بناء اختبار تحصيلي محكي المرجع في القياس والتقويم باستخدام النموذج اللوغاريتمي ثلاثي المعلم لنظرية الاستجابة للمفردة الاختبارية

الملخص:

هدفت الدراسة الحالية إلى بناء اختبار تحصيلي محكي المرجع في القياس والتقويم التربوي، وتقدير خصائصه السيكومترية استنادا إلى النظرية الحديثة في القياس، وقد اشتمل الاختبار على (٤٠) مفردة من نوع الاختيار من متعدد ذات أربعة بدائل، وطبّق على أفراد عينة الدراسة البالغ عددهم (١٠٠٠) طالباً وطالبة في جامعة تبوك. وتوصلت نتائج الدراسة إلى أن افتراضات نظرية الاستجابة للمفردة الاختبارية قد تحققت في بيانات الدراسة، ومطابقة الاستجابات لعدد (٣٣) مفردة من مفردات الاختبار لافتراضات النموذج اللوغاريتمي ثلاثي المعلم حيث كانت الاستجابات لتلك المفردات بصورتها النهائية ضمن حدود المطابقة، وتم التأكد من تمتع الاختبار بمظاهر متعددة من الصدق، وحذْف (٧) مفردات اختبارية لم تطابق النموذج، كما أظهرت النتائج المتعلقة بتقديرات معالم المفردة (الصعوبة، التمييز، التخمين) أنها مقبولة ضمن المعايير التي أوردتها أدبيات القياس والتقويم في هذا المجال.

الكلمات المفتاحية: اختبار محكي المرجع، النموذج اللوغاريتمي ثلاثي المعلم، نظرية الاستجابة للمفردة الاختبارية

Constructing an achievement criterion– referenced test for measurement and evaluation by using a logistic three parameter model of the Item Response Theory

Abstract:

The study aimed at constructing achievement criterion–referenced test for measurement and evaluation. The study estimates the psychometric characteristics for this test by using a logistic three parameter model of the Item Response Theory. The test consisted of (40) multiple choice items with four alternatives. The test was administrated to a sample of (1000) students at the University of Tabuk. The results indicated (33)items satisfied the assumptions of logistic three parameter model.

The responses to those items were found within the fitting boundaries، and they also have several validity indicators، as well as deleting (7) items which did not satisfy the model assumptions Also the results relating to the estimates of item parameters (difficulty, discrimination, guessing) show that it is acceptable within the parameters reported by of measurement and evaluation literature in this field.

Keywords:criterion reference test، logistic three parameter mode item response theory.

مقدمة:

تعتمد عملية القياس التربوي والنفسي على أدوات ومقاييس تقدم بيانات كمية تتيح للتربويين فهم الظاهرة التربوية بشكل محدد، وتعد الاختبارات التحصيلية جزءاً مهما من العملية التعليمية، فمن خلالها يمكن الحكم على مدى تحقق الأهداف التعليمية كما تفيد الطالب في التعرف على مستوى تحصيله الدراسي.

ومن خلال متابعة التطورات في بناء المقاييس التربوية والنفسية نلاحظ اتجاها نحو استخدام القياس معياري المرجع و بالمقابل هناك اتجاه آخر نحو استخدام القياس محكي المرجع، واستخدام نظرية استجابة المفردة الاختبارية، وتجدر الإشارة إلى أن كثيرا من الدراسات تناولت أوجه الشبه والاختلاف بين القياس المحكي المرجع والمعياري المرجع، فالقياس معياري المرجع يعتمد على تفسير الدرجات في ضوء معايير محددة، أي تنسب درجات الطالب في الاختبار إلى جداول معيارية، ومن ثم تحديد مستوى أداء الطالب النسبي بناء على موقع درجته مقارنةً بمتوسط تحصيل المجموعة المعيارية في الاختبار. ويركز هذا الاتجاه على الفروق الفردية بين الطلبة بحيث يوضح مدى الفارق بين مستوى طالب وآخر أي أنه يقيس مقدرة طالب معين مقارنة بمقدرة طلبة آخرين. وتعرضت المقاييس معيارية المرجع لبعض أوجه النقد، حيث تبين أن تلك الاختبارات تعتمد على مقارنة أداء الطالب بأفراد مجموعته الصفية، وبالتالي فإنه من المحتمل أن يختلف موقع الطالب باختلاف خصائص مجموعته المعيارية. بالإضافة إلى أنه ليس بالضرورة أن يمثل متوسط المجموعة المعيارية الأداء المطلوب للنجاح.

وتهتم هذه الاختبارات أيضا بالفروق الفردية بين الطلبة دون النظر إلى مدى إتقان الطالب للمهارات والمعلومات المراد قياسها وذلك لتأهيله لبرامج تدريبية وتعليمية جديدة، وبالتالي فهي لا تساعد على تقويم وتشخيص كل من كفاية الطالب وبعض

البرامج التدريبية القائمة على الكفايات بالقدر المطلوب، أي أنها لا تزود المعلم بمعلومات دقيقة تساعد في اتخاذ القرارات التربوية المناسبة حول مستوى كفاية الطالب والمنهج الدراسي، وهذه الانتقادات لعبت دوراً في ظهور تصورات حديثة في منهجية الاختبارات والمقاييس التربوية أدت إلى ظهور اتجاه آخر أطلق عليه القياس محكي المرجع.

فالواقع الذي يجب أن لا نغفله هو أن التعليم لم يعد مقتصراً على مجرد التمييز بين الطلبة في السمة المقاسة، وإنما يجب أن يركز على اكتسابهم مهارات معينة وتحقيق أهداف محددة، بل وإتقان المهارات والمعلومات، وبالتالي مقارنة أداء الفرد بمستوى أداء معين في مجال معين من السلوك وهذا ما تهدف إليه الاختبارات المحكية. (علام، ٢٠٠٥).

كما أن الاختبارات محكية المرجع تعد أكثر أنواع الاختبارات التحصيلية ملاءمة لقياس تحصيل الطالب وتقويمه، لأنها تحدد المهارات والكفايات المطلوب إتقانها بدقة، حيث يمكن قياسها وملاحظتها بشكل مباشر، مما يساعد على التشخيص، وتصنيف الطلبة إلى متقنين وغير متقنين، وتحديد البرامج العلاجية المناسبة. فالحاجة الماسة لمعلومات أكثر تفصيلاً عن أداء المتعلمين، وظهور فكرة التعلم بإتقان أدت إلى ظهور الاختبارات محكية المرجع، حيث أخذت حركة القياس المحكي حيزاً مهماً عند علماء القياس التربوي. ومع تطور الاختبارات محكية المرجع، تباين تحديد مفهوم الاختبار محكي المرجع، حيث عرفَّه ثورندايك (Thorndaike, 1997, p56) أنه الاختبار الذي يغطي نطاقاً ضيقاً ويستخدم لأجل اتخاذ قرارات التمكن. كما عرَّفه على أنه طريقة لتحديد مستوى الفرد بالنسبة إلى مجال سلوكي محدد. ويرى النبهان (٢٠٠٤، ص ٩٧) أن الاختبار محكي المرجع هو طريقة للتحقق من اكتساب المتعلم الكفايات أو المهارات الأساسية التي تُعَّبر عن نواتج تعليمية محددة. أما العزاوي (٢٠٠٧) فيشير للاختبارات المحكي المرجع على أنها تهتم بمقارنة أداء الطالب بمستوى أداء معين بغض النظر عن أداء المجموعة.

إلا أن عملية التصنيف تتم بناء على مقارنة أداء المتعلم بالنقطة المحددة مسبقاً على متصل التحصيل، والتي تقابل ما يسمى بدرجة القطع التي تفصل بين المتقنين وغير المتقنين للمحتوى الدراسي المقيس. وعرَّف تراتيرا (Tiratira,2009, p41) درجة القطع أنها نقطة على متصل يعد عندها الحد الأدنى المقبول الذي يفصل بين التصنيفات الايجابية والسلبية. وعرَّفها كلاين وفيدلين (Klein, 2009 & Velden) بأنها نقطة على مقياس الدرجات الملاحظة تمُيز بين الأشخاص المتقنين وغير المتقنين. أما بيدل (Biddle,1993) فيشير إلى أن درجة القطع تُستخدم للفصل بين الأشخاص الذين نجحوا في الاختبار والذين لم ينجحوا.

وتعد نظرية الاستجابة للمفردة (Item Response Theory) من التطورات الحديثة في مجال القياس والتقويم التربوي والنفسي، حيث قدمت أساليب سيكومترية ذات فاعلية كبيرة في بناء المقاييس التربوية والنفسية وطريقة تفسير درجات تلك المقاييس مقارنة بالنظرية التقليدية في القياس. وتستند هذه النظرية الاستجابة للمفردة الاختبارية إلى افتراض وجود متصل للسمة بحيث يمكن تقدير احتمال الإجابة الصحيحة للفرد عن المفردة إذا علم موقعه على هذا المتصل.

(Warm ,2008 .p56)

ولنظرية الاستجابة للمفردة الاختبارية مزايا منها: يكون تقدير قدرة الفرد مستقلاً عن عينة المفردات التي تطبق عليه أي أن تقدير قدرات الأفراد يكون متحرراً من خصائص المفردات. وتقديرات معالم المفردات تكون مستقلة إحصائياً عن عينة الأفراد التي استُخدمت في تقدير هذه الخصائص، وأنه يمكن الحصول على إحصائي (الخطأ المعياري في التقدير) لتقدير درجة الدقة في قياس القدرة لكل فرد ويختلف الإحصائي من فرد لآخر

(Hambleton ،swaminathan &Rogers,1999.p11)

وتعد استقلالية القياس بمثابة الفرق الجوهري بين النظرية التقليدية ونظرية الاستجابة للمفردة، وهو ما عبَّر عنه لورد (Lord,1980) بخاصية اللاتغاير Invariance (Zumbo &Rup,2006)، هذا وقد بُنيت هذه النظرية على افتراضات ينبغي تحققها في البيانات لكي تؤدي إلى نتائج يمكن الوثوق بها، ومن أهم الافتراضات:

- أحادية البعد (Unidimensionality): يعني أن هناك قدرة واحدة تفسر أداء الفرد في الاختبار، لذا تسمى بالنماذج أحادية البعد، أما النماذج التي تفترض وجود أكثر من قدرة واحدة تكمن وراء هذا الأداء تسمى متعددة الأبعاد.
- الاستقلال الموضعي Local Independence: أي أن تكون استجابات الفرد على المفردات المختلفة في الاختبار مستقلة استقلالاً إحصائياً عند مستوى معين من القدرة، أي أن استجابة الفرد على مفردة معينة ما يجب أن لا تؤثر سلباً أو إيجاباً على استجابته لمفردة اختبارية أخرى.

(Croker & Algina,1986)

- منحنى خصائص المفردة (Item Characteristic Curve (ICC): وهو علاقة رياضية تربط بين احتمال استجابة الفرد استجابة صحيحة على مفردة الاختبار، وبين القدرة التي تقيسها مجموعة المفردات التي تحتوي على تلك المفردة. ويشير بوضوح إلى أن احتمال الإجابة عن المفردة إجابة صحيحة

يزداد بازدياد قدرة الفرد، بسبب أن المنحنى تراكمي صاعد، وتوصف هذه المنحنيات في نماذج الاختبارات المصممة لقياس سمة واحدة (أحادية البُعد) بدلالة معلم واحد، أو معلمين، أو ثلاثة معالم.

- السرعة في الأداء Speediness: تفترض أن عامل السرعة ليس له دور في الإجابة عن المفردة، أي أن خصائص الفرد في الإجابة على مفردات الاختبار يرجع إلى انخفاض قدرته وليس إلى تأثير عامل السرعة في إجابته

(Allen & Yen, 1989 .p97)

وتوجد مجموعة من الافتراضات خاصة بكل نموذج من نماذج الاستجابة المفردة، حيث إن النموذج الثلاثي يعد الحالة العامة لهذه النماذج، وذلك لأنه يتم فيه تقدير معلم الصعوبة (b)، ومعْلم التمييز (a)، ومعْلم التخمين (c). أما النموذج الأحادي، فيتم فيه تقدير معلم الصعوبة (b) فقط، ويفترض أن التمييز متساوٍ لجميع المفردات الاختبارية، ويساوي واحد، والتخمين (c) متساوٍ لجميع المفردات الاختبارية، ويساوي صفر. أما النموذج الثنائي، فيتم فيه تقدير معلم الصعوبة (b)، ومعلم التمييز (a)، ويفترض أن معلم التخمين (c) متساوٍ لجميع المفردات الاختبارية، ويساوي صفراً، وهذه الافتراضات يصعب تحققها، أيٍ وجود مجموعة من المفردات تمييزها متساوٍ، ويساوي واحداً واحتمالية التخمين تساوي صفراً للمفردات من نوع الاختيار من متعددٍ.

وقد بينت دراسة ستوكنج (Stocking,1990) أن تقدير معالم المفردة يعد قضية مهمة عند استخدام نظرية الاستجابة للمفردة وخاصة في التطبيقات العملية التي تعتمد على تلك التقديرات، وقد أشارت إلى أن التخطيط لعينات التعيير يتم اختيارها بالطريقة العشوائية، ويفترض القائمون ببناء الاختبارات أن السمات تتوزع توزيعاً اعتدالياً في المجتمع، وكلما كان حجم العينة كبيراً وممثلاً لمجتمع الدراسة يكون توزيع السمة في العينة قريباً من التوزيع الاعتدالي، وأن حجم العينة يلعب دورا كبيراً في دقة تقديرات معالم المفردات وقدرة الأفراد عند استخدام نظرية الاستجابة للمفردة. إلا أن الأدبيات التي تناولت نظرية الاستجابة للمفردة لا تشترط التوزيع الاعتدالي كما هو الحال في النظرية التقليدية في القياس (Baker, 2001. p27).

وتتضمن نظرية الاستجابة للمفردة نماذج متعددة من أهمها:

نموذج راش (Rash Model): ويعد من أكثر هذه النماذج شيوعاً، حيث يسمح باختلاف المفردات في صعوبتها، ويعَّبر عنه رياضياً بالعلاقة الآتية:

النموذج ثنائي المعْلم: يسمح باختلاف مفردات الاختبار في صعوبتها وتمييزها، ويعبَّر عنه بالعلاقة الآتية:

والنموذج ثلاثي المعْلم: (الصعوبة، والتمييز، التخمين)، وهذا النموذج هو المستخدم في الدراسة الحالية، و يمثل بالعلاقة الآتية:

حيث ترمز احتمال إجابة الفرد الذي اختير عشوائيا من مستوى القدرة () على المفردة (i) إجابة صحيحة، b: معْلم الصعوبة، a: معلم التمييز ، c معْلم التخمين، معْلم القدرة، D = ١,٧ وتمثل عامل التدريج (Scaling Factor).

وتلعب دالة معلومات الاختبار دوراً مهماً في نظرية الاستجابة للمفردة' إذ يمكن من خلالها تحديد الخطأ المعياري في التقدير. وتتميز دالة معلومات الاختبار بأنها تمثل مجموع دوال معلومات المفردات عند مستوى معين من القدرة، ودالة معلومات الاختبار مستقلة عن عينة المختبرين وبذلك تقدم نظرية الاستجابة للمفردة ميزات إضافية فيما يتعلق بدقة تقدير أخطاء القياس، ويرتبط الخطأ في التقدير ارتباطاً عكسياً بالجذر التربيعي لدالة معلومات الاختبار وفق العلاقة التالية:

دالة معلومات الاختبار:

الخطأ المعياري في التقدير:

وهذا يعني أن الخطأ المعياري في التقدير يكون أقل ما يمكن عند مستويات القدرة التي تناظر أقصى تقدير للمعلومات. ومن الاستخدامات المفيدة لدالة معلومات المفردة إمكانية معرفة مدى مساهمة كل مفردة في دالة معلومات الاختبار بشكل مستقل عن المفردات الأخرى للاختبار، فإذا علمنا قدرات مجموعة من المختبرين فيمكن انتقاء مفردات الاختبار التي تعمل على زيادة المعلومات التي يقدمها الاختبار في المدى الذي تتوزع فيه قدراتهم في الاختبار. (Brannick,2003)

وأجرى عبابنة (٢٠٠٤) دراسة هدفت لتعرُّف أثر حجم العينة وطريقة انتقائها على دقة تقدير معالم المفردة والقدرة لاختبار قدرة عقلية باستخدام أداة مكونة من (٧١) مفردة طُبقت على عينة مؤلفة من (١٠٠٠) فرد وتم استخدام النموذج ثلاثي المعْلم في تعيير مفردات الاختبار. وتوصلت الدراسة إلى أن الدقة في تقدير معْلم التمييز تزداد عند زيادة تباين قدرات المفحوصين، كما تزداد الدقة فبي تقدير معْلمّي القدرة والصعوبة عندما يكون مدى القدرة للمختبرين متوافقاً مع مدى صعوبة مفردة الاختبار وكذلك تزداد الدقة في تقدير معْلم التخمين عند استخدام عينة من ذوي القدرة المنخفضة في تعيير مفردة الاختبار.

وأجرى الشريفين (٢٠٠٦) دراسة هدفت إلى تقدير الخصائص السيكومترية لاختبار محكي المرجع في القياس والتقويم التربوي وفق النظرية الحديثة في القياس. ولتحقيق هدف الدراسة تم بناء اختبار يشتمل على (٥٠) مفردة من نوع الاختيار ذات أربعة بدائل طُبق على عينة الدراسة البالغ عددها (٢٢٢) طالبا وطالبة. وأشارت النتائج إلى مطابقة الاستجابات عن (٣٧) مفردة من مفردات الاختبار لافتراضات نموذج راش، وحذفْ (١٣) مفردة لعدم مطابقتها النموذج، حيث كانت جميع مفردات الاختبار في صيغتها النهائية ضمن حدود المطابقة لمؤشرّي متوسطات المربعات الداخلية والخارجية، وبلغ معامل الثبات للأفراد (٠،٩٠)، ومعامل الثبات للاختبار (٠،٩٥)، وتم التأكد من تمتع الاختبار بمظاهر متعددة من الصدق، وأن الاختبار يقدم أكبر قدر من المعلومات عند مستويات القدرة المتوسطة، ويقدم أقل قدر من المعلومات عند مستويات القدرة العالية والمنخفضة.

وأجْرى حمادنه (٢٠٠٩) دراسة هدفت لاستخدام نظرية الاستجابة للمفردة في بناء اختبار تحصيلي محكي المرجع في مقرر الإحصاء لطلبة الصف الأول الثانوي في مرحلة التعليم العام، حيث أعد الباحث (٢٨) مفردة من نوع الاختيار من أربعة بدائل، وتكونت عينة الدراسة من (٤١١) طلبة ذلك الصف من مدارس محافظة المفرق في الأردن. وأظهرت نتائج الدراسة إلى تحقق افتراضات نظرية الاستجابة للمفردة بيانات الدراسة، ومطابقة الاستجابات لعدد (٢٤) مفردة من مفردات الاختبار لتوقعات النموذج ثلاثي المعْلم وحذف (٤) مفردات لم تطابق النموذج، كما بينت النتائج المتعلقة بتقديرات معالم المفردة (الصعوبة، والتمييز، والتخمين) أنها كانت مقبولة ضمن المحكات التي أوردتها أدبيات القياس.

وأجْرى العيثاوي (٢٠٠٩) دراسة هدفت لاستخدام نموذج راش في بناء اختبار تحصيلي محكي المرجع في مقرر علم نفس الفروق الفردية، ولتحقيق أهداف الدراسة، وأعد الباحث اختباراً مكونا من (١٧٣) مفردة من نوع الاختيار من متعدد ذات أربعة بدائل تغطي مستويات الأهداف المعرفية الثلاثة الأولى بحسب تصنيف بلوم (المعرفة، والفهم، التطبيق) . واشتملت عينة البحث على (٣٣٦) طالباً وطالبة من الجامعات العراقية، حيث تم اختيارهم بالطريقة الطبقية العشوائية، واستخدم الباحث برنامج راسكال (Rascal) لتحليل البيانات. وتضمنت الصورة النهائية للاختبار (١٠٨) مفردة مطابقة لافتراضات النموذج، فقد تم حذف (٢٨) مفردة لا تحقق أحادية البُعد من خلال التحليل العاملي للمفردات بطريقة المكونات الأساسية، فيما حذف (٢٩) مفردة لعدم المطابقة مع افتراضات نموذج راش، كما تم حذف (٨) مفردات لا تحقق استقلالية القياس. وتم تعيير (١٠٨) مفردة وفقاً لصعوبة المفردات وقدرات الأفراد للحصول على الاختبار بصيغته النهائية، وبلغت قيمة معامل ثباته (٠،٩٥٢) .

وأجرى السامرائي والخفاجي (٢٠١٢) دراسة هدفت لبناء اختبار تحصيلي محكي المرجع في مادة علم نفس لطلبة أقسام العلوم التربوية والنفسية باستخدام نموذج راش، اشتمل الاختبار على (١٦٠) مفردة من نوع الاختيار من متعدد ذات أربعة بدائل، لقياس (١٦٠) هدفاً سلوكياً تغطي مفردات المادة. وطُبق الاختبار على عينة طبقية عشوائية اشتملت على (٣٤٩) طالباً وطالبة. وقام الباحث بتحليل بيانات الاختبار وفقا لنموذج راش وأُجرى تحليل عاملي بطريقة المكونات الأساسية، الذي أظهر أن مفردات الاختبار تقيس عاملاً عاماً واحداً ما عدا (٢٢) مفردة اختبارية تم استبعادها من الاختبار. وأظهرت النتائج وجود (١٣) مفردة كانت قيمة مربع كاي لعشر مفردات منها دالة إحصائياً، لذا تم استبعادها من الاختبار، فيما استبعد برنامج RASCAL ثلاث مفردات لعدم مطابقتها لنموذج راش، واقتراب معامل التمييز للاختبار من (١)، كما طابقت المفردات الاختبارية المتبقية والبالغة (١٢٥) مفردة اختبارية متطلبات استقلالية القياس، كما بلغت قيمة ثبات الاختبار (٠،٩٥٧) .

وأجْرت علي (٢٠١٢) دراسة هدفت للكشف عن فاعلية استخدام نموذج راش في بناء اختبار تحصيلي محكي المرجع لمقرر القياس والتقويم في التربية ولتحقيق أهداف البحث قامت الباحثة ببناء اختبار محكي المرجع للمقرر المذكور، وتكون من ٩٩ مفردة اختبارية من نوع اختيار من متعدد، أما عينة البحث فقد تكونت من (٤٢٦) طالبا وطالبة من طلبة كلية التربية قسم معلم الصف في جامعة دمشق، واستخدم برنامج (Winsteps)، وتم حذف (١٠) مفردات بسبب وقوعها خارج حدود المطابقة، وبالتالي أصبح العدد النهائي لمفردات الاختبار (٨٩) مفردة تقع جميعها ضمن حدود المطابقة. وبينت النتائج أن معامل ثبات المفردات قد بلغ (٠,٩٨) ومعامل ثبات الأفراد (٠,٩٤) وكليهما معامل مرتفع، ولحساب معامل ثبات الاختبار محكي المرجع عند تعييره بنموذج راش تم حساب معامل الثبات المتعلق بالأفراد والثبات المتعلق بمفردات الاختبار من خلال معامل الفصل (G) Item Separation Index لكل من الأفراد والمفردات الاختبارية وقد بلغ معامل الفصل المتعلق بالأفراد(٤,٠١) ومعامل الثبات للأفراد (٠,٩٤) في حين وصل معامل الفصل المتعلق بالمفردات (٦,٤٩) ومعامل الثبات للمفردات (٠,٩٨). كما بينت النتائج أن قيمة دالة المعلومات للاختبار تقدم أكبر قدر من المعلومات عند مستويات القدرة المتوسطة، وكانت أكثر المفردات التي تقدم معلومات عن مقرر القياس والتقويم في التربية هي المفردات (٤٥، ٤٦) وتبلغ كمية المعلومات (٢٠،٣٩) في حين تنخفض كمية المعلومات عند مستويات القدرة العالية والمتدنية.

وأجرى الجبوري (٢٠١٢) دراسة هدفت لبناء اختبار تحصيلي محكي المرجع لمادة الادارة والاشراف التربوي وفق نموذج راش، واشتمل الاختبار في صيغته الأولية على (١٥٣) مفردة، وطبق الاختبار على (٣١٠) من الطلبة اختيروا بالأسلوب الطبقي

العشوائي من (٩) معاهد في محافظة بغداد. وتم التأكد من افتراضات النموذج أحادي البعد باستخدام التحليل العاملي الاستكشافي، وحذفت (١٠) مفردات لانخفاض تشبعاتها بالعامل، واستخدم الباحث برنامج RASCAL لإجراء مطابقة المفردات لنموذج راش، واستبعد (١٤) مفردة ، كما أُجري اختبار مطابقة استقلالية الفقرات حيث استبعدت(١٨) مفردة لعدم استقلاليتها وتكون الاختبار بصورته النهائية من (١١١) مفردة، وكان معامل ثبات الاختبار (٠،٩٥).

وأجرت حجازي والخطيب (٢٠١٤) دراسة هدفت للكشف عن التوافق بين النظرية الكلاسيكية والنموذج ثنائي المعْلم في مطابقة مفردات اختبار محكي المرجع في المادة النظرية لأحكام التلاوة والتجويد. واشتمل الاختبار على (٤١) مفردة من نوع الاختيار، وطبق على عينة مكونة من (٤٠٤) طالبا وطالبة تم اختيارهم بالطريقة العشوائية الطبقية من(١٦) مركزاً قرآنياً (٩ ذكور، ٧ إناث) في العاصمة عمان. وأظهرت النتائج مطابقة ٤٠ مفردة للنظرية الكلاسيكية و٣٩ مفردة اختبارية للنموذج ثنائي المعلم، وكان معامل الثبات وفق النظرية الكلاسيكية ٠،٩٣ ووفق النموذج ثنائي المعلم ٠,٩٤

وعلى الرغم من كثرة الدراسات التي تناولت بناء اختبارات محكية المرجع باستخدام النظرية الكلاسيكية ونظرية الاستجابة للمفردة في مقررات دراسية متنوعة في مرحلة التعليم الجامعي ومرحلة التعليم العام، إلا أن معظمها أهتم ببناء اختبار تحصيلي محكي المرجع باستخدام نموذج راش في مقررات متنوعة لذا جاءت هذه الدراسة لتتناول بناء اختبار تحصيلي محكي المرجع في مقرر القياس والتقويم التربوي الذي يدرس في الجامعات وفق النموذج ثلاثي المعْلم.

أهمية الدراسة:

تكمن أهمية هذه الدراسة في سعيها للتوصل إلى بناء اختبار تحصيلي محكي المرجع في القياس والتقويم التربوي يمتاز بالموضوعية والدقة في قياس السمة، وذلك من خلال إيجاد مفردات متحررة من خصائص الأفراد بحيث يصبح ذا قيمة في استخدامه في الكشف عن مستوى تحصيل طلبة كلية التربية في مجال القياس والتقويم في الجامعات.

وتكمن أهمية هذه الدراسة كذلك من استخدامها للنموذج اللوغاريتمي ثلاثي المعْلم في تطوير الاختبارات محكية المرجع، والتي من شأنها أن تفتح الأبواب لإجراء دراسات لاحقة تبرز التطبيقات العملية للنماذج البارامترية في القياس النفسي والتربوي.

أهداف الدراسة:

بناء اختبار محكي المرجع في مقرر القياس والتقويم التربوي يتمتع بخصائص سيكومترية مقبولة وفق النموذج ثلاثي المعْلم كأحد نماذج نظرية الاستجابة للمفردة الاختبارية.

أسئلة الدراسة:

تسعى الدراسة إلى الإجابة على الأسئلة الآتية:

١ – ما مدى تحقق افتراضات النموذج ثلاثي المعْلم في بيانات الدراسة الحالية؟

٢– ما مدى مطابقة استجابات أفراد العينة للنموذج ثلاثي المعْلم؟

٣– ما معالم الصعوبة والتمييز والتخمين للمفردة الاختبارية؟

مصطلحات الدراسة:

الاختبار المحكي المرجع reference Test Criterion: هو الاختبار الذي يُستخدم في تقييم أداء الفرد بالنسبة إلى محك(مستوى أداء مطلق) دون الحاجة إلى مقارنة أدائه بأداء الأفراد الآخرين. (علام، ٢٠٠١، ص ١٢٩).

النموذج ثلاثي المعْلم Three–Parameter Model: هو أحد نماذج نظرية الاستجابة للمفردة الاختبارية ثنائية التدريج حيث يمكن لهذا النموذج تقدير ثلاثة معالم للمفردة الاختبارية: الصعوبة، التمييز، التخمين.

الطريقة والإجراءات:

مجتمع الدراسة:

تكون مجتمع الدراسة من طلاب وطالبات الدبلوم العام في التربية المسجلين خلال العامين الدراسيين ١٤٣٤–١٤٣٦ هـ في جامعة تبوك في المملكة العربية السعودية.

عينة الدراسة:

تكونت عينة الدراسة من جميع طلاب وطالبات الدبلوم العام في التربية المسجلين في مقرر القياس والتقويم خلال العامين الدراسيين ١٤٣٤–١٤٣٦هـ البالغ عددهم (١٠٠٠) طالبا وطالبة، حيث بلغ عدد الطلبة (٤٥٠) طالبا وعدد الطالبات (٥٥٠) طالبة موزعين على (٢٨) شعبة.

أداة الدراسة:

ومر بناء أداة الدراسة بالمراحل الآتي:

- تحديد الغرض من الاختبار.

- تحديد محتوى الاختبار وذلك بالرجوع إلى وصف المقرر والأدبيات الموائمة، وقد انحصر المحتوى بالموضوعات الآتية: مفاهيم أساسية في القياس والتقويم، وأنواع الاختبارات، وبناء الاختبارات، وتحليل نتائج الاختبار.

- صياغة الأهداف السلوكية التي تناولتها الموضوعات والمتوقع اتقانها من قبل الطلبة.

- قام الباحثان بصياغة (٤٥) مفردة من نوع الاختيار من متعدد ذات أربعة بدائل أحدها يمثل الإجابة الصحيحة لقياس كل هدف من الأهداف السلوكية، وقد تطلب قياس بعضها أكثر من مفردة، وقد روعيت شروط بناء هذا النوع من المفردات الاختبارية، ومطابقتها للهدف السلوكي الذي تقيسه من حيث المحتوى والمستوى المعرفي.

صدق الأداة:

وللتأكد من صدق المحتوى، عُرضت مفردات الاختبار على خمسة محكمين من أعضاء هيئة التدريس في الجامعة الذين يدّرسون مقرر القياس والتقويم التربوي وطُلب من كل منهم أن يبدي رأيه في: درجة الاتفاق بين الأهداف السلوكية والهدف الأساسي للاختبار وتمثيلها لها، ومدى قياس المفردات للهدف السلوكي المتعلق بها، ومدى اتساق الاختبار ككل مع الهدف الأساسي منه، وبناء على ملاحظاتهم حُذفت خمس مفردات، و أجمع المحكمون أن هناك مفردات تؤدي غرضها، وأنها عينة ممثلة للمجال السلوكي الذي يقيسه الاختبار وبلغ عدد مفردات الاختبار (٤٠) مفردة.(ملحق ١) .

ثبات الأداة: تم استخراج معامل ثبات الاتساق الداخلي للاختبار طبقاً لمعادلة كرونباخ ألفا (Cronbach Alpha) باستخدام برنامج SPSS15، وقد بلغ معامل الثبات ٠,٩٠، وتشير هذه القيمة إلى أن الاختبار يتمتع بدرجة مرتفعة في قياس السمة.

إجراءات الدراسة:

- **تم تطبيق الاختبار في صيغته النهائية (٤٠) مفردة على الطلبة قبل نهاية الفصل الدراسي الثاني للعام ١٤٣٥-١٤٣٦هـ، وبعد إتمام عملية التطبيق وجمع أوراق الاستجابات، قام الباحثان بتصحيح أوراق الإجابة، وتم التصحيح بإعطاء درجة واحدة (١) للإجابة الصحيحة والدرجة صفر (٠) للإجابة الخاطئة، وتمثل الدرجة الكلية للطالب التي يحصل عليها في الاختبار المجموع الكلي لعدد مفردات الاختبار والتي أجاب عنها إجابة صحيحة.**

- تم تفريغ استجابات المختبرين على مفردات الاختبار في ملفات خاصة في ذاكرة الحاسوب تمهيدا لإجراء المعالجات الإحصائية اللازمة باستخدام برامج خاصة (SPSS20، BILOG-MG3) وللإجابة على أسئلة الدراسة.

نتائج الدراسة:

السؤال الأول: ما مدى تحقّق افتراضات نظرية الاستجابة للمفردة الاختبارية وفق النموذج ثلاثي المعْلم؟

أولا. افتراض أحادية البعد Unidimensionality:

تم التحقق من هذا افتراض باستخدام التحليل العاملي (Factor Analysis) من خلال برنامج SPSS باستخدام طريقة المكونات الرئيسة Principle Component Analysis، ويُظهر جدول (١) قيم الجذور الكامنة Eigen Values ونسب التباين المفسرExplained Variance للعوامل الأربعة الأولى وناتج قسمة الجذر الأول على العامل الثاني.

جدول (١)

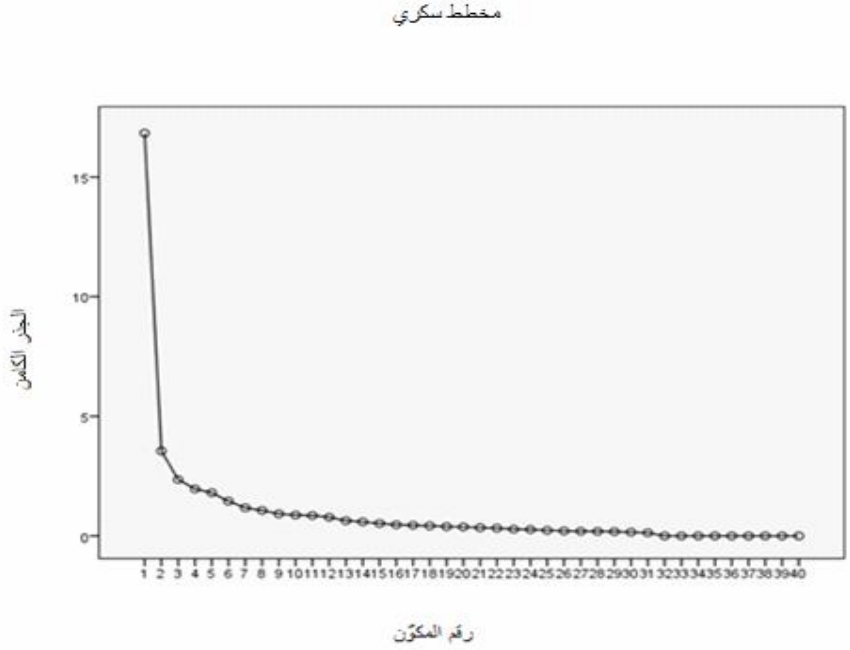
قيم الجذر الكامن و التباين المفسر والتباين المفسر التراكمي للعوامل الأربعة الأولى

العامل	الجذر الكامن	النسبة المئوية للتباين المفسر	النسبة المئوية للتباين المفسر التراكمي	الجذر الكامن الأول / الجذر الكامن الثاني
١	١٦,٨٣٢	٤٢,٠٨١	٤٢,٠٨١	٤,٧٤١
٢	٣,٥٥٠	٨,٨٧٤	٥٠,٩٥٥	

٣	٢,٩٦٧	٥,٨٩١	٦١,٨٤٦
٤	١,٩٦٧	٤,٩١٨	٦١,٧٦٤

يلاحظ من جدول (١) أن قيمة التباين المفسَّر للعامل الأول تخطت الـ ٢٠% كمؤشر على أحادية البعد، كما تخطى ناتج قسمة الجذر الكامن الأول على الجذر الكامن الثاني القيمة(٢) كمؤشر ثانٍ على تحقق أحادية البعد (Hattie,1985)، ويتعزز تحقق افتراض أحادية البعد من خلال تمثيل الجذور الكامنة بيانياً، باستخدام برنامج (SPSS) ، وهو ما يعرف مخطط سكري Scree Plot الذي يظهر في شكل (١).

(Hambelton & Swaminathan,1985)



شكل (١)

شكل بياني يوضح نتائج التحليل العاملي لبيانات عينة الدراسة

ثانيا. افتراض الاستقلال الموضعي Local Independence:

بما أن افتراض أحادية البعد متحقق وفقا لما أشير إليه سابقا، فهذا يعني أن افتراض الاستقلال الموضعي متحقق.

(Hambleton & Swaminathan, 1985)

السؤال الثاني: ما مدى مطابقة استجابات مفردات الاختبار للنموذج ثلاثي المعْلم؟

للتحقق من مطابقة الأفراد والمفردات للنموذج ثلاثي المعْلم تم استخدام برنامج BILOG-MG3 وأظهرت نتائج التحليل أن جميع استجابات الأفراد جاءت متطابقة مع توقعات النموذج باستثناء أربعة طلاب أرقامهم (٣٠٧، ٤٠٥، ٥٠١، ٦٠٥) حيث كانت قيم مربع كاي ذات دلالة إحصائية عند مستوى دلالة ($\alpha \leq$ ٠,٠٥). وبهدف التحقق من درجة مطابقة المفردات الاختبارية للنموذج ثلاثي المعْلم، تم استخدام البرنامج نفسه، وبناءً على مؤشر مربع كاي عند مستوى الدلالة(٠,٠٥ ≤ α) الذي يعتبر المفردة الاختبارية غير مطابقة للنموذج إذا كان مقدار الاحتمال لهذه المفردة أقل أو يساوي ٠,٠٥، وقد تم إيجاد قيم مربع كاي، ومستوى الدلالة وحالة مطابقة المفردة للنموذج ثلاثي المعْلم، ويوضح جدول (٢) قيم هذه المؤشرات.

جدول (٢)

قيم مربع كاي، ومستوى الدلالة للنموذج ثلاثي المعلم

رقم المفردة	كا^2 لحسن المطابقة	الدلالة الإحصائية	رقم المفردة	كا^2 لحسن المطابقة	الدلالة الإحصائية
١	٩,٤	٠,٣٣٠	٢١	٣,٤	٠,٢٠٤
٢	١٢,٧	٠,١٥٠	٢٢	٨٦,٥	٠,٠٠٢
٣	١١,٧	٠,١٦٥	٢٣	١٤٥,٨	٠,٠٠٠
٤	٣٦,٦	٠,٢٨٧	٢٤	١٤٨,٨	٠,٠٠٠
٥	٣٠,٦	٠,١٧٢	٢٥	٦٤,٩	٠,٢٣٦
٦	٩,١	٠,١٢٠	٢٦	٤٧,٤	٠,٠٩٢
٧	٢٦,٦	٠,١٦٨	٢٧	١١,٥	٠,٢٤٤
٨	١٥,٩	٠,٢٦١	٢٨	١٤,١	٠,١٧٦
٩	٩,٣	٠,١٨٣	٢٩	٥٢	٠,١٧٣
١٠	١٢	٠,٢٩٦	٣٠	٢٩,٦	٠,٢١٨
١١	٧,٥	٠,٧٦٠	٣١	٢٢,٧	٠,٠٠٦
١٢	٨,٦	٠,٣١٧	٣٢	٨,٦	٠,٢٢١
١٣	٢١	٠,١٨٩	٣٣	١٠,١	٠,١٥١

١٤	١٧,٥	٠,٢٧٨	٣٤	٢٦	٠,١٥٥
١٥	٣٢,٦	٠,٠٩٥	٣٥	١٢,٣	٠,٤١٨
١٦	٣٢,٧	٠,٢٥٧	٣٦	١١٧,٨	٠,٠٠٠
١٧	٤,٧	٠,١٧١	٣٧	٢٦,٢	٠,٠٠١
١٨	١٠,٢	٠,١٢٠	٣٨	١١,٤	٠,١٨٢
١٩	٤,٤	٠,٢٤٠	٣٩	١٣,٢	٠,٠٩٠
٢٠	١٩,٨	٠,٢٦١	٤٠	٢٦,٢	٠,٠٠١

مستوى الدلالة الإحصائي ٠،٠٥ ≥ α

يظهر من جدول (٢) قيم اختبار مربع كاي ودلالتها الإحصائية عند مستوى ٠,٠٥ ≥ α؛ وأن (٣٣) مفردة من أصل (٤٠) مفردة غير دالة إحصائيًا عند مستوى ٠,٠٥ ≥ α، مما يشير إلى مطابقتها للنموذج ثلاثي المعْلم، و المفردات هي: (١، ٢، ٣، ٤، ٥، ٦، ٧، ٨، ٩، ١٠، ١١، ١٢، ١٣، ١٤، ١٥، ١٦، ١٧، ١٨، ١٩، ٢٠، ٢١، ٢٥، ٢٦، ٢٧، ٢٨، ٢٩، ٣٠، ٣٢، ٣٣، ٣٤، ٣٥، ٣٨، ٣٩) . ملحق (١)

أما بالنسبة للمفردات الاختبارية (٢٢، ٢٣، ٢٤، ٣١، ٣٦، ٣٧، ٤٠) فقد دلت قيمة مربع كاي على وجود دلالة إحصائية عند مستوى ٠,٠٥ ≥ α، وهذا يشير إلى عدم مطابقة هذه المفردات للنموذج ثلاثي المعْلم.

السؤال الثالث: ما قيم تقديرات معالم (التمييز، والصعوبة، والتخمين) للاختبار والأخطاء المعيارية في تقديرها وفق النموذج ثلاثي المعلم؟

وللإجابة عن سؤال الدراسة تم إيجاد معالم مفردات الاختبار: الصعوبة، والتمييز، والتخمين كما في جدول (٣).

جدول (٣)

قيم معالم مفردات الاختبار (الصعوبة والتمييز والتخمين) والأخطاء المعيارية

رقم المفردة	(التمييز)	الخطأ المعياري للتمييز	(الصعوب)	الخطأ المعياري للصعوبة	(التخمين)	الخطأ المعياري للتخمين
١	١,٢٦٦	٠,٢١٧	-٠,٩٥٢	٠,٣٢٧	٠,٣١٨	٠,١٠٣
٢	١,٦٤١	٠,٢٩٦	-٠,٢٦٢	٠,١٩٨	٠,٢٩٢	٠,٠٦٨

رقم المفردة	(التمييز)	الخطأ المعياري للتمييز	(الصعوب)	الخطأ المعياري للصعوبة	(التخمين)	الخطأ المعياري للتخمين
٣	١,٢٦٥	٠,٤٢٩	٢,٤٠٨	٠,٤٩٦	٠,٢١٣	٠,٠٣٣
٤	١,٤٣١	٠,٢٧٢	-٠,٠٤١	٠,٢٠٤	٠,٢٧٣	٠,٠٦٨
٥	١,٥٢٧	٠,٢٣٥	-٠,٩٧٨	٠,٢٥٣	٠,٢٨٨	٠,٠٩٢
٦	١,٨٩٢	٠,٢٧٦	-٠,٢٠٩	٠,١٢٩	٠,١٧٥	٠,٠٥٠
٧	٠,٨٥٦	٠,٢٧٥	١,٠٢٥	٠,١٠٢	٠,٤٥٢	٠,٠٦١
٨	٠,٥٥١	٠,٠٨٦	٠,٦٣٦	٠,٢٩٢	٠,٢٤١	٠,٠٨٣
٩	٠,٩٩٢	٠,٢١٥	١,٧٦١	٠,١١٨	٠,٢٣٩	٠,٠٣٨
١٠	١,١٨٠	٠,١٠١	-٠,٣٥٩	٠,١١٨	٠,١٦٩	٠,٠٤٧
١١	١,٠٧٣	٠,٠٧٨	٠,١٧٥	٠,٠٨٤	٠,٢٢٤	٠,٠٣٩
١٢	٠,٨٩٤	٠,٢١٨	١,٥٥٦	٠,١٤٦	٠,٢٩٣	٠,٠٢٤
١٣	٠,٧٧٨	٠,١٥٦	٠,٦١٣	٠,٢٠٤	٠,٢٤٧	٠,٠٦٧
١٤	١,٠٩١	٠,٤٤٦	٠,٨٨٢	٠,١١٣	٠,٢٩٦	٠,٠٣٩
١٥	١,٤٣١	٠,٢٠٥	-١,٦٣٥	٠,٢٧٢	٠,٢٨٣	٠,١٠٣
١٦	٠,٨٤٣	٠,١٨٤	١,٠٠٦	٠,١٤٧	٠,٣٥٤	٠,٠٤٩
١٧	٠,٦٢٠	٠,١٨٨	٠,١٦٠	٠,٢٥١	٠,٤٧٩	٠,٠٦٢
١٨	٠,٤٧١	٠,٠٧٢	٠,٠٤١	٠,٣٧١	٠,٢٤٤	٠,٠٤٨
١٩	١,٧٢٨	٠,٢٨٩	-٠,٢٨٠	٠,١٧٣	٠,٢٥٦	٠,٠٦٣
٢٠	١,٨٨٩	٠,٤٣٣	٠,٧٩٢	٠,١٣٧	٠,٣٧٠	٠,٠٤٤
٢١	١,٩٤٣	٠,٣١١	٠,٠٠٨	٠,١٢٣	٠,٢٠٨	٠,٠٤٩
٢٥	١,٢٠٣	٠,٤٣١	٠,٤٣٤	٠,١٥٥	٠,٣١٤	٠,٠٥٢
٢٦	٠,١٠٧	٠,٠٣٨	-٠,٨١٢	١,٥٩٦	٠,٢٥٤	٠,٠٦٧

رقم المفردة	(التمييز)	الخطأ المعياري للتمييز	(الصعوب)	الخطأ المعياري للصعوبة	(التخمين)	الخطأ المعياري للتخمين
٢٧	١,٩٣٣	٠,٣٧٩	٠,١٥٥-	٠,١٨٢	٠,٥٩٤	٠,٠٦١
٢٨	٠,٠٩١	٠,٠٣١	٢,٤٥٧	١,٥٧٠	٠,١٢٣	٠,٠٤٥
~~٢٩~~	٠,٤٧٠	٠,١٤٩	١,٦١١	٠,٥٦٢	٠,١٤٩	٠,٠٦٢
٣٠	١,٩١٨	٠,٣٥٢	٠,٦٤٠	٠,١١٠	٠,٢٤٣	٠,٠٤٣
٣٢	١,٠١٠	٠,٣٥٣	١,٣١٤	٠,١٠٥	٠,٢٥٢	٠,٠٢٧
٣٣	١,٢٥١	٠,٣٠٥	٠,٨٥٣	٠,١٩٦	٠,٢٨١	٠,٠٥٥
٣٤	٠,٩٢٨	٠,٢٢٤	٠,٢٢٣	٠,٢٥٧	٠,١٧٥	٠,٠٧٤
٣٥	٠,٩١٨	٠,٢٢٢	١,١٢٥	٠,٢٣٩	٠,١١٥	٠,٠٥٠
٣٨	٠,١٠٤	٠,٠٣٧	١,٣٦٢-	٠,٠٦٧	٠,١٦٣	٠,٠٦١
٣٩	١,٣٦٨	٠,١٥٨	٠,٧٤٨	٠,٠٣٤	٠,٢٤٠	٠,٠١٢
المتوسط الحسابي	1.111	0.23	0.41	0.28	0.27	0.06
الانحراف المعياري	0.54	0.12	0.99	0.35	0.10	0.02

يتضح من جدول (٣) أن قيمة معْلم التمييز تراوحت بين (٠,٠٩١ – ١,٩٤٣) بمتوسط قدره (١,١١١) وانحراف معياري قدره (٠,٥٤)، وأن قيمة معْلم الصعوبة تراوحت بين (- ١,٦٣٥ – ٢,٤٥٧) بمتوسط قدره (٠,٤١) وانحراف معياري قدره (٠,٩٩)، وأن قيمة معْلم التخمين تراوحت بين (٠,١١٥ – ٠,٥٩٤) بمتوسط قدره (٠,٢٧) وانحراف معياري قدره (٠,١٠).

وفي ضوء تلك النتائج يوصي الباحثان بإجراء المزيد من الدراسات حول الخصائص السيكومترية للاختبار الذي تم بناؤه باستخدام نماذج سمات كامنة أخرى لنظرية الاستجابة للمفردة الاختبارية مثل نموذج راش وغيره. وذلك للتعرف على مدى مطابقة مفردات الاختبار لنماذج أخرى من نماذج نظرية الاستجابة للمفردة

الاختبارية، والتعرف على تقديرات معالم (التمييز، والصعوبة، والتخمين) لمفردات الاختبار والأخطاء المعيارية في ضوء هذه النماذج.

المراجع العربية:

- الجبوري، رشيد. (٢٠١٢) . بناء اختبار تحصيلي لطلبة الإدارة والإشراف التربوي وفق نظرية السمات الكامنة، معاهد إعداد المعلمين، مجلة الأستاذ، (٢٠٣)، ١٣٩٢-١٤٢١
- حجازي، تغريد والخطيب، عبد الله (٢٠١٤) التوافق بين النظرية الكلاسيكية والنموذج ثنائي المعلمة في مطابقة فقرات اختبار محكي المرجع في أحكام التلاوة والتجويد. مجلة جامعة النجاح (العلوم الإنسانية)، ٢٨(١٠)، ٢٢٤١ - ٢٢٦٦.
- حمادنة، إياد. (٢٠٠٩). استخدام نظرية الاستجابة للمفردة في بناء اختبار محكي المرجع في الرياضيات وفق النموذج اللوجستي ثلاثي المعلم. مجلة العلوم التربوية والنفسية، جامعة البحرين، ١٠(٢) ، ٢١٥-٢٣٨.
- السامرائي، محمد والخفاجي، أحمد (٢٠١٢) بناء اختبار تحصيلي محكي المرجع في مادة علم النفس الخواص، أقسام العلوم التربوية والنفسية. مجلة الأستاذ، (٣)، ٦٤-١٠٠٢
- الشريفين، نضال. (٢٠٠٦). الخصائص السيكومترية لاختبار محكي المرجع في القياس والتقويم التربوي وفق النظرية الحديثة في القياس. مجلة العلوم التربوية والنفسية، جامعة البحرين، ٧ (٤)، ٧٩-١١٠.
- عبابنة، عماد. (٢٠٠٤). أثر حجم العينة وطريقة انتقائها وعدد الفقرات وطريقة انتقائها على دقة تقدير معالم المفردة والقدرة لاختبار قدرة عقلية باستخدام نظرية

استجابة المفردة. رسالة دكتوراه غير منشورة، جامعة عمان العربية للدراسات العليا، الأردن.

- العزاوي، رحيم يونس. (٢٠٠٨). المنهل في العلوم التربوية: القياس والتقويم في العملية التربوية، ط١. بغداد: دار دجلة للنشر والتوزيع.
- علام، صلاح الدين. (٢٠٠١). الاختبارات التشخيصية مرجعية المحك في المجالات التربوية والنفسية والتدريبة. القاهرة: دار الفكر العربي.
- علام، صلاح الدين. (٢٠٠٥). نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس التربوي والنفسي. القاهرة: دار الفكر العربي.
- علي، نداء بهاء الدين. (٢٠١٢). فاعلية استخدام نموذج راش في بناء اختبار محكي المرجع لمقرر القياس والتقويم في التربية. رسالة دكتوراه غير منشورة، جامعة دمشق.
- العيثاوي، أحمد. (٢٠٠٩). استخدام نموذج راش وفق نظرية السمات الكامنة في بناء اختبار تحصيلي في مادة علم نفس الفروق الفردية. رسالة ماجستير غير منشورة، جامعة بغداد، كلية التربية ابن رشد.
- النبهان، موسى.(٢٠٠٤). أساسيات القياس في العلوم السلوكية، ط١، جامعة مؤتة. عمان: دار الشروق للنشر والتوزيع، الأردن.

المراجع الأجنبية:

−Allen, M& . Yen, W. M., (1989): Introduction to Measurement theory. California, Broks Cole publishing company.

-Baker, F.B. (2001). The basics of item response theory (2nd Ed). College Park, MD: ERIC Clearing House on Assessment and Evaluation.

− Biddle, R. (1993). How to Set Cutoff Scores for Knowledge Tests Used In Promotion, Training, Certification, and Licensing. Public Personnel Management, 22(1), 69-73.

- Brannick, M. (2003). "Basics of IRT one – linefile": //a:/item response theory.htm.

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

-Hambleton, R & Swaminathan, H. (1985). Item response theory principles and applications. Boston: Nijhoff Publishing.

-Hambleton, R, K. Swaminathan, H& . Rogers, H. (1999). Fundamentals of item response theory. Sage Publication, Newbury Park, CA.

-Hattie, 1985. J. Hattie, Methodology review: assessing unidimensionality of tests and items. Applied Psychological Measurement 9 (1985), pp. 139–164.

- Klein, M & Velden, U. (2009). Intuition who will pass dental OSCE? Comparison of the Angoff and borderline regression standard setting methods. European Journal of Dental Educational, 13(1). 162-171

- Lord M.F. (1980). Application of item response theory to practical testing problems, Hillsdale, NJ. Er/ Baum.

-Rup, A.A& . Zumbo, B.D. (2006). Understanding parameter invariance in unidimensional IRT models. Educational and Psychological Measurement, 66, 63-84.

-Stocking, M.K. (1990): Specifying Optimum examinees for item parameter estimation in item response theory, Psychometrika. Vol. 55, No. 3, 461-475

-Thorndike، M. (1997). Measurement and Evaluation in Psychology and Education٦ ،rd.ed. Upper Saddle River, New Jersey, Merrill.

− Tiratira, N. (2009). Cutoff Scores- The Basic Angoff Method and the Item Response Theory Method. The International. Journal of Educational and Psychological Assessment, 1 (1), 39-47.

-Warm, T.(٢٠٠٨) . A primer of Item Response Theory، Coast Guard Institute, Oklahoma City, U. S

ملحق (١)

أداة الدراسة (اختبار في القياس والتقويم)

السؤال الأول: ضع دائرة حول رمز البديل الصحيح فيما يأتي: عدد الأسئلة (٤٠)

١-يقع متغير المراحل التعليمية (رياض أطفال، أساسي، متوسط، ثانوي.) على المقياس. .....

أ- الاسمي ب-الرتبي ج- الفئوي د- النسبة

٢-مجموعة السلوكيات التي تظهر كأنها وحدة واحدة هو تعريف. .....

أ- السمة ب- المهارة ج- القدرة د- القياس

٣- التقويم التربوي الذي يتم في نهاية الحصة الصفية هو تقويم.

أ- تمهيدي ب- تجميعي ج- بنائي د-تكويني

٤. الغرض الأساسي من عملية القياس الصفي هو.....

أ- الكشف عن الفروق الفردية ب- رصد الدرجات ج- التشجيع على المذاكرة د- تطوير التعليم

٥- محاولة الطالب إعطاء أي إجابة على السؤال المقالي (دون معرفته لمحتوى المفردة) للحصول على عدد من الدرجات تسمى. ....

أ- تخمين عشوائي ب- تورية ج- إسقاط د - تخمين ذكي

٦ – تشير كلمة موضوعية في الاختبارات الموضوعية إلى موضوعية. .....

أ- السؤال ب- التصحيح ج- الإجابة د- الطالب

٧ - إحدى الخصائص التالية خاصية بسيطة (أحادية البعد) .....

أ- القدرة العددية ب- الصحة ج- الذكاء د- الصحة النفسية

٨- إطار مرجعي يستخدم لمقارنة وتفسير أداء الفرد على الاختبار بهدف تحديد معنى ودلالة الدرجة "هو تعريف.

أ- المعيار ب- الثبات ج- الصدق د- الموضوعية. ....

٩- المقياس الذي يملك خصائص أي مقياس آخر(أكثر المقاييس دقة)، هو مقياس.

أ- نسبة ب- فئوي ج- اسمي د- رتبي

١٠- إن تحديد قيم عددية للسمات حسب قواعد محددة هو تعريف. .....

أ- الاختبار ب- القياس ج- السمة د- التقويم

١١- عندما نحكم على درجة امتلاك فرد لخاصية معينة من خلال درجة امتلاك مجموعته لهذه الخاصية فإن قياسنا يوصف بأنه قياس. ...

أ- غير مباشر ب- نسبي ج- غير تام د - نسبة

١٢- عندما يجيب بعض الطلبة عن فقرات مقياس نفسي، قد يعطون صورة سلبية أو إيجابية عن أنفسهم، يسمى هذا

أ- الهالة ب- التزييف ج- التورية د- الإذعان

١٣- من الفروق بين القياس والتقويم. ....

أ-القياس أعم وأشمل من التقويم ب- القياس يصف السلوك وصفا كميا بينما التقويم يصفه وصفا نوعيا

ج- التقويم أعم وأشمل من القياس د- القياس يصف السلوك كميا ونوعيا بينما التقويم يصفه نوعيا

١٤- "طريقة منظمة لقياس مستوى امتلاك الفرد لمهارات معينة في مادة دراسية كان قد تعلمها بصفة رسمية من خلال إجابته عن عينة من الأسئلة التي تمثل محتوى المادة " هو تعريف. .

أ-التقويم ب- القياس ج- الاختبار د- التقدير

١٥- تعتبر اختبارات الكفايات للمعلمين كشرط لتعيينهم معلمين في وزارة التربية والتعليم في المملكة العربية السعودية، اختبارات. .....

أ-معيارية المرجع ب- محكية المرجع ج- محكية ومعيارية د- معياري القدرة

١٦- يعد المعدل التراكمي للطالب في الجامعة متغيراً.. .....

أ- متصلاً ب- منفصلاً ج- نوعياً د- وثّاباً

١٧ -الاختبارات المقننة التي تصمم للتنبؤ بنجاح المتعلم في تعلم لاحق أو بمحك معين، هي اختبارات. ..

أ- استعداد ب- تحصيلية ج- قدرات عقلية د- استعداد وتحصيلية

١٨- اختبار القبول في الجامعات لبرنامج تعليمي يعد اختباراً:

أ- جمعيا ب- فرديا ج- غير لغوي د- تحصيليا

-19يفضل في المرحلة الأساسية الأولى أن تكون الاختبارات التي يعدها المعلم من نوع. .....

أ- اختيار من متعدد ب- الصواب والخطأ ج- المقالي د- التكميل

٢٠- الاختبار الذي يتضمن فقرات مصاغة لفظيا أو على شكل رسومات غير محددة تماما من حيث المطلوب، هو..

أ- إسقاطي ب- معياري المرجع ج- محكي المرجع د- استعداد

٢١- أحد اختبارات الذكاء التالية اختباراً فرديا. ..

أ-اختبار بينيه ب- اختبار الفا ج- اختبار بيتا د- اختبار رافن

٢٢- اختبار الذكاء الجمعي لمن لا يجيدون القراءة والكتابة الذي صمم بعد الحرب العالمية هو اختبار. ....

أ- بينيه ب- بيتا ج- الفا د- كاتل

٢٣- تعد الاختبارات التالية اختبارات أقصى أداء.. .....

أ- الاستعداد ب- العادات الدراسية ج- الاتجاهات د- الميول

٢٤ - الفر ق الأساسي بين الاختبارات المعيارية المرجع والمحكية المرجع يكون في.....

أ- وضوح الفقرات ب- تفسير النتائج ج- تحديد الإجابة د- موضوعية التصحيح

٢٥ – إذا اعتبر المدرس أن جميع طلبة الصف راسبون في امتحان معين فهذا يعني أن تفسير المعلم لعلامات طلابه.....

أ- معياري المرجع ب- محكي المرجع ج- مرجعي النمو د- مرجعي القدرة

٢٦ -يعد التجريب الأولي لفقرات الاختبار وتحليلها قبل إعدادها بالصورة النهائية صفة مميزة للاختبارات. ..

أ- التحصيلية ب- المقننة ج- الاستعداد د- القدرات

٢٧- إجابة طالب عن فقرات في أداة قياس مصممة لقياس الطموح، فان الاختبار يقيس. .....

أ- الأداء العادي ب- أقصى أداء ج- سرعة الأداء د- قدرة أداء. ...

٢٨-محاولة الطالب إعطاء أي إجابة على السؤال الاختياري(دون معرفته للمحتوى) للحصول على عدد من الدرجات تسمى. ....

أ- تخمين عشوائي ب- تورية ج- إسقاط د - تخمين ذكي

٢٩- أمام الطالب احتمال أكبر أن يخمن الإجابة الصحيحة على المفردة من نوع. .....

أ- الصواب والخطأ ب- اختيار من متعدد (٣) بدائل

ج- اختيار من متعدد (٤) بدائل د- اختيار من متعدد (٥) بدائل

٣٠-أكثر أنواع الاختبارات عرضة لتزييف الاستجابة من قبل المفحوص، هي الاختبارات. ....

أ- الاسقاطية ب- التقدير الذاتي ج- القدرات د- الاستعداد

٣١-كلما زاد عدد البدائل في سؤال اختيار من متعدد. .....

أ- زادت دقة القياس ب- قلت فرص التخمين العشوائي ج- زادت فرص التخمين العشوائي د- قلت دقة القياس

٣٢ –تتحدد الأوزان النسبية لموضوع دراسي معين في لائحة المواصفات لبناء اختبار تحصيلي من خلال. .....

أ – عدد صفحاته ب– عدد حصصه ج– آراء الخبراء د – كل ما سبق

٣٣– يتحدد عدد الأسئلة في كل خلية من خلايا لائحة المواصفات على ضوء. .....

أ– الأهداف المراد تحقيقها ب– الأهمية النسبية للموضوع ج– أ، ب صحيحان د– الزمن المخصص للاختبار

٣٤ –كل مما يلي من جوانب الضعف في الأسئلة الموضوعية(الاختيارية) فيما عدا..

أ–لا تتيح للطالب الفرصة للتعبير عن أفكاره

ب– تمكن الطالب من الوصول إلى الإجابة الصحيحة بالتخمين العشوائي.

ج– لا تحتاج من الطالب أن يسأل عن طريقة الإجابة.

د– تحتاج إلى مهارة وخبرة كبيرة في صياغتها وإعدادها

٣٥– ما يلي جدول مواصفات لبناء اختبار في الرياضيات

الموضوع /الأهداف	تذكر ١٠%	فهم ٢٠%	تطبيق ٧٠%	المجموع ١٠٠%
أ ٢٠%				
ب ٣٠%				
ج٥٠ %			*	
المجموع				٦٠

في ضوء ما سبق فإن عدد الأسئلة الخاصة بمستوى التطبيق للموضوع (ج) هو. .....

أ– ١٤ ب– ٢١ ج– ٣٠ د– ١٨

٣٦– لرفع قيمة معامل صعوبة فقرة اختيار من متعدد يفضل زيادة. .....

أ– تجانس البدائل ب– عدد البدائل ج– طول متن السؤال د– قيم فاعلية المموهات

٣٧–معامل الصعوبة الأمثل لاختبار تحصيلي معياري المرجع هو. .

●٠،٦٠ ب- ٠،٨٠ ج- ٠،٥٠ د- ٠،٢٠

٣٨ – معامل التمييز المقبول لاختبار تحصيلي معياري المرجع. ...

● ≤ ٠،١٢ ب- ≤ ٠،٢٠ ج- ≤ ٠،١٤ د- ≤ ٠،١٠

٣٩- إذا كان مجموع درجات طلاب على سؤال(٨٠) درجة والدرجة المستحقة للسؤال (١٠) درجات وعدد الطلبة الذين أجابوا عن السؤال(١٢) طالبا، فإن معامل صعوبة السؤال....

أ- ٠,٢٥ ب- ٠,٦٦ ج- ٠,٥٠ د- ٠,٧٥

٤٠- كان مجموع درجات طلاب الفئة العليا على سؤال (٨٠) درجة، وعدد طلاب الفئة العليا(١٠) طلاب ومجموع درجات طلاب الفئة الدنيا على السؤال (٣٠) درجة، وعدد طلاب الفئة الدنيا (١٠) طلاب، والدرجة المستحقة للسؤال (١٠) درجات، فإن معامل تمييز السؤال. .....

أ- ٠,٤٠ ب- ٠,٥٠ ج- ٠,٣٠ د- ٠،٢٠

انتهت الأسئلة