياته الحسابية بأنها أكثر صعوبة من نموذج راش والصيغة الرياضية لهذا النموذج هي:

$$P_i(\theta) = \frac{e^{a_i(\theta_j - b_i)}}{1 + e^{a_i(\theta_j - b_i)}}$$

حيث إن $P_i(\theta)$ احتمال أن يجيب فرد (j) إجابة صحيحة عن المفردة (i)، و(θ_j) هي درجة قدرة الفرد (j)، و (b_i) هي بارامتر صعوبة المفردة (i)، و(a_i) هي بارامتر تمييز المفردة (i).

فبالإضافة إلى صعوبة المفردة يوجد القوة التمييزية للمفردة بين مستويات القدرة، التي تتناسب طرديًّا مع ميل المنحنى المميز لها عند نقطة انقلابه؛ فالميل البسيط يدل على تمييز منخفض، والميل الحاد يدل على تمييز مرتفع، حيث يُعرف معامل تمييز المفردة بأنه ميل المنحنى المميز للمفردة "ظل الزاوية التي يصنعها المماس مع المنحنى" عند مركزه؛ أي عند النقطة التي تحدد

بارامتر الصعوبة (Bachman, 1990: 205)، حيث يوضح ميل المنحنى إلى أي درجة تفصل المفردة بين الأفراد خلال متصل القدرة، ويُرمز إلى بارامتر تمييز المفردة بالرمز (ai)، وينتمي نظريًا إلى الفترة من سالب ما لا نهاية إلى موجب ما لا نهاية (van der Linden & Hambleton, 1996: 5)، ويُعرف تمييز المفردة بأنه معدل التغير في احتمال الإجابة الصحيحة تبعًا للتغير في مستوى القدرة (Thorndike, 1997: 500).

٣. النموذج ثلاثي البارامتر "المَعْلَم" Three – Parameter Logistic Mode

ويقوم على ثلاثة معالم هي: الصعوبة، والتمييز، والتخمين، ويتميز عن النموذج ثنائي المَعْلَم بأنه أضاف بارامتر التخمين (الخط التقاربي السفلي لمنحنى خصائص المفردة)، ويمثل قيمة الاحتمال المقابلة لأدنى مستوى من القدرة؛ أي قيمة احتمال نجاح الأفراد متدني القدرة في الإجابة عن المفردة، وبكلمات أخرى هو قيمة الحد الأدنى الذي يحصل عليها الفرد عندما تكون قدرته أو السمة الكامنة لديه تساوي صفرًا (Weiss & Yoes, 1991: 76).

والصيغة الرياضية لهذا النموذج هي:

$$P_i(\theta) = c_i + (1-c_i)\frac{e^{a_i(\theta_j - b_i)}}{1+e^{a_i(\theta_j - b_i)}}$$

حيث إن $P_i(\theta)$ احتمال أن يجيب فرد (j) إجابة صحيحة على المفردة (i).

و(θ_j) هي بارامتر قدرة الفرد (j)، و(b_i) هي بارامتر صعوبة المفردة (i).

و(a_i) هي بارامتر تمييز المفردة (i)، و(c_i) هي بارامتر التخمين.

ثالثًا: الدراسات السابقة

لقد أجريت العديد من البحوث والدراسات التي استخدامت نماذج نظرية الاستجابة للمفردة في بناء بنوك أسئلة، ومنها: دراسة محمود (٢٠١٢)، التي سعت إلى التحقق من استخدام نموذج راش في بناء بنك أسئلة في مقرر سيكولوجية التعلم، فتم اختيار خمسة موضوعات من المقرر وتم إعداد اختبار تحصيلي في كل موضوع من خلال تحليل المحتوى، وتم اختيار تصميم رابطة المفردات لإجراء ضم الاختبارات الخمسة، فاستخدمت رابطة مفردات داخلية، ومن ثم كان هناك أربع صور اختبارية، تم تطبيقها على العينة، وقد ضمت (١٢٤٧) ممتحنًا، وتم استخدام برنامج ConQuest 2.0 لإجراء التحليل الإحصائي في ضوء نموذج راش، وتم تقييم ملاءمة المفردات للنموذج بفحص معاملات تمييز المفردات وفحص إحصاءات الملاءمة الخارجية والداخلية، وتم استبعاد المفردات التي تجاوزت قيمة الملاءمة الخارجية أو الداخلية

الخاصة بها حدود الدلالة، وأعيد تحليل المفردات، فأظهرت جميع المفردات المتبقية ملاءمة جيدة.

ودراسة العديلات (٢٠١٢)، التي هدفت إلى بناء بنك أسئلة في الرياضيات لطلبة الصف الرابع وفقًا لنظرية الاستجابة للمفردة (النموذج ثلاثي المَعْلَم) والنظرية الكلاسيكية في القياس، حيث تم بناء تجمع من ١٠٢ مفردة، من خلال تحليل محتوى منهاج الرياضيات للصف الرابع الأساسي، وبناء جدول المواصفات لتمثيل الأهداف المعرفية، ووزعت هذه المفردات على ثلاثة اختبارات، تكونت كل منها من ٤٢ مفردة؛ منها ١٢ مفردة مشتركة في كل اختبار، وتم تطبيق الاختبارات على ١٢٠٠ طالب وطالبة، وتم استخدام برنامج BILOG-Mg في تحليل البيانات باستخدام النموذج ثلاثي المَعْلَم، وبينت نتائج التحليل أن ٢٥ من ١٠٢ مفردة لم تطابق مع النموذج الثلاثي المَعْلَم، وأن مَعْلَمة صعوبة المفردات تراوحت من -١,٨٣٩ إلى ١,٣٢ لوجيت وبمتوسط ٠,٢٥٥ لوجيت، وأن مَعْلَمة تمييز المفردات تراوحت ما بين ٠,٨٦٤ و٢,٩٦٢ بمتوسط ٢,٠١٦، أما مَعْلَمة التخمين للمفردات فقد تراوحت ما بين ٠,٠٦ و٠,٤٥١ بمتوسط ٠,٢٤. وتم استخدام برنامج SPSS في الحصول على معاملات الصعوبة والتمييز وفقًا للنظرية الكلاسيكية في القياس، وبينت نتائج التحليل أن معاملات الصعوبة للمفردات تراوحت بين ٠,٢٦ و ٠,٩١ بمتوسط صعوبة ٠,٥٨، وأن معاملات التمييز تراوحت بين ٠,٢٨ و ٠,٩٠ بمتوسط تمييز للمفردات ٠,٥٥، وبنيت قاعدة بيانات لتخزين ٧٧ مفردة وتم تخزين مع كل مفردة البيانات التالية: الوحدة الدراسة، والإجابة الصحيحة، ومَعْلَم الصعوبة، ومَعْلَم التمييز، ومَعْلَم التخمين، ودالة المعلومات.

ودراسة الأبرط (٢٠١٢)، التي هدفت إلى اشتقاق معايير لاختبارات مختارة من مفردات بنك أسئلة في مقرر الرياضيات للثانوية العامة في اليمن، استنادًا إلى تقديرات القدرة حسب نظرية الاستجابة للمفردة، حيث تم بناء جدول مواصفات اشتمل على المواضيع المدرجة في منهاج الرياضيات للصف الثالث الثانوي العلمي في الجمهورية اليمنية، واعتمادًا على جدول المواصفات تم بناء تجمع من المفردات يتكون من ٢٢٠ مفردة من نوع الاختيار من متعدد، تكرر منها ٢٠ مفردة مشتركة، وقد توزعت هذه المفردات على أربعة اختبارات؛ اختباران متكافئان من حيث المحتوى، يتكون كل منهما من ٦٠ مفردة، منها ١٠ مفردات مكررة في كل صورة للوحدات الخمس الأولى، واختباران متكافئان من حيث المحتوى، يتكون كل منها من ٣٠ مفردة؛ منها ١٠ مفردات مكررة في كل صورة للوحدات السادسة والسابعة. طبقت الاختبارات الأربعة على عينة

مؤلفة من ١٢٠٠ طالب من طلبة الصف الثالث الثانوى العلمي في محافظة ذمار في الجمهورية اليمنية، ثم تم تحليل البيانات باستخدام برامج BILOG-MG , SPSS، ومن خلال نتائج التحليل تم حذف ٢٦ مفردة لعدم مطابقتها للنموذج من أصل ٢٢٠ فقرة، فأصبح عدد المفردات المتبقية ١٩٤ مفردة كونت بنك الأسئلة، وتم تخزينها في برنامج FastTest v2.0؛ منها ٩٨ مفردة خاصة بالوحدات الخمس الأولى، و٩٦ مفردة خاصة بالوحدتين السادسة والسابعة.

ودراسة محمد (٢٠١٦) التي استهدفت المقارنة بين النماذج اللوغاريتمية الثلاثة لنظرية الاستجابة للمفردة (الأحادي، الثنائي، الثلاثي) البارامتر في تدريج بنك للأسئلة الموضوعية في مادة الرياضيات للصف الخامس الابتدائي، وذلك من حيث دقة القياس التي تتمثل في انخفاض قيمة الخطأ المعياري لصعوبة المفردات، وانخفاض قيمة الخطأ المعياري لقدرة الأفراد، وارتفاع مؤشرات الثبات لبنك الأسئلة، وارتفاع قيم دوال المعلومات، وتم استخدام ١٠ صور اختبارية، وهذه الصور تكون بنك أسئلة، يتكون من ٧٨ مفردة اختبارية، وبلغت عينة الدراسة ١٥٠٠ تلميذ وتلميذة، من الصف الخامس الابتدائي، وأسفرت النتائج عن أن أكثر النماذج اللوغاريتمية الثلاثة ملاءمة للبيانات هو النموذج ثلاثي البارامتر، وهو الأفضل في دقة القياس.

ودراسة عبدالفتاح (٢٠١٧) التي اهتمت بالتحقق من أثر تقدير الدرجات وطول الاختبار وحجم العينة لاختبار تحصيلي في مادة العلوم على مَعْلَمي (قدرة الفرد – صعوبة المفردة) ودالة معلومات الاختبار في ضوء نموذج راش، وبلغت العينة النهائية للدراسة (٤٧٦) تلميذًا وتلميذة من تلاميذ الصف الثاني الإعدادي، وتضمنت أدوات الدراسة: اختبارين محكي المرجع (موضوعي – مقالي) في مادة العلوم، وأسفرت نتائج الدراسة عن أن مَعْلَم صعوبة المفردة ومَعْلَم قدرة الفرد ودالة معلومات الاختبار تختلف باختلاف كل من نمط مفردة الاختبار، وحجم عينة التحليل، وطول الاختبار الموضوعي، وطول الاختبار المقالي.

ودراسة الجمل وآخرين (٢٠١٧)، التي هدفت إلى تدريج بنك أسئلة مادة الرياضيات لتلاميذ الصف الثالث الابتدائي باستخدام النموذج اللوغاريتمي ثلاثي المَعْلَم، والكشف عن مدى اختلاف تقديرات قدرات الأفراد باختلاف الاختبارات المسحوبة من البنك "الاختبار السهل، الاختبار الصعب"، كما تهدف الدراسة إلى الكشف عن مدى اختلاف دالة معلومات الاختبار باختلاف مستوى العينة "منخفضة القدرة، مرتفعة القدرة". وتم تدريج بنك أسئلة مكون من ٢٢٥ مفردة على عينة تدريج بلغت ٥٢٩ تلميذًا وتلميذة بالصف الثالث الابتدائي، في حين تمثلت عينة الدراسة النهائية من ١٦٦٨، وخلصت الدراسة إلى أنه لا توجد فروق دالة إحصائيًا في تقديرات قدرات الأفراد باختلاف صعوبة الاختبار "الاختبار السهل، والاختبار الصعب" كما توصلت الدراسة

إلى أنه لا توجد فروق ذات دلالة إحصائية في دالة معلومات الاختبار باختلاف مستوى القدرة "منخفضة القدرة، ومرتفعة القدرة".

ودراسة حمزة (٢٠١٧)، التي هدفت إلى تدريج مقياس التدفق النفسي باستخدام نظرية الاستجابة للمفردة، مستعينة بنموذج راش لمقياس التقدير على ميزان تدريج خطي واحد له صفر واحد، وله وحدة قياس ثابتة ومعرفة، وذلك بعد اختبار كفاءة مقياس التقدير المستخدم والتحقق من توافر أحادية البعد في بيانات التحليل، وتمثلت عينة البحث في (٤٠٠) طالبة من طالبات الفرقة الأولى بكلية البنات جامعة عين شمس، وتكون المقياس في صورته الأولية من (٧٢) مفردة تغطي (٩) مكونات، وبلغ عدد مفردات المقياس النهائية ٦٦ مفردة بعد حذف المفردات غير الملاءمة وأمكن تدريج مفردات المقياس وحساب تقديرات الأفراد المقابلة لكل درجة خام.

ودراسة سخيم (٢٠١٨)، التي هدفت إلى بناء بنك أسئلة في مقرر العلوم لطلاب الصف السادس، حيث تم بناء تجمع من ١٢٠ مفردة من خلال تحليل محتوى مقرر العلوم لطلاب الصف السادس، وبناء جدول المواصفات لتمثيل الأهداف المعرفية بمفردات، ووزعت هذه المفردات من حيث المحتوى في أربعة اختبارات، تكون كل اختبار من ٣٠ مفردة، وتكونت عينة الدراسة من ٣٩٨ طالبًا وطالبة، وتم استخدام برنامج SPSS في الحصول على معاملات الصعوبة والتمييز وفقًا للنظرية الكلاسيكية في القياس، وبينت نتائج التحليل أن متوسط معاملات الصعوبة للنماذج الاختبارية تراوحت بين ٠,٠٣٢ و٠,٤١، وأن متوسط معاملات التمييز للنماذج الاختبارية تراوحت بين ٠,٣٣ و٠,٣٨ وتم استخدام برنامج R في الحصول على مَعْلَم الصعوبة وفقًا للنموذج أحادي البارامتر، وبينت نتائج التحليل أن متوسطات معالم الصعوبة للنماذج الاختبارية تراوحت بين -١,١٧ إلى -٠,٨٢، وأنه وفقًا لنموذج راش تم حذف ٤ مفردات، وأصبح العدد النهائي المتبقي ١١٦ مفردة. ووفقًا لنظرية القياس الكلاسيكية فإنه يتم استبعاد ٢٣ مفردة والمتبقي يكون ٩٧ مفردة.

دراسة الربابعة (٢٠١٩)، التي هدفت إلى بناء بنك أسئلة في مستويات أحكام التلاوة والتجويد الثلاثة في القرآن الكريم باستخدام النموذج أحادي المَعْلَمة، واستخراج المؤشرات الإحصائية لمفردات البنك، حيث تم بناء جدول المواصفات، ثم تم بناء تجمع من المفردات عددها الإجمالي ٤٥٩ مفردة، منها ١٤٧ مفردة للمستوى الأول، والمستوى الثاني، ١٦٠ مفردة والمستوى الثالث ١٥٢ مفردة. وتكونت عينة الدراسة من ١١٠٥ طالبة من المستويات الثلاثة لأحكام التلاوة والتجويد في مراكز تحفيظ القرآن الكريم في محافظة عجلون، حيث إن عدد طالبات

المستوى الأول ٤٠٧ طالبة، المستوى الثاني ٤٠٣ طالبة، والمستوى الثالث ٣٩٥ طالبة، وتم تحليل البيانات ببرنامج BILOG-MG باستخدام النموذج أحادي المَعْلَمة، حيث تم مطابقة البيانات للنموذج المستخدم، وبلغ متوسط الصعوبة لنموذجي المستوى الأول ٢,١٣، ٢,٢ ولنموذجي المستوى الثاني ٢,١٣، ١,١٨ ولنموذجي المستوى الثالث -١,١٤، -١,٢٧ . وتم حذف المفردات غير المطابقة للنموذج أحادي المَعْلَمة، ثم تم استخدام برنامج Item Bank Management Platform 2019؛ لتخزين الفقرات المطابقة في قواعد بينات بنوك أسئلة المستويات الأول والثاني والثالث، وكان عددها للمستوى الأول ١٤١ مفردة، وللمستوى الثاني ١٥٢ مفردة، وللمستوى الثالث١٢٩ مفردة.

دراسة كوان وآخرين (Kwan & Others, 2019)، التي هدفت إلى بناء وتدريج بنك أسئلة جديد للفكر الإيجابي لقياس جودة الحياة المرتبطة بالصحة في سنغافورة، حيث أُعِدَّ تجمع مبدئي من المفردات يتكون من ٤٨ مفردة مستمدة من الأدوات المعدة مسبقًا والصالحة للاستخدام في سنغافورة، وصيغت لتطبيقها باللغة الإنجليزية على عينة تتكون من ٤٩٣ مريضًا في المستشفي، وبعد جمع البيانات تم التحقق من افتراض أحادية البعد باستخدام التحليل العاملي الاستكشافي والتحليل العاملي التوكيدي، وكذلك التحقق من افتراض استقلال الموضع باستخدام مؤشر Q3، ثم استخدم برنامج Xcalibre في تدريج مفردات بنك الأسئلة، بعد مطابقة البيانات للنموذج المستخدم وأصبح عدد المفردات المكونة لبنك الأسئلة النهائي ٣٦ مفردة، وبلغ معامل الثبات ألفا كرونباخ ٠,٩٧، وتراوحت قيم مَعْلَمة الصعوبة للمفردات بين -٣,٦ إلى ٢,٤ لوجيت، وقيم مستويات القدرة تراوحت بين -٣,٥ إلى ٠,٥ لوجيت، وكانت المعلومات القصوى عند مستوى قدرة (ثيتا)= -٢,٩.

ودراسة عالم وأبو هادي(٢٠١٩)، التي هدفت إلى بناء بنك أسئلة في مادة الرياضيات للصف التاسع الأساسي في ضوء نظرية الاستجابة للمفردة باستخدام النموذج اللوغاريتمي ثلاثي المَعْلَمة، والتحقق من جودة توظيف برنامج حاسوبي لإدارة مفردات بنك الأسئلة من حيث: التخزين، والاستدعاء، واختيار مفردات منه لتشكيل اختبار جديد. ولتحقيق الهدف الأول تم تحليل المحتوى وتحديد مخرجات التعلم المستهدفة، ثم تم بناء تجمع من المفردات، وتشكيل جدول لتمثيل كل مخرجات التعلم بمفردات اختبارية. حيث تم صياغة ٢١٠ مفردة اختبارية من نوع الاختيار من متعدد موزعة على ستة نماذج متكافئة من حيث المحتوى، تكون كل منها من ٤٠ مفردة، منها ٦ مفردات مكررة في كل نموذج. وطبقت الاختبارات على عينة ١٢٠٠ طالب وطالبة من طلبة الصف التاسع من التعليم الأساسي بمدارس صنعاء، حيث طبق كل نموذج على ٢٠٠

طالب وطالبة. ثم تم تحليل البيانات باستخدام برنامج BILOG-MG، وقد بينت نتائج التحليل فيما يتعلق بمطابقة المفردات أن ٤ مفردات من أصل ٢١٠ مفردة لم تطابق النموذج ثلاثي المَعْلَمة، في حين نتائج المطابقة للأفراد أظهرت أن ٤ أفراد من أصل ١٢٠٠ فرد لم يطابق النموذج ثلاثي المَعْلَمة. ولتحقيق الهدف الثاني للدراسة تم توظيف برنامج حاسوبي لإدارة بنك الأسئلة مصمم من قبل الباحثين، حيث تم تخزين ٢٠٦ مفردة مع المعلومات الخاصة بها، وبلغة حاسوبية تعطي المخرجات حسب المطلوب من الاختبار المراد تكوينه. وعند انتقاء مفردات من البنك بمواصفات معينة استجاب البرنامج لذلك، وقدم المفردات التي تحقق الشروط المطلوبة مع المعلومات الخاصة بكل مفردة، كما تم ملاحظة أن المفردات المنتقاة توزعت على الوحدات الدراسية بشكل يستجيب لرغبة المستخدم، وتبعًا للائحة المواصفات التي بُني عليها الاختبار.

ودراسة ليونج وآخرين (Leung & Others, 2020)، التي هدفت إلى بناء وتدريج بنك أسئلة لقياس الوظائف البدنية في سنغافورة، حيث أعد تجمع من المفردات يتكون من ٦١ مفردة وطبق على عدد ٤٩٦ حالة، وبعد جمع البيانات تم التحقق من افتراض أحادية البعد باستخدام التحليل العاملي الاستكشافي والتحليل العاملي التوكيدي، ومن افتراض استقلال الموضع باستخدام مؤشر Q3، وتم تدريج مفردات بنك الأسئلة بعد مطابقة البيانات للنموذج المستخدم، وأصبح عدد المفردات المكونة لبنك الأسئلة النهائي ٥٥ مفردة، وبلغت قيمة معامل الثبات ألفا كرونباخ ٠,٩٠، وتراوحت قيم مَعْلَمة الصعوبة للمفردات بين -٤,٢٥ إلى -٠,٧٦ لوجيت، وقيم مَعْلَمة التمييز للمفردات بين ١,٠ إلى ١,٧٦، وقيم مستويات القدرة تراوحت بين -٣,٥ إلى -١,٥ لوجيت.

منهجية البحث وإجراءاته

إجراءات البحث:

لبناء بنك الأسئلة تم اتباع الخطوات الآتية:

أولاً: التخطيط

أثناء مرحلة التخطيط لبناء بنك أسئلة، يتم تحديد المحتوى المستهدف بالقياس، والإعداد لبناء تجمع من المفردات لقياس هذا المحتوى، ويتم تحديد طريقة توزيع المفردات؛ لبناء نماذج متعددة من الاختبارات، بعد اختيار تصميم جمع البيانات؛ حيث إن توزيع المفردات على النماذج الاختبارية يعتمد عليه، وتحديد عينة جمع البيانات في ضوء الإمكانات المتاحة، مع تقدير الفترة الزمنية اللازمة لتطبيق كل نموذج اختباري، مع الوضع في الاعتبار أن لا يقل عدد

الممتحنين عن (١٠٠) ممتحن للإجابة عن كل نموذج من النماذج الاختبارية (Thorndike, 1982:15)، وأيضًا يتم تحديد برمجيات الكمبيوتر التي سوف تستخدم في المعالجات الإحصائية وفي إدارة بنك الأسئلة، وفي مرحلة التخطيط ولتنفيذ هذا البحث تم تحديد:

* المحتوى المستهدف بالقياس: الباب الأول بمادة الكمبيوتر للصف الأول الإعدادي، ويتضمن موضوعات (أساسيات نظام الكمبيوتر، وأنظمة التشغيل).

* تصميم جمع البيانات: تصميم المجموعة الواحدة Single Group Design لبناء بنك الأسئلة.

* عدد المفردات لكل هدف لا تقل عن ثلاث مفردات.

* عدد النماذج الاختبارية: ثلاثة نماذج.

* الفترة الزمنية اللازمة لتطبيق النموذج الاختباري تقريبًا ٤٥ دقيقة.

برامج الكمبيوتر المستخدمة في المعالجات الإحصائية وإدارة بنك الأسئلة:

١. برنامج SPSS V26 لإدخال البيانات بعد تصحيحها ولإجراء التحليل العاملي الاستكشافي.

٢. برنامج R للحصول على تقديرات لمؤشر Q3 للكشف عن استقلال الموضع بين المفردات.

٣. برنامج Xcalibre لإجراء المعالجات الإحصائية المتعلقة بنظرية الاستجابة للمفردة.

٤. برنامج FastTEST professional testing system V2.3 كنظام متكامل لإدارة بنك الأسئلة.

ثانيًا: إعداد الاختبارات

تم اتباع الخطوات التالية لإعداد الاختبارات:

١. تحديد الهدف العام من الاختبار:

يهدف الاختبار إلى قياس تحصيل طلاب الصف الأول الإعدادي في موضوعات (أساسيات نظام الكمبيوتر، وأنظمة التشغيل) بمادة الكمبيوتر وتكنولوجيا المعلومات والاتصالات.

٢. تحليل المحتوى:

تم تحليل محتوى موضوعات في ضوء الأهداف السلوكية وفق تصنيف بلوم على مستويات (التذكر، الفهم) وللتحقق من صدق وثبات تحليل المحتوى قام الباحث بالإجراءات الآتية:

أ. صدق المحكمين: حيث تم عرض تحليل المحتوى على بعض الخبراء والمتخصصين في مجال المناهج وطرق التدريس وتكنولوجيا التعليم وعلوم الحاسب لإبداء الرأي حول ما ورد بالتحليل، وقد تم إجراء التعديلات اللازمة حسب اقتراحاتهم.

ب. ثبات تحليل المحتوى: قام الباحث بتحليل محتوى موضوعات التعلم المختارة، وقامت باحثة أخرى بتحليل نفس الموضوعات في ذات الوقت، وبحساب معامل ثبات التطابقي (الصياد، ١٩٨٨) معامل الثبات التطابقي = عدد الأهداف المتطابقة لدى الباحثين/ عدد الأهداف الكلي.

معامل الثبات التطابقي = ٣٢/٣٥ = ٠,٩١

وكذلك تم تقدير معامل الثبات بطريقة هولستي "معامل الاتفاق" (Holisti, p.141) التي تنص على:

$$C.R = \frac{2M}{N_1 + N_2}$$

حيث إن M عدد الأهداف المتفق عليها بين الباحثين = ٣٢

و $N_1 + N_2$ مجموع الأهداف لدى الباحثين = ٣٥ + ٣٥ =٧٠

وقد أشارت النتائج إلى أن معامل الثبات التطابقي وكذلك معامل الثبات لهولستي = ٠,٩١ وهو ما يشير إلى أن ثبات تحليل المحتوى عبر الباحثين جيد ومقبول.

٣. إعداد جدول مواصفات الاختبار:

تم إعداد جدول المواصفات وفقًا للخطوات الآتية:

أ. تحديد الأهمية النسبية لكل موضوع من موضوعات التعلم، وذلك في ضوء: عدد الحصص، وعدد الصفحات.

ب. تحديد الأهمية النسبية لأهداف تدريس كل موضوع على حدة.

ج. الصورة النهائية لجدول المواصفات.

ولقد اقترح المحكمون أن لا يزيد عدد الأسئلة على (٣٥) في نموذج الاختبار، ولكي يأخذ جدول المواصفات شكله النهائي، تم تحديد عدد المفردات لمستوى التذكر، ومستوى الفهم؛ لكل موضوع كما يلي:

جدول رقم (١)

يوضح الصورة النهائية لجدول مواصفات الاختبار التحصيلي

الموضوع	الأهداف		المجموع
	تذكر	فهم	
الأول – أساسيات نظام الكمبيوتر	٢٥	٥	٣٠
الثاني – أنظمة التشغيل	٤	١	٥
المجموع	٢٩	٦	٣٥

٤. صياغة مفردات الاختبار:

تم صياغة المفردات في صورة اختبارات موضوعية معتمدة على نوعين من الأسئلة؛ الاختيار من متعدد، والصواب والخطأ؛ حيث تتميز هذه الاختبارات بارتفاع معدل صدقها وثباتها بالإضافة إلى تمتعها بدرجة عالية من الموضوعية، ولأنها ملائمة لقياس التحصيل في مختلف الأهداف التعليمية (محاسنة، ٢٠١٣: ١٦٣). وقد تم صياغة ٣ مفردات؛ لقياس كل هدف على حدة، وتم تكوين تجمع من المفردات عدده ١٠٥ مفردات. وللتحقق من صدق محتوى المفردات المصاغة قام الباحث بعرضها على بعض الخبراء والمتخصصين في المناهج وطرق التدريس وتكنولوجيا التعليم؛ لإبداء الرأي في مدى مناسبة ودقة المفردات لغويًا وعلميًا لعينة الدراسة (طلاب الصف الأول الإعدادي)، ومدى مناسبة كل مفردة في قياس الهدف التي صيغت من أجله، ومدى خلو المفردة من أي مؤشرات للإجابة الصحيحة عنها أو عن المفردات الأخرى. وتم عمل التعديلات اللازمة في ضوء آرائهم واقتراحاتهم.

أدوات البحث

تم توزيع تجمع المفردات وتكوين ثلاثة نماذج اختبارية Test Forms تقيس نفس الأهداف، يتكون كل نموذج اختباري من (٣٥) مفردة، وفقًا لجدول المواصفات المعد سابقًا بإجمالي ١٠٥ مفردات؛ لتشكل بنك الأسئلة، وتم استخدم تصميم المجموعة الواحدة Single Group Design لجمع البيانات ولضم "ربط" مفردات النادج الاختبارية، حيث تم تطبيق النماذج الثلاث الاختبارية على نفس الطلاب.

عينة البحث

تكونت عينة جمع البيانات من طلاب الصف الأول الإعدادي بمدرسة آمون الخاصة في محافظة الإسماعيلية، وبلغ عددهم (١٧٩) طالبًا وطالبة.

التحقق من افتراضات نظرية الاستجابة

١. التحرر من السرعة في الإجابة:

أتيح للطلاب الوقت الكافي للإجابة عن أسئلة كل نموذج من النماذج الاختبارية، حيث أعطيت لهم الفرصة للتفكير في كل مفردة على حدة ثم الإجابة، ومن ثم التقليل من التخمين، وأيضًا تحقق افتراض أن عامل السرعة في الإجابة ليس له تأثير في الإجابة عن مفردات الاختبار.

٢. استقلال الموضع:

للتحقق من عدم انتهاك افتراض استقلال الموضع للمفردات تم استخدام مؤشر Q3.

جدول رقم (٢)

ملخص نتائج مؤشر Q3

المؤشر	الاختبار الأول	الاختبار الثاني	الاختبار الثالث	بنك الأسئلة
نسبة عدد أزواج المفردات التي تظهر ارتباطًا موضعيًا (Q3 أكبر من 0.22) إلى العدد الكلي لمعاملات ارتباط البواقي بين أزواج المفردات Percentage of residual correlations above 0.22	$\frac{5}{595} = 0.0084$	$\frac{2}{595} = 0.0033$	$\frac{3}{595} = 0.0050$	$\frac{35}{5460} = 0.0064$
القيمة المطلقة لمتوسط معاملات ارتباط البواقي بين أزواج المفردات Average absolute residual correlations < 0.10	0.0621	0.0631	0.0627	0.0639

يتبين من الجدول رقم (٢) أن القسمة المطلقة لمتوسط معاملات ارتباط البواقي بين أزواج المفردات أقل من ٠,١ لكل من نماذج الاختبارات الثلاث: الأول والثاني والثالث، وأيضًا لبنك الأسئلة، كما أن نسبة عدد أزواج المفردات التي أظهرت ارتباطًا موضعيًّا (Q3 أكبر من 0.22) إلى العدد الكلي لمعاملات ارتباط البواقي بين أزواج المفردات تراوحت بين ٠,٠٠٣٣ إلى ٠,٠٠٨٤ وهذه النتائج تشير إلى تحقق شرط استقلال الموضع.

٣. أحادية البعد:

تم التحقق من افتراض أحادية البعد باستخدام التحليل العاملي الاستكشافي وفقًا لما ورد في ريكاس (Reckase,1979)

جدول رقم (٣)

نسبة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني للنماذج الاختبارية الثلاثة ولبنك الأسئلة

العامل	الجذر الكامن (الاختبار الأول)	الجذر الكامن (الاختبار الثاني)	الجذر الكامن (الاختبار الثالث)	الجذر الكامن (بنك الأسئلة)
الأول	٦,٧٨٠	٦,٥٦٠	٧,١٨٠	٢٠,٩٧٣
الثاني	١,٠٤١	١,٦١٥	١,٦٥٨	٢,٥٣٨
نسبة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني	٦,٥١	٤,٠٦	٤,٣٣	٨,٢٦

يتبين من الجدول رقم (٣) أن شرط أحادية البعد قد تحقق في كل من مفردات نموذج الاختبار الأول والثاني والثالث وكذلك لبنك الأسئلة، حيث إن نسبة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني أكبر من (٢).

نتائج البحث

وللإجابة عن السؤال الأول من أسئلة البحث، الذي نصه: ما دلالات مطابقة البيانات لأحد نماذج الاستجابة للمفردة؟

تم استخدام برنامج إكس كاليبر Xcalibre للحصول على نتائج مطابقة النموذج ككل لكل من(الاختبار الأول- الاختبار الثاني- الاختبار الثالث- مفردات بنك الأسئلة) وفقًا للنماذج الثلاثة (النموذج أحادي البارامتر – النموذج ثنائي البارامتر – النموذج ثلاثي البارامتر) وذلك من أجل التعرف على أنسب نماذج الاستجابة للمفردة للاعتماد عليه في تدريج مفردات بنك الأسئلة.

جدول رقم (٤)

نتائج مطابقة النموذج ككل Overall Model Fit

نموذج الاستجابة للمفردة	الاختبار	عدد المفردات	قيمة مربع كاي المحسوبة	درجة الحرية	P-Value
نموذج أحادي البارامتر	الاختبار الأول	٣٥	٧٧٦,١٦١	٤٩٠	٠,٠٠٠
	الاختبار الثاني	٣٥	٦٢٨,٩٦١	٤٩٠	٠,٠٠٠
	الاختبار الثالث	٣٥	٧١٠,٥٥٩	٤٩٠	٠,٠٠٠
	مفردات بنك الأسئلة	١٠٥	٢٠٥٣,٦	١٤٧٠	٠,٠٠٠
نموذج ثنائي البارامتر	الاختبار الأول	٣٥	٤٨٤,٤٧	٤٥٥	٠,١٦٤
	الاختبار الثاني	٣٥	٤٤٥,٩٤٤	٤٥٥	٠,٦١٠
	الاختبار الثالث	٣٥	٤٤٧,٥٦٧	٤٥٥	٠,٥٨٩
	مفردات بنك الأسئلة	١٠٥	١٣٣٨,٥٨٤	١٣٦٥	٠,٦٩
نموذج ثلاثي البارامتر	الاختبار الأول	٣٥	٤٧٠,٨١٤	٤٢٠	٠,٠٤٤
	الاختبار الثاني	٣٥	٤٧٣,٩٧٤	٤٢٠	٠,٠٣٥
	الاختبار الثالث	٣٥	٤٣٣,٥٣٣	٤٢٠	٠,٣١٤
	مفردات بنك الأسئلة	١٠٥	١٣٦١,٨٨٩	١٢٦٠	٠,٠٢٣

يتبين من الجدول رقم (٤) أن اختبار مربع كاي غير دال إحصائيًّا عند استخدام النموذج ثنائي البارامتر، ودال إحصائيًّا عند استخدام النموذج أحادي وثلاثي البارامتر، ما عدا في حالة استخدام النموذج ثلاثي البارامتر في معالجة بيانات الاختبار الثالث، وهذه النتائج تشير إلى ملاءمة بيانات الاختبارات الثلاثة ومفردات بنك الأسئلة للنموذج ثنائي البارامتر؛ ولذلك فإنه تمت المعالجات الإحصائية باستخدام النموذج ثنائي البارامتر.

والجدول التالي يعرض نتائج مطابقة مفردات بنك الأسئلة للنموذج ثنائي البارامتر.

جدول رقم (٥)

نتائج مطابقة مفردات بنك الأسئلة للنموذج ثنائي البارامتر

المفردة	قيمة مربع كاي المحسوبة	P-Value	المفردة	قيمة مربع كاي المحسوبة	P-Value	المفردة	قيمة مربع كاي المحسوبة	P-Value
١	٩,٩٨٠١	٠,٦٩٧٤	٣٦	٥,٣٩١٩	٠,٦٠٧٣	٧١	٧,٦٨٠٦	٠,٦١٥٦
٢	٢٢,٩١٥٠	٠,٤٩٠٥	٣٧	١٢,٣٣٢٢	٠,٢٣٥٦	٧٢	١٣,٦٣٥٤	٠,٦٦٢٠
٣	١١,٩٣٠٤	٠,٦٠٧٧	٣٨	١١,٤٣٧٦	٠,٦٨٨٤	٧٣	١٣,٢٣٠٧	٠,٦٤٥٢
٤	٢٣,٩٨٥٢	٠,١١٩٣	٣٩	١٥,٩٦٠٠	٠,٥١٥٩	٧٤	١٨,٩٢٦٤	٠,١٩١٨
٥	٩,٣٥٤٧	٠,٤٨٣٥	٤٠	٢٣,١٩٩٠	٠,١٧٠٥	٧٥	٢٢,٣٦٦١	٠,٣٥٨٢
٦	١٠,٠٠٠٨	٠,٧٥١٨	٤١	١٠,٧٩٥٧	٠,٤٤٤٢	٧٦	٨,٧٦٥١	٠,٦٥٠٠
٧	٩,٢٧٠٠	٠,٦٥٦٤	٤٢	٩,٨٧٥٦	٠,٦٤٨٩	٧٧	٤,٨٤٤٧	٠,٧٦٠١
٨	١٥,٣٢٧٨	٠,٦٧٨٣	٤٣	١٢,٢٠٩٤	٠,٤٣٠٦	٧٨	٨,١٢٣٧	٠,٦٣٨٥
٩	٢٣,٤٠٠٠	٠,٢٣٨٥	٤٤	٢٧,١٤٧٨	٠,١٩١٧	٧٩	١٢,١١٧٢	٠,٣٣٩٠
١٠	١٤,١٧١٤	٠,٥٥٤٨	٤٥	١٧,٩٦٥٣	٠,٣٦٦٠	٨٠	١٣,٢٦١٢	٠,٤٩٠٢
١١	٥,٨٥٨٤	٠,٦٥٠٣	٤٦	١١,٥٣٨٩	٠,٨٣٢٩	٨١	٩,٤١٦٦	٠,٧٢٩٠
١٢	١٥,١٣٠٤	٠,٤٢٢٧	٤٧	١٩,٢١٢٤	٠,٣٥٧٣	٨٢	١٢,٤٣١٦	٠,٦٠٨٤
١٣	١٤,٣٢٤٠	٠,٥٨٠٥	٤٨	١٠,٨٧٧٧	٠,٤٥٨٩	٨٣	٧,٩٨٠٩	٠,٥٦٣٤
١٤	١٠,٨٩٩٩	٠,٣٧٩٧	٤٩	١٨,٤٢١٦	٠,٢٣١١	٨٤	٩,٥٩١٥	٠,٥٥٩٩
١٥	١٢,٧٧٢٨	٠,٢٥٦٣	٥٠	٩,٢٥٠١	٠,٣٣٠٨	٨٥	١٢,٠٢١٥	٠,٧٠٥٢
١٦	٩,١٢٩٧	٠,٦٠١٦	٥١	١٤,٦١٤٦	٠,٦٢٥١	٨٦	١٠,١٢٦٧	٠,٦٧٣٧
١٧	١٩,٦٧٩٥	٠,٦٤٠٢	٥٢	٩,٧٣٢٤	٠,٥٨٠٣	٨٧	١١,٣٢٧٠	٠,١٩٣١
١٨	٧,٧٧٨٥	٠,٠٩٧٥	٥٣	٦,٩٤٥٩	٠,٦٠٠٨	٨٨	٩,١٠١٠	٠,٥٩٤٥
١٩	٩,٨٠٦١	٠,٦٤٦١	٥٤	٧,٧٩٨٣	٠,٥٧٩٨	٨٩	١٠,٤٢٤٠	٠,٤٢٣٨

المفردة	قيمة مربع كاي المحسوبة	P-Value	المفردة	قيمة مربع كاي المحسوبة	P-Value	المفردة	قيمة مربع كاي المحسوبة	P-Value
٢٠	١٧,٧٣٢٠	٠,٤٧٣٥	٥٥	٩,٩٨٦٧	٠,٤٥١٢	٩٠	٩,٨٦٥٦	٠,٦٨٤٨
٢١	١٦,٠١٢٨	٠,٣٤٤٤	٥٦	٢١,٦٨٠٧	٠,٤٠٤٩	٩١	١٧,٤١٧٣	٠,٢٤٠١
٢٢	٩,٨٩٤٨	٠,٧٩١١	٥٧	٦,٦٩٨٨	٠,٤٥١٧	٩٢	٩,٤٧٤٧	٠,٥٥٦٧
٢٣	١٤,٠٥٠٢	٠,١٩٥٦	٥٨	١٠,٥٩٥٣	٠,٦٢٠٧	٩٣	٩,٥٥٦٠	٠,٤٦٧٠
٢٤	١١,٤٩٤٨	٠,٥٧٣٩	٥٩	٨,٨٥٢٣	٠,٦٣٤٤	٩٤	١٠,٣٨٢١	٠,٤٥٦١
٢٥	١٣,٦٧٠٨	٠,٢٩٨٤	٦٠	١٠,٨٤٦٦	٠,٥٢٠٦	٩٥	٩,٨٣٤٢	٠,٤٧٢٨
٢٦	١٣,٢٩٣٤	٠,٦٠٥٨	٦١	١٢,٤٦٤٩	٠,٣١٧٣	٩٦	١٢,٢٦٥٥	٠,٧٠٠٥
٢٧	٨,٨٤٤٤	٠,٦٢٩٣	٦٢	١٧,٩٩٢٥	٠,٥٨١٩	٩٧	٤,٨٩٠٠	٠,٧٣٨٠
٢٨	١٤,٨٣٧٥	٠,٤٢٢٢	٦٣	٩,٧٩٢٣	٠,٥٦٩٤	٩٨	١٢,٤١٦٤	٠,٤٢٤٧
٢٩	١١,٦٢٤٧	٠,٣٧٤٦	٦٤	٧,٢٩٠٧	٠,٥٢٤٩	٩٩	٢٢,٥٨٥٧	٠,٣٢٢٨
٣٠	١٩,٦٤٥٦	٠,٥٣١٢	٦٥	٩,٠١٢٧	٠,٣٢٨٠	١٠٠	١٧,٠١٥٥	٠,٥٢٧٩
٣١	١٦,٩٤٤٤	٠,٣٤٦٠	٦٦	٢٢,٥٨٧٤	٠,٦١٧٥	١٠١	١٢,٧١١٤	٠,٥٣٤٣
٣٢	٦,١٥١٠	٠,٧٦٨٨	٦٧	١٧,٢٣٠٨	٠,٣٥٤٦	١٠٢	١٠,١١٦٨	٠,٤٦٦٨
٣٣	١٠,٨٦٢٤	٠,٤٢٤٧	٦٨	٥,٤٣٧٧	٠,٤٥٤٤	١٠٣	١١,٦٧٢٧	٠,٥٦٤٢
٣٤	٧,٣٤٧٧	٠,٣٠٨٢	٦٩	١٦,٨٩٨٧	٠,٤٤١٠	١٠٤	٦,٦٧٣٠	٠,٣٨٧٩
٣٥	٢٥,٧١٣٢	٠,٣٤٩٣	٧٠	٩,٠٤٦٦	٠,٥١١٥	١٠٥	١٥,٠٢٠٥	٠,٣٣١٩

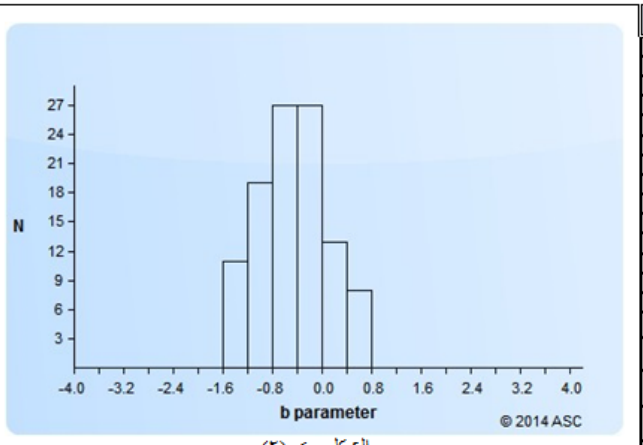
يتبين من الجدول رقم (٥) أن نتائج مطابقة مفردات بنك الأسئلة للنموذج ثنائي البارامتر تحقق مطابقة جيدة، حيث إن قيم اختبار مربع كاي غير دال إحصائيًّا لجميع المفردات.

للإجابة عن السؤال الثاني الذي نصه: ما قيم بارامترات الصعوبة والتمييز ونتائج مطابقة مفردات بنك الأسئلة وفقًا لنموذج نظرية الاستجابة للمفردة المناسب للبيانات؟

تم استخدام برنامج إكس كاليبر Xcalibre للحصول على تقديرات قيم بارامترات الصعوبة والتمييز اعتمادًا على نموذج الاستجابة المفرد ثنائي البارامتر.

جدول رقم (٦)

التوزيع التكراري لفئات بارامتر الصعوبة لمفردات بنك الأسئلة



الشكل رقم (٢)

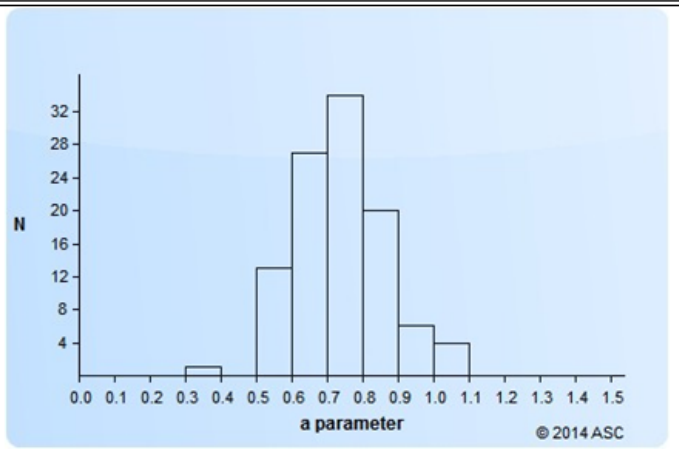
المدرج التكراري لتقديرات بارامتر الصعوبة (b) لمفردات بنك الأسئلة

المدى	التكرار
-4.0 to -3.6	0
-3.6 to -3.2	0
-3.2 to -2.8	0
-2.8 to -2.4	0
-2.4 to -2.0	0
-2.0 to -1.6	0
-1.6 to -1.2	11
-1.2 to -0.8	19
-0.8 to -0.4	27
-0.4 to 0.0	27
0.0 to 0.4	13
0.4 to 0.8	8
0.8 to 1.2	0
1.2 to 1.6	0
1.6 to 2.0	0
2.0 to 2.4	0
2.4 to 2.8	0
2.8 to 3.2	0
3.2 to 3.6	0
3.6 to 4.0	0

من الجدول التكراري رقم (٦) يتبين أن بارامتر الصعوبة يتراوح بين -١,٦ إلى ٠,٨ وهذا يشير إلى أن مفردات بنك الأسئلة قادرة على قياس قدرة الطلاب بدرجة مرتفعة لذوي القدرة المتوسطة في المحتوى العلمي الذي تقيسه مفردات بنك الأسئلة.، ومن الجدول التكراري رقم (٧) يتبين أن قيم تقديرات بارامتر التمييز لمفردات بنك الأسئلة تقع ما بين ٠,٤ و١,١.

جدول رقم (٧)

التوزيع التكراري لفئات بارامتر التمييز لمفردات نواة بنك الأسئلة



المدى	التكرار
0.00 to 0.10	0
0.10 to 0.20	0
0.20 to 0.30	0
0.30 to 0.40	1
0.40 to 0.50	0
0.50 to 0.60	13
0.60 to 0.70	27
0.70 to 0.80	34
0.80 to 0.90	20
0.90 to 1.00	6
1.00 to 1.10	4
1.10 to 1.20	0
1.20 to 1.30	0
1.30 to 1.40	0
1.40 to 1.50	0

الشكل رقم (٣)
المدرج التكراري لتقديرات بارامتر التمييز(a) لمفردات بنك الأسئلة

للإجابة عن السؤال الثالث الذي نصه: ما خصائص بنك الأسئلة من حيث مدى القدرة ودالة المعلومات؟

تم استخدام برنامج إكس كاليبر Xcalibre للحصول على قيم تقديرات القدرة (θ) ودالة المعلومات لبنك الأسئلة اعتمادًا على نموذج الاستجابة المفرد ثنائي البارامتر.

جدول رقم (٨)

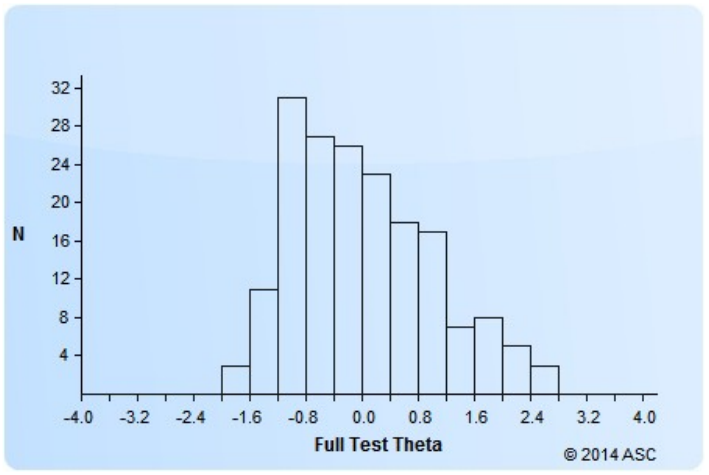
ملخص تقديرات القدرة (θ)

ن	المتوسط	الانحراف المعياري	الالتواء	أقل قيمة	الإرباعي الأدنى	الوسيط	الإرباعي الأعلى	أعلى قيمة
١٧٩	٠	١	٠,٥٩١	-١,٨٨٥	-٠,٨٣١	-٠,١٢٦	٠,٥٨٧	٢,٧٨٤

يتضح من الجدول رقم (٨) أن تقديرات القدرة تراوحت بين -٠,١٨٨٥ إلى ٢,٧٨٤ بمتوسط صفر وانحراف معياري واحد، وأن الإرباعي الأعلى لتوزيع القدرة هو ٠,٥٨٧ في حين بلغ الأدنى -٠,٨٣١.

جدول رقم (٩)

الجدول التكراري لفئات تقديرات القدرة (θ)



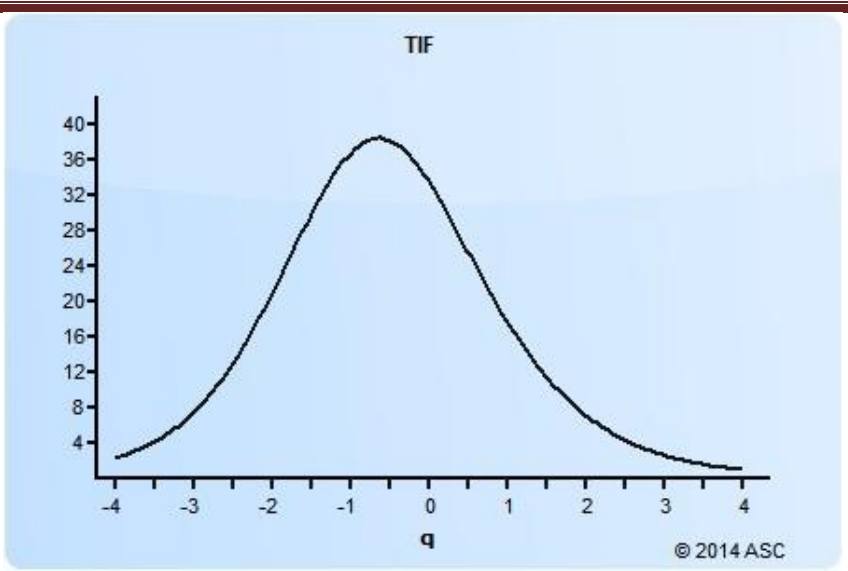
الشكل رقم (٤)

المدرج التكراري لتقدير القدرة (θ)

المدى	التكرار
Below -4	0
-4.0 to -3.6	0
-3.6 to -3.2	0
-3.2 to -2.8	0
-2.8 to -2.4	0
-2.4 to -2.0	0
-2.0 to -1.6	3
-1.6 to -1.2	11
-1.2 to -0.8	31
-0.8 to -0.4	27
-0.4 to 0.0	26
0.0 to 0.4	23
0.4 to 0.8	18
0.8 to 1.2	17
1.2 to 1.6	7
1.6 to 2.0	8
2.0 to 2.4	5
2.4 to 2.8	3
2.8 to 3.2	0
3.2 to 3.6	0
3.6 to 4.0	0
Above +4	0

نلاحظ من الجدول (٨)، (٩) والشكل (٤) أنه من خلال تطبيق مفردات بنك الأسئلة على عدد (١٧٩) طالبًا، فكانت نتائج تقدير القدرة بين -٠,١٨٨٥ إلى ٢,٧٨٤ بمتوسط صفر وانحراف معياري واحد، وأن نسبة الطلاب ذوي القدرات المتوسطة فأقل بلغت تقريبًا ٥٥% في حين أن ٤٥% من الطلاب قدراتهم متوسطة فأعلى.

دالة المعلومات لبنك الأسئلة:



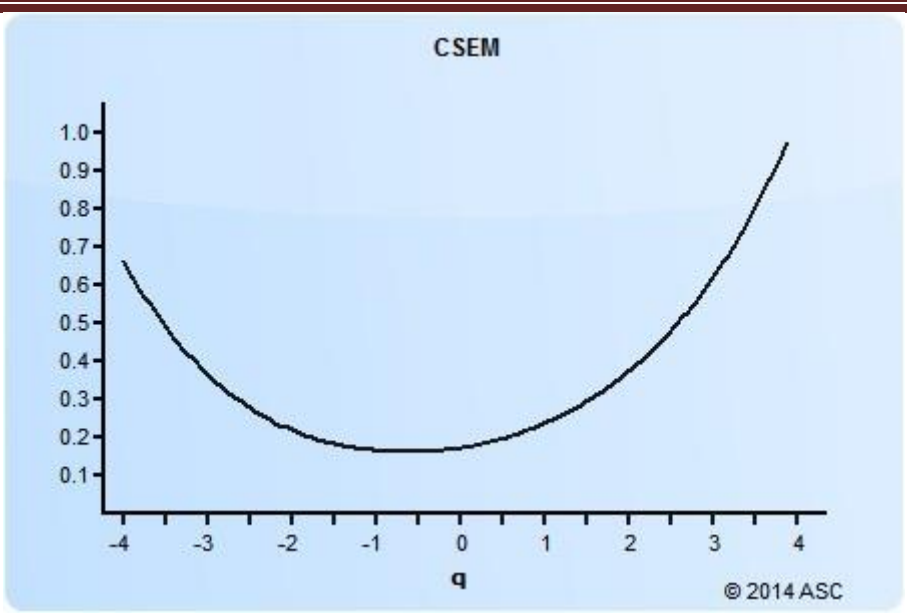
الشكل رقم (٥)

دالة المعلومات لبنك الأسئلة

يعرض الشكل رقم (٥) الرسم البياني لدالة معلومات بنك الأسئلة (TIF) (جميع المفردات التي تم تدريجها لبناء بنك الأسئلة)، ودالة المعلومات تعتبر تمثيلًا بيانيًّا لكمية المعلومات التي يوفرها الاختبار في كل مستوى من القدرة (θ). وكانت المعلومات القصوى = ٣٨,٤٤١ عند مستوى قدرة (θ)= -٠,٦٥

ويعرض الشكل رقم (٦) الرسم البياني لدالة الخطأ المعياري الشرطي للقياس (CSEM) لبنك الأسئلة. وCSEM هو دالة مقلوب دالة معلومات بنك الأسئلة TIF، ويقدر مقدار الخطأ في تقدير القدرة (θ) لكل مستوى من مستويات القدرة، وكان أدنى قيمة لـ CSEM = ٠,١٦١ عند θ = -٠,٦٥.

وتشير نتائج دالة المعلومات لبنك الأسئلة إلى أن مقدار المعلومات التي تفسرها مفردات بنك الأسئلة تختلف باختلاف قدرة الطلاب، حيث إن مقدار المعلومات التي قدمتها مفردات البنك بلغت أقصي درجة(٣٨,٤٤) عند مستوى قدرة (-٠,٦٥) لوجيت وبأدنى خطأ (٠,١٦).

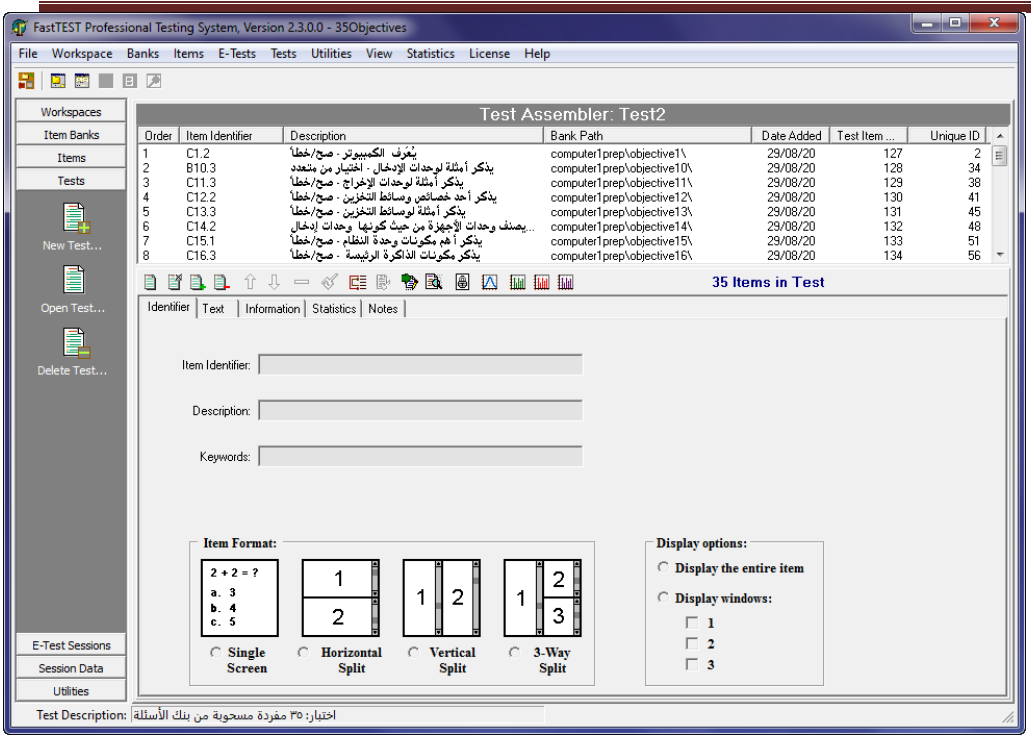


الشكل رقم (٦)

الخطأ المعياري للقياس المشروط لبنك الأسئلة

للإجابة عن السؤال لرابع الذي نصه: ما خصائص اختبار مسحوب مفرداته من بنك الأسئلة من حيث مدى القدرة ودالة المعلومات؟

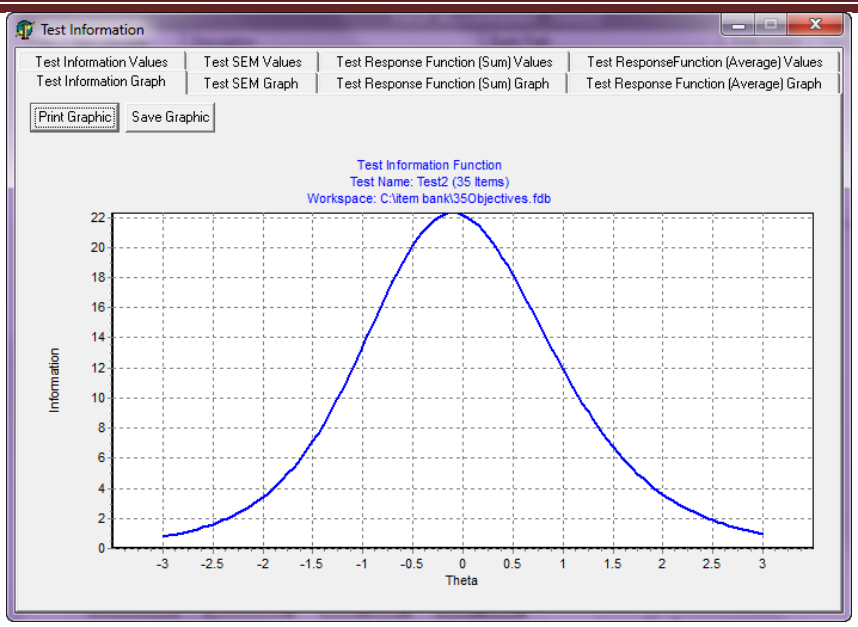
تم استخدام برنامج FastTEST professional testing system V2.3 كنظام متكامل لإدارة بنك الأسئلة، حيث تم إنشاء "ملف قاعدة بيانات" مؤمن بكلمة سر ويتضمن مفردات بنك الأسئلة وما يرتبط بها من معلومات سبق الإشارة إليها، وللحصول على اختبار جديد من بنك الأسئلة، تم إنشاء اختبار وتسميته باسم يعبر عن الهدف منه، ثم اختيار ٣٥ مفردة لتكون الاختبار في ضوء الأهداف السلوكية التي تم إعداده لقياسها، والشاشة التالية توضح واجهة البرنامج.



الشكل رقم (٧)

واجهة برنامج نظام إدارة بنوك الأسئلة FastTEST professional testing system

وفيما يأتي المعلومات التي وفرها برنامج إدارة بنك الأسئلة عن الاختبار المسحوب منه "خصائص الاختبار":

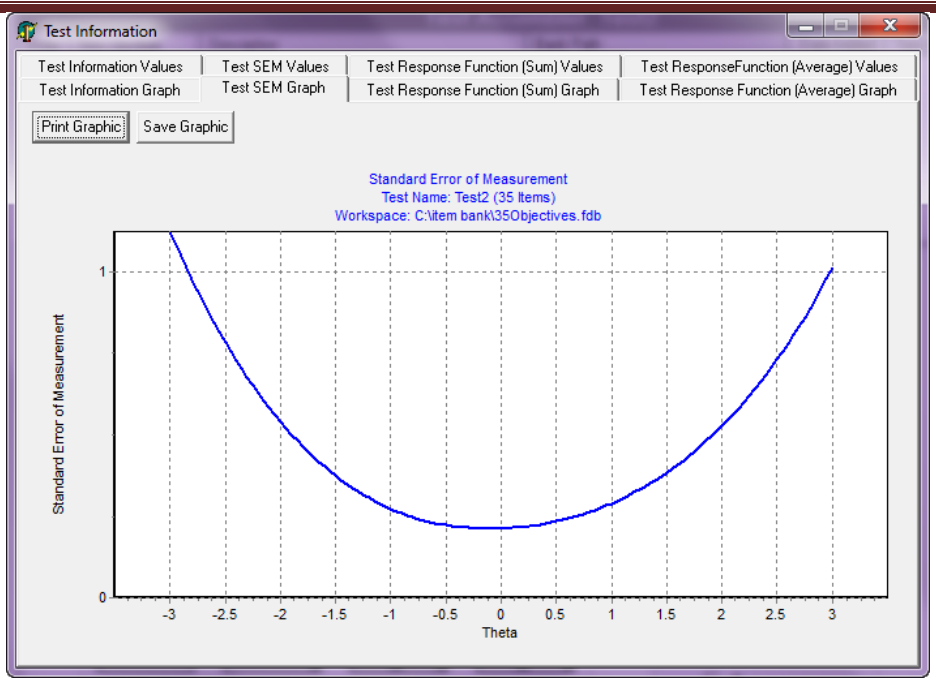


الشكل رقم (٨)

دالة المعلومات الاختبار المسحوب من بنك الأسئلة

ومن خلال الشكل (٨) فإن مستوى دالة معلومات الاختبار الأعلى= ٢٢,٢٨ عند مستوى قدرة=- ٠,١٠

والشكل التالي يعرض الرسم البياني لدالة الخطأ المعياري الشرطي للقياس (CSEM).



الشكل رقم (٩)

الخطأ المعياري للقياس المشروط للاختبار المسحوب من بنك الأسئلة

من الشكل رقم (٩) فإن أدنى قيمة لدالة الخطأ المعياري الشرطي للقياس CSEM = ٠,٢١١٨ عند مستوى قدرة=-٠,١٠.

وتشير نتائج دالة المعلومات للاختبار المسحوب من بنك الأسئلة إلى أن مقدار المعلومات التي تفسرها مفردات الاختبار تختلف باختلاف قدرة الطلاب، حيث إن مقدار المعلومات التي يقدمها الاختبار بلغت أقصي درجة (٢٢,٢٨) عند مستوى قدرة (-٠,١٠) لوجيت وبأدنى خطأ (٠,٢١١٨).

مناقشة النتائج:

سعت الدراسة الحالية إلى التعرف على أنسب نماذج الاستجابة للمفردة؛ ليتم من خلاله تقدير بارامترات الصعوبة والتمييز. وأسفرت نتائج مطابقة النموذج ككل، لكل من (الاختبار الأول، الاختبار الثاني، الاختبار الثالث، مفردات بنك الأسئلة) وفقًا للنماذج الثلاثة (أحادي البارامتر، ثنائي البارامتر، ثلاثي البارامتر) أن اختبار مربع كاي غير دال في حالة استخدام النموذج ثنائي البارامتر، وهو ما يشير إلى ملاءمة بيانات الاختبارات الثلاثة ومفردات بنك الأسئلة لهذا النموذج، ولكن اختبار مربع كاي دال في حالة استخدام النموذج أحادي وثلاثي البارامتر، وهذا يشير إلى

عدم ملاءمة البيانات لهذه النماذج، ما عدا في حالة استخدم النموذج ثلاثي البارامتر في معالجة بيانات الاختبار الثالث. ولذلك فإنه من الأنسب عند إجراء المعالجات الإحصائية لبيانات مفردات بنك الأسئلة وكذلك للاختبارات الأول والثاني والثالث استخدام النموذج ثنائي البارامتر؛ أي أنه في الدراسة الحالية لم يتم تحديد النموذج الذي سوف يتم استخدامه في المعالجات الإحصائية مسبقًا، وتجدر الإشارة إلى أن معظم الدراسات السابقة حددت نموذج الاستجابة للمفردة، ففي دراسة محمود (٢٠١٢)، ودراسة عبدالفتاح (٢٠١٧)، ودراسة حمزة (٢٠١٧)، ودراسة سخيم (٢٠١٨)، ودراسة الربابعة (٢٠١٩)، ودراسة كوان وآخرين (Kwan & Others, 2019)، ودراسة ليونج وآخرين (Leung & Others, 2020)، تم تحديد واستخدام نموذج راش، في حين أنه في دراسة العديلات (٢٠١٢)، ودراسة عالم وأبو هادي(٢٠١٩) تم تحديد واستخدام نموذج ثلاثي البارامتر.

كما توصلت نتائج تدريج مفردات بنك الأسئلة وفقًا لنظرية الاستجابة للمفردة (النموذج ثنائي البارامتر) إلى تحقق مطابقة جيدة مع البيانات، حيث إن قيم اختبار مربع كاي غير دال لجميع المفردات. وتتراوح بارامترات الصعوبة لمفردات بنك الأسئلة بين -١,٦ و٠,٨ وهذا يشير إلى أن مفردات بنك الأسئلة قادرة على قياس قدرة الطلاب بدرجة مرتفعة لذوي القدرة المتوسطة في المحتوى العلمي للبنك. كما أسفرت النتائج عن أن أكثر من (٩٣,٣%) من معاملات تمييز المفردات يقع ما بين ٠,٥ إلى ١,١. وأن تقديرات القدرة تراوحت بين -٠,١٨٨٥ و٢,٧٨٤ لوجيت، وأن نسبة الطلاب ذوي القدرات المتوسطة فأقل بلغت تقريبًا ٥٥% في حين أن ٤٥% من الطلاب قدراتهم أعلى من المتوسط. وأن أعلى قيمة لدالة معلومات بنك الأسئلة= ٣٨,٤٤١ عند مستوى قدرة (ثيتا)= -٠,٦٥.

ومن خلال البحوث والدراسات، تبين أن معظم هذه البحوث والدراسات اعتمدت في جمع البيانات وضم "ربط" المفردات على تصميم المفردات المشتركة بين نماذج اختبارية متعددة تطبق على أفراد مختلفين، وباستخدام أحد نماذج الاستجابة للمفردة يتم تدريج مفردات بنك الأسئلة ثم تخزين المفردات والمعلومات المتعلقة بها في إحدى قواعد البيانات، ولكن في البحث الحالي تم الاعتماد على تصميم المجموعة الواحدة، وتم التحقق من قدرة نظام إدارة بنك الأسئلة في بناء اختبارات معلومة الخصائص.

التوصيات:

في ضوء نتائج البحث الحالية يمكن التوصية بما يأتي:

١. التوسع في إنشاء وحدات متخصصة في بنوك الأسئلة تكون تابعة لوزارة التربية والتعليم، ووزارة التعليم العالي لتوفير بنوك الأسئلة اللازمة للمراحل التعليمية المختلفة وفي مجالات الدراسة المتنوعة، بحيث يسهل سحب الاختبارات اللازمة لأغراض التقويم المختلفة سواء التقويم المرحلي أو النهائي بشكل يضمن تلبية احتياجات المعلمين من ناحية، ويضمن كفاءة الاختبارات ودقتها ومصداقيتها من ناحية أخرى.

٢. تشجيع العمل الفريقي المؤسسي لإنشاء بنوك أسئلة على درجة عالية من الكفاءة والسعي إلى اعتمادها من الجهات المختصة؛ لتيسير عملية الاستفادة الفعلية التطبيقية منها.

٣. ترسيخ فكرة أن بنك الأسئلة عملية مستمرة؛ أي أنها تحتاج إلى تطوير مستمر لتحقيق الأهداف المرجوة منها.

٤. المراجعة الدقيقة لأي اختبارات يتم إضافتها لبنوك الأسئلة بهدف التطوير، وذلك من خلال المتخصصين وفي ضوء معايير محددة.

٥. دعم الأبحاث التطبيقية التي تهدف إلى توفير بنوك أسئلة معدة وفق النظريات الحديثة والحرص على الاستفادة منها.

٦. عقد دورات وورش عمل تدريبية لمسئولي وأعضاء وحدات القياس والتقويم التابعة لوزارة التربية والتعليم، ووزارة التعليم الجامعي.

٧. عقد ندوات توعوية وإرشادية للمعلمين عن أهمية بنوك الأسئلة، وحثهم على المشاركة في وضع مفردات اختبارية تقيس المستوى الحقيقي لأداء طلابهم في المراحل التعليمية المختلفة.

المراجع

الأبرط، محمد عبدالله صالح (٢٠١٢): اشتقاق معايير لاختبارات مختارة من فقرات بنك أسئلة في مقرر الرياضيات للثانوية العامة في اليمن استنادًا إلى تقديرات القدرة حسب نظرية الاستجابة للفقرة، رسالة دكتوراه، كلية العلوم التربوية والنفسية، جامعة عمان العربية، الأردن.

الجمل،هبة محمد إبراهيم (٢٠١٧): تدريج بنك أسئلة مادة الرياضيات لتلاميذ الصف الثالث الإبتدائي باستخدام النموذج اللوغاريتمي الثلاثي المَعْلَم، مجلة كلية التربية جامعة بورسعيد، العدد(٢١)، ٨٣٧ – ٨٦٧ .

الربابعة، صفاء محمد أحمد (٢٠١٩). بناء بنك أسئلة في أحكام التلاوة والتجويد باستخدام النموذج أحادي المعلم، (رسالة ماجستير غير منشورة). جامعة اليرموك.

السرحان، صباح جميل فدعوس(٢٠١٥) : برنامج مقترح لبناء بنك أسئلة للرياضيات لتلاميذ مرحلة التعليم الأساسية الدنيا بالأردن (محافظة المفرق – أنموذجا)، رسالة دكتوراه، معهد البحوث والدراسات الإستراتيجية، جامعة أم درمان الإسلامية، السودان.

الصراف، قاسم (٢٠٠٢). القياس والتقويم في التربية والتعليم، القاهرة، درا الكتاب الحديث.

الصياد، عبد العاطي أحمد(١٩٨٨). نحو بناء مؤشر إحصائي جديد لقياس الثبات بطريقة التطبيق وإعادة التطبيق. بحوث المؤتمر الرابع لعلم النفس في مصر، ٢٥-٢٧ يناير ١٩٨٨، القاهرة: مركز التنمية البشرية والمعلومات.

العديلات، تقى أحمد عبدالله(٢٠١٢) : بناء بنك أسئلة في الرياضيات لطلبة الصف الرابع وفقًا لنماذج نظرية الاستجابة للفقرة، رسالة ماجستير، كلية التربية، جامعة اليرموك، الأردن.

الوليلي، إسماعيل حسن فهيم (٢٠٠٥) . تكافؤ درجات الاختبارات في ضوء نظريتي القياس الكلاسيكية والحديثة دراسة سيكومترية مقارنة، مجلة كلية التربية، جامعة بنها، كلية التربية، المجلد (١٥) العدد(٦٣)، ٩٨- ١٤٩.

إلياس، فوزي (١٩٩٣). مشروع بنوك المفردات الاختبارية مسار جديد تربوي في سلطنة عمان، رسالة التربية، ١٦٥-١٨٤.

أبو جلالة، صبحي (١٩٩٩). اتجاهات معاصرة في التقويم التربوي وبناء الاختبارات وبنوك الأسئلة، الكويت، مكتبة الفلاح للنشر.

أبو هاشم، السيد محمد (٢٠٠٧). التوجهات المستقبلية للتقويم النفسي والتربوي وتطبيقاتها في مجال التربية الخاصة، المجلة العربية للتربية الخاصة، العدد ١١، ١٥٧ص ص-١٨٢.

بخيت، عبدالرحيم، ياركندي، هانم (٢٠٠١). برمجة التقويم وبنوك الأسئلة، المجلة المصرية للدراسات النفسية، العدد (٣٠)، ١-٢٢ .

حمزة، منى إمام محمود (٢٠١٧): تدريج مقياس التدفق النفسي باستخدام نظرية الاستجابة للمفردة، مجلة البحث العلمي في التربية جامعة عين شمس، المجلد(٤)العدد(١٨)، ١٩٣-٢١٦.

سخيم، خالد أحمد حسن، و محمد، محمد حمد النيل. (٢٠١٨) . تصور مقترح لإعداد بنك أسئلة في ضوء جدول المواصفات وخصائص الاختبار الجيد: تطبيقا علي مقرر العلوم الصف السادس بمرحلة التعليم الأساسي باليمن (رسالة ماجستير غير منشورة). جامعة القرآن الكريم والعلوم الإسلامية.

عالم، توفيق علي، و أبو هادى، سوسن حسن. (٢٠١٩). بناء بنك أسئلة في مادة الرياضيات للصف التاسع الأساسي في ضوء نظرية استجابة الفقرة والتحقق من وجود توظيف برنامج حاسوبي لإدارته. أبحاث: جامعة الحديدة – كلية التربية بالحديدة، ع١٥، ١٦٠-١٩٧.

عبدالفتاح، إيمان عبدالفتاح السيد(٢٠١٧) : أثر تقدير الدرجات وطول الاختبار وحجم العينة لاختبار تحصيلي في مادة العلوم على مَعْلَمي (قدرة الفرد – صعوبة المفردة) ودالة معلومات الاختبار في ضوء نموذج راش، رسالة دكتوراه، كلية التربية، جامعة بنها.

علام، صلاح الدين محمود. (١٩٨٦). تطورات معاصرة في القياس النفسي والتربوي، الكويت: جامعة الكويت، إدارة النشر والتأليف والترجمة.

علام، صلاح الدين محمود. (١٩٩٧). استخدام النماذج الإحصائية السيكومترية في تصميم أنظمة بنوك الأسئلة وبنائها. المجلة العربية للتربية –تونس، المجلد (١٧) العدد(٢)، ٨-٣٩.

علام، صلاح الدين محمود. (٢٠٠٠). القياس والتقويم التربوي والنفسي: أساسياته وتطبيقاته وتوجهاته المعاصرة، القاهرة: دار الفكر العربي.

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

عودة، أحمد سليمان. (٢٠٠٢). القياس والتقويم في العملية التدريسية، الأردن، أربد: دار الأمل.

كاظم، أمينة محمد. (١٩٨٨). دراسة نظرية نقدية حول القياس الموضوعي للسلوك " نموذج راش"، الكويت: مؤسسة الكويت للتقدم العلمي، إدارة التأليف والترجمة والنشر.

كاظم، أمينة محمد. (٢٠٠٤).بنوك الأسئلة في التعليم الجامعي، المؤتمر القومي السنوي الحادي عشر، التعليم الجامعي العربي .. آفاق الإصلاح والتطوير، الجزء(١)، ٥٢ ـ ٦٨، القاهرة.

محاسنة، إبراهيم محمد(٢٠١٣):القياس النفسي في ظل النظرية التقليدية والنظرية الحديثة، الطبعة الأولى، عمان، الأردن، دار جرير للنشر والتوزيع.

محمد، محمد جاد محمد (٢٠١٦) : دراسة سيكومترية في دقة بناء بنك أسئلة في مادة الرياضيات باستخدام بعض نماذج نظرية الاستجابة للمفردة، المؤسسة العربية للاستشارات العلمية وتنمية الموارد، المجلد (١٧) العدد(٥٤)، ٢٦٩ – ٢٦٣.

محمود، سومية شكري محمد (٢٠١٢) : استخدام نموذج راش في بناء بنك أسئلة لقياس التحصيل في مقرر سيكولوجية التعلم لدى طلاب كلية التربية بالمنيا، رسالة دكتوراه، كلية التربية، جامعة المنيا.

Anzaldua, R. M. (2002). Item Banks: what, Where, Why and How, Paper Presented at the Annual Meeting of the southwest Educational Research Association (25th, Austin, Tx, Feb.pp 14-16.

Assessment Systems Company (2020). Item banking . . Retrieved from https://www.assess.com/item-banking-can-improve-assessment/

Bachman, L. (1991). Fundamental considerations in language testing. Oxford: Oxford Uni. Press.

Bandalos, D. L. (2018). Measurement Theory and Applications for the Social Sciences. Guilford Publications.

Bennett, R. E. (2002). Inexorable and inevitable: The continuing story of technology and assessment. Journal of Technology, Learning, and Assessment, 1(1). [Available from http://www.jtla.org]

Burghof, K. L. (2001). Assembling an Item Bank for Computerized Linear and Adaptive testing in geography. International Education , v2, n4. [Available from http://www.flinders.edu.au/education/iej]

Choppin, B.(1990). "Evaluation, Assessment and Measurement". In International Encylopedia of Educational Evaluation, Edited by Walberg, H.& Haertel, G., N.Y.:Pergamon Press.

Crocker, L. & Algina, J. (1986). Introduction to classical and modern test theory. New York: Holt, Rinehart & Winston.

De Ayala, R. J. (2009). The theory and practice of item response theory. New York, NY: Guilford Press.

De Boeck, P., & Wilson, M. (2004). Explanatory Item Response Models: A Generalized Linear and Nonlinear Approach. Springer.

Retrieved from https://books.google.com.eg/books?id=pDeLy5L14mAC

Demars, C. (2010). Item response theory. New York: Oxford University Press, Inc.

Hambleton R. & Swaminathan H. (1990). Item response theory: Principles and applications. Boston: Kluwer Academic Publishers.

Hambleton R., Swaminathan H. & Rogers, H. (1991). Fundamentals of item response theory. London: Sage Publications.

Hambleton, R. K., & Swaminathan, H. (1985). Item Response Theory: Principles and Applications. Springer Netherlands.

Holisti, O. (1996): Content Analysis for the Social Sciences and Humanities, London Massachusetts, Addison - Wesley Publishing Company.

Irwing, P., Booth, T., & Hughes, D. J. (2018). The Wiley Handbook of Psychometric Testing, 2 Volume Set: A Multidisciplinary Reference on Survey, Scale and Test Development. Wiley.

Kwan, Y. H., Uy, E. J., Bautista, D. C., Xin, X., Xiao, Y., Lee, G. L., Subramaniam, M., Vaingankar, J. A., Chan, M. F., Kumar, N., Cheung, Y. B., Chua, T., & Thumboo, J. (2019). Development and calibration of a novel social relationship item bank to measure health-related quality of life (HRQoL) in Singapore. Health and quality of life outcomes, 17(1), 82. https://doi.org/10.1186/s12955-019-1150-9.

Leung, Y. Y., Uy, E., Bautista, D. C., Pua, Y. H., Kwan, Y. H., Cheung, Y. B., Xiao, Y., Chua, T., & Thumboo, J. (2020). Calibration of a physical functioning item bank for measurement of health-related quality of life in Singapore. Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation, 10.1007/s11136-020-02535-0. Advance online publication. https://doi.org/10.1007/s11136-020-02535-0.

Linden, W. J. van der (2016). Handbook of Item Response Theory, Volume One: Models. CRC Press.

Millman, J., & Arter, J. (1984). Issues in Item Banking. Journal of Educational Measurement,v21,n4,pp 315-330. Retrieved from http://www.jstor.org/stable/1434584

Nakamura, Y. (2001). Rasch Measurement and Item Banking: Theory and Practice., Educational Resources Information Center (ERIC) ,pp 1-15.

Osterlind, S. J. (2012). Constructing Test Items. Springer Netherlands.

Paek, I., & Cole, K. (2019). Using R for Item Response Theory Model Applications. Taylor & Francis.

Reckase, M. D. (1979). Unifactor Latent Trait Models Applied to Multifactor Tests: Results and Implications. Journal of Educational Statistics, 4(3), 207–230. https://doi.org/10.3102/10769986004003207.

Robitzsch A, Kiefer T, Wu M (2020). TAM: Test Analysis Modules. R package version 3.5-19, https://CRAN.R-project.org/package=TAM.

Rudner, Lawrence M. (1998). Item Banking, ERIC/AE Digest Series, Available at https://eric.ed.gov/?q=ED423310&ft=on&id=ED423310.

Sijtsma, K., & Molenaar, I. W. (2002). Introduction to Nonparametric Item Response Theory. SAGE Publications. Retrieved from https://books.google.com.eg/books?id=xxJaRuLC_4sC.

Smetherham , D. (1979). Banking school knowledge. British Journal of Educational Studies, v27,n1,pp 57-68.

Squires, P. (2003). Concept paper on an item bank approach to testing, applied skills and knowledge (pp. 1-7) Retrieved from www.appliedskills.com.

Thorndike, R. (1982). Applied psychometrics. London: Houghton Mifflin Company.

Thorndike, R. (1997). Measurement and evaluation in psychology and evaluation. London: Prentice-Hall.

Van Der Linden, W & Eggen, T. (1986). An Empirical Bayesian Approach to Item Banking, Applied Psychological Measurement, v10, n 4, pp 345-354.

Van Der Linden, W. J., & Hambleton, R. K. (1996). Handbook of Modern Item Response Theory. Springer New York. Retrieved from https://books.google.com.eg/books?id=aytUuwl4ku0C.

Wainer, H., Dorans, N. J., Flaugher, R., Green, B. F., & Mislevy, R. J. (2000). Computerized Adaptive Testing: A Primer. Taylor & Francis. Retrieved from https://books.google.com.eg/books?id=73d9AwAAQBAJ.

Weiss, D. & Yoes, M. (1991). Item response theory. In R. Hambleton & J. Zaal (Eds.), Advances in educational and psychological testing: Theory and applications (pp. 69- 96). London: Kluwer Academic Publishers.

Wilson, M. (2005). Constructing Measures: An Item Response Modeling Approach. Lawrence Erlbaum Associates. Retrieved from https://books.google.com.eg/books?id=S35iH040YLQC.

Zhu, W.; Fox, C.; Park, Y.; Fisette, L.; Dyson, B.; Graber, C.; Avery, M.; Franck, M.; Placek, H.; Rink, J.; Raynes, D.,(2011). Development and Calibration of an Item Bank for PE Metrics Assessments: Standard 1, Measurement in Physical Education and Exercise Science, v15, n2, pp119-137.

الملخص:

هدف البحث الحالي إلى بناء بنك أسئلة محوسب في ضوء نظرية الاستجابة للمفردة فى موضوعات الباب الأول من مقرر الكمبيوتر وتكنولوجيا المعلومات والاتصالات للصف الأول الإعدادي، وتحليل محتواها وبناء جدول المواصفات، ثم بناء تجمع من المفردات يتكون من ١٠٥ مفردة، وزعت على ثلاثة نماذج اختبارية كل نموذج يحتوى على ٣٥ مفردة وطبقت النماذج الثلاث الاختبارية على نفس الطلاب وعددهم ١٧٩ طالب وطالبة، حيث تم الاعتماد على تصميم المجموعة الواحدة Single Group Design لجمع البيانات ولضم الاختبارات، وتم التحقق من افتراض استقلال الموضع باستخدم مؤشرQ3، وافتراض أحادية البعد باستخدام التحليل العاملي الاستكشافي، وتم تقييم مطابقة النموذج ككل، لكل من (الاختبار الأول، والاختبار الثاني، والاختبار الثالث، ومفردات بنك الأسئلة) وفقًا لنماذج نظرية الاستجابة للمفردة الثلاثة (أحادي البارامتر، وثنائي البارامتر، وثلاثي البارامتر) باستخدام برنامج إكس كاليبر Xcalibre وأسفرت النتائج عن ملائمة البيانات للنموذج ثنائي البارامتر، وتم تقييم ملاءمة المفردات للنموذج ثنائي البارامتر، فأظهرت جميع المفردات ملاءمة جيدة، وتم تدريج وضم مفردات الاختبارات الثلاثة؛ لتكوين بنك الأسئلة، وباستخدام برنامج FastTEST professional testing system V2.3 كنظام متكامل لإدارة بنك الأسئلة، تم إنشاء قاعدة بيانات مؤمنة بكلمة سر، وتتضمن شجرة الأهداف السلوكية للمحتوى، حيث تم حفظ مفردات بنك الاسئلة، وكل مفردة في ملف قاعدة البيانات معرفة برمز وحيد خاص بها يرتبط به معلومات المفردة، بحيث يمكن التعرف على الهدف الذي تقيسه المفردة وفي أي مستوى معرفي، ونص المفردة وطريقة عرضها ونوعها والإجابة الصحيحة ومعاملي الصعوبة والتمييز وفقًا لنظرية القياس الكلاسيكية، وقيمة بارامتر الصعوبة "b" و بارامتر التمييز "a" وقيمة اختبار حسن المطابقة ودلالته، وتم التأكد من قدرة نظام إدارة بنك الأسئلة على سحب مفردات من مستودعه وإنشاء وتكوين اختبارات معلومة الخصائص.

الكلمات المفتاحية: بنك أسئلة – نظرية الاستجابة للمفردة – تصميم المجموعة الواحدة

Developing of a Computerized Item Bank in Accordance with Item Response Theory

Abstract

The research aims to develop a Computerized Item Bank in Accordance with Item Response Theory to the item for assessing the academic achievement of the topics of the first unit of the Computer, Information and Communication Technology course of the first-year preparatory stage students. The target content was analyzed addressing the specified topics; a specification table was developed; a group of items consisting of 105 were gathered; and such items were distributed to three test forms as each form contained 35 items. The three tests were administered to the same students totaling 179 making use of single group design for collecting data, linking tests and calibrating their items to form the Item bank. The position independence assumption was verified using the Q3 index and the one-dimensional assumption was also verified using the exploratory factor analysis. The conformity of the model as a whole was evaluated for each of (the first test, the second test, the third test, the fourth test, and the Item bank) according to the three item response theory models (one-parameter, two-parameter, and three-parameter) using the Xcalibre program. The results revealed the suitability of the data with the two-parameter model. Furthermore, the items of the three tests have been calibrated and linked to create the Item bank, then the new items were linked, calibrated and added to the Item bank by using the Fast TEST professional testing system V2.3 as an integrated system for managing the Item bank. Added to that, a secure database file was created with a password to save all the items within the instructional objectives tree for the content so that each item was defined by a single symbol to the following information and characteristics of the item taking into account the item text, its presentation, type, correct answer, the objective that it measure and the level of knowledge, the factors of difficulty and distinction according to the Classical Test Theory, the value of the difficulty parameter "b", the discrimination parameter "a", the conformity test value and its significance. The ability of the Item bank management system to withdraw items from its repository and create tests with known characteristics was also verified.