بناء اختبار محكي المرجع في مناهج البحث العلمي للدراسات العليا والتحقق من دلالات فاعلية الاختبار وفقراته وفق النظريتين الكلاسيكية والحديثة في القياس

د. هانى سليمان أحمد الخالدى

ملخص البحث :

هدفت الدراسة الحالية إلى بناء اختبار محكي المرجع في مناهج البحث العلمي للدراسات العليا والتحقق من دلالات الفاعلية للاختبار وفقراته وفق النظريتين الكلاسيكية والحديثة في القياس(وفقاً للنموذج اللوجستي ثنائي المعلم)، ولتحقيق هذا الهدف تم بناء اختبار في مادة مناهج البحث، يقيس تحصيل الطالب في مادة المناهج، تألف الاختبار من (٤٠) فقرة من نوع الاختيار من أربعة بدائل، تكونت عينة الدراسة من (٢٥٣) طالب من طلبة الدراسات العليا في جامعة نايف العربية للعلوم الأمنية. أظهرت نتائج الدراسة ما يأتي: بالنسبة للنظرية الكلاسيكية فقد تم حذف فقرتين فقط كان تمييزهما منخفض ليصبح الاختبار مكوناً من (٣٨) فقرة، وكان المتوسط العام لصعوبة الفقرات يساوي (٠.٧٩)، ومتوسط تمييز الفقرات يساوي (٠.٤١). أما فيما يتعلق بالنظرية الحديثة في القياس فقد تم التحقق من افتراضاتها. وأظهرت النتائج عدم مطابقة (١٣) فقرة للنموذج اللوجستي ثنائي المعلم، حيث كانت قيمة احتمالية المطابقة لكل منها أقل من (٠.٠١)، لذلك تم حذف هذه الفقرات غير المطابقة للنموذج، وعليه تم الحصول على اختبار محكي المرجع في مادة مناهج البحث العلمي والذي تكون في صورته النهائية من (٢٧) فقرة، حيث تراوحت قيم الصعوبة بين (٠.٢٣٨) و (٣.١٧٢) وبمتوسط حسابي قدره (١.٦٤٩)، وقيم التمييز تراوحت بين (٠.٧٥) و(٥.٣٣٦) وبمتوسط حسابي قدره (٢.١١).

الكلمات المفتاحية: *النظرية الحديثة، النظرية الكلاسيكية، اختبار محكي المرجع، أحادية البعد*

Abstract

Building a Criterion Test for Research Curriculums in Post Graduate Studies and Investigating its Effectiveness and items according to Classical and Modern Theories of Measurement This study aimed to create a criterion test for in academic research curriculums in post graduate studies, and to investigate its effectiveness and items according to classical and modern theories of measurement (according to the binary logistic model parameter). To achieve this goal, a test was created for a research methodologies study unit to measure the extent of learning by students in this subject. The subject was made up from 40 items, of a kind choose from four alternatives . The study sample was made up from 253 students in post graduate study programs in Naif Arab University for Security Sciences. Results of the study revealed the following: According to the classical theory only 2 item were removed. Their differential was low which resulted in the test being only 38 item. The general average for the difficulty of the items was equal to (0.79) and the average differential items was equal to (0.41). As for the modern theory of measurement, its hypotheses were conformed. Results revealed a lack of conformity in 13 item of the binary logistic model parameter. The potential value of conformity between them was less than (0.01). Therefore these items which were not in conformity with the model were removed. Based on this, a criterion test was created in the academic research subject which was made up at the end of the study from 27 item . The difficulty values ranged between (0.238) and (3.172) and a calculated average of (1.649), the differential values ranged between (0.75) and (5.336) and a calculated average of (2.11)).

Keywords: Modern theory, Classical theory, Criterion test, One-dimensional

المقدمـــة:

تعتمد عمليتي القياس والتقويم على أدوات ومقاييس تقدم بيانات رقمية تتيح للتربويين فهم الظاهرة، ولذلك تعتبر الاختبارات والمقاييس التربوية جزءاً هاماً من أجزاء العملية التربوية، فمن خلال الاختبارات يستطيع عضو هيئة التدريس الحكم على مدى تحقق أهدافه التدريسية، ويستطيع أن يحكم على الأسلوب المتبع في شرحه من خلال نتائج الطلبة، وهل هو بحاجة إلى تغيير أسلوبه أم لا. كما تفيد الاختبارات الطالب أيضاً في التعرف على مستوى تحصيله العلمي من خلال نتائجه على هذا الاختبار. كما تتعدد استخدامات الاختبارات في المؤسسات التربوية وتتباين أهدافها. ولكي نحصل على معلومات مفيدة ونتائج دقيقة ينبغي أن يستند تفسير درجاتها إلى نظام محدد مثل نظام اختبارات مرجعية المحك.

وتعد عمليتي القياس والتقويم من المجالات المهمة في العلوم التربوية، حيث أنهما يهدفان إلى تطوير كل من المعلم وتحصيل المتعلم، وهما يعتمدان على تصميم وبناء أدوات تشخيصية أو تكوينية لقياس التحصيل. ومن هذه الأدوات بناء اختبار محكي المرجع للتحقق من

كفاءة المتعلمين واكتسابهم للمهارات الرئيسية والمعارف والاتجاهات المرجوة(علام،٢٠٠١).

وتستخدم الاختبارات محكية المرجع لقياس جوانب محددة في معرفة الطالب تكون مرتبطة بأهداف برنامج معين، لذلك يمكن استخدامها في تقويم التعلم الإتقاني، إذ يمكن تحديد مستوى الأداء المقبول على شكل تقدير عددي مثل أن لا تقل درجة الطالب عن ٨٥% حتى يصنف أنه طالب متقن، كما يمكن استخدام الاختبارات محكية المرجع لاتخاذ قرارات تعليمية وذلك من خلال الاختبارات القبلية التي يتم اجراؤها في بداية البرنامج التعليمي. ولذا فإن هذه الاختبارات تقيس مستويات فنقارن أداء الطالب بكمية المادة التي تم تعلمها أو بمستوى الإتقان دون الحاجة إلى مقارنة أدائه بأداء زملائه كما هو الحال في الاختبارات معيارية المرجع. لذلك فإن غرض الاختبار محكي المرجع اتخاذ قرارات حول ما إذا كان الطالب يتقن مجال سلوكي معين، أما في الاختبارات معيارية المرجع يكون غرضنا الأساسي ترتيب مجموعة من الطلبة على متصل القدرة أو السمة، وتبرز في هذا الإطار اختلافات عديدة بين نوعي الاختبارات، ففي حالة الاختبارات محكية المرجع تكون المقارنة المستهدفة مع المعايير أو مستوى الإتقان الذي يحدد سلفاً، فيما تكون المقارنة المستهدفة في حالة الاختبارات معيارية المرجع مع أداء المجموعة المعيارية، ويتم في الاختبارات محكية المرجع تصنيف الطلبة إلى طلبة متقنين وطلبة غير متقنين(عبابنة، ٢٠٠٩).

إن دراسة العملية التعليمية وتطويرها مرتبطة ارتباطاً وثيقاً بتقييم المخرجات التعليمية، كما أن الاعتماد على المدلولات الرقمية في شتى مجالات التقييم ركن أساسي في هذه العملية، ويعد القياس النفسي والتربوي المستند إلى الأسس الإحصائية من الركائز التي يعتمد عليها التقويم. وللقياس نظريتان هما النظرية الكلاسيكية في القياس والنظرية الحديثة في القياس (الطواها، ٢٠١١).

إن ما يميز النظرية الكلاسيكية أن تقدير الأداء على الاختبار يعتمد على الاستجابات الصحيحة للمفحوص بغض النظر عن خصائص الفقرات التي كانت الاستجابة لها صحيحة وشكلت العلامة الخام، وتتعامل مع التوزيع للمجموعة التي تتقدم للامتحان ويتم فيها وصف تحصيل الطالب من خلال موقعه في هذا التوزيع(التقي،٢٠٠٩). وقد قدمت هذه النظرية مفاهيماً وطرقاً في معالجة الاختبارات الموضوعية ثنائية التدريج وأمكن باستخدامها التوصل إلى دلالات مقبولة عن دقة القياس وموضوعيته، كما يتمثلان في مفاهيم الصدق والثبات للاختبار، وإحصائيات الفقرة المتمثلة في معاملي الصعوبة والتمييز، فصعوبة الفقرة تمثل نسبة المفحوصين الذين أجابوا على الفقرة إجابة صحيحة، أما تمييز الفقرة تمثل معامل الارتباط بين العلامة على الفقرة والعلامة الكلية (Crocker & Algina,1986).

ومع ذلك يرى بعض المختصين أن النظرية الكلاسيكية لم تقدم طرقاً ومفاهيماً مقنعةً يمكن

بها تقييم الاختبارات المقالية ومن ثم التعرف على إمكانات تطويرها وتحسينها. وقد أشار هاملتون وسوامنثان إلى وجود قصور فيها قد لا يمكنها من معالجة بعض القضايا المهمة في القياس والتقويم، ومن أهم تلك الجوانب أن خصائص الاختبار وفقراته تتحدد بالنسبة لمجموعة معينة من المفحوصين، كما أن خصائص المفحوصين أنفسهم تتحدد أيضاً بالنسبة لاختبار معين، كما أنها لا تبين مدى تحصيل الطالب على مستوى الفقرة، إنما على الاختبار ككل، وتفترض أن الخطأ المعياري في القياس متساو لكل المفحوصين(Hamblton, Swarninathan &Rogers,1991). لذلك وجدت النظرية الحديثة في القياس التي تعد ثورة في القياس النفسي والتربوي، وذلك كما يعدها كثير من علماء النفس (علام،١٩٨٦؛ عبد السلام وكاظم والشرقاوي والشيخ، ١٩٩٦).

ولقد نال هذا المدخل الجديد اهتمام الكثيرين من علماء القياس المعاصر أمثال"Hambleton " "Right" & "Lord" وغيرهم، فهذا المدخل يساعد في التغلب على كثير من المشكلات الناجمة عن استخدام النظرية الكلاسيكية في بناء وتصميم الاختبارات وتحليل وتفسير نتائجها، ومواجهة القصور فيها، كما أنه يحقق موضوعية القياس (الوليلي، ٢٠٠٢).

ويعتبر الهدف الأساسي لنماذج الاستجابة للفقرة هو تحديد العلاقة بين استجابات الأفراد على اختبار معين والسمة الكامنة وراء تلك الاستجابات، لكي يمكن تحديد مقدار السمات الكامنة وراء أداء الأفراد في الاختبارات المختلفة للإفادة من ذلك في التنبؤ بسلوكهم في مواقف مماثلة واتخاذ قرارات بشأنهم في ضوء هذا التقدير الكمي للسمات. وأصبحت نظرية الاستجابة للفقرة وسيلة أساسية وشائعة في بناء وتطوير الاختبارات، وقد شجع توفر البرامج الاحصائية على تطبيق النماذج المختلفة لهذه النظرية في مجال الاختبارات والمقاييس، وتتلخص فكرة هذه النماذج في ربطها خصائص الفقرات بمعلم أو أكثر، فهي تقدم بديلاً لنظرية القياس الكلاسيكية في تقديرها لمعالم الفرد والفقرة بأقل قدر من الخطأ ودون اللجوء إلى عينة فقرات للاختبار عشوائية من النطاق المقاس أو الحصول على عينة كبيرة جداً من الفقرات مماثلة لهذا النطاق(علام، ٢٠٠٥). وقد بنيت هذه النظرية والنماذج الرياضية المتعلقة بها على افتراضات قوية ينبغي تحققها في البيانات؛ لكي تؤدي إلى نتائج يمكن الوثوق بها، ومن أهم هذه الافتراضات: أحادية البعد (Unidimensionality)، والاستقلال الموضعي للفقرات (Local Item Independence)، والوتيرية (Montonicity)، والتحرر من السرعة (Hampleton & Swarninathan, 1985). ونتج عن هذه النظرية مجموعة من النماذج تعرف باسم نماذج السمات الكامنة (Latent Trait Models)، تهدف إلى تحديد العلاقات بين أداء الفرد في الاختبار والسمة التي تكمن وراء هذا الأداء وتفسيره. وتعد النماذج اللوجستية ذات المعلم الواحد (نموذج راش) والمعلمين (نموذج

بيرنيوم) والثلاثة معالم (نموذج لورد) من أهم النماذج واسعة الانتشار في هذا المجال. وتختلف هذه النماذج في عدد معالم الفقرة التي تقدرها، إذ يمثل النموذج اللوجستي ثلاثي المعلم الشكل العام للنماذج اللوجستية؛ لأنه يضم المعالم الثلاثة الممكنة للفقرة وهي: الصعوبة، والتمييز، والتخمين(difficult parameter, discrimination parameter, guessing parameter)(bi,ai,ci) على الترتيب، والصيغة الرياضية للنموذج ثلاثي المعلم كالآتي:

$$Pi(Ø)=ci+(1-ci)[1+exp(-Dai(Ø-bi)]^{-1}$$
$$i=1,2,3\ldots\ldots\ldots\ldots,n$$

حيث Pi(Ø) : احتمال إجابة المفحوص الذي اختير عشوائياً من مستوى القدرة (Ø) على الفقرة(i) إجابة صحيحة .

Ci: معلم التخمين ، ai : معلم تمييز الفقرة، bi: معلم صعوبة الفقرة، وØ قدرة المفحوص. أما في النموذج اللوجستي ثنائي المعلم: فيفترض أن جميع قيم التخمين للفقرات تقترب من الصفر، ولذلك الصيغة الرياضية لهذا النموذج كالآتي:

$$Pi(Ø)=[1+exp(-Dai(Ø-bi)]^{-1}$$
$$i=1,2,3\ldots\ldots\ldots\ldots\ldots,n$$

وفي النموذج اللوجستي ذي المعلم الواحد يفترض أن جميع قيم التخمين للفقرات تقترب من الصفر، وأن جميع الفقرات تميز بالقدر نفسه بين الأفراد، ولكنها تتباين فقط في صعوبتها، فتأخذ الصيغة الرياضية الآتية:

$$Pi(Ø)=[1+exp(-D(Ø-bi)]^{-1}$$
$$i=1,2,3\ldots\ldots\ldots\ldots\ldots,n$$

ويعد النموذج اللوجستي ذي المعلم الواحد أكثر نماذج السمات الكامنة في عدد الافتراضات اللازم توافرها في البيانات (Camilli & Shepared, 1994).

وتعد دالة المعلومات من المفاهيم الأساسية في نظرية الاستجابة للمفردة الاختبارية سواء كانت المفردة ثنائية الدرجة، أو تتطلب استجابات متعددة. وتساعد هذه الدالة في بناء أدوات قياس تجعل دقة القياس أو المعلومات المطلوبة أكبر ما يمكن. إن مقدار المعلومات المستمدة من مفردة اختبارية معينة يكون أكبر ما يمكن حول قيمة معلم صعوبة المفردة؛ لكي تسهم المفردات بأكبر قدر من المعلومات فيما يتعلق بمستوى قدرة الأفراد، لذلك ينبغي أن تتمركز صعوبة كل مفردة حول قدرة الفرد الذي يختبر بها، وذلك في النموذجين أحادي وثنائي المعلم. أما في حالة النموذج ثلاثي المعلم فإن أكبر قدر من المعلومات التي تسهم بها مفردة معينة يحدث عند مستوى قدرة أقل قليلاً من قيمة معلم صعوبة المفردة، وذلك بسبب احتمال تأثير معلم التخمين في الاستجابات(علام، ٢٠٠٥).

من جهة أخرى ينبغي التحقق من مطابقة مجموعة الفقرات التي تشكل الاختبار كوحدة واحدة لنموذج معين، وكذلك التحقق من مطابقة كل مفردة على حدة للنموذج، وخاصة إذا كان الاختبار يشمل على أنواع مختلفة من الفقرات التي تتطلب نماذج مختلفة لتقدير معالمها. ويمكن التعرف على الفقرات الضعيفة من خلال مؤشرات التمييز والصعوبة، فإذا كانت قيمة معلم التمييز سالبة أو موجبة ضعيفة تعتبر الفقرة ضعيفة، وكذلك تكون الفقرة ضعيفة إذا كانت

الفقرة سهلة جداً، أو صعبة جداً. ومن الأسباب التي تجعل فقرات الاختبار ضعيفة عدم مطابقة النموذج مع فقرات الاختبار(Chernyshenko, Stark,Chan, Drasgow & Williams, 2001).

وهذا ولم يحظ موضوع بناء اختبار محكي المرجع في مناهج البحث العلمي للدراسات العليا وفقاً لنظرية الاستجابة للمفردة باهتمام الكثير من الباحثين، لكن هناك الكثير من الدراسات العربية والأجنبية ناقشت موضوع بناء أنواع أخرى من الاختبارات وفقاً لنظرية الاستجابة للمفردة والنظرية الكلاسيكية في القياس، ففي دراسة علام (١٩٩٠) التي هدفت إلى استعمال نموذج راش في بناء اختبار تشخيصي محكي المرجع لقياس المعارف الأساسية في إعداد خطة البحوث التربوية والنفسية لدى طلبة الدراسات العليا بكليات التربية في مصر. تألف الاختبار في صورته الأولية من (٣٢) فقرة من نوع الاختيار من متعدد لكل فقرة خمسة بدائل، تكون الاختبار من (٣٠) هدفاً سلوكياً تتعلق بالمعارف الأساسية في إعداد خطة البحوث التربوية والنفسية، وقد تألفت عينة التحليل الإحصائي من (٤٨) طالباً ممن أنهوا مقرر مناهج البحث ويعدون خطة رسالة الماجستير في أحد فروع العلوم التربوية والنفسية. تم تحليل البيانات وفقاً لنموذج راش وباستعمال طريقة (Prox) التي اقترحها كل من رايت (Wright) ودوكلاس (Douglas). وأظهرت النتائج عدم ملائمة (٤) فقرات لافتراضات نموذج راش، وأيضاً عدم ملائمة بيانات فردين من العينة لمتطلبات النموذج، وتراوحت القيم التقديرية لصعوبة الفقرات بين(٢.٧٠) لوجيت و (٢.٣٠) لوجيت بمتوسط مقداره (صفر)، وانحراف معياري مقداره (١.٢٣) لوجيت.

وقام هوانج (Hwang,2002) بدراسة هدفت إلى المقارنة بين النظرية الكلاسيكية ونظرية الاستجابة للمفردة بنماذجها الثلاثة من حيث تقديرات مؤشرات صعوبة الفقرات، وتمييزها، ومستويات القدرة للطلاب، ومن حيث استقرار تقديرات الفقرات الاختبارية باختلاف مستويات قدرة الطلاب. وقد بلغ حجم عينة الدراسة (٦٠٠) طالب من طلاب الصف الثامن، وقد استخدم اختبار تحصيلي في الرياضيات والذي يشمل على (١٥) فقرة. وقد توصلت نتائج هذه الدراسة إلى أن معاملات الارتباط المرتفعة (معظمها أكبر من ٠.٩٠) بين المؤشرات الإحصائية المستمدة من النظريتين تشير إلى تشابه كبير بينهما. كما كان هناك تشابه واضح بين النظريتين في تقدير مؤشرات الصعوبة للفقرات، إلا أن هذا التشابه يقل فيما يتعلق بتقدير مؤشرات التمييز للفقرات.

وأجرى ياسين (٢٠٠٤) بدراسة هدفت إلى معرفة الخصائص السيكومترية لاختبار محكي المرجع في الكيمياء لطلاب الصف الأول الثانوي مقدرة وفقاً للنظريتين الكلاسيكية والحديثة في القياس. وقد تكون الاختبار من (٥٢) فقرة من نوع الاختيار من متعدد من أربعة بدائل. وتكون مجتمع الدراسة من طلاب وطالبات الصف الاول الثانوي العلمي للعام

الدراسي ٢٠٠٤م في الاردن، وكان عددهم (٢٤٦٢) طالباً وطالبة. وتكونت عينة الدراسة من (٤٨١) طالباً وطالبة، تم اختيارهم بالطريقة العشوائية العنقودية.

وأظهرت نتائج الدراسة مطابقة (٤٣) فقرة وفقاً للنموذج أحادي المعلم، ومطابقة (٤٨) فقرة وفقاً للنموذج ثلاثي المعلم. وبلغ معامل الصدق المرتبط بمحك(٠.٨٤)، ومعامل الاتساق الداخلي بطريقة الفا كرونباخ (٠.٩٠)، كما تم تقدير معامل الثبات للاختبار وفقاً للنظرية الحديثة في القياس حيث بلغت قيمته (٠.٩٩). وأظهرت النتائج الخاصة بالنظرية الحديثة في القياس أن متوسط صعوبة الفقرات وفق النموذج اللوجستي أحادي المعلم صفر لوجيت، ومتوسط الخطأ المعياري (٠.١١٩)، ومتوسط تمييز الفقرات (٠.٧٧)، ومتوسط الخطأ المعياري لمعاملات التمييز بلغ (٠.١٢٥)، في حين بلغ متوسط تخمين الفقرات (٠.٢٤).

وقام العيثاوي (٢٠٠٩) بدراسة هدفت إلى استعمال نموذج راش وفق نظرية السمات الكامنة في بناء اختبار تحصيلي في مادة علم نفس الفروق الفردية. ولتحقيق هدف الدراسة تم بناء اختبار مستنداً إلى المفردات المقررة من قبل الهيئة القطاعية لوزارة التعليم العالي العراقية. وقد أعد الباحث (١٧٣) فقرة من نوع الاختيار المتعدد تغطي مستويات الاهداف المعرفية الثلاث الاولى بحسب تصنيف بلوم وهي (المعرفة، الفهم، والتطبيق). واشتملت عينة الدراسة على (٣٣٦) طالباً وطالبة من طلبة كلية التربية للجامعات العراقية جرى اختيارهم بالطريقة العشوائية الطبقية. وكانت أهم النتائج أن الاختبار أصبح في صورته النهائية (١٠٨) فقرة مطابقة لافتراضات النموذج اللوجستي المستخدم ومتطلباته، حيث جرى حذف (٢٨) فقرة لا تحقق أحادية البعد من خلال التحليل العاملي لفقرات الاختبار بطريقة المكونات الاساسية، فيما حذفت (٢٩) فقرة باستخدام دلالة قيم مربع كاي (غير مطابقة للنموذج المستخدم)، وقد اقتربت قيمة معامل التمييز من (١) صحيح حيث كانت القيمة ضمن المدى المقبول (٠.٦٢٩)، كما تم استبعاد (٨) فقرات لا تحقق استقلالية القياس. وأخيراً بلغت قيمة معامل الثبات للاختبار بصورته النهائية (٠.٩٥٢)، ومعامل تمييزه (٠.٧١٣).

وأجرى السامرائي والخفاجي (٢٠١٢) دراسة هدفت لبناء اختبار تحصيلي محكي المرجع في مادة علم نفس الخواص لطلبة أقسام العلوم التربوية والنفسية لكليات التربية في الجامعات العراقية، وتألف الاختبار من (١٦٠) فقرة اختبارية من نوع الاختيار من متعدد، وذلك لقياس (١٦٠) هدفاً سلوكياً تغطي مفردات المادة. وقد جرى تطبيق الاختبار على عينة عشوائية طبقية بلغت (٣٤٩) طالباً وطالبة وقد اعتمد الباحث على نموذج راش وهو أحد نماذج نظرية السمات الكامنة في تحليل بيانات الاختبار. وقد أخضعت بيانات فقرات الاختبار إلى التحليل العاملي بطريقة المكونات الاساسية الذي أظهر أن فقرات الاختبار تقيس بعداً واحداً

عدا (٢٢) فقرة تم استبعادها من الاختبار. وأظهرت نتائج الدراسة وجود (١٣) فقرة كانت قيمة مربع كاي لعشرة فقرات منها دالة إحصائياً لذا تم استبعادها من الاختبار، فيما استبعد برنامج راسكال ثلاث فقرات لعدم مطابقتها لنموذج راش. وقد اقترب معامل تمييز الاختبار من (١) إذ كانت قيمة معامل تمييز الاختبار ضمن المدى المقبول والتي بلغت (١ــ ١.٦٠)، وقد بلغت قيمة ثبات الاختبار (٠.٩٥٧).

ومن الدراسات التي قارنت بين النظريتين الكلاسيكية والحديثة في القياس، دراسة ثيسن وزيموسكي (Thissen & Zimowski,1997) التي هدفت إلى مقارنة دقة نظرية الاستجابة للمفردة، مع دقة النظرية الكلاسيكية في القياس في تقديرات علامة المجال لاختبار متعدد الأبعاد (في كتابة الكلمات) في الفيزياء والكيمياء والأحياء. حيث كانت الكلمة تقرأ ثم تستعمل في جملة، ثم تقرأ ثانية، ويعطى هذا المفحوص ١٠ ثوان لكتابتها. تكونت عينة الدراسة من تسجيل استجابات ١٠٠٠ طالب من طلبة الدراسات العليا في جامعة تكساس في الولايات المتحدة الامريكية. دلت نتائج الدراسة على تفوق نظرية الاستجابة للمفردة على النظرية الكلاسيكية في الدقة في تقدير علامة المجال السلوكي.

كذلك أجرى فان (Fan, 1998) دراسة هدفت إلى مقارنة الخصائص السيكومترية لفقرات اختبار محكي المرجع، وفق النظريتين الكلاسيكية والحديثة في القياس. وتكون الاختبار من اختبارين فرعيين: أحدهما في القراءة والآخر في الرياضيات، وتألف اختبار القراءة من ٤٨ فقرة، واختبار الرياضيات من ٦٠ فقرة. تكونت عينة الدراسة من ٨٠ عينة فرعية، سحبت عشوائياً من مجتمع الدراسة المكون من أكثر من (١٩٣٠٠٠) طالب في الصف الحادي عشر في ولاية تكساس. أظهرت نتائج الدراسة أن معاملات الارتباط بين تقديرات صعوبة الفقرة في النظريتين عالية، وكانت أفضل الارتباطات بين النظرية الكلاسيكية، والنموذج أحادي المعلم، وكذلك معاملات الارتباط بين تقديرات تمييز الفقرة في النظريتين عالية.

أما دراسة واغنر وهارفي (Wagner & Harvey, 2003) فقد هدفت إلى تطوير أداة جديدة لأوجه التفكير الناقد لاختبار واتسون ــ جيلسر باستخدام نظرية الاستجابة للمفردة. حيث قام الباحثان بزيادة عدد البدائل، لتكوين اختبار جديد. تكونت عينة الدراسة من ٤٠٧ طالب وطالبة في المستوى الجامعي. وقد دلت نتائج التحليل على نسبة تخمين أقل، وتحسين في دالة معلومات الفقرة للاختبار المطور، وكانت أخطاء القياس في اختبار واتسون ـــ جليسر أكبر من أخطاء القياس للاختبار الجديد بمقدار ٥٠% على الأقل.

التعقيب على الدراسات السابقة:

يلاحظ من خلال العرض السابق للدراسات السابقة أنها تناولت العديد من الجوانب التي بحثت المقارنة بين النظريتين الكلاسيكية والحديثة في القياس، وتقدير معالم الفقرات باستخدام النماذج اللوجستية المختلفة في نظرية الاستجابة للمفردة، وأحياناً المقارنة بين النماذج

اللوجستية الثلاث في نظرية الاستجابة للمفردة، ولكن لم تركز الدراسات السابقة على بناء اختبار محكي المرجع خاص بطلبة الدراسات العليا في مقرر مناهج البحث العلمي باستخدام أحد النماذج اللوجستية المناسبة؛ لذلك انفردت هذه الدراسة ببناء اختبار محكي المرجع لطلبة الدراسات العليا في مادة مناهج البحث العلمي باستخدام النموذج اللوجستي ثنائي المعلم أحادي البعد وذلك باستخدام البرنامج الإحصائي Bilog-MG3 في الحصول على العديد من المؤشرات الإحصائية الخاصة بنظرية الاستجابة للمفردة. بالإضافة إلى استخدام البرنامج الإحصائي SPSS للحصول على بعض المؤشرات الإحصائية الأخرى الخاصة بالنظرية الكلاسيكية للقياس.

مشكلة البحث وتساؤلاته :

رغم الاهتمام الكبير بنظرية القياس التقليدية والاعتماد عليها لفترة طويلة من الزمن، والتي ما زالت مستمرة إلى يومنا هذا؛ إلا أنها تعرضت للكثير من الانتقادات التي تتهمها بالاعتماد على الدرجة الحقيقية وإهمال تحقيق الموضوعية في القياس. لذلك ظهرت الاختبارات معيارية المرجع، ومن ثم الاختبارات محكية المرجع، إلا أنها لم تحقق الدقة والموضوعية في القياس، مما دفع الباحثين والمهتمين في مجال القياس للبحث والتقصي إلى أن ظهرت النظرية الحديثة في القياس والتي أطلق عليها نظرية الاستجابة للمفردة، وتتميز هذه النظرية بقدرتها على تحقيق الموضوعية والدقة في القياس لذلك استخدمت في تصميم مختلف أنواع الاختبارات والمقاييس بأنواعها المختلفة.

إن التعليم الجامعي في الوطن العربي بحاجة إلى توافر اختبارات تحصيلية محكية المرجع تتمتع بدرجة مقبولة من الصدق والثبات يمكن من خلالها قياس تحصيل الطلبة والتحقق من مدى تمكن أو إتقان الطلبة للمادة الدراسية. كما يمكن أن تمثل هذه الاختبارات أدوات مناسبة لتشخيص ومعرفة الأهداف التربوية التي تحققت والتي لم تتحقق واقتراح الأسباب التدريسية والتدريبية في ضوء نتائج عملية التشخيص. وبناء على ما تقدم تتحدد مشكلة البحث الحالي في عدم وجود اختبار تحصيلي محكي المرجع في مادة مناهج البحث العلمي للدراسات العليا مبني وفقاً لنظرية الاستجابة للمفردة، ونماذج القياس المنبثقة عنها، لما يتمتع به هذا الاختبار من دقة في القياس من خلال التحرر من أثر قدرة الفرد على معالم فقرات الاختبار، والتحرر من أثر معالم الفقرات على قدرة الفرد الذي يجيب عنها. وبشكل أكثر تحديداً تحاول هذه الدراسة الإجابة عن الأسئلة التالية:

السؤال الأول " ما أنسب نموذج (الأحادي أو الثنائي أو الثلاثي المعلم) من نماذج نظرية الاستجابة للمفردة ثنائية التصحيح ؟ "

السؤال الثاني "ما مدى تحقق افتراضات نظرية الاستجابة للمفردة في بيانات الدراسة وفقاً للنموذج اللوجستي الأنسب ؟ "

السؤال الثالث " ما الخصائص السيكومترية لفقرات اختبار مناهج البحث العلمي من حيث

دلالات الصعوبة والتمييز حسب النظرية الكلاسيكية في القياس ؟ "

السؤال الرابع "ما درجة مطابقة استجابات الأفراد لتوقعات النموذج اللوجستي الأنسب لنظرية الاستجابة للمفردة؟"

السؤال الخامس "ما الخصائص السيكومترية لاختبار مناهج البحث حسب النموذج اللوجستي الأنسب في نظرية الاستجابة للمفردة ؟"

السؤال السادس "ما الخصائص الإحصائية لتقديرات قدرات الطلاب والاخطاء المعيارية في تقديرها؟"

أهداف الدراسة:

هدفت هذه الدراسة إلى استخدام نظرية الاستجابة للمفردة والنظرية الكلاسيكية في القياس في بناء اختبار محكي المرجع في مادة مناهج البحث العلمي بحيث يتمتع هذا الاختبار بدلالات صدق وثبات مقبولة، وبمعايير تتفق مع متطلبات النموذج اللوجستي الأحادي أو الثنائي أو الثلاثي المعلم، من أجل قياس تحصيل طلبة الدراسات العليا في مادة مناهج البحث العلمي.

أهمية البحث:

تتضح أهمية هذه الدراسة من خلال بناء الاختبارات محكية المرجع وذلك لكونها تسعى إلى تقويم المتعلمين تقويماً تشخيصياً يحدد جوانب القوة والتميز، ومواطن الضعف للتغلب عليها من أجل تحسين نواتج البرامج التعليمية وتحقيق الأهداف التعليمية والتدريسية المرجوة منها لدى المتعلمين (عبد السلام وآخرون، ١٩٩٦؛ علام، ٢٠٠١).

ومن المسلم به إنه كلما تقدمت أساليب القياس ودقة التقدير الكمي في علم من العلوم كان ذلك مظهراً من مظاهر رقي ذلك العلم وتقدمه وارتفاع مكانته بين العلوم الأخرى، فعملية تكميم الظواهر وتفسيرها خاصية من خصائص البحث العلمي الجيد (بركات، ١٩٨٣). لذا فقد ركز أصحاب الاتجاه المعاصر في القياس والتقويم جهودهم للتوصل إلى أعلى مستويات الدقة، والموضوعية في القياس بحيث تتحقق أدق علاقة بين أداة القياس والسمة المراد قياسها (Nunnally, 1978). وبدأ اهتمامهم ينصب نحو الاختبارات محكية المرجع، حيث ارتبط هذا النوع من الاختبارات بمدخل من مداخل القياس أطلق عليه نماذج السمات الكامنة (Hulin, et al, 1983).

وبناء عليه يكتسب البحث الحالي أهميته بأنه يهدف إلى بناء اختبار تحصيلي محكي المرجع وفقاً لنظرية الاستجابة للمفردة، وتوفير أداة تمكن من قياس تحصيل الطلبة في مادة مناهج البحث العلمي بمستوى عال من الدقة والموضوعية، وبما يحقق أدق علاقة بين هذه الأداة وقدرة التحصيل التي تكمن وراء استجابة الفرد على كل فقرة من فقرات الاختبار.

محددات الدراسة:

١. اقتصرت الدراسة الحالية على مادة مناهج البحث العلمي التي تدرس في جامعة نايف العربية للعلوم الأمنية.

٢.تم جمع البيانات في نهاية الفصل الدراسي الأول للعام ٢٠١٤/٢٠١٥م

٣.استخدام أحد النماذج اللوجستية (النموذج اللوجستي ثنائي المعلم) في نظرية الاستجابة للمفردة.

٤.محددات في أداة البحث: يمكن تعميم نتائج هذه الدراسة بالقدر الذي تتوافر فيه لهذه الأداة دلالات الفاعلية (حسب نظريتي القياس الكلاسيكية والحديثة).

مصطلحات الدراسة :

الاختبار:" يعني أداة أو وسيلة أو طريقة تقدم للمفحوص سلسلة من المهمات عليه أن يستجيب لها، بحيث أن طريقة استجابته تدل على مقدار ما عنده في تلك الصفحة"(عبابنة، ٢٠٠٩، ٢٦).

الاختبار محكي المرجع:" الاختبار الذي يرد إلى محك هو ذلك الاختبار الذي يذكر شيئاً عن الفرد بدون الرجوع إلى أداء فرد آخر" (عبدالسلام وآخرون، ١٩٩٦، ٢٧).

الاختبار مرجعي المحك هو " ذلك الاختبار الذي يستخدم في تقدير أداء الفرد في نطاق سلوكي محدد تحديداً دقيقاً " (علام، ٢٠٠١، ٢٤).

ويعرف إجرائياً في هذا البحث على أنه " مجموعة فقرات من نوع الاختيار المتعدد بواقع أربعة بدائل (أحدهما صحيح والثلاثة الباقية خاطئة) أعدت وفق افتراضات نظرية السمات الكامنة وتحليل نتائج استجابات الطلبة باستخدام أحد النماذج اللوجستية بما يحقق دقة وموضوعية قياس تحصيل طلبة الدراسات العليا في مادة مناهج البحث العلمي".

مادة مناهج البحث العلمي: " هي مادة دراسية مقررة لطلبة الدراسات العليا في جامعة نايف العربية للعلوم الأمنية".

النظرية التقليدية(الكلاسيكية):" هي واحدة من نظريات القياس التي تستخدم بغرض تحديد العوامل التي تؤثر على الدرجة التي يحصل عليها الفرد في الاختبار، وترتكز هذه النظرية على مفهوم الدرجة الحقيقية والدرجة الخطأ، والذى يفترض أنه لو أمكن أن نجري الاختبار عدة مرات على الفرد بعناصر جديدة وتحت ظروف مختلفة ، فإننا نحصل على درجات ملحوظة مختلفة متوسطها هو أقرب تقدير غير متحيز لقدرة الفرد أو درجته الحقيقية" **(Randall,1998:4)** .

نظرية السمات الكامنة:" هي مجموعة من الطرق الإحصائية التي تستخدم في حساب معالم الفقرات والأفراد الخاصة بالاختبار من صعوبة وتمييز وتخمين ودالة معلومات، والمكونة من نماذج أحادية البعد هي (النموذج أحادي المعلم وثنائي المعلم وثلاثي المعلم)".(Linden & Hamblton, 1997)

منهجية الدراسة وإجراءاتها :

استخدم الباحث المنهج الوصفي التحليلي في دراسته، من خلال التحقق تجريبياً من مدى تحقق افتراضات النظرية الحديثة في القياس، وكذلك تم تحليل فقرات الاختبار باستخدام النموذج اللوجستي ثنائي المعلم، واستخدم الباحث العديد من المؤشرات الإحصائية للتحقق من صلاحية الاختبار للتطبيق، بالإضافة إلى إجراء

مقارنة بين النظرية الحديثة والنظرية الكلاسيكية في القياس.

مجتمع الدراسة:

تكون مجتمع الدراسة من جميع طلاب الماجستير والدبلوم العالي سنة اولى والذين درسوا مقرر مناهج البحث العلمي خلال العام الدراسي للفصل الدراسي الأول ٢٠١٤/٢٠١٥م في جامعة نايف العربية للعلوم الأمنية بالرياض، وقد بلغ عدد الطلاب (٥٢٣) طالب.

عينة الدراسة:

تكونت عينة الدراسة من (٢٥٣) طالب، تم اختيارهم بصورة عشوائية من مجتمع الدراسة، لذلك كانت العينة عشوائية بسيطة وشكلت ما نسبته (٤٨%) من مجتمع الدراسة.

أداة الدراسة:

ولتحقيق أهداف الدراسة قام الباحث بإعداد اختبار محكي المرجع في مادة مناهج البحث العلمي للدراسات العليا، وقد اقتضى بناء أسئلة الاختبار ووضعه في صيغته النهائية الاسترشاد بالأسس المتبعة ببناء اختبار محكي المرجع، وفيما يلي الإجراءات التي اتبعها الباحث.

تم تحديد محتوى مادة مناهج البحث العلمي من خلال حصر جميع مفردات المادة، وبعد الاطلاع على الأهداف العامة للمادة بدأ الباحث في صياغة الأهداف السلوكية لفصول المادة وبحسب المستويات المعرفية الأولى ضمن تصنيف بلوم للأهداف التعليمية في البعد المعرفي، وبعد صياغة الاهداف السلوكية تم إعداد جدول للمواصفات ببعديه (المحتوى، ومستويات الاهداف في المجال المعرفي) للتعرف على طبيعة توزيع الاهداف بحسب الاهمية النسبية لكل فصل وفقاً لعدد الصفحات الفعلية التي تحتوي على فقرات كل فصل والوزن النسبي لكل مستوى. ثم قام الباحث بصياغة (٥٠) فقرة من نوع الاختيار من متعدد بأربعة بدائل واحدة منها يمثل الإجابة الصحيحة. وقد روعيت الشروط الواجب توفرها في مثل هذا النوع من الفقرات، وحرص الباحث على أن تقيس كل فقرة هدفاً محدداً وفقاً لقائمة الاهداف الخاصة بكل فصل دراسي، إذ تم تحليلها تحليلاً مفصلاً، ثم عرضت الصورة الأولية للاختبار على مجموعة من المحكمين والمختصين في القياس والتقويم؛ وذلك لأخذ آرائهم في مدى صدق الفقرة في قياس الهدف السلوكي المحدد، ومدى مناسبة الفقرة.

صدق الاختبار

تم التأكد من صدق الاختبار باستقصاء المؤشرات التالية:

أولاً: صدق المحتوى: تم التأكد من هذا المؤشر من خلال عرض نتائج تحليل الاهداف والمحتوى ببناء جدول المواصفات.

ثانياً: الصدق الظاهري: للتحقق من صلاحية الفقرات وصدقها في قياس الاهداف السلوكية عرض الباحث فقرات الاختبار على مجموعة من المحكمين المختصين حيث بلغ عددهم (٥) حسب ملحق (١). وقد تبين من خلال آراء المحكمين صلاحية أغلب الفقرات وصدقها في قياس ما وضعت لقياسه فيما عدا بعض الفقرات

وعددها (٦) حيث تم حذفها؛ ليبلغ عدد فقرات الاختبار (٤٤) فقرة، بعد أن كان عدده (٥٠) فقرة.

ثالثاً: الصدق العاملي: تم التحقق من الصدق العاملي باستخدام تحليل المكونات الاساسية (Principal Component Analysis)، وحساب قيم الجذر الكامن(Eigen Value)، وملاحظة نسبة التباين المفسر (Explained Variance) لكل عامل من العوامل.

ثبات الاختبار

تم تقدير ثبات الاختبار الامبريقي /التجريبي (Empirical Reliability) بصورته النهائية وفقاً للنموذج ثنائي المعلم باستخدام Bilog-MG3، وكذلك تم تقدير ثبات الاتساق الداخلي (كودر رتشاردسون KR-20) للاختبار باستخدام برنامج الحزم الاحصائية للعلوم الاجتماعية (SPSS) كما هو مبين بالجدول (١).

الجدول (١)

معاملات الثبات التجريبي والاتساق الداخلي وفق النموذج اللوجستي ثنائي المعلم لنظرية الاستجابة للمفردة

ثبات الاختبار	
التجريبي/ الامبريقي	الاتساق الداخلي
٠.٨٣٥٢	٠.٨٦٤

يتبين من الجدول رقم أن هناك شبه تطابق بين قيمتي معامل الثبات التجريبي ومعامل الاتساق الداخلي للاختبار، وتشير هذه القيم إلى أن الاختبار يتمتع بدرجة مقبولة من الثبات.

الإجراءات

تم تطبيق الاختبار في صورته الأولية (٤٤) فقرة على عينة استطلاعية من خارج عينة الدراسة بلغ عددها (٥٦) طالب وقد انحصر الغرض من التطبيق على: تحديد الزمن الذي سوف يستغرقه الطالب للإجابة على الفقرات، التعرف على مستوى صعوبة وتمييز الفقرات، ومدى وضوح الصياغة اللغوية للفقرات وفاعلية المموهات. وبعد دراسة إجابات العينة الاستطلاعية ودراسة معاملات الصعوبة والتمييز، تم حذف الفقرات التي كان تمييزها أقل من (٠.٢٠)، وبلغ عدد الفقرات المحذوفة (٤) فقرات، وبذلك أصبح عدد فقرات الاختبار بعد تطبيق العينة الاستطلاعية (٤٠) فقرة.

جمع البيانات:

بعد إعداد الاختبار وتطبيقه على عينة استطلاعية كان عدد فقراته (٤٠)، جرى تطبيقه على أفراد عينة الدراسة والبالغ عددهم (٢٥٣) طالب بشكل جماعي داخل القاعات التدريسية، وقد تم تقدير الوقت اللازم للإجابة على الاختبار بـــــساعة وربع، ثم قام الباحث بتصحيح فقرات الاختبار، ومن ثم تم إدخال البيانات في ذاكرة الحاسوب، واستخدم البرنامج الإحصائي (SPSS)، وبرنامج (Bilog-MG3) لإجراء التحليلات الإحصائية اللازمة للإجابة عن تساؤلات الدراسة.

المعالجات الإحصائية:

للإجابة عن اسئلة الدراسة تم التحقق من افتراض احادية البعد للاختبار من خلال اجراء

التحليل العاملي بطريقة المكونات الاساسية باستخدام برنامج(SPSS)، واستخدم برنامج (Bilog-MG3) لتقدير معالم الفقرات (الصعوبة، التمييز) باستخدام طريقة الارجحية العظمى الهامشية (MML).

نتائج الدراسة ومناقشتها :

أولاً : النتائج المتعلقة بالسؤال الأول

نص السؤال الأول " ما أنسب نموذج (الأحادي أو الثنائي أو الثلاثي المعلم) من نماذج نظرية الاستجابة للمفردة ثنائية التصحيح ؟ "

تم تحديد النموذج الأنسب من خلال اختبار مطابقة النموذج model fit. ولأن الاختبار المعد في هذه الدراسة من نوع الاختيار من متعدد من أربع بدائل، فمن المتوقع أن يتضمن الاختبار معلم التخمين، بمعنى آخر فإن النموذج الأنسب هو النموذج ثلاثي المعلم. إلا أن تضمين النموذج معالم إضافية سيقلل من دقة التقدير لأنه سيتطلب تقدير معالم ليست بحاجة إلى تقدير(Jiao & Lau, 2003). ولأن النموذج الأحادي متضمن في النموذج الثاني كونه يتضمن معلماً إضافياً عن النموذج الأحادي (تمييز الفقرة)، والنموذج الثنائي متضمن في النموذج الثلاثي لأنه يتضمن معلماً إضافياً عن النموذج الثنائي (التخمين) (Zimowski, Muraki, Mislevy & Bock, 2003)، فقد تم اختبار الفرضيتين الصفريتين الآتيتين:

الفرضية الأولى: تضمين النموذج الأحادي معالم إضافية لا يؤدي إلى تحسن دال إحصائياً عند مستوى دلالة (α=٠.٠٠٥) في مطابقة هذا النموذج للاستجابات عن الفقرات.

الفرضية الثانية: تضمين النموذج الثنائي معالم إضافية لا يؤدي إلى تحسن دال إحصائياً عند مستوى دلالة (α=٠.٠٠٥) في مطابقة هذا النموذج للاستجابات عن الفقرات.

ولاختبار كل من الفرضيتين فقد استخدم الباحث برمجية Bilog-MG3 وحساب الإحصائي التالي ((the negative of 2 loglikelihood (-2LLH) لكل من النماذج الثلاث ولنفس فقرات الاختبار، ومن ثم إيجاد الفرق بين قيمتي(-2LLH) لكل من أزواج النماذج الثلاثي والثنائي، وكذلك الثنائي والأحادي. ولأن الفرق الموجب بين هذه القيم له توزيع كاي تربيع بدرجات حرية تساوي الفرق بين عدد معالم النموذجين (Zimowski et al.,2003)، فقد تم حساب القيمة الناتجة من الفرق بين قيمتي (-2LLH) للنموذجين الأحادي الثنائي المعلم وكانت تساوي (٨٧.٥٣). كما تم حساب الفرق بين قيمتي (-2LLH) للنموذجين الثنائي والثلاثي المعلم وكانت تساوي (١٣.٢٢). ومن ثم تم مقارنة كل من القيمتين مع قيمة مربع كاي بأربعين درجة حرية ومستوى دلالة ٠.٠٥ والتي تساوي (٣٥.١٧)، ووفقاً لقيمة مربع كاي فقد تم رفض الفرضية الصفرية الأولى وقبول الثانية. أي أن تضمين النموذج الأحادي معالم إضافية يؤدي إلى تحسين الدالة إحصائياً في مطابقة النموذج للاستجابات عن فقرات الاختبار، في حين تضمين النموذج الثنائي معالم إضافية لم يؤد إلى تحسين الدالة في

مطابقة النموذج للاستجابات على فقرات الاختبار. وهذا يدل إلى أن النموذج ثنائي المعلم هو النموذج الأنسب للاستجابات على فقرات الاختبار.

ثانياً: النتائج المتعلقة بالسؤال الثاني

نص السؤال "ما مدى تحقق افتراضات نظرية الاستجابة للمفردة في بيانات الدراسة وفقاً للنموذج اللوجستي ثنائي المعلم ؟ "

أ-افتراض أحادية البعد Unidimensionality

تم التحقق من الافتراض باستخدام المؤشرات التالية:

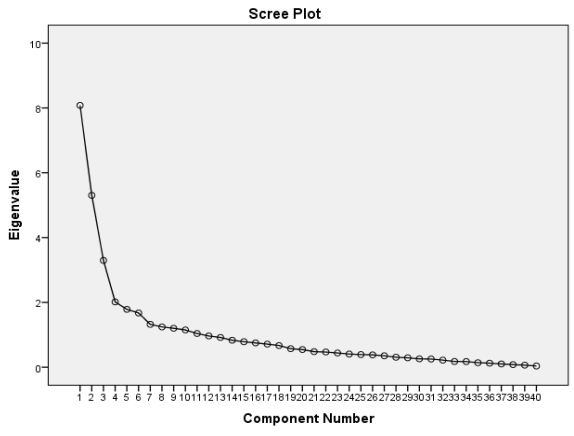
١. تم إجراء تحليل عاملي لبيانات الاختبار النهائي المتعلقة باستجابات (٢٥٣) طالب عن (٤٠) فقرة تمثل الاختبار النهائي باستخدام برنامج (SPSS)، وأفرز التحليل العاملي (١١) عامل قيمة الجذر الكامن لكل منها تزيد عن واحد، حيث فسر العامل الأول منها(٢٠.١٩٨%) من التباين، وفسرت جميع العوامل (٧٠.٢٦٨%) الكلي، ويبين الجدول (٢) قيم الجذور الكامنة، ونسبة التباين المفسر لكل عامل، وكذلك نسبة التباين المفسر الكلية.

رقم العامل	الجذر الكامن	نسبة التباين المفسر %	نسبة التباين المفسر التراكمية %
1	8.079	20.198	20.198
2	5.303	13.257	33.455
3	3.295	8.238	41.692
4	2.014	5.035	46.727
5	1.785	4.462	51.190
6	1.674	4.185	55.375
7	1.323	3.307	58.682
8	1.244	3.109	61.791
9	1.201	3.004	64.794
10	1.149	2.873	67.668
11	1.040	2.601	70.268

ويلاحظ من نتائج الجدول (٢) أن قيمة الجذر الكامن للعامل الاول (٨.٠٧٩) ويفسر ما نسبته (٢٠.١٩٨%) من التباين الكلي، ويعد الاختبار أحادي البعد إذا كان نسبة ما يفسره العامل الأول أكثر من (٢٠%) (Reckase, 1979: 207-230)، وبهذا المعيار يتبين أن الاختبار يحقق افتراض أحادية البعد.

٢.تمثيل الجذور الكامنة بيانياً باستخدام ما يعرف باختبار فحص العوامل (Scree Plot)

نلاحظ في الشكل (١) أن الجذر الكامن للعامل الأول يظهر بشكل أكبر من الجذور الكامنة لبقية العوامل، وباستخدام هذا الشكل فإن القرار حول عدد العوامل وفقاً لوهلن (Loehlin, 1987) يتم التوصل إليه اعتماداً على النقطة التي يتغير فيها ميل منحنى الجذور الكامنة بسرعة من منحنى يعامد تقريباً محور السينات إلى منحى أفقي تقريباً، وهذا يحصل عند العامل الأول، بمعنى أنه يمكن افتراض تحقق أحادية البعد.



الشكل (١)

الجذور الكامنة لعوامل الاختبار

٣.ثبات الاتساق الداخلي

وأظهرت القيمة العالية لمعامل ثبات الاتساق الداخلي (KR-20) والذي بلغت قيمته (٠.٨٦) لنموذج الاختبار مؤشراً على تحقق افتراض أحادية البعد وفقاً لرأي كرونباخ (Cronbach,1951)، إذ يرى أن معامل (KR-20) يعد مؤشراً جيداً للتحقق من أحادية البعد؛ لأنه يمثل القيمة المتوقعة لنسبة التباين المفسر من العوامل المشتركة بين الفقرات عند ارتباط عينتين عشوائيتين من تجمع للفقرات، فقيمة هذا العامل المرتفعة تقدم مؤشراً جيداً للتحقق من افتراض أحادية البعد للاختبار.

ب-افتراض الاستقلال الموضعي Local Independence.

اكتفى الباحث بالتحقق من افتراض أحادية البعد للاستدلال على تحقق افتراض الاستقلال الموضعي؛ نظراً لأن افتراض أحادية البعد يكافئ افتراض الاستقلال الموضعي(Hambleton and Swaminathan,1985, Hulin, Drasgow and Parson,1983).

ج-افتراض خاصية المنحنى المميز للمفردة Item Characteristic Curve (ICC).

إن احتمال إجابة الطالب لفقرة على الاختبار إجابة صحيحة يزداد بزيادة قدرة الطالب.

د-افتراض التحرر من السرعة Non Speededness

تأكد الباحث أن إخفاق الطلاب في الإجابة عن فقرات الاختبار يعود إلى انخفاض قدراتهم، وليس إلى تأثير عامل السرعة في إجاباتهم وذلك عن طريق إعطاء الوقت الكافي لهم للإجابة عن فقرات الاختبار ، بالإضافة إلى أنه لم يشتكي أي طالب من ضيق الوقت أثناء الإجابة على الاختبار.

ثالثاً: النتائج المتعلقة بالسؤال الثالث:

نص السؤال " ما الخصائص السيكومترية لفقرات اختبار مناهج البحث العلمي من حيث دلالات الصعوبة والتمييز حسب النظرية الكلاسيكية في القياس ؟ "

تم استخراج بيانات صعوبة الفقرات وتمييزها بحساب نسبة الإجابة الصحيحة لكل فقرة للدلالة على صعوبتها ومعامل الارتباط الثنائي بين الفقرة والعلامة الكلية للدلالة على تمييز الفقرة ويبين الجدول (٣) هذه النتائج.

جدول (٣)

بيانات الصعوبة والتمييز لفقرات اختبار مناهج البحث العلمي

رقم الفقرة	صعوبة الفقرة	تمييز الفقرة	رقم الفقرة	صعوبة الفقرة	تمييز الفقرة
١	.78	.122	٢١	.82	.248**
٢	.77	.266**	٢٢	.90	.596**
٣	.95	.420**	٢٣	.60	.370**
٤	.90	.441**	٢٤	.87	.649**
٥	.96	.288**	٢٥	.75	.531**
٦	.81	.376**	٢٦	.89	.619**
٧	.75	.203**	٢٧	.89	.535**
٨	.90	.306**	٢٨	.90	.378**
٩	.70	.697**	٢٩	.90	.321**
١٠	.83	.593**	٣٠	.87	.436**
١١	.87	.311**	٣١	.94	.459**
١٢	.93	.440**	٣٢	.52	.394**
١٣	.90	.249**	٣٣	.86	.617**
١٤	.95	.317**	٣٤	.83	.180**
١٥	.72	.373**	٣٥	.72	.511**
١٦	.81	.424**	٣٦	.90	.280**
١٧	.65	.580**	٣٧	.55	.472**
١٨	.87	.598**	٣٨	.25	.417**
١٩	.76	.350**	٣٩	.66	.382**
٢٠	.89	.320**	٤٠	.66	.380**

يتبين من الجدول أن قيم الصعوبة كنسب للإجابة الصحيحة للفقرة تراوحت ما بين

(٠.٢٥) و (٠.٩٦) وبمتوسط (٠.٧٩) وقيم التمييز تراوحت بين (٠.١٢٢) و (٠.٦٩٧) وبمتوسط (٠.٤١) وجميع هذه القيم ذات دلالة إحصائية عند مستوى دلالة (α=٠.٠٥) وتعد هذه القيم في مستوى مقبول لأغراض بناء اختبار محكي المرجع في هذه الدراسة حسب النظرية الكلاسيكية في القياس، ما عدا العبارتين (١) و (٣٤) حيث يتبين من الجدول السابق أن العبارة رقم (١) كانت غير دالة إحصائياً وكان تمييزها أقل من (٠.٢٠)، والعبارة رقم (٣٤) كانت دالة إحصائياً ولكن كان تمييزها أقل من (٠.٢٠) مما يعني وجوب حذف هاتين الفقرتين حسب مفهوم النظرية الكلاسيكية في القياس. ليصبح الاختبار في صورته النهائية (٣٨) فقرة حسب النظرية الكلاسيكية.

رابعاً: النتائج المتعلقة بالسؤال الرابع :

نص السؤال "ما درجة مطابقة استجابات الأفراد والفقرات لتوقعات النموذج اللوجستي ثنائي المعلم لنظرية الاستجابة للمفردة؟"

جرى فحص حسن مطابقة البيانات حسب النموذج اللوجستي ثنائي المعلم؛ وذلك لاختيار الطلاب المطابقين للنموذج الثنائي، وكذلك اختيار الفقرات المطابقة لهذا النموذج، وقد أفرزت نتائج التحليل الأولي باستخدام اختبار مربع كاي عند مستوى دلالة (α=٠.٠١)، مطابقة استجابات جميع الطلاب للنموذج اللوجستي ثنائي المعلم. وبعد ذلك قام الباحث بإعادة التحليل لاختبار مدى مطابقة الفقرات للنموذج اللوجستي الثنائي، حيث أظهرت نتائج التحليل في المرة الثانية الخاصة بمدى مطابقة الفقرات للنموذج اللوجستي ثنائي المعلم من خلال اختبار مربع كاي عند مستوى دلالة (α=٠.٠١) عدم مطابقة (١٣) فقرة للنموذج، حيث كانت قيمة احتمالية المطابقة لكل منها أقل من ٠.٠١ كما يبينه الجدول (٤).

الجدول (٤): نتائج فحص مطابقة البيانات حسب النموذج اللوجستي ثنائي المعلم

رقم الفقرة	مستوى دلالة اختبار مربع كاي	رقم الفقرة	مستوى دلالة اختبار مربع كاي	رقم الفقرة	مستوى دلالة اختبار مربع كاي	رقم الفقرة	مستوى دلالة اختبار مربع كاي
١	٠.٠٠٠٠	١١	٠.٠٠٠٠	٢١	٠.٠٠٠٠	٣١	٠.٧٧٢٠
٢	٠.٠٠٣٧	١٢	٠.٣٥٠٦	٢٢	٠.١٩٨٥	٣٢	٠.٠٠٠٠
٣	٠.٩٦٧٦	١٣	٠.٩٧٣٦	٢٣	٠.١٠٠٣	٣٣	٠.٢٢٩٣
٤	٠.٠٩٠٦	١٤	٠.٩٩٦٢	٢٤	٠.١٨١٥	٣٤	٠.٠٠١٥
٥	٠.٩٨٢	١٥	٠.٠٠٢٣	٢٥	٠.٦٥٦٨	٣٥	٠.٤٣٥٠
٦	٠.٠٠٠٤	١٦	٠.٠٠٠٦	٢٦	٠.٨٢٧٢	٣٦	٠.٠١٧٩
٧	٠.٠٠٢٦	١٧	٠.٣٧٩٧	٢٧	٠.٩٦٠٧	٣٧	٠.٠١١٦
٨	٠.٧٣٦٦	١٨	٠.٤٩٦٤	٢٨	٠.٠٨٥٥	٣٨	٠.٠٠٠٠
٩	٠.٤٣٣٧	١٩	٠.٠٠٠٠	٢٩	٠.٠١٥١	٣٩	٠.٦٧٧٣
١٠	٠.٩٦٦٧	٢٠	٠.٠٣٣٢	٣٠	٠.٠١١٩	٤٠	٠.٠٠٠٠

بناءً على النتائج الموضحة في الجدول (٤) سيتم حذف جميع الفقرات غير المطابقة للنموذج اللوجستي ثنائي المعلم التي مستوى دلالة مربع كاي أقل من ٠.٠١ وهي الفقرات التي تسلسلها(٤٠، ٣٨، ٣٤ ،٣٢، ٢١ ،١٩ ،١٦، ١٥، ١١ ،٧، ٦، ٢ ،١)، ثم إعادة التحليل للحصول على تقديرات نهائية لكل من معالم الفقرات وقدرات الأفراد على النموذج اللوجستي ثنائي المعلم. فبعد حذف الفقرات غير المطابقة للنموذج والتي كان عددها (١٣) فقرة تم الحصول على اختبار محكي المرجع في مادة مناهج البحث العلمي للدراسات العليا وفقاً للنموذج اللوجستي ثنائي المعلم لنظرية الاستجابة للمفردة، والذي تكون في صورته النهائية من (٢٧) فقرة، كما هو موضح في الملحق (٢).

خامساً: النتائج المتعلقة بالسؤال الخامس:

نص السؤال "ما الخصائص السيكومترية لاختبار مناهج البحث حسب النموذج اللوجستي ثنائي المعلم في نظرية الاستجابة للمفردة من حيث: معالم الصعوبة والتمييز؟"

ففي نظرية الاستجابة للمفردة تحسب خصائص الفقرات من خلال الفقرات التي يتضمنها الاختبار، إذ أن الإجابة عن الفقرات تعتمد على عاملين: هما القدرة وراء الاستجابة التي يظهرها الطالب، والمعالم التي تتحلى فيها الفقرة، والتي تتمثل في درجة صعوبتها في حال اعتماد النموذج اللوجستي أحادي المعلم، ودرجة الصعوبة والتمييز في النموذج ثنائي المعلم، والصعوبة والتمييز والتخمين في حال اعتماد النموذج ذي الثلاثة معالم. وقد تم استخدام البرنامج الإحصائي Bilog-MG3 لتقدير قيم الصعوبة والتمييز لفقرات الاختبار حسب نظرية الاستجابة للمفردة، وهذا ما تبرزه نتائج الجدول (٥) من تطبيق برنامج Bilog-MG3 لفقرات الاختبار في صورته النهائية المعتمدة (٢٧ فقرة) بعد حذف جميع الفقرات غير المطابقة للنموذج ثنائي المعلم لنظرية الاستجابة للمفردة.

الجدول (٥)

تقدير معالم الصعوبة والتمييز حسب نظرية الاستجابة للمفردة حسب النموذج ثنائي المعلم

رقم الفقرة في الاختبار	صعوبة الفقرة	الخطأ المعياري للصعوبة	تمييز الفقرة	الخطأ المعياري للتمييز	رقم الفقرة في الاختبار	صعوبة الفقرة	الخطأ المعياري للصعوبة	تمييز الفقرة	الخطأ المعياري للتمييز
٣	٢.١٣٥	٠.٣٠٣	٢.٠١٥	٠.٢٥٢	٢٤	١.٢٥٢	٠.١٠٩	٤.١٠٦	٠.٢٢٥
٤	٢.١٩٤	٠.٣٦٣	١.٢٢٧	٠.١٧٧	٢٥	١.٠٤٤	٠.١٣١	١.٦١٣	٠.١٥٠
٥	٢.٦١٥	٠.٥٥٦	١.٥١٥	٠.٢٨٠	٢٦	١.٣٢٠	٠.٠٩٨	٤.٩١٥	٠.٢٠٠
٨	٢.٤٤١	٠.٤٧٦	١.٠٧٧	٠.١٨٥	٢٧	١.٣٩٧	٠.٠٩٥	٣.٨٦٦	٠.١٨٣
٩	٠.٦٨٤	٠.٠٧٦	٢.٨٢٩	٠.٢١٥	٢٨	٢.٠٨٩	٠.٣٦٨	١.٣٦٤	٠.٢٢٧
١٠	١.١٩٤	٠.١٠٢	٢.٤٨٦	٠.١٦٧	٢٩	٢.٥٢٥	٠.٥٧٤	١.٠٢٨	٠.٢١٩
١٢	٢.٠٤٦	٠.٣٠٧	١.٧٨٦	٠.٢٤٣	٣٠	١.٩٢٩	٠.٣٣٤	١.٢٤٤	٠.٢١٥
١٣	٣.١٧٢	٠.٨٢٧	٠.٧٥٠	٠.١٧٥	٣١	١.٩٩١	٠.٢٦٨	٢.١٤٩	٠.٢٧٥
١٤	٢.٣٤٢	٠.٥٠٤	٢.١٤٨	٠.٤٥٣	٣٣	١.٢٣١	٠.١٠٠	٣.٧٦٤	٠.١٨١
١٧	٠.٥٦٣	٠.٠٩١	١.٨٨٨	٠.١٤٣	٣٥	٠.٩٠٥	٠.١٢٢	١.٤٧١	٠.١٤٨
١٨	١.٣١٩	٠.٠٨٣	٣.٥٠٠	٠.٢١٥	٣٦	٢.٥٦٩	٠.٦٣٢	١.٠٠٤	٠.٢٣٢
٢٠	٢.٥٨٥	٠.٥٩٥	٠.٩٢٥	٠.١٩٢	٣٧	٠.٢٣٨	٠.١٢٥	١.١٧٢	٠.١٢١
٢٢	١.٣٧١	٠.٠٧٦	٥.٣٣٦	٠.٢٥٨	٣٩	٠.٨٣٢	٠.١٧٨	٠.٩٤٢	٠.١١٨
٢٣	٠.٥٤٩	٠.١٦٣	٠.٩٠٢	٠.١٢٦					

يتبين من الجدول (٥) أن قيم معلم الصعوبة تتراوح بين ٠.٢٣٨ ــ ٣.١٧٢ بمتوسط حسابي قدره ١.٦٤٩، وتراوحت قيم الخطأ المعياري لتقدير معلم الصعوبة ما بين ٠.٠٧٦ ــ ٠.٨٢٧، في حين تراوحت قيم معلم التمييز ما بين ٠.٧٥ ــ ٥.٣٣٦ وبمتوسط حسابي قدره ٢.١١ ،وتراوحت قيم الخطأ المعياري لتقدير معلم التمييز ما بين ٠.١١٨ ــ٠.٤٥٣، في حقيقة الأمر لا يوجد معيار واضح ومحدد لصعوبة الفقرة أو تمييزها في نظرية الاستجابة للمفردة، فالصعوبة قيمة نسبية تتغير من اختبار الى آخر إلا أن الفقرات التي يزيد معامل صعوبتها عن (١.٥) تعد صعبة، والتي يقل معامل صعوبتها عن (-١.٥) تعد فقرة سهلة، لذلك حسب النتائج الموضحة في الجدول (٥) فإن بعض الفقرات في الاختبار كانت صعبة، إما لوجود غموض في بنية هذه الفقرات أو ربما كانت المموهات جيدة فأضلت الطالب عن الإجابة الصحيحة، إضافة إلى عدم قدرة الطلاب على استغلال بعض المؤشرات في بنية الفقرة، وكذلك افتقار الطلاب للمعرفة اللازمة للإجابة عن الفقرة مما جعل بعض الفقرات في الاختبار تعد فقرات صعبة. أما بالنسبة لمعلم التمييز فقد أشارت النتائج الموضحة في الجدول (٥) إلى أن جميع قيم التمييز كانت موجبة، وهذا يعني أن احتمال إجابة هذه الفقرة من قبل الطلاب الذين لهم تحصيل عالي أكبر من احتمال إجابة الطلاب الذين لهم تحصيل منخفض (Richichi, 1996)، وكلما زادت قيمة تمييز الفقرة كان ذلك أفضل ودليلاً على قيمة وجودة تلك الفقرة. وحسبما ذكر كوي وبكمان (Choi & Bachman, 1992) إلى أنه إذا كانت قيمة تمييز الفقرة أكبر من (١.٠)

فهذا دليل على أن الفقرة جيدة، أما إذا كانت قيمة التمييز للفقرة بين (٠.٧٠-١.٠٠) فتمييز الفقرة مقبول، وإذا قل تمييز الفقرة عن (٠.٧٠) تقل جودة تلك الفقرة. حسب هذا التصنيف يمكن القول أن جميع فقرات الاختبار تتمتع بتمييز جيد، ودليل على أن الفقرات كان بنائها جيد حسب مفهوم نظرية الاستجابة للمفردة .

سادساً: النتائج السؤال السادس:

نص السؤال "ما الخصائص الإحصائية لتقديرات قدرات الطلاب والأخطاء المعيارية في تقديرها؟"

تم تقدير معلم القدرة باستخدام طريقة الارجحية العظمى الهامشية(Method Maximum (Likelihood)(MML) من خلال برمجية Bilog-MG3، كما يوضحه الجدول (٦).

جدول (٦): قيم تقديرات القدرة للطلاب وفقاً لنظرية الاستجابة للمفردة

مدى قدرات الطلاب	التكرارات
-١.٦٩٧ــ -٠.٦٤٢	٥٢
-٠.٦٤٣ ــ ٠.٠٠٠	٧٦
٠.٠٠٠ ــ ٠.٦٤٢	٤٦
٠.٦٤٣ــ ١.٩٥٥	٧٩
المجموع	٢٥٣

يتبين من الجدول (٦) أن قيم تقديرات القدرة للطلاب الناتجة من تطبيق اختبار مناهج البحث العلمي قد تراوحت ما بين (-١.٦٩٧ إلى ١.٩٩٥). وتناقصت قيم الخطأ المعياري لمعلم القدرة كلما انتقلنا من مستويات القدرة الدنيا إلى مستويات القدرة العليا.

مناقشة النتائج:

بالنسبة للنتائج المتعلقة بأنسب النماذج (الأحادي أو الثنائي أو الثلاثي المعلم) من نماذج نظرية الاستجابة للمفردة للاختبار ثنائي التصحيح تم تحديد النموذج الأنسب من خلال اختبار مطابقة النموذج model fit. فقد استخدم الباحث برمجية Bilog-MG3 وحساب الإحصائي التالي ((-) the negative of 2 loglikelihood 2LLH) لكل من النماذج الثلاث ولنفس فقرات الاختبار، ومن ثم ايجاد الفرق بين قيمتي(-2LLH) لكل من أزواج النماذج الثلاثي والثنائي، وكذلك الثنائي والأحادي. فتبين أن تضمين النموذج الأحادي معالم إضافية يؤدي إلى تحسين الدالة إحصائياً في مطابقة النموذج للاستجابات عن فقرات الاختبار. وهذا يدل إلى أن النموذج ثنائي المعلم هو النموذج الأنسب للاستجابات على فقرات الاختبار على الرغم من أن الاختبار يحتوي على معلم التخمين.

أما فيما يتعلق بمناقشة النتائج المتعلقة بمدى تحقق افتراضات نظرية الاستجابة للمفردة وفق النموذج ثنائي المعلم، فقد تبين من خلال النتائج في الجدول رقم (٢) والمتعلقة بالتحليل العاملي للبيانات ظهور (١١) عامل، فسر العامل الأول منها (٢٠.١٩٨%) من التباين، وفسرت العوامل الباقية ما قيمته (٥٠.١%) من التباين الكلي، وهذا دليل على تحقق افتراض أحادية البعد وفقاً للتحليل العاملي المعتمد على المكونات الاساسية، حيث أشار عدد من الباحثين ومنهم لورد (Lord, 1980) أن الفقرات تحقق أحادية البعد إذا كان

الجذر الكامن للعامل الأول أكبر بكثير من بقية الجذور الكامنة لبقية العوامل. وقد تأكد الباحث من هذا الافتراض كذلك بمؤشرات أخرى منها فحص العوامل البياني من خلال الرسم، حيث ظهر عامل واحد أكثر من بقية العوامل وهذا ما بينه الشكل (١). وكذلك استخدم الباحث مؤشر ثبات الاتساق الداخلي، حيث تدل القيمة العالية لثبات الاختبار على أحادية البعد للاختبار.

أما فيما يتعلق بمناقشة الخصائص السيكومترية لفقرات اختبار مناهج البحث العلمي من حيث دلالات الصعوبة والتمييز حسب النظرية الكلاسيكية في القياس، فقد أظهرت النتائج وجود فقرتين تمييزهما أقل من ٠.٢٠، مما يعني وجوب حذف هاتين الفقرتين. وتبين من النتائج أن قيم الصعوبة تراوحت بين ٠.٢٥ و ٠.٩٦ وبمتوسط ٠.٧٩ وقيم التمييز تراوحت بين ٠.١٢٢ و ٠.٦٩٧ وبمتوسط ٠.٤١ ، وتعد هذه القيم في مستوى مقبول لأغراض بناء اختبار محكي المرجع.

أما فيما يتعلق بمناقشة النتائج المتعلقة بمطابقة الفقرات للنموذج اللوجستي ثنائي المعلم لنظرية الاستجابة للمفردة، فقد تبين وجود (١٣) فقرة لم تطابق توقعات النموذج اللوجستي ثنائي المعلم، حيث كانت قيمة احتمالية المطابقة لكل منها أقل من ٠.٠١ كما بينه الجدول (٤). فبعد حذف الفقرات غير المطابقة للنموذج والتي كان عددها (١٣) فقرة تم الحصول على اختبار محكي المرجع في مادة مناهج البحث العلمي للدراسات العليا وفقاً للنموذج اللوجستي ثنائي المعلم لنظرية الاستجابة للمفردة، والذي تكون في صورته النهائية من (٢٧) فقرة. وتتفق هذه النتيجة مع دراسة السامرائي والخفاجي (٢٠١٢).

وبالنسبة لمناقشة النتائج المتعلقة بتقدير معلمي الصعوبة والتمييز وفقاً للنموذج اللوجستي ثنائي المعلم لنظرية الاستجابة للمفردة، فقد تراوحت قيم الصعوبة بين ٠.٢٣٨ و ٣.١٧٢ وبمتوسط حسابي قدره ١.٦٤٩، وتراوحت قيم التمييز بين ٠.٧٥ و ٥.٣٣٦ وبمتوسط حسابي قدره ٢.١١، وعند المقارنة بين النظرية الكلاسيكية ونظرية الاستجابة للمفردة فقد تبين أن المؤشرات الإحصائية المستمدة من النظريتين تشير إلى اختلاف بينهما في تقدير مؤشرات الصعوبة والتمييز ولصالح النظرية الحديثة في القياس، لذلك تختلف الدراسة الحالية مع دراسة (Hwang, 2002).

أما بالنسبة لمناقشة النتائج المتعلقة بالخصائص الإحصائية لتقديرات قدرات الطلاب نلاحظ أن قدرات الطلاب تراوحت ما بين – ١.٦٩٧ إلى ١.٩٥٥ ، وكانت قيم الخطأ المعياري تتناقص كلما انتقلنا من مستويات القدرة الدنيا إلى مستويات القدرة العليا.

وتختلف نتائج هذه الدراسة من ناحية النموذج اللوجستي المستخدم مع دراسة كل من (علام، ١٩٨٥؛ العيثاوي،٢٠٠٩؛ السامرائي والخفاجي،٢٠١٢؛ Fan,1998) إذ استخدمت هذه الدراسات النموذج اللوجستي أحادي المعلم.

وقد خرج الباحث باختبار تحصيلي في مادة مناهج البحث العلمي لطلبة الدراسات العليا مبني وفق النظريتين الكلاسيكية والحديثة في القياس، وبما أن المؤشرات الإحصائية المستمدة من النظرية الحديثة كانت الأفضل، فقد تكون الاختبار في صورته النهائية من (٢٧) فقرة، بعد حذف الفقرات غير المطابقة للنموذج اللوجستي ثنائي المعلم.

المقترحـــات:

١.إجراء دراسات مقارنة بين النماذج اللوجستية (الأحادي، الثنائي، الثلاثي) المعلم الخاصة بنظرية الاستجابة للمفردة في هذا المجـــال، بزيادة حجم العينة وفقرات الاختبار للتوصل للنموذج الأمثل.

٢.استخدام الاختبار الذي تم إعداده وفقاً لنظرية الاستجابة للمفردة، والذي تكون في صورته النهائية من (٢٧) فقـــرة؛ لقيـــاس تحصـــيل الطلاب في مناهج البحث العلمي لأنه يتمتع بمؤشرات إحصائية مقبولة.

٣.إجراء دراسة لبناء اختبار تحصيلي في مـــواد دراسية أخرى لطلبة الدراسات العليـــا وفقـــاً لنظرية القياس الحديثة.

المراجـــع:

بركات، محمد (١٩٨٣). **علم النفس التعليمي، القياس النفسي والتقويم التربوي،ج٢.** الكويت: دار القلم.

التقي، احمد(٢٠٠٩). **النظرية الحديثة في القياس،** ط١. عمان: دار المسيرة للنشر والتوزيع.

السامرائي، محمد؛ والخفاجي، احمد(٢٠١٢). **بناء اختبار تحصيلي محكي المرجع في مادة علم نفس الخواص لطلبة أقسام العلوم التربوية والنفسية.** مجلة الاستاذ: العدد ٢٠٣. ص:٩٦٤-١٠٠٢.

عبابنة، عماد(٢٠٠٩).**الاختبارات محكية المرجع.** عمان: دار المسيرة للنشر والتوزيع.

عبد السلام، ونادية؛ و الشرقاوي، أنور؛ و الشيخ، سليمان؛ وكاظم، امينة (١٩٩٦). **اتجاهات معاصرة في القياس والتقويم النفسي والتربوي،** القاهرة: مكتبة الانجلو المصرية.

علام، صلاح الدين (١٩٨٦). **تطورات معاصرة في القياس النفسي والتربوي.** الكويت: كلية الآداب.

ـــــ (١٩٩٠).**استخدام نموذج راش في بناء مقياس هدفي المرجع للمعارف الأساسية في إعداد خطة البحوث النفسية والتربوية،** مجلة كلية التربية، جامعة الأزهر: القاهرة.

ـــــ (٢٠٠١). **الاختبارات التشخيصية مرجعية المحك في المجالات التربوية والنفسية،** القاهرة: دار الفكر العربي.

ـــــ (٢٠٠٥). **نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعدد الأبعاد وتطبيقاتها في القياس النفسي والتربوي،** القاهرة: دار الفكر العربي.

العيثاوي، محمد (٢٠٠٩). **استخدام نموذج راش وفق نظرية السمات الكامنة في بناء اختبار**

تحصيلي في مادة علم النفس الفروق الفردية، رسالة ماجستير غير منشورة، جامعة بغداد، العراق.

الطواها، سعيد (٢٠١١). **بناء مقياس للمهارات فوق المعرفية في اللغة العربية والتحقق من دلالات الفاعلية للمقياس وفقراته وفق النظريتين الكلاسيكية والحديثة في القياس**، رسالة دكتوراه غير منشورة، جامعة عمان العربية للدراسات العليا، الاردن.

الوليلى، إسماعيل (٢٠٠٢). **دراسة سيكومترية مقـــارنة لبعض نماذج الاستجابة للمفردة في انتقاء مفردات الاختبارات مرجعية المحك**، رسالة دكتوراه غير منشورة، جامعة الأزهر، مصر.

ياسين، عمر(٢٠٠٤). **الخصائص السيكومترية لاختبار محكي المرجع في الكيمياء لطلاب الصف الأول الثانوي العلمي مقدرة وفق النظريتين الكلاسيكية والحديثة في القياس.** رسالة ماجستير غير منشورة ، جامعة عمان العربية للدراسات العليا، الاردن.

Camilli, G., & Shepared, L.(1994).**Methods for Identifying biased test items**. Thousand Oaks, CA:Sage Publications.

Chernyshenko, o., Stark,S., Chan,K., Drasgow,F.,& Williams, B.(2001).**Fitting item response theory models to two personality inventories**: Issues and insights. Multivariate Behavioral Research,36(4),523-562.

Choi, I., & Bachman, F.(1992). **An investigation into the adequacy of three IRT models for data from two EFL reading tests Language Teasting**.9: 51-78.

Crocher, L. & Algina, J(1986). **Introduction to classical and modern test theory**, New York G.B.S college publishing.

Cronbach, L. J.(1951). **Coefficient alpha and the internal structure of tests**. Psychometrika. 16: 297-334.

Fan, X.(1998). **Item Response Theory and Classical Theory**: An Empirical Comparison of Their Item /Person Statistics. Educational and Psychological Measurement, 58(3),357-381.

Hambleton, R.K.& Swaminathan, H.(1985). **Item Response Theory Principles and Application**. Kluwir. Nij hof Publishing. Boston .

Hambleton,R.K.& Swanminathan, H.& Rogers,H.(1991). **Fundamentals of Item Response Theory**. New York: Sage Puplications the International Professional Puplishers.

Hulin, C.;Drasgow,F.; & Parson,K.(1983).**Item response theory: Applications to psychological measurement**. Homewood, Illinois:Dow Jones-Irwin.

Hwang, D.Y.(2002). **Classical Test Theory and Item Response Theory**: Analytical and Empirical Comparison. Paper Presented at the Annual Meeting of the Southwest Educational Research Association, Austin, TX.(ERIC Document Reproduction Service No. ED 466779).

Jiao, H. & Lau, A. (2003). **The effects of model misfit in computerized classification test**. Paper presented at the annual meeting of the National Council of Educational Measurement, Chicago, IL, April.

Linden,W.J.V.D: Hambieton,R.K.(1997). **Item Response Theory ,in Handbook of Modern Item Response Theory** , New York, Springer-Verlag.

Loehlin,J.C.(1987). **Latent variable models**. New Jersey: Lawrence Erlbaum Associates.

Lord,F.M .(1980).Applications of item response theory to practical testing problems. Hillsdale, N.J: Lawrence Erlbaum.

Nunnally, J.C(1978). **Psychological theory, 2nd cd, New York, Mcgraw-Hill**.

Randall ,S.(1998). **Comparing Measurement Theories** . *Paper Present at the Annual Meeting of the American Educational Research Association (San Diego , CA , April 13-17)*

Richichi, R.(1996). **An analysis of test bank multiple choice items using item response theory**. Author, SLD.(ERIC Document Reproduction Service NO. ED 405367.

Thissen, D., & Zimowiski, M.(1997). **IRT estimation of domain scores**. Journal of Education Measurement,34(3),179-211.

Wagner, A., & Harvey,R.(2003). **Developing a new critical thinking test using item response theory** . Paper Presented at the Annual Conference of the Society for Industrial Organizational Psychology. Orlando Retrieved May,4,from www.criticalthinking.org.

Zimowiski, M., Muraki, E., Mislevy, R. & Bock, R. (2003**). BILOG-MG3. In Mathilda du Tiot (Ed).** IRT from SSI: BILOG-MG, MULTILOG, PARSCALE, and TESTFACT. Chicago: Scientific Software International (SSI), Inc.

ملحق (١)

قائمة بأسماء المحكمين وجهة عملهم

الاسم	التخصص	جهة العمل
أ. د.عبد العاطي الصياد	تصميم البحوث والاحصاء والقياس والتقويم	عميد كلية العلوم الاستراتيجية في جامعة نايف
أ.د. خالد ابراهيم الكردي	علم النفس	رئيس قسم علم النفس في جامعة نايف الأمنية
أد. عبد الحفيظ سعيد مقدم	علم النفس الاجتماعي التطبيقي	عميد مركز الدراسات والبحوث في جامعة نايف
أ.د. السيد فهمي محمد	علم النفس الإكلينيكي	عضو هيئة التدريس في جامعة نايف
د. محمد هندية	علم النفس الإكلينيكي	عضو هيئة التدريس في جامعة نايف

ملحق (٢)

الصورة النهائية لاختبار مناهج البحث العلمي وفقاً لنظرية الاستجابة للمفردة (٢٧) فقرة

اختر الإجابة الصحيحة في كل مما يلي:

3. أحد المصادر التالية من مصادر البيانات الثانوية في المنهج التاريخي وهي:

أ) السجلات المشاهدة ب) المقابلات ج) شاهد العيان د) تقارير مشاهدة من قبل الآخرين

4. إذا طلب منك إجراء دراسة للتعرف على مدى اختلاف الاحداث المنحرفين من غير المنحرفين في الوازع الديني ، فإن المنهج المناسب لذلك هو :

أ) السببي المقارن ب) المسحي ج) التجريبي د) الارتباطي

٥. عدد من الحالات التي تؤخذ من عناصر المجتمع الاصلي للوصول إلى نتائج يمكن تعميمها تعني :

أ) البحث ب) العينة ج) النظرية د) الفرضية

٨. من عيوب المنهج الارتباطي أنه لا يدرس :

أ) قوة العلاقة بين المتغيرات ب) العلاقة التنبؤية ج) العلاقة السببية د) العلاقة بين متغيرين أو أكثر

٩. أي وسائل جمع البيانات التالية لا تصلح في حالة الابحاث النوعية (الكيفية) :

أ) الاختبارات ب) المقابلات ج) الملاحظات د) الوثائق

١٠. العينة التي تشترط أن يكون مجتمع البحث معروفاً هي:

أ) العشوائية ب) غير العشوائية ج) كرة الثلج د) القصدية

١٢.أي مما يلي يعد من طرق تقدير ثبات الاختبار:

أ) الطريقة التلازمية ب) طريقة التقنين ج) طريقة التجزئة النصفية د) كل ما ذكر صحيح

١٣. في المنهج التجريبي يمكن الحصول على مجموعات متكافئة بإحدى الطرق التالية :

أ) العشوائية ب) المزاوجة ج) التوائم د) كل ما ذكر صحيح

١٤. عندما يقوم الباحث بمراقبة جماعة والانخراط معهم بغرض دراسة حياتهم وسلوكهم نكون امام :

أ) ملاحظة مشاركة ب) ملاحظة غير مشاركة ج) ملاحظة غير مقصودة د) ملاحظة بسيطة

١٧. الصدق الخارجي للتصميم يعني :

أ) دقة النتائج التي يحصل عليها الباحث ب) إمكانية تعميم النتائج ج) ثبات أدوات القياس د) صدق البناء

١٨. العينات التالية لا يمكن تعميم نتائجها عدا واحدة هي :

أ) العينة القصدية ب) العينة الحصصية ج) العينة الطبقية د) عينة الفروق القصوى

٢٠. من فوائد كون الباحث يأخذ دوره كمشارك بالكامل (ملاحظة بالمشاركة) :

أ) أن المشاهدة تكون أكثر واقعية ب) أن الأفراد يحافظون على تصرفهم الطبيعي

ج) أن البيانات تكون أكثر صدقاً د) جميع ما ذكر صحيح

٢٢. المقابلات المعمقة

أ) تشبه الاستجواب ب) تتقيد بأسئلة محددة مسبقاً ومتتالية

ج) تشجع الاشخاص الذين تتم مقابلتهم على التحدث والإسهاب والحرية في الإجابة

د) تستخدم الاسئلة ذات الإجابة الثنائية " نعم أو لا "

٢٣. العبارات التالية تعتبر من مزايا الاستبيان عدا واحدة

أ) الموضوعية في الإجابة ب) الحصول على المعلومات من عدد كبير

ج) المقارنة بين الإجابات المختلفة د) القدرة على قراءة ردود فعل المستجيبين

٢٤. "مدى الدقة أو الاتساق أو استقرار درجاته فيما لو طبق على عينة من الافراد في مناسبتين مختلفتين " يقصد به

أ) ثبات الاختبار ب) صدق الاختبار ج) إعادة الاختبار د)تكافئ الاختبار

٢٥. في العينة العشوائية المنتظمة يكون :

أ) الشخص الاول اختير عشوائياً ب) الشخص الثاني اختير عشوائياً

ج) الشخص الاخير اختير عشوائياً د) جميع الاشخاص اختيروا عشوائياً

٢٦. من مزايا الملاحظة بالمشاركة

أ) انها تعطي الباحث معلومات وفيرة وغزيرة ب) أكثر مصداقية لأنها مأخوذة من الواقع الحقيقي الغير مصطنع
ج) يدرس الباحث الظاهرة موضع الدراسة عن كثب د) أ + ب صحيحتان

٢٧. تتميز المقابلة كأداة لجمع البيانات بالمزايا التالية عدا واحدة :

أ) ارتفاع نسبة المردود مقارنة بالاستبيان
ب) المقابلة من أنسب اساليب جمع البيانات في المجتمعات الامية أو الاطفال
ج) يتعذر إجراء المقابلة احياناً مع بعض الشخصيات المهمة
د) هي الاسلوب الانسب حين يكون المبحوثين غير راغبين في الأدلاء بآرائهم كتابةً

٢٨.العينات التالية هي خاصة في البحث النوعي عدا واحدة هي :

أ) العينة الشاملة ب) العينة المتيسرة ج) الفروق القصوى د) الحالات الخاصة

٢٩. إذا أراد باحث اختيار خمسة طلاب من ٣٠ طالباً بالطريقة المنتظمة ، فإن ارقام الطلاب التي ينبغي اختيارها:

أ) ٣ ، ٩ ، ١٥ ، ٢١ ، ٢٧ ب) ٧، ١٣ ، ١٩ ، ٢٥ ، ٣١ ج)٤ ، ١٠ ، ١٦ ، ٢٢ ، ٢٩

٣٠. العينة التي يمكن تعميم نتائجها لا بد أن تكون :

أ) كبيرة ب) شاملة ج) متنوعة د) احتمالية

٣١. في حالة أن يكون دور الباحث ملاحظ بالمشاركة فإنه يحتاج إلى أن :

أ) يتعرف إلى كل فرد من أفراد الدراسة ب) يخفي حقيقة دوره على أفراد الدراسة
ج) يجعل احتكاكه بالآخرين انتقائياً د) يحدد مسبقاً قائمة بالأشياء التي سوف يشاهدها

٣٣. من عيوب المقابلة الجماعية

أ) يمكن أن يساعد المبحوثين بعضهم بعضاً على تذكر المعلومات أو مراجعتها
ب) قد يسيطر أحد أفراد الجماعة على جو المقابلة
ج) قد يحجم البعض عن ذكر مشاكلهم الشخصية امام افراد الجماعة
د) ب + ج صحيحتان

٣٥. أراد باحث القيام بدراسة اتجاهات طلاب جامعة الملك سعود نحو الدراسة بالجامعة، فاختار لذلك عينة عشوائية تمثل كلاً من طلاب الدبلوم والبكالوريس والماجستير والدكتوراه. ما نوع العينة العشوائية التي ينبغي اختيارها؟

أ)طبقية ب) عنقودية ج) بسيطة د) منتظمة

٣٦. أي أشكال التوثيق التالية صحيحة للمرجع وفقاً للجمعية الأمريكية لعلم النفس:

أ)العساف، صالح(٢٠٠٩). المدخل إلى البحث في العلوم السلوكية. مكتبة العبيكان :الرياض.
ب) العساف، صالح(٢٠٠٩). المدخل إلى البحث في العلوم السلوكية. الرياض: مكتبة العبيكان.
ج) المدخل إلى البحث في العلوم السلوكية (٢٠٠٩). العساف، صالح. مكتبة العبيكان: الرياض.
د) المدخل إلى البحث في العلوم السلوكية (٢٠٠٩). العساف، صالح. الرياض: مكتبة العبيكان.

٣٧.منهج البحث الذي فيه صعوبة في التعميم هو:

أ) الارتباطي ب) السببي المقارن ج) المسحي د) التجريبي

٣٩. أي الفروض التالية يتطلب استعمال المنهج التجريبي:

أ) لا يوجد فروق في معدل التحصيل في الرياضيات بين الذكور والإناث
ب) يوجد علاقة ارتباطية بين العمر والوزن
ج) طريقة العصف الذهني أكثر تأثيراً من طريقة المحاضرة في تحصيل الطلاب
د) معدل تحصيل الطلاب في المدينة أفضل من تحصيل الطلاب في الهجر.