بناء اختبار محكي المرجع لمناهج البحث لطلبة كلية التربية باستخدام نظرية الاستجابة للمفردة – نموذج راش

د. طلال سعد المطيري

(دكتوراه في القياس والتقويم التربوي – دولة الكويت)

الملخص

تعد الاختبارات محكية المرجع من أهم المقاييس التي تساعدنا في قياس الأهداف السلوكية بغرض التعرف على مدى إتقان أو كفاءة الفرد في مهارة ما. هدفت هذه الدراسة إلى بناء اختبار مرجعي المحك لقياس الكفاءة والمهارة البحثية لدى طلبة كلية التربية باستخدام نظرية استجابة الفقرة. حيث تم بناء اختبار محكي المرجع تتمحور فقراته حول مجالات عدة متعلقة بالبحث العلمي مثل: أنواع البحث حسب المنهج والتصميم، عناصر البحث، مصادر المعلومات للبحوث التربوية، المتغيرات، أنواع العينات، أدوات البحث التربوي، الخصائص السيكومترية للبحث التربوي، والإحصاء. طبق الاختبار المكون من (٤٠) فقرة على عينة مكونة من (٣٦٢) طالب وطالبة من كلية التربية جامعة الكويت وكلية التربية الأساسية بدولة الكويت. توصلت هذه الدراسة من خلال استخدام النموذج الأحادي إلى اختبار يتكون من (٣٧) فقرة لقياس المهارات البحثية، حيث يتمتع بالصدق من خلال تطابقه مع فرضيات نموذج راش، بالإضافة إلى انه يتمتع بدرجة ثبات مرتفعة لكل من صعوبة الفقرة وقدرة الفرد. وتوصي هذه الدراسة بالاستفادة من الاختبار المصمم من خلال تطبيقه على الطلبة المتقدمين للدراسات العليا في الكليات التربوية من أجل تحديد كفاءة الطالب البحثية وما إذا كان بحاجة إلى ساعات مكثفة من مادة مناهج البحث قبل البدء بالمواد التخصصية.

الكلمات المفتاحية: اختبار محكي المرجع، مناهج البحث، نظرية استجابة المفردة، نموذج راش،

Constructing a Criterion-Reference Test in Research Methods for Students in the School of Education by Using Item Response Theory – Rasch Model

Dr. Talal S. Almutairi

(PhD. in Educational Evaluation, Assessment and Effectiveness – Kuwait)

Criterion-referenced tests are administered to identify what students have learnt as compared to what they expected to learn and to identify how well they are progressing toward a specified performance level in a specific set of aims. The purpose of the present study was to construct a criterion-referenced test to measure the research knowledge and skills of students in the school of education. A 40-question test that had 4 answer choices for each question was administered to 362 students at Kuwait University d The Public Authority for Applied Education and Training. The test was designed using Item Response Theory, and its results showed that the responses to 37 of its questions fit the Rasch model's assumptions and questions were located within the range of the Rasch model's infit and outfit indicators. In addition, both the persons and items were highly reliable. Finally, this study indicates that the designed test is appropriate for measuring the research knowledge and skills of students applying for

postgraduate studies and that such test can inform placement decisions or identify applicants who should be required to complete an intensive research-methods module.

Keywords: Criterion-Reference test, Research methods, Item Response Theory, Rasch model.

بناء اختبار محكي المرجع لمناهج البحث لطلبة كلية التربية باستخدام نظرية الاستجابة للمفردة – نموذج راش

د. طلال سعد المطيري

(دكتوراه في القياس والتقويم التربوي – دولة الكويت)

المقدمة

الاختبار يشار إليه بأنه مجموعة من الأسئلة، الهدف منها جمع إجابات الفرد للوصول إلى خاصية سلوكية أو عدد من الخصائص السلوكية. ومن أهم مواصفات الاختبار الجيد لقياس سلوك الأفراد واستجاباتهم: الصدق، الثبات، والموضوعية. لذلك، يسعى الباحثون في مجال القياس والتربية وعلم النفس إلى الوصول دائماً إلى تقديرات للأمر المراد قياسه من خلال اختبار يتسم بالصدق والثبات والموضوعية.

وقد نشأت فكرة الاختبارات التحصيلية الحديثة (الموضوعية) نتيجة الحاجة الماسة لقياس الأهداف السلوكية أما لتصنيف أو لمقارنة بين المفحوصين أو لتحديد مدى الإتقان أو الكفاءة، مثل تلك الاختبارات المستخدمة لمنح ترخيص لمهنة ما أو اجتياز متطلبات اللغة. بالإضافة إلى تقويم البرامج التعليمية واتخاذ بعض القرارات ذات الصلة بها أو لإجراءات استطلاعية لبعض الإنجازات التربوية (Mehrens & Lehmann, 1991). على عكس الاختبارات التحصيلية التقليدية (المقالية) التي لا تتسم في غالب الأحيان بالموضوعية في الأعداد والتصحيح ولا توصف بالصادقة والثابتة (مراد وسليمان، ٢٠١٢).

وقبل التطرق إلى أهداف هذه الدراسة والإطار النظري، نجد انه لا بد من توضيح الفرق بين الاختبارات المقننة standardised وغير المقننة non-standardised. ويوضح Thorndike & Thorndike-Christ (٢٠١٠) أن الاختبارات المقننة تطبق على سياق كبير مثل مجموعة من المدارس أو فئة مجتمعية معينة، لقياس أهداف عامة مشتركة تشمل المعارف والمهارات من خلال بعض العبارات التي تقيس كل مهارة أو

موضوع على حدة. وكل عبارة من هذه العبارات تم تحليلها ومعالجتها إحصائيا لتأكد من أنها صادقة وثابتة قبل أن تكون جزء من الاختبار، لتمدنا بمعايير الأداء لمجموعات مختلفة على نطاق واسع لتشمل المجتمع أو الدولة. أما الاختبارات غير المقننة أو كما يطلق عليها في بعض الأحيان اختبارات من صنع المعلم teacher-made test، أنما تقيس أهداف معينة خاصة بالفصل أو بالمدرسة أو مجموعة أفراد في برنامج معين للحصول على معايير أداء عينة محددة. حيث تبنى هذه الاختبارات على اختيار عدة عبارات لأجزاء معينة من المحتوى، وهذه العبارات من النادر جداً أن تكون تم مراجعتها ومعالجتها إحصائيا وبالتالي تفتقر إلى الثبات.

والاختبار التحصيلي يختلف باختلاف كيفية تفسير نتائجه، أما محكية المرجع criterion-referenced test أو معيارية المرجع Norm-referenced test. الاختبارات محكية المرجع هي تلك الاختبارات التي تقوم على تحديد محك معين (درجة معينة– درجة القطع) وتقارن مستوى المفحوص بالنسبة إلى ذلك المحك وفي الغالب تستخدم لقياس مستوى الكفاءة الفردية أو لتحديد ما يمتلك المفحوص من معارف ومهارات. أما الاختبارات معيارية المرجع فهي تلك الاختبارات التي تحدد مستوى أداْء المفحوص وتسمح بمقارنته مع متوسط تحصيل قرنائه أي المجموعة التي طبق عليها الاختبار وفي الغالب يركز هذا الاتجاه إلى تعيين الفروق الفردية بين المجموعة المفحوصة (النجار، ٢٠١٠).

لذلك، فإن الأسس التي يقوم عليها الاختبار التحصيلي لحكم بناءه واستخدامه تسهم إسهامً مباشراً في تحسين التعلم والتدريس، لأن الخلل في بناءه وفي صحة استخدامه يمكن أن يؤثر على توجيه المفحوص وأيضاً على القرارات التي ممكن أن تتأخذ استناداً عليه. على سبيل المثال، ممكن للاختبار أن يساعد على توجيه انتباه المفحوصين نحو كل المجالات الهامة، ومن الممكن أن يوجه تركيزهم على محتوى المنهج والاكتفاء بالدراسة السطحية دون الفهم العميق. وتوضح عيد (٢٠٠٦) بعض أهم الأسس الهامة التي تجعل الاختبار مقياساً صادقاً وثابتاً، أن يكون الاختبار قياساً لنواتج محددة

ومتنوعة للتعلم، أن تكون العينة ممثلة أي بمعنى عينة محددة للسلوك المراد قياسه، إن يكون نوع الأسئلة مناسب لاستدعاء الاستجابة المرغوبة، وأخيراً تحديد الغرض من الاختبار مثل للقبول بداية العام الدراسي، أو اختبارات بنائية أثناء عملية التعليم أو نهائية، أو تشخصية للتعرف على الصعوبات.

مشكلة الدراسة

تعد التجارب العربية في مجال تطوير الاختبارات محكية المرجع محدودة نوعاً ما من حيث بناء الاختبارات والتحليلات الإحصائية وكيفية التفسير والاستخدام كما أشار الحموري والكحلوت (٢٠٠٩). وإننا نعاني كمجتمعات عربية من بعض المشاكل التي يكون سببها البناء التراكمي للمعرفة. فمن واقع التجربة الشخصية وجدنا عدداً لا يستهان به من طلبة الماجستير في كافة العلوم التربوية يعانون من ضعف الحصيلة المعرفية تجاه مهارات البحث العلمي وتطبيقاته على الرغم من إن مقرر مناهج البحث يعتبر مقرراً إلزامياً في مرحلة البكالوريوس لطلبة التربية في شتى البلدان العربية. لذلك تولدت الفكرة ببناء اختبار محكي المرجع لمادة مناهج البحث حتى يتم التمييز بين الطلبة ذوي الخلفية البحثية الجيدة وبين الطلبة الذين يفتقرون للأساسيات البحثية بقصد إعطائهم بعض الساعات أو المقررات التكثيفية لعلاج المشاكل ولتفادي وقوع الأخطاء البحثية في المراحل المتأخرة في برنامج الماجستير (إعداد الاطروحة)، أو من الممكن استخدامه في انتقاء الطلبة في مرحلة القبول في برامج الدراسات العليا.

وسعت هذه الدراسة للإجابة عن الأسئلة التالية:

١- ما مدى تحقق بيانات الاختبار المصمم لافتراضات نموذج راش؟

٢- ما تقديرات معالم صعوبة فقرات الاختبار المصمم وفق نموذج راش؟

٣- ما الخصائص السيكومترية للاختبار المصمم؟

هدف من الدراسة

تهدف الدراسة الراهنة إلى بناء اختبار مرجعي المحك لمادة مناهج البحث لطلبة كلية التربية المقبلين على الدراسات العليا باستخدام نظرية استجابة الفقرة – نموذج راش، ذي درجة عالية من الصدق والثبات يمكن استخدامه من قبل كلية التربية لتحديد مستوى الطلبة بمهارات البحث العلمي.

حيث تصميم اختبار محكي المرجع لمناهج البحث قبل البدء في برنامج تعليمي معين (الماجستير) يعد الخيار الأمثل من حيث نوعية الاختبار، كما أشار كل من النجار (٢٠١٠) و مراد وسليمان (٢٠١٢) إلى أن الاختبارات محكية المرجع تعد أفضل الخيارات في عدة حالات من ضمنها مرحلة القبول في برنامج تعليمي بغرض التأكد من توافر الحد الأدنى من المستوى المطلوب للبدء لدى المتقدم.

أهمية الدراسة

تكتسب هذه الدراسة أهميتها، في إنها ستوفر للقائمين على برامج الدراسات العليا في الكليات التربوية معلومات تساعد في اتخاذ القرارات المناسبة تجاه الطلبة، أما من خلال انتقاء الأكفاء منهم أو التصنيف بين الطلبة حسب الحاجة إلى ساعات مكثفة لمناهج البحث لتفادي أي أخطاء غير مقبولة من طالب الماجستير نتيجة ضعف الخلفية البحثية.

كما إنها تكتسب أهميتها في كونها تتناول جانباً علمياً مهماً في العملية التربوية وذلك من خلال بناء اختبار محكي المرجع وفقاً للنظرية الحديثة في مناهج البحث لطلبة كلية التربية، حيث يمكن إن يستخدم في تشخص مواطن القوة ونقاط الضعف وتوفير معلومات عن مستويات الطلبة.

الإطار النظري والدراسات السابقة

الاختبارات أما محكية المرجع أو معيارية المرجع. واتجه البعض من الباحثين وعلماء القياس إلى توجيه الانتقادات تجاه نوع من التفسير مقابل أخر. فنجد البعض نقل أو تحدث عن الاختبارات معيارية المرجع أنها تصنف الأفراد حسب مراكزهم النسبية

وتركز على تبيان الفروق الفردية بغض النظر عن مدى تمكن أو إتقان الفرد للهدف السلوكي المراد قياسه. بينما محكية المرجع تقوم على قياس المفحوصين بشكل فردي وصولا إلى تقويم البرنامج أو سياق معين.

بينما نرى إن الاختبار كان محكياً أو معيارياً لكل منهما إيجابيات بحد ذاته، حيث إن كل اختبار يخدم الهدف المراد تحقيقه أو قياسه مثل عندما نريد إصدار أحكام على الطلبة للوصول إلى تصنيف أو ترتيب (الأول فالثاني) أو مقارنة مستوى الطلبة بين بعضهم البعض نجد لا بديل عن الاختبار مرجعي المعيار. لذلك، نرجح الرأي القائل لكل منهم فوائد حسب الغرض من الاختبار واستخدامه والتفسير المراد الوصل إليه.

ولعل التوصيف المذكور من قبل Bond (١٩٩٦) أدق وأوضح في هذا الخصوص حيث توضح أن الاختبارين المحكي والمعياري عبارة عن قسمين منفصلين ولكل منهما غرض مختلف ومحتوى مختلف وتوصيف للنتائج مختلف ولكل منهما مميزات حسب الاستخدام وكل منهما يمكن أن يصمم على إنه اختبار مقنن Standardised. حيث توضح الباحثة أن أهم الأسباب لاختيار الاختبار معياري المرجع هو تصنيف الطلبة وتسليط الضوء على نتائج الطلبة بمقارنتهم مع بعضهم البعض. فعلى سبيل المثال، عندما يريد المعلم اختيار المتميزين من بين الطلبة للمشاركة في برامج معينة تتطلب قدرات عالية مثلاً في القراءة أو الرياضيات فإن المعياري المرجع يمكن يخدم هذا الغرض. بينما على الجانب الآخر الاختبار المحكي يعطي تفاصيل عن الطالب أو المفحوص لوصول لحكم حول كفاءته ومدى تحقيق الأهداف السلوكية التي يهدف الاختبار لنظر إلى تحققها عند كل طالب على حده. وتشير أيضاً إن محتوى الاختبار هو اهم جزء من الاختلافات بينهم حيث المعياري المرجع يبنى على اختيار محتوى الذي يركز على تصنيف وترتيب مستويات الطلبة من الأعلى إلى الأدنى، بينما المحكي يعتمد على اختيار محتوى يعكس إلى أي مدى نتائج التعلم والمعرفة لدى المفحوص تحققت ويتم اختيارها وفقها لأهميتها في المنهج أو

الموضوع. وبالتالي لك منهما اختيار وللتفضيل بينهم يجب يراعى الغرض والمحتوى والتفسير للاختبار.

ويتوافق راي Bond مع راي من سبقه Ornstein (١٩٩٣) الذي تطرق إلى توضيح إن لكل اختبار فوائد مرجوة وأضاف ما يعد ميزة أو عيب للاختبارين على حد سواء، هو أن الاختبار معياري المرجع لابد إنه يطبق على عينة كبيرة للمقارنة بينهم بينما الاختبار محكي المرجع يمكن إن يطبق على عينة صغيره أو يطبق بشكل فردي على المفحوص. ولذلك ترى Bond (١٩٩٦) أن الاختبار معياري المرجع قد يعاب عليه إن إجراءاته التطبيقية تتطلب تكلفه مادية اعلى من الاختبار محكي المرجع الذي يمكن إن يطبق بشكل فردي أو عينة بسيطة.

ويدعم ذلك كل من Bachman (2004)، Brown (٢٠٠٥ ؛ ٢٠١٢) ، Brown & Hudson (٢٠٠٢) إنه الاختبارات المعيارية والمحكية تختلف من التصميم الذي يشمل اختيار المحتوى وصياغة الفقرات وترتيب الفقرات إلى أجراء الاستطلاع على الاختبار ومن ثم الاختلاف يستمر في تحليل الفقرات وقياس صدقها وثباتها وصولاً إلى النتائج النهائية. ونستنتج من آرائهم، مفاضلة نوع على الآخر يكون استناداً على التفسير المراد الوصل إليه، ولا نرجح القول إن المعياري أفضل من المحكي ولا عكس ذلك.

والسؤال المطروح بعد حسم المميزات لكل النوعين، هو ماهي استراتيجية بناء اختبار محكي المرجع للوصول إلى اختبار جيد. يوضح علام (٢٠٠١) إن هناك مجموعة من الخطوات لا بد أن تؤخذ بعين الاعتبار عند القيام ببناء اختبار محكي المرجع. أولا، لا بد من إن يكون النطاق السلوكي الذي يهدف الاختبار لقياسه محدد حتى يكون هذا بمثابة الإطار المرجعي لتشار إليه علامة الفرد في الاختبار. ثانياً، تحليل النطاق السلوكي إلى مكوناته من معارف ومهارات أي بمعنى تحديد المكونات أو المتطلبات السلوكية التي لا بد إن تكون متوفرة لدى المفحوص لكي يشار إليه بالكفاءة أو امتلاكه للمهارات الرئيسية المقاسة من الاختبار. ثالثاً، تحديد الأهداف السلوكية

المتعلقة بتلك المكونات أو المتطلبات بحيث يمكن قياسها من خلال مفردات الاختبار. رابعاً، تكوين مواصفات الاختبار بشكل تفصيلي ودقيق التي تشمل مفردات الاختبار وخصائص الاستجابات والتعليمات ودرجات المحك للوصل إلى كتابة مفردات متجانسة وبناء متوازي ومتكافئ. خامساً، تجميع المفردات وتنسيقها في ضوء الأهداف المراد قياسها. سادساً، تطبيق استطلاعي للاختبار لمعرفة مدى ملائمة الاختبار للمفحوصين. سابعاً، المعالجة الإحصائية للتحقق من صدق الاختبار وثبات درجاته والحكم على أن المفردات تفي بالخصائص المرجوة.

ويتفق كلاً من الحموري والكحلوت (٢٠٠٩) مع هذه الخطوات من حيث المضمون من خلال دراسة تم إجراءها لتقديم استراتيجية محددة لبناء اختبارات محكية المرجع. أولاً، إعداد مواصفات تفصيلية شاملة للاختبار تشمل تحديد الغرض والمحتوى والمجموعة التي يطبق عليها الاختبار. ثانياً، إعداد الفقرات الاختبارية وفقاً للخطوات التالية بناء الفقرات بوصفها أمثله على مواصفات الفقرات وكل هدف من مجموعة الأهداف، ومن ثم عرض هذه الفقرات ومواصفاتها على فريق مختص للحكم عليها، ويلي تلك الخطوة تقويم وضوح تلك الفقرات من خلال تجريبها على عينة بسيطة. ثالثاً، مراجعة الفقرات الاختبارية وتصديقها من خلال المراجعة المنطقية والمعالجة الإحصائية، المنطقية من خلال التأكد من ملائمة الفقرات لمعايير البناء بالأدب التربوي والتحقق من اتساقها مع مواصفات الفقرة للاطلاع على أن فقرات الاختبار تضم الأهداف المراد قياسها، ومراجعة الوزن النسبي في قياس المجال السلوكي وصدق المحتوى من خلال تقديرات المحكمين. أما المعالجة الإحصائية في تبيان الثبات والوصول إلى مؤشرات الصدق والتأكد من معامل الصعوبة والتمييز وفاعلية الاختبار.

وبالانتقال إلى أساليب القياس المستخدمة في الاختبارات محكية المرجع نجد أن هناك طريقتين وهما النظرية التقليدية أو الكلاسيكية Classical Test Theory والنظرية الحديثة المعروفة بنظرية الاستجابة للمفردة Item Response Theory.

النظرية الكلاسيكية يمكن اختصارها بأنها توزيع الدرجات التي يحصل عليها المفحوص على بعد القدرة التي يقيسها الاختبار على المنحى الاعتدالي، الذي يعتمد تماماً على مدى صعوبة أو سهولة الفقرات بالنسبة للمفحوصين وعلى مدى تمييز كل فقرة بين المستويات المختلفة للقدرة التي يهدف الاختبار لقياسها (النجار، ٢٠١٠). بينما النظرية الحديثة تقوم على الافتراض بأن إجابة المفحوص على الفقرة تتحدد في ضوء السمات الكامنة التي تمثل قدرة المفحوص، وتلك العلاقة بين أداء المفحوص وقدرته ويمكن أن توصف من خلال منحى خصائص الفقرة الذي يعتمد على عدة أنواع من المعالم وهي الصعوبة والتمييز والتخمين للفقرة (النجار، ٢٠١٠ ؛ مراد وسليمان، ٢٠١٢). ومن الافتراضات التي تستند عليها نماذج النظرية الحديثة–السمات الكامنة: فرض أحادية البعد والتي تعني وجود قدرة واحد فقط ممكن إن تفسر أداء المفحوص بالاختبار. استقلالية الفقرة تعني إن استجابة المفحوص للفقرة لا تتأثر باستجاباته على فقرات أخرى. منحنى خاصية الفقرة المتعلق بوصف العلاقة التي تربيط بين احتمالية إجابة المفحوصين بجميع مستوياتهم المختلفة إجابة صحيحة عن الفقرة وبين القدرة التي تقيسها مجموعة فقرات الاختبار، ويتم توصيف المنحى من خلال معامل الصعوبة والتمييز والتخمين. وأخيراً عامل التحرر من السرعة أي بمعنى السرعة لا تعد عامل مؤثر في إخفاق المفحوص في اختيار الإجابة الصحيحة بل قدرته هي المؤثرة (Hambleton & Swaminathan, 1985). ومن نماذج هذه النظرية: النموذج اللوجستي ثلاثي المعالم، ثنائي المعلم، وأحادي المعلم كما أوضحتها الحاج (٢٠١٦):

أحادي المعلم: هو النموذج المعروف باسم العالم الدنماركي راش والذي يشير إلى إن فقرات الاختبار تختلف في مدى صعوبتها فقط ولكن تتجانس وتتساوى في التمييز بين مستويات المفحوصين ولا يفترض وجود التخمين.

ثنائي المعلم: وهذا النموذج يفترض إن فقرات الاختبار تختلف في نتائج معلم الصعوبة ومعلم التمييز ويفترض هذا النموذج إن المفحوصين لا يلجأون التخمين وسيلة عند الإجابة.

ثلاثي المعلم: في هذا النموذج يتم وضع معادلة تضم ثلاثة معالم لدرجة المفحوص أي بمعنى الفقرات تختلف في كل معلم، تختلف في الصعوبة والتميز والتخمين.

والجدير بالذكر هنا، إن النظرية الكلاسيكية تعرضت لكثير من النقد والذي تمثل معظمه الإشارة إلى وجود قصور في الصدق والثبات وظهور بعض المشكلات المصاحبة للاختبار مثل حذف فقرات الاختبار غير مميزة بين أفراد العينة بغض النظر عن قيمتها (مراد وسليمان، ٢٠١٢ و عيد، ٢٠٠٤). ونجد ذلك في الدراسات التي قامت على مقارنة بين النظرية الكلاسيكية والنظرية الحديثة على سبيل المثال أبو هاشم (٢٠٠٦) وجد إن النظرية الحديثة متمثلة بنموذج راش أدق في تحقق خطية التدرج وأحادية القياس واستقلاليته، ونظرا لما تتميز به أداة القياس من خلال استخدام نموذج راش بالموضوعية، يدعم الباحث الرأي المتمثل بإن استخدام نماذج السمات الكامنة يمكن إن تتغلب على النقد المصحوب للنظرية الكلاسيكية. وتظهر دراسة Magno (٢٠٠٩) إن هناك بعض القيود على النظرية الكلاسيكية مقابل بعض المزايا للنظرية الحديثة تتمثل في إن معامل الصعوبة من الممكن إن لا يتغير عبر العينات ومعامل الصعوبة أكثر استقراراً عبر أشكال الاختبارات المختلفة باستخدام النظرية الحديثة، بالإضافة إلى إن ثبات الاتساق الداخلي للاختبار مستقر عبر العينات من خلال النظرية الحديثة مقارنة بالنظرية الكلاسيكية. وأيضاً وجد كلا من Alkhadher & Albursan (٢٠١٧) من خلال دراسة شملت ٢٦٧٤ طالب للمقارنة بين دقة النظرية الكلاسيكية والحديثة ووجدوا إن النظرية الحديثة أدق من النظرية الكلاسيكية في قياس قدرة المشاركين وإن معامل الصعوبة يختلف بين النظريتين لذلك يؤصون بتطوير الاختبارات استناداً على النظرية الحديثة من خلال نماذج الاستجابة للفقرة. بينما يرى كلاً من Jabrayilov, Emons & Sijtsma (٢٠١٦) من خلال مقارنتهم لنظرية الكلاسيكية والحديثة إن نظرية الاستجابة للفقرة تتفوق على النظرية الكلاسيكية شرطي إن يكون الاختبار متكون من ٢٠ فقرة على الأقل لكن بالنسبة للاختبارات القصيرة فالنظرية الكلاسيكية أفضل في تحديد التغير أو القدرة لدي المفحوص.

وهناك عدة دراسات تتشابه مع موضوع الدراسة الحالية المتعلق ببناء اختبار محكي المرجع للمرحلة الجامعية. أولاً، دراسة قام بها أبو هاشم (٢٠٠٦) التي هدفت إلى المقارنة بين النظرية الحديثة (نموذج راش) والكلاسيكية لبناء اختبار محكي المرجع لقائمة المداخل الدراسية لدى طلاب المرحلة الجامعية. توصل من خلال دراسته إلى اختبار محكي يشتمل في صورته النهائية على (٢١) فقرة لقياس مداخل الدراسة تشمل مدخل العميق والسطحي والاستراتيجي بالإضافة إلى بعض الجوانب ذات التأثير في التعلم مثل ترابط الأفكار، استخدام الأدلة والبراهين، والدافعية الداخلية. توصل من خلال الاختبار إلى إن هناك فروق بين الطلبة ذوي التحصيل المتدني والتحصيل المرتفع في المرحلة الجامعية في مداخل الدراسة لصالح ذوي التحصيل المرتفع. وعالج السامرائي والخفاجي (٢٠١٢) المشكلة المترتبة على عدم وجود اختبار تحصيلي محكي المرجع لمادة علم النفس الخواص لطلبة كلية التربية في القسمين التربوي والنفسي. حيث قاما بعمل اختبار محكي المرجع باستخدام نموذج راش وتوصلا إلى اختبار مكون من (١٢٥) فقرة ذات استقلالية قياسية وحقق درجة عالية من الثبات والاختبار قادر على التمييز بين الطلبة من خلال عينة شملت ٣٤٩ طالبة وطالبة في الجامعات العراقية. ودراسة كلاً من هادي ومراد (٢٠١٣) التي هدفت إلى بناء اختبار لكفايات اللغة العربية لقبول الطلبة في برنامج الدراسات العليا في الكليات الإنسانية بجامعة الكويت حيث تضمن الاختبار ثلاثة مجالات شملت القدرة على الفهم، والتعبير، والتفكير. طبق الاختبار على عينة تجريبية وعينة الدراسة التي شملت ٣٣٥ طالب وطالبة وأسفرت النتائج باستخدام نموذج راش إلى اختبار جيد يتمتع بالثبات ويتكون من (٨١) فقرة لقياس قدرات الأفراد.

وقام الشريفين (٢٠٠٦) بدراسة هدفت إلى تقدير الخصائص السيكومترية لاختبار محكي المرجع في القياس والتقويم التربوي باستخدام النظرية الحديثة كإطار تحليلي (نموذج راش). شملت العينة على ٢٢٢ طالب وطالبة موزعين على أربعة شعب دراسية. توصل الباحث إلى اختبار مكون من (٣٧) فقرة بصورته النهائية، وبينت

نتائج الدراسة أن قيمة دالة المعلومات للاختبار تقدم أكبر كمية من المعلومات عند مستويات القدرة المتوسطة وأقل منها بالنسبة للمعلومات عن مستويات القدرة العالية والمتدنية. وقامت أيضاً علي (٢٠١٢) بدراسة تهدف إلى بناء اختبار تحصيلي محكي المرجع لمقرر القياس والتقويم في التربية باستخدام نموذج راش وشملت عينة ٤٢٦ طالب وطالبة من كلية التربية جامعة دمشق. توصلت الباحث إلى اختبار تحصيلي لمقرر القياس والتقويم يشمل (٨٩) فقرة يتمتع بدرجات عالية من ثبات المفردة وثبات الأفراد وبينت الباحثة أن الاختبار يقدم أكبر كمية معلومات عن مقرر القياس والتقويم في التربية عند مستويات القدرة المتوسطة وتنخفض عند المستويات العالية والمتدنية من خلال استخراج قيمة دالة المعلومات. وعمل أيضاً Almusawi (٢٠١٦) على بناء اختبار محكي المرجع لمادة القياس والتقويم التربوي على عينة ٣٤٨ طالباً في جامعة البحرين، ومن خلال توظيف نظرية الاستجابة للمفردة، توصل من خلاله إلى اعتماد اختبار محكي يتكون من (٣١) فقرة وفقاً لنموذج راش يتمتع بدرجة من الثبات للإفراد ولفقرات الاختبار ويقدم أكبر كمية من المعلومات عن مستويات القدرة المتوسطة.

ونجحت الكساسبة (٢٠١٣) في بناء اختبار محكي المرجع لقياس مدى إتقان طلبة الدراسات العليا في جامعة مؤتة– الأردن لكفايات البحث العلمي. حيث قامت الباحثة بإعداد اختبار مكون من (٤٠) فقرة طبق على ٣٠٠ طالب وطالبة في الجامعة. ووجدت الباحثة من خلال الاختبار إن هناك فروق بالكفاءة بين الطلبة لصالح طلبة الدكتوراه ثم الماجستير ثم الدبلوم ونتج عن البحث تحديد القدرة البحثية لطلبة الجامعة على أنها متدنية حيث لم تتجاوز نسبة الطلبة الذين تجاوزوا درجة القطع ٣١%. وعمل الطراونة (٢٠١٦) على بناء اختبار محكي المرجع في الإحصاء التحليلي لطلبة الدراسات العليا في الجامعات الأردنية الحكومية باستخدام النظرية الحديثة (نموذج راش) وشملت الدراسة على عينة ٣١٩ طالب وطالبة في مرحلة الماجستير. توصل الباحث إلى اختبار مكون من (٣٦) فقرة يتمتع بدلالات صدق وثبات مقبولة، حيث أوصى الباحث على اعتماد على الاختبار في تقويم برامج الدراسات العليا والتحقق

أيضا من خصائصه باستخدام النموذج الثنائي والثلاثي المعلمة. بينما قام Alhabashneh & Najjar (٢٠١٧) بدراسة تهدف إلى بناء اختبار محكي المرجع لقياس الكفايات البحثية والإحصائية شملت عينة ٢٧٥ في الجامعات الأردنية الحكومية، حيث توصل الباحثين إلى اختبار يتكون من (٤٥) فقرة على درجة من الثبات، وأظهر الاختبار فروق بين امتلاك الطلبة للكفايات حسب المرحلة الدراسية لصالح طلبة الدكتوراه وأوضحت الدراسة إلى وجود تدني واضح في إتقان طلبة الدراسات العليا في الجامعات الأردنية للكفايات البحثية والإحصائية.

قامت الحاج (٢٠١٦) ببناء اختبار تحصيلي محكي المرجع لمقرر علم النفس التربوي وفق نموذج راش الأحادية لنظرية الاستجابة للمفردة. حيث صممت الباحثة اختبار مكون من (٥٠) فقرة حذف منها (٢٥) فقرة غير مطابة مع النموذج المستخدم استناداً على عينة قوامها ٢٤٦ طالبة من جامعة أم القرى- السعودية. حيث توصلت الباحثة من خلال الاختبار بصورته النهائية إلى دلالات من الثبات ومستوى صعوبة متوسط.

ومن خلال عرض هذه الدراسات نجد البعض منها اهتمت ببناء اختبار محكي المرجع لعدة مقررات علمية خاصة بكلية التربية ذات أهمية كبرى لمخرجاتها مثل مقررات القياس والتقويم وعلم النفس التربوي، بينما نجد البعض الاخر من هذه الدراسات اتجه نحو بناء اختبار محكي المرجع يخدم الكليات الجامعية في التعرف على الكفاءة والتميز بين الطلبة مثل اختبار الكفاءة اللغوية – اللغة العربية والمهارات البحثية والاحصائية.

وتختلف تلك الدراسات في نوع القياس المستخدم لبناء الاختبارات المحكية فنجد على سبيل المثال دراسة هادي ومراد (٢٠٠٦) والحاج (٢٠١٦) استخدما النظرية الحديثة متمثلة بنموذج راش لبناء الاختبار، وعلى الجانب الأخر نجد كلاً من الكساسبة (٢٠١٣) و Alhabashneh & Najjar (٢٠١٧) استخدما النظرية الكلاسيكية للوصول الى اختبار محكي المرجع.

ونجد الاختلاف بين الدراسة الحالية وما يشابهها من الدراسات السابقة مثل دراسة الكساسبة (٢٠١٣) التي تطرقت إلى الكفاءة البحثية باستخدام أسلوب القياس الكلاسيكي في بناء الاختبار المحكي بينما الدراسة الحالية اعتمدت نموذج راش–النظرية الحديثة لبناء اختبارها المتعلق بمناهج البحث. وهذا الاختلاف أيضاً ينطبق على دراسة Alhabashneh & Najjar (٢٠١٧) التي اهتمت ببناء اختبار محكي المرجع للكفايات البحثية والاحصائية معاً وفقاً للنظرية الكلاسيكية.

المنهج وإجراءات الدراسة

مجتمع الدراسة وعينتها:

يتكون مجتمع الدراسة من طلبة وطالبات الكليات التربوية في دولة الكويت (كلية التربية–جامعة الكويت وكلية التربية الأساسية في الهيئة العامة للتعليم التطبيقي والتدريب). حيث هؤلاء الطلبة والطالبات المتوقع تخرجهم سوف يكون توجههم إلى السلك التعليمي وإتمام مرحلة الدراسات العليا لاحقا. ونظراً للعدد الكبير من الطلبة والطالبات في تلك الكليات يصعب تطبيق الاختبار عليهم جميعاً، لذلك تم اختيار عينة من الطلبة المتوقع تخرجهم (أي بالسنوات الأخيرة). وتنقسم العينة إلى قسمين: العينة التجريبية شملت ٥٢ طالب وطالبة من كلية التربية جامعة الكويت في الفصل الدراسي الثاني من العام ٢٠١٤–٢٠١٥ لطلبة وطالبات مقرر حلقة بحث (المتطلب الأخير قبل التخرج). وعينة المطابقة مع نموذج راش التي تم تطبيق الاختبار عليها بعد تجريبه وحذف بعض الفقرات وتتكون من ٣٦٢ طالب وطالبة في كلية التربية جامعة الكويت وكلية التربية الأساسية بالتعليم التطبيقي (متوقع تخرجهم– مقرر حلقة بحث) في فصلين دراسيين مختلفين في العام ٢٠١٤–٢٠١٥ والعام ٢٠١٦–٢٠١٧. وهذا التباين الزمني البسيط يعود لصعوبة تطبيق البحث نظراً لعدم تفرغ الباحث وانشغاله في عدة أمور بحثية أخرى منها أعداد أطروحة الدكتوراه.

أداة الدراسة:

لبناء الاختبار المستخدم في هذه الدراسة تم الاسترشاد بالدراسات السابقة والباحثين الذين تطرقوا إلى خطوات بناء اختبار محكي المرجع ووضحوا الأسس لذلك مثل علام (٢٠٠١)، عيد (٢٠٠٦)، و الحموري والكحلوت (٢٠٠٩). بالبداية تم تحديد الغرض من الاختبار وهو قياس قدرة الأفراد المقبلين على الدراسات العليا في المهارات والكفايات البحثية. ومن ثم قام الباحث بتحديد المحتوى المراد قياسه وذلك من خلال طريقتين، الأولى بالرجوع إلى توصيف مقرر مناهج البحث والكتب المعتمدة لتدريس الطلبة في كلية التربية–جامعة الكويت وكلية التربية الأساسية للتحديد محتوى الاختبار ومعرفة الأهداف السلوكية التي لا بد إن تتوفر في الفرد الذي اجتاز المقرر بنجاح. أما الطريقة الثانية، تم التواصل مع بعض أساتذة مقرر مناهج البحث في الكليتين في دولة الكويت وأيضا بعض الأستاذة في جامعة الدمام (سابقاً) والملك فيصل والملك سعود في المملكة العربية السعودية للأخذ بآرائهم تجاه المحتوى وتحديد المحاور الأساسية التي يتضمنها الاختبار وهي (أنواع البحث حسب المنهج والتصميم، عناصر البحث، مصادر المعلومات للبحوث التربوية، المتغيرات، أنواع العينات، أدوات البحث التربوي، الخصائص السيكومترية للبحث التربوي، الإحصاء).

ومن ثم صاغ الباحث الأهداف السلوكية لكل تلك المحاور الأساسية التي تتعلق بالمحتوى المراد قياسه، ومن ثم تم صياغة (٥٠) فقرة اختبارية من نوع الاختيار المتعدد التي تتكون من أربع بدائل استناداً على الأهداف السلوكية الرئيسية والفرعية (المساعدة). ومن ثم قام الباحث بالتحقق من صدق المحتوى من خلال عرضة على مجموعة مكونه من (٦) أعضاء هيئة تدريس في الكليتين من الذين لديهم الخبرة البحثية والبعض منهم متخصص في القياس والتقويم بالإضافة إلى الخبرة في تدريس الطلبة لمقرر مناهج البحث لإبداء رأيهم في فقرات الاختبار، من حيث وضوح فقراته وسلامتها اللغوية ومدى ملاءمتها لقياس أهداف السياق. وقدم قام الباحث بإجراء بعض التعديلات على متن بعض الفقرات وبدائلها لتكون أكثر ملائمة من الناحية الفنية

والعلمية وذلك استناداً على ملاحظات المحكمين التي شملت فقط بعض التعديلات لأغراض السلامة اللغوية والتغيرات لبعض بدائل الفقرات للوصول الى صياغة أوضح وأدق. ومن خلال رأي المحكمين أيضاً أتضح إن الاختبار لم يحتوي على عبارات مكررة أو غير مناسبة للأهداف السلوكية.

التجربة الاستطلاعية

طبق الاختبار بصورته الأولية التي تحتوي على (٥٠) فقرة تشمل المحاور الرئيسية سالفة الذكر على عينة ٥٧ طالب وطالبة في كلية التربية في جامعة الكويت من خارج عينة الدراسة والمقيدين في مقرر حلقة بحث المتطلب الأخير للتخرج. وتلخص الهدف من هذه التجربة الاستطلاعية ما يلي: التعرف على مدى ملائمة موضوعات الاختبار وأسئلة للمفحوصين، التعرف على مدى وضوح تعليمات الاختبار والبيانات الوارد فيه للمفحوصين، معرفة الوقت الكافي لتطبيق الاختبار، حساب الخصائص السيكومترية.

بعد الحصول على إجابات الطلبة تم استبعاد خمس طلبة تحتوي أوراق إجاباتهم على اختيار أكثر من بديل لكل فقرة، وبالتالي تكون العينة التجريبية النهائية ٥٢ طالب وطالبة. وقدم تم استخدام برنامج SPSS لتحليل البيانات لاختبار (٥٠) فقرة لتوضيح معامل الصعوبة ومعامل التمييز وفقاً للنظرية الكلاسيكية. تم حذف (١٠) فقرات من الاختبار لان الباحث احتفظ بالفقرات التي معامل صعوبتها يتراوح بين (٠,٣٤ – ٠,٨٤)، والاحتفاظ بالفقرات التي كان قيم معامل التمييز لها أكثر من (٠.٣٠)، مثلما أوضح كلاً من مراد وسليمان (٢٠١٢) و عيد (٢٠٠٦) معايير معامل الصعوبة والتمييز للحكم على ملائمة السؤال. وبالتالي تم الاحتفاظ (٤٠) فقرة، كما لم تشكو العينة من بعض النقاط المتعلقة بالصياغة اللغوية للاختبار أو عدم وضوح الاختبار بشكل عام. كم أن متوسط زمن الإجابة على الاختبار بلغ ٥٠ دقيقة.

بعد أن أصبح عدد الفقرات (٤٠) فقرة تم حساب معاملات الارتباط بين المفردات والدرجة الكلية للإشارة إلى صدق الاتساق الداخلي للاختبار. ووجد إن جميع معاملات

ارتباط المفرد بالدرجة الكلية للاختبار دالة عند مستوى (٠,٠١) مما يدل على أن جميع المفردات تنتمي له. وفيما يتعلق بحساب ثبات الاختبار تم استخدام معامل الفا كرونباخ (٠,٨٨) الذي يوضح إن الاختبار يتمتع بدرجة ثبات عالية.

نتائج الدراسة

بعد تصحيح الاختبار وإعطاء الإجابة الصحيحة (١) والإجابة غير الصحيحة عن الفقرة (٠). وفيما يلي استعراض للبيانات الإحصائية التي أسفر عنها تحليل استجابات العينة على فقرات الاختبار وذلك باستخدام برنامج SPSS v.24 والخاصية المضافة Essential extensions for Rasch in SPSS.

السؤال الأول: ما مدى تحقق بيانات الاختبار المصمم لافتراضات نموذج راش؟

كخطوة من خطوات تطبيق نموذج راش قبل التحليل يتوجب حذف المفحوصين الحاصلين على الدرجة النهائية أو صفر لعدم مناسبة الاختبار لقدراتهم، ثم يتم حذف الفقرات التي أجابها جميع الأفراد إجابة صحيحة أو أخطى فيها جميع المفحوصين (كاظم، ١٩٨٨). ومن خلال هذا الاختبار لم يحصل أي مفحوص على العلامة الكاملة (٤٠) ولم يحصل أي مفحوص على صفر وكذلك لم يكن هناك أي فقرة جاوب عليها المفحصين جميعهم بشكل صحيح أو خاطئ وبالتالي لم يتم استبعاد أي إجابات أو أفراد العينة وجميعهم دخلوا في التحليل الأولي.

تم حساب التقديرات النهائية لمعلمي صعوبة الفقرة وقدرة الأفراد، للتحقق من مطابقة قدرات الأفراد للنموذج ومطابقة فقرات الاختبار للنموذج من خلال استخدام الملائمة الداخلية (Infit) والخارجية (Outfit) عند مستوى +٢، وتم تحليل المطابقة وفق ثلاث مراحل:

أولاً، وجدت نتائج التحليل إن متوسط القدرة للأفراد ليس ببعيد عن (٠) وهو ما يفترضه النموذج، وبانحراف معياري قريب من (١) وبالتالي تم تطبيق الملائمة الداخلية والخارجية حيث إن الأفراد التي تجاوزت قدرتهم (٢) تم حذفهم من العينة والسبب يعود إن هؤلاء غير مطابقين للنموذج وتبتعد استجاباتهم عن التوقعات حيث

مستوى الصعوبة أكبر من قدراتهم والتخمين قد يكون سبباً وراء ذلك، وكان عددهم (٥) أفراد. بينما الذين تعدت استجاباتهم (–٢) تم حذفهم ويبلغ عددهم (٤) والسبب المتوقع لعدم مطابقتهم هو قد يكون الحرص الشديد أو البط في الإجابة أو الغش.

ثانياً، وتم الإبقاء على (٣٥٣) بنفس عدد فقرات الاختبار (٤٠) لأجراء التحليل الثاني المتعلق بتطابق الفقرات لنموذج راش. أسفرت النتائج اقتراب المتوسط الحسابي للصعوبة الفقرة من الصفر الافتراضية و بانحراف معياري قريب من (١) وهو ما يفترضه النموذج. ومن خلال الملائمة الداخلية والخارجية +٢ تم الإبقاء على جميع الفقرات باستثناء الفقرات التالية (٣٠ / ٣٩/ ٤٠) لعدم ملاءمتها، حيث (٢) تعني أما عيب في الصياغة أو أن الفقرة تعرف متغيراً آخر يختلف عن باقي، وأما (–٢) قد تعني إن الفقرة تقيس متغيراً آخر مرتبط بالمتغير موضع القياس.

ثالثاُ، تم أجراء التحليل على عينة مكونة من (٣٥٣) وأيضاً اختبار مكون من (٣٧) فقرة ولم يجد الباحث أي من قدرات الأفراد أو صعوبة الفقرات من تجاوز +٢ وإنما تراوحت القيم داخل هذا النطاق والذي يدل على تجانس فقرات الاختبار على قياس القدرة المراد قياسها وتدرج قدرات الأفراد. وكان تقدير المتوسط الحسابي لصعوبة الفقرات (٠,٠٨) بانحراف معياري (.٩٢)، والمتوسط الحسابي لقدرة الأفراد بلغ (٠,٢٠) بانحراف معياري (.٧٥).

وتم التحقق من افتراضات نموذج راش كالتالي:

أحادية البعد: للتحقق من أحادية البعد تم استخدام التحليل العاملي وقد أسفر التحليل العاملي عن: (٠,٨٤) لاختبار Kasier-Meyer-Olkin، وأسفرت نتيجة Bartlett's Test عن دلالة إحصائية، وبالتالي تم تحقق شروط تطبيق التحليل العاملي على البيانات (Coakes & Steed, 2009). تم استخدام طريقة المكونات الأساسية (Principal Components Analysis)، وتم تدوير المحاور بطريقة (Varimax) عن أثني عشر عامل.

ويتضح من خلال جدول (١) الذي يوضح الجذر الكامن، ونسبة التباين المفسر والتراكمي المفسر لجميع العوامل. ويتضح أن نسبة التباين المفسر للعامل الأول لم تقل عن (٤٠%)، وان نسبة قيمة الجذر الكامن للعامل الأول إلى قيمة الجذر الكامن للعامل الثاني تزيد عن (١٠) حيث يتضح إنه العامل الرئيسي، وبالتالي يتبين إن هناك سمة كامنة واحدة هي المسؤولة عن تفسير التباين في درجات الاختبار كما أشار Hambelton & Swaminathan (١٩٨٥). ونتيجة لذلك يمكن قبول افتراضية نموذج راش المتعلق بأحادية البعد.

جدول (١) قيم الجذر الكامن والتباين المفسر للعوامل المستخلصة من التحليل العاملي

رقم العامل	الجذر الكامن	نسبة التباين المفسر	نسبة التباين التراكمي المفسر
١	١٩,٤٣٢	٤٨,٥٨١	٤٨,٥٨١
٢	٧,٧٩١	١٩,٤٧٩	٦٨,٠٦٠
٣	٣,٩٢٠	٩,٨٠٠	٧٧,٨٦٠
٤	٢,٢٩٥	٥,٧٣٨	٨٣,٥٩٩
٥	١,٨٧٢	٤,٦٨٠	٨٨,٢٧٨
٦	١,٣٦١	٣,٤٠٢	٩١,٦٨١
٧	.٩٥٢	٢,٣١١	٩٣,٩٩٢
٨	.٦٦٤	١,٦٦٠	٩٥,٦٥٣
٩	.٥١٧	١,٢٩٣	٩٦,٩٤٦
١٠	.٤٥١	١,١٢٨	٩٨,٠٧٤
١١	.١٩٩	.٤٩٨	٩٨,٥٧٢
١٢	.١٦٣	.٤٠٧	٩٨,٩٧٩

الاستقلال الموضعي: تم تحقق افتراض الاستقلال الموضعي، حيث بمجرد تحقق أحادية البعد في الاختبار يتحقق الاستقلال الموضعي أو كفيلة بتحقق الاستقلال الموضع (Hambelton & Swaminathan, 1985).

تساوي معاملات التمييز: تم حساب معاملات التمييز لفقرات الاختبار للتحقق من تساوي معاملات التمييز من خلال معامل الارتباط الثنائي (Point Biserial Correlation Coefficient). ونلاحظ من جدول (٢) إن متوسط قيم المعاملات التمييز (٠.٤٤) ومعاملات التمييز تراوحت بين (٠,٣٠ – ٠,٦٨)، أي أن مدى معامل التمييز التي تقع فيها الفقرات (٠,٣٨). وللتحقق من تساوي معاملات التمييز ومطابقتها لنموذج راش يجب أن تكون جميع قيم معاملات التمييز تقع بين المدى (متوسط معاملات التمييز± ٠,٣٨). وبالتالي فإن المدى الواقع ضمن حدود المطابقة يتراوح بين (٠,٠٦ – ٠,٨٢) ومن خلال الجدول (٢) نلاحظ إن جميع قيم معاملات تمييز الفقرات وقعت ضمن هذا المدى، وهذا دلالة على تحقق افتراض تساوي معاملات التمييز كما أشار Hambleton & Swaminathan (١٩٨٥).

جدول (٢) معاملات تميز أسئلة الاختبار باستخدام معامل الارتباط الثنائي.

الفقرة	rpbis	الفقرة	rpbis
١	.٥٠	٢٠	.٤٢
٢	.٥١	٢١	.٣١
٣	.٤٧	٢٢	.٥٦
٤	.٣٢	٢٣	.٣٦
٥	.٣٧	٢٤	.٣٣
٦	.٥٠	٢٥	.٣٩
٧	.٣٠	26	.٣٣
٨	.٣٢	٢٧	.٤٠
٩	.٥٦	٢٨	.٥٧

الفقرة	rpbis	الفقرة	rpbis
١٠	.٣٠	٢٩	.٥١
١١	.٥٩	٣٠	.٤٢
١٢	.٣٩	٣١	.٦٨
١٣	.٥١	٣٢	.٣٦
١٤	.٣٠	٣٣	.٤٠
١٥	.٥٧	٣٤	.٥١
١٦	.٣٧	٣٥	.٤٩
١٧	.٥٦	٣٦	.٥١
١٨	.٥٣	٣٧	.٣٣
١٩	.٣٥		

عدم وجود التخمين: تم التأكد من عدم وجود تخمين من خلال فحص جميع الفقرات ذات الصعوبة العالية وملاحظة إجابات المفحوصين من الفئة الدنيا لتلك الفقرات. حيث يتضح إن أدائهم في الفقرات الصعبة متدني وبشكل ملحوظ، وبالتالي استناداً على إجابات هذه الفئة الدنيا الخاطئة للأسئلة الصعبة يمكن تحقيق شرط عدم وجود التخمين.

السرعة: الاختبار ليس اختبار سرعة حيث تم الحصول على إجابات المفحوصين كاملة دون ترك أي سؤال دون أجابه. ولم يشهد الباحث أي من مفحوص اختار بديل معين من البدائل لجميع الأسئلة، ناهيك عن أنه لم يسلم أي مفحوص الاختبار بأقل من منتصف الوقت المقرر للاختبار الذي بلغ ما يقارب ٥٠ دقيقة بالتجربة الاستطلاعية.

السؤال الثاني: ما تقديرات معالم صعوبة فقرات الاختبار المصمم وفق نموذج راش؟

للإجابة على هذا السؤال تم تقدير معالم صعوبة فقرات الاختبار بوحدة اللوجيت وتم عرض الخطأ المعياري لكل فقرة، كما كان المتوسط الحسابي لصعوبة الفقرات قريب

من الصفر الافتراضية حيث بلغ (٠,٠٨). ويوضح جدول (٣) أن مدى قيم معامل الصعوبة للفقرات يتراوح بين (-١,٤٢ / ١,٤٧) والتي تعد داخل نطاق الصعوبة المقبولة الذي يقع بين (-٢ /٢) كما وضح ذلك كلاً من Hambelton & Swaminathan (١٩٨٥).

جدول (٣) معاملات صعوبة الأسئلة بعد تدرجها باستخدام نموذج راش

الفقرة	الصعوبة باللوجيت	الخطأ المعياري	الفقرة	الصعوبة باللوجيت	الخطأ المعياري
١	-٠,٠٥	.١٠٧	٢٠	.٦٣	.١٠٨
٢	-.٣٥	.١١٠	٢١	-.١٢	.١٠٧
٣	-.٧٥	.١١٧	٢٢	-.٢٩	.١٠٩
٤	.٤١	.١٠٦	٢٣	.٦٠	.١٠٨
٥	-.٣١	.١٠٩	٢٤	.٤٧	.١٠٧
٦	.٣٩	.١٠٧	٢٥	١,٣٠	.١٢٠
٧	-١,٠٧	.١٢٦	٢٦	-١,٤٢	.١٣٩
٨	-١,٣٤	.١٣٦	٢٧	-.٨٨	.١٢٠
٩	.٨٣	.١١٠	٢٨	.٤٢	.١٠٧
١٠	-١,٣٣	.١٣٢	٢٩	-١,١٩	.١٣٠
١١	.٣٧	.١٠٦	٣٠	.٦٠	.١٠٨
١٢	.٣٥	.١٠٦	٣١	.٧٧	.١٠٩
١٣	.٣٥	.١٠٥	٣٢	.٦٤	.١٠٩
١٤	.٣٩	.١٠٨	٣٣	-.٠٥	.١٠٧
١٥	.١٣	.١٠٦	٣٤	-.١٨	.١٠٨
١٦	.٣٧	.١٠٦	٣٥	.٦١	.١٠٨
١٧	-.٠٧	.١٠٦	٣٦	١,٢٢	.١٢١
١٨	-.٢١	.١٠٨	٣٧	١,٤٧	.١٢٥
١٩	.٤١	.١٠٦			

السؤال الثالث: ما الخصائص السيكومترية للاختبار المصمم؟

إن تحقيق افتراضات نموذج راش يعد في جوهرة تقييم لمدى صدق الاختبار (كاظم،١٩٨٨) وبالتالي هذا الاختبار يتمتع بدرجة صدق، نتيجة لتطابق بيانات الاختبار مع افتراضات نموذج راش بعد حذف الاستجابات والفقرات غير المطابقة. بالإضافة إلى ذلك، صدق المحتوى الذي تم تطبيق في هذه الدراسة من خلال الاستعانة ببعض المختصين لعرض الاختبار عليهم وتم ذكر ذلك بالتفصيل سابقاً.

وفيما يتعلق بثبات الاختبار، تم حساب معامل الثبات المتعلق بالفقرات (٠,٨٩) وتم حساب معامل الثبات المتعلق بقدرة الأفراد (٠,٩١) وهي قيم تعتبر عن ارتفاع معامل الثبات والذي يدل على إن فقرات الاختبار يمكن لها التمييز بين مستويات المفحوصين وإن عينة الأفراد كانت كافية للتحديد وتمييز مستويات صعوبة الفقرات، وبالتالي يمكن الاستفادة من هذا الاختبار للتحقق الهدف المنشود منه.

التوصيات:

توصي الدراسة بالتالي:

تطبيق اختبار مناهج البحث على الطلبة المتقدمين للدراسات العليا في الكليات التربوية من أجل تحديد كفاءة الطالب البحثية وما إذا كان بحاجة إلى ساعات مكثفة من مادة مناهج البحث قبل البدء بمواد التخصص. حيث هذا الاختبار يتمتع بالصدق من خلال تطابقه مع فرضيات نموذج راش، بالإضافة إلى انه يتمتع بدرجة ثبات مرتفعة لكل من صعوبة الفقرة وقدرة الفرد.

تم تحديد درجة القطع للاختبار في هذه الدراسة باستخدام Angoff Mehtod من خلال أخذ متوسط المحكمين وكانت تقريباً (٢٠ درجة). لكن قد تكون هناك الحاجة إلى دراسة للتأكد من صلاحية درجة قطع الاختبار من خلال أسلوب قائم على الأهداف السلوكية لتحديد درجة القطع.

التحقق الإضافي من الاختبار من خلال استخدام نماذج أخرى للنظرية الحديثة (الثنائي والثلاثي المعلمة)

توصي الدراسة باستخدام النظرية الحديثة في تصميم اختبارات أخرى لقياس قدرات الأفراد المقبلين على دراسة الماجستير والدكتوراه (مثل اللغة العربية، القدرات الإحصائية) وعدم الاعتماد الكلي على المعدل التراكمي في البكالوريوس في قبول الطالب.

المراجع

أبو هاشم، حسن. (٢٠٠٦). دراسة مقارنة بين النظرية التقليدية ونموذج راش في اختيار فقرات مقياس مداخل الدراسة لدى طلاب الجامعة. مجلة كلية التربية – جامعة الزقازيق، يناير (٥٢)، ١ – ٥٢.

الحاج، فريال. (٢٠١٦). الخصائص السيكومترية لاختبار تحصيلي محكي المرجع في مقرر علم النفس التربوي وفق نموذج راش. المجلة التربوية – جامعة الكويت، ٣٠ (١٢٠)، ١٧١ – ٢٠٠.

الحموري، هند و الكحلوت، أحمد (٢٠٠٩). استراتيجية مقترحة لبناء اختبارات تشخيصية محكية المرجع. مجلة العلوم التربوية والنفسية – جامعة البحرين، ١٠ (٢)، ١٨٥ – ٢١٤.

السامرائي، محمد والخفاجي، أحمد. (٢٠١٢). بناء اختبار تحصيلي محكي المرجع في مادة علم نفس الخواص لطلبة أقسام العلوم التربوية والنفسية. مجلة الأستاذ، (٢٠٣)، ٩٦٤ – ١٠٠٢.

الشريفين، نضال. (٢٠٠٦). الخصائص السيكومترية لاختبار محكي المرجع في القياس والتقويم التربوي وفق النظرية الحديثة في القياس التربوي والنفسي. مجلة العلوم النفسية والتربوية– جامعة البحرين، ٧ (٤)، ٨٠ – ١٠٩.

الطراونة، صبري. (٢٠١٦). بناء اختبار محكي المرجع في الإحصاء التحليلي لطلبة الدراسات العليا في كليات العلوم التربوية في الجامعات الأردنية الحكومية باستخدام نظرية استجابة الفقرة. مجلة العلوم التربوية والنفسية–جامعة البحرين، ١٧ (١)، ٥١٥– ٥٤٢.

الكساسبة، حنان. (٢٠١٣). بناء اختبار محكي المرجع لقياس مدى إتقان طلبة الدراسات العليا في جامعة مؤتة لكفايات البحث العلمي. (رسالة ماجستير، جامعة مؤتة، الأردن). متوفرة على: mandumah.com/Record/786787

النجار، نبيل. (٢٠١٠). القياس والتقويم: منظور تطبيقي مع تطبيقات برمجية SPSS. عمان: دار الحامد للنشر والتوزيع.

علام، صلاح الدين محمود (٢٠٠١). الاختبارات التشخيصية مرجعية المحك في المجالات التربوية والنفسية والتدريبية. القاهرة: دار الفكر العربي.

علي، نداء. (٢٠١٢). فاعلية استخدام نموذج راش في بناء اختبار تحصيلي محكي المرجع لمقرر القياس والتقويم في التربية. (رسالة دكتوراه، جامعة دمشق، سوريا). متوفرة على: http://damascusuniversity.edu.sy

عيد، غادة. (٢٠٠٤). الدرجة الحقيقية المقدرة باستخدام نظرية السمات الكامنة والنظرية الكلاسيكية: دراسة سيكومترية. مجلة أم القرى للعلوم التربوية والاجتماعية والإنسانية، ١٦ (٢)، ٢٢٨ – ٢٨٤.

عيد، غادة. (٢٠٠٦). القياس والتقويم التربوي مع تطبيقات برنامج SPSS. الكويت: الفلاح للنشر والتوزيع.

كاظم، أمينة. (١٩٨٨). استخدام نموذج راش في بناء اختبار تحصيلي في علم النفس وتحقيق التفسير الموضوعي للنتائج. الكويت: مطبوعات جامعة الكويت.

مراد، صلاح و سليمان، أمين. (٢٠١٢). الاختبارات والمقاييس في العلوم النفسية والتربوية: خطوات إعدادها وخصائصها. الكويت: دار الكتاب الحديث.

هادي، فوزية و مراد، صلاح. (٢٠١٣). استخدام نموذج راش في إعداد وتقنين اختبار في اللغة العربية لقبول طلاب الدراسات العليا بكليات العلوم الإنسانية بجامعة الكويت. المجلة التربوية –جامعة الكويت، ٢٧ (١٠٦)، ١٥ – ٥٣.

Alhabashneh, M., & Najjar, N. (2017). Constructing a criterion reference test to measure the research and statistical competencies of graduate students at the Jordanian governmental universities. Journal of Education and Practice, 8 (2), 67 – 81.

ALkhadher, M., & Albursan, I. (2017). Accuracy of Measurement in the Classical and the Modern Test Theory: An Empirical Study on a Children Intelligence Test. International Journal of Psychological Studies, 9(1), 71.

Almusawi, N. (2016). Using item response models to develop a criterion referenced test to measure the students' achievement in educational evaluation. Journal of Educational and Psychological studies – Sultan Qaboos University, 10 (4), 727 – 736.

Bachman, L. F. (2004). Statistical analyses for language assessment. Cambridge: Cambridge University .

Bond, L. (1996). Norm- and Criterion-referenced testing. Practical Assessment, Research & Evaluation, 5 (2), 1-3.

Brown, J. D. (2005). Testing in language programs: A comprehensive guide to English language assessment. New York: McGraw-Hill .

Brown J. D. (2012a). Choosing the right type of assessment. In C. Coombe, S. J. Stoynoff, P. Davidson, & B. O'Sullivan (Eds.). The Cambridge guide to second language assessment (pp. 133-139). Cambridge: Cambridge University.

Brown, J. D., & Hudson, T. (2002). Criterion-referenced language testing. Cambridge: Cambridge University Press.

Coakes, S., & Steed, L. (2009). SPSS analysis without anguish: Version 16.0 for windows. Milton-Brisbane: Wiley.

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

Jabrayilov, R., Emons, W., & Sijtsma, K. (2016). Comparison of classical test theory and item response theory in individual change assessment. Applied Psychological Measurement, 40(8), 559-572.

Kean, J., & Reilly, J. (2014). Item response theory. Handbook for Clinical Research: Design, Statistics and Implementation. New York, NY: Demos Medical Publishing .

Magno, C. (2009). Demonstrating the difference between classical test theory and item response theory using derived test data. The International Journal of Educational and Psychological Assessment, 1 (1), 1-11 .

Mehrens, W. & Lehmann, I. (1991). Measurement and evaluation in education and psychology (4th ed.). Fort Worth : Holt, Rinehart and Winston.

Ornstein, A. (1993). Norm-referenced and Criterion-referenced tests: An overview. NASSP Bulletin, 77 (555) 28-39.

Thorndike, R., & Thorndike-Christ, T. (2010). Measurement and evaluation in psychology and education (8th ed.). London: Pearson Education.

ملحق (١)

الاختبار بصورته النهائية بعد تطبيق نموذج راش

اختبار مناهج البحث

إعداد: د. طلال سعد المطيري

أجب عن الأسئلة الأتية

اختر الإجابة الصحيحة من البدائل التي تلي كل عبارة من العبارات التالية:

١- نوع البحث الذي يتم من خلاله عرض البيانات باستخدام الكلمات والصور ونادراً ما يتم استخدام لغة الأرقام، البحث:

أ- الأساسي. ب- الكمي.

ج- النوعي. د- الكمي النوعي.

٢- البحث الذي تتم فيه معالجة البيانات بأساليب إحصائية رقمية بهدف تعميم النتائج على مجتمع الدراسة، البحث:

أ- الأساسي. ب- الكمي.

ج- النوعي. د- المختلط/المندمج.

٣- دراسة موضوع في فترة زمنية معينة لفهم الحاضر والتنبؤ بالمستقبل من خلال بعض المصادر الأولية والثانوية، البحث:

أ- الإجرائي. ب- التاريخي.

ج- الوصفي. د- التجريبي.

٤- قام باحث ما بعمل دراسة لطرق التدريس الحديثة لمادة العلوم، وقام بإجراء اختبار أولي لقياس مدى فهم الطلبة للمادة ومن ثم عمل على تطبيق برنامج إلكتروني لتدريس العلوم لمدة شهر وبعد ذلك طبق اختبار لقياس مدى التغير في فهم الطلبة. يعتبر هذا النوع من البحوث:

أ- الوصفية. ب- الكمية.

ج- الإجرائية. د- التجريبية.

٥- البحث الذي يعتمد على دراسة الواقع أو ظاهرة توجد بالواقع ويهتم بوصفها من خلال بعض الدراسات المسحية أو المقارنة أو الدراسات الارتباطية.

أ- التجريبي. ب- الإحصائي.

ج- الوصفي. د- الإجرائي.

٦- قام مدير مدرسة بعمل بحث عن التسرب الدراسي في مدرسته حيث كان يشعر بوجود هذه المشكلة التي لها انعكاسات سلبية على ميدان العمل مما يتطلب حل هذه المشكلة. يعتبر هذا النوع من البحوث:

أ- الإجرائية. ب- التجريبية.

ج- الوصفية. د- استنتاجية.

٧- ظاهرة يواجها الأفراد سواء معلمين، طلبة، أولياء أمور تتمثل عن معرفة يجهلها الباحث وتحتاج إلى تقصي علمي عنها.

أ- مشكلة البحث. ب- فرضيات البحث.

ج- أسئلة البحث. د- عينة البحث.

٨- من مصادر اختيار مشكلة البحث ما يلي:

أ- الخبرة. ب- النظريات والدراسات السابقة.

ج- القضايا المجتمعية والعلمية. د- جميع ما سبق صحيح.

٩- العلاقة بين مهارات التفكير النقدي ونتائج الطلبة، تمثل هذه العبارة صياغة ..... للمشكلة.

أ- تصريحية. ب- استفهامية.

ج- فرضية. د- جميع ما سبق غير صحيح.

١٠- بعد الشروع في تحديد المشكلة البحثية يتوجب على الباحث طرح تساؤل عام ينتهي بأسئلة أكثر دقة.

أ- منهج الدراسة. ب- فرضيات الدراسة.

ج- أسئلة الدراسة. د- جميع ما سبق.

١١- حل مؤقت لمشكلة ما أو تخمين من قبل الباحث لحل مشكلة أو التنبؤ بالنتائج.

أ- تفسير الدراسة. ب- فرضية الدراسة.

ج- نتائج الدراسة. د- حقائق بحثية.

١٢- مستوى القلق لدى طلبة الثانوية العامة اعلى من مستوى القلق لدى طلبة التعليم الجامعي، هذه العبارة تمثل الفرضية .....

أ- الصفرية. ب- غير المتجهة.

ج- المتجهة. د- جميع ما سبق غير صحيح.

١٣- العبارة التالية " لا توجد فروق بين الطلبة في الصف الخامس ". تمثل الفرضية .....

أ- الصفرية. ب- غير المتجهة.

ج- المتجهة. د- كلاً من أ – ب صحيح.

١٤- إجراء يلجأ إليه الباحث قبل الاستقرار على خطة البحث وتنفيذها بشكل كلي .....

أ- الدراسة المسحية. ب- الدراسة الاستطلاعية.

ج- الدراسة الوصفية. د- جميع ما سبق صحيح.

١٥- ..... عبارة عن وصف دقيق للدراسة ومتطلباتها حيث يعمل على توجيه خطوات الدراسة ومراحل تنفيذها.

أ- الإطار النظري. ب- الدراسات السابقة.

ج- خطة البحث. د- كلاً من ب – ج صحيح.

١٦- الحالة التي لا بد أن يصرح الباحث بها والتي تشمل عوامل في إجراءات الدراسة تكون غير مناسبة أو غير متوقعة.

أ- أغراض الدراسة. ب- محددات الدراسة.

ج- منهجية البحث. د- التوصيات.

١٧- المرحلة التي تقود الباحث إلى جملة من الاستنتاجات التي تعزز قيامة بالدراسة وتحدد جوانب النقص أو عدم ثبات النتائج.

أ- المنهجية البحثية. ب-التوصيات والمقترحات.

ج- الدراسات السابقة. د- جميع ما سبق غير صحيح.

١٨- عملية الرصد والتوثيق لجميع المعلومات والبيانات التي استخدمها الباحث في دراسته.

أ- ملاحق الدراسة. ب- منهج الدراسة.

ج- مراجع الدراسة. د- فهرس الدراسة.

١٩- دراسة مستوى أداء المعلمين وفقاً لمتغير الجنس (ذكر / أنثى) يعتبر.....

أ- متغير مستقل. ب- متغير تابع.

ج- متغير معدل. د- متغير ضابط.

٢٠- دراسة مستوى تحصيل طلبة الثانوية العامة بناءً على التوزيع الجغرافي (محافظات / مناطق)، الدرجات التي تعبر عن تحصيل الطلبة تعتبر متغير .....

أ- مستقل. ب- تابع.

ج- معدل. د- ضابط

٢١- ..... عبارة عن جميع الأفراد الذين يشكلون موضوع مشكلة البحث.

أ- مجتمع الدراسة. ب- فريق الباحثين.

ج- عينة الدراسة. د- مجموعة المحكمين.

٢٢- ..... مجموعة جزئية التي يمكن تعميم النتائج من خلالها.

أ- مجتمع الدراسة. ب- فريق الباحثين.

ج- عينة الدراسة. د- مجموعة المحكمين.

٢٣- ..... هي يكون فيها لكل فرد من مجتمع الدراسة الفرصة المماثلة للمشاركة في البحث.

أ- العينة العشوائية. ب- العينة الغرضية (القصدية).

ج- العينة الحصصية. د- العينة المتيسرة (الصدف).

٢٤- ..... هي التي يستخدم فيها الباحث الحكم الشخصي على أساس أنها هي الأفضل لتحقيق ما تهدف إليه الدراسة. مثال: اختيار المعلمين ذوي الخبرة التي تتجاوز عشرون عاماً وأكثر.

أ- العينة العشوائية. ب- العينة الغرضية (القصدية).

ج- العينة الحصصية. د- العينة المتيسرة (الصدف).

٢٥- العينة التي يتم اختيارها من خلال تقسيم المجتمع إلى مجموعات أو مستويات ومن ثم اختيار أفراد من كل مجموعة أو مستوى اختياراً غير عشوائي.

أ- الحصصية. ب- الغرضية (القصدية).

ج- المتيسرة (الصدفة). د- جميع ما سبق غير صحيح.

٢٦- من مميزات الاستبيان في البحوث ما يلي:

أ- التكلفة المادية المنخفضة. ب- شمولية تطبيقها على عدد كبير.

ج- الحصول على المعلومات الحساسة. د- جميع ما سبق صحيح.

٢٧- يستطيع الباحث من خلال ..... التأكد من مدى صدق المفحوص في إجاباته عن طريق أسئلة مرنة مرتبطة بالمجالات التي يبحث عنها أو يشك فيها.

أ- الاستبيان. ب- المقابلة.

ج- الملاحظة. د- أ – ج صحيح.

٢٨- الملاحظة كأداة بحثية ومن أنواعها:

أ- المباشرة وغير المباشرة. ب- المقصودة وغير المقصودة.

ج- بالمشاركة ودون المشاركة. د- جميع ما سبق صحيح.

٢٩- من عيوب الاستبيان ما يلي:

أ- عدم المقدرة المشاركين على التعبير اللفظي.

ب- قد يتجاهل المشاركين أسئلة معينة.

ج- ليس لدى المشاركين الحرية بالبوح بالمعلومات.

د- جميع ما سبق صحيح.

٣٠- من الطرائق المستخدمة في التحقق من الخصائص السيكومترية لأداة البحث.

أ- الموضوعية. ب- الصدق.

ج- الثبات. د- جميع ما سبق صحيح.

٣١- يمكن الحصول على عدد من المؤشرات التي تعزز صدق الأداة من خلال .....

أ- المحتوى. ب- البناء.

ج- دلالة المحك. د- جميع ما سبق صحيح.

٣٢- من الطرق التي تستخدم للتحقق من الثبات .....

أ- الاتساق الداخلي. ب- التحليل العاملي.

ج- الصدق التنبؤي. د- جميع ما سبق صحيح.

٣٣- من العوامل المؤثر في ثبات الاختبار ما يلي:

أ- طول الاختبار. ب- مستوى صعوبة الاختبار.

ج- زمن الاختبار. د- جميع ما سبق صحيح.

٣٤- من العوامل المؤثرة في صدق أداة القياس ما يلي:

أ- لغة الأداة. ب- غموض الفقرات.

ج- عوامل القلق لدى المفحوص. د- جميع ما سبق صحيح.

٣٥- قيام الباحث ببعض العمليات الحسابية لعرض كلاً من المتوسط الحسابي، الوسيط، التكرارات، والنسبة المئوية، تسمى هذه العملية بالإحصاء .....

أ- الاستدلالي. ب- الوصفي.

ج- الاستنتاجي. د- كلاً من أ-ج صحيح.

٣٦- قيام الباحث بالاستنتاجات حول خصائص مجتمع، مثل إيجاد الفروق بين الذكر والأنثى أو بين المراحل التعليمية، تسمى هذه العملية بالإحصاء .....

أ- الاستدلالي. ب- الوصفي.

ج- الاستنتاجي. د. جميع ما سبق غير صحيح.

٣٧- ..... يعتبر من أدق مقاييس التشتت لمجموعة من الدرجات فهو لا يعتمد على الدرجات المتطرفة ولا يغفل إشارات الفروق عن المتوسط.

أ- الوسيط. ب- الانحراف المعياري.

ج- المدى الربيعي. د- المدى المطلق.