الخصائص السيكومترية لاختبار محكي المرجع في البحث التربوي وفق نظرية الاستجابة للفقرة

د. محمد عبد المجيد المصري

جامعة عمان العربية/ عمان

الملخص:

هدفت هذه الدراسة إلى استخدام نظرية الاستجابة للفقرة في بناء اختبار محكي المرجع في البحث التربوي وفق النموذج اللوجستي أحادي المعلم. تم إعداد اختبار لقياس تحصيل الطلبة في مساق البحث التربوي مؤلف من (48) فقرة من نوع الاختيار من أربعة بدائل، وبلغت عينة الدراسة (220) طالباً وطالبة من طلبة كلية التربية المحويت باليمن، وأظهرت النتائج تحقق افتراضات النموذج أحادي المعلم ومطابقة (42) فقرة من فقرات الاختبار لتوقعات النموذج، وتم حذف (4) فقرات غير مطابقة للنموذج كما بينت النتائج أن تقديرات معلمي الصعوبة والتمييز ضمن المحكات المقبولة في أدبيات القياس التربوي والنفسي.

كما أظهرت نتائج تقديرات قدرات الأفراد استبعاد(7) أفراد غير مطابقين للنموذج وتراوحت معالم القدرة لأفراد عينة الدراسة بين (−2.659)لوجيت و(2.812)لوجيت .

مقدمة:

يتطلب تقويم الطلبة استخدام أدوات قياس مختلفة تشكل الإختبارات ركنها الأساسي، للحصول على بيانات يمكن الاستناد عليها في اتخاذ قرارات وأحكام ذات الصلة بتصنيف الطلبة، وكشف نقاط القوة والضعف لديهم.

ويتطلب تصنيف الطلبة قياس مهارات محددة يتم تحليلها على صورة أهداف تعليمية اجرائية قابلة للملاحظة والقياس، وبالتالي يتم التحقق من كفاءة المتعلمين واكتسابهم للمعارف والمهارات والاتجاهات المرجوة، وقد حرصت كليات التربية على تنمية المهارات الأساسية لطلبتها، وتوافر اختبارات للكشف عن مستوى تمكن الطلبه لهذه المهارات وقد ظهر نوعان من الاختبارات :

الاختبارات معيارية المرجع هدفها الأساس هو: ترتيب مجموعة من الطلبة على متصل القدرة أو السمة المراد قياسها، ويتم فيها مقارنة أداء الفرد مع أداء المجموعة المعيارية له.

أما هدف الاختبارات محكية المرجع: فهو اتخاذ قرارات حول مستوى اتقان الطالب لمجال سلوكي محدد ويعرف الاختبار بأنه مجموعة من نتاجات التعلم المستهدفة(عبابنة ،2009) .

إضافة لذلك فإن بناء اختبار محكي المرجع يتطلب درجة عالية من الاتساق بين السؤال والهدف الذي يقيسه، كما أن حجم عينة الاختبار فيها أعلى من الاختبارات معيارية المرجع(عودة ،2005)

وقد تعددت الطرق التي اقترحها علماء القياس التربوي والنفسي التي يمكن استخدامها في تحديد مستويات الأداء في اتخاذ القرارات التعليمية المتعلقة بالأفراد (علام،2001).

وتستخدم الاختبارات محكية المرجع في تقدير أداء الفرد بالنسبة إلى محك أو مستوى أداء محدد مسبقاً.

وتتميز الاختبارات محكية المرجع (Criterion-Referenced Tests) بأن لها قيمة تشخيصية في تحديدجوانبالقوة والضعف في تحصيل المتعلمين،ويمكن استخدامها لتقويم الأفراد وما يمتلكون من معارف ومهارات وظيفية مكتسبة أو تدريبية محددة الأهداف والنواتج(علام،2000).

والهدف الأساس لأي نظريه في القياس هو التمكن من الوصول إلى تنبؤات حول القدرات أو السمات التي تقاس بفقرات الاختبار(عودة ،2005) .

ويبرز الأدب التربوي نظريتين لتحقيق هذا الهدف بالحصول على تقديرات موضوعية لقدرات الأفراد الأولى النظرية الكلاسيكية : التي شاع استخدامها على نطاق واسع في بناء الاختبارات والمقاييس التربوية والنفسية ويتم التعبير في إطار هذه النظرية عن مفهوم القدرة بالعلامة الحقيقية وهي توقع العلامة المشاهدة التي يمكن الحصول عليها عند تطبيق الاختبار على الفرد المفحوص مرات عدة (Crocker&Aligina,1986)

أما النظرية الثانية فهي النظرية الحديثة في القياس والتي تعرف بنظرية الإستجابة للفقرة (Item Response Theory) (IRT)

إذ يعدها العلماء ثورة في القياس النفسي والتربوي (علام،1986)،كاظم والشرقاوي والشيخ وعبد السلام 1996). والتي تستند على أساس مختلف عن النظرية الكلاسيكية، إذ تفترض أنه من الممكن التنبؤ بأداء المفحوصين أو تفسير أدائهم على اختبار عقلي أو نفسي محدد على ضوء خاصية أوخصائص مميزة لهذا الأداء تعرف بالسمات (Traits) وهي خفية يصعب ملاحظتها مباشرة ويمكن تقديرها أو الاستدلال عليها من أداء الأفراد على مجموعة من فقرات الإختبار المعد لقياسها (Embretsom&Reise,2000,sheng, 2005).

وبالتالي يمكن التغلب على الانتقادات التي واجهت النظرية الكلاسيكية في القياس التي أشار اليها هامبلتونوسوامنثيان(Hambleton &swaminathan,1985)ومن أهمها:
إن الخصائص السكومترية للأختبار يتم تحديدها بالنسبة لمجموعة من المفحوصين، كما أن خصائص المفحوصين تتحد من استخدام اختبار معين وليس بشكل مطلق، كما أنها تفترض تساوي الخطأ المعياري لجميع المفحوصين وأنها لاتوضح مدى تحصيل المفحوصين وفق مستوى الفقرة وإنما على الاختبار ككل (Hambleton & Swaminathan & Rogers,1999)

وعليه فأن نظرية الاستجابه للفقرة يمكن أن تسهم في حل المشكلات التي عجزت عن حلها النظرية الكلاسيكيةكالقياس التكيفي (Adaptive Measurement) . وبنوك الأسئلة في مفهومها الجديد وتقديرات القدرة المستقلة للمفحوصين عن عينة الفقرات التي تعرضوا لها.

وتستند نظرية الإستجابة للفقرة على مجموعة من الافتراضات الأساسية التي حددها. (Hambleton & Jones 1993 ,Hambleton &swaminathan,1985) المشار اليها في (الشريفين ،2006).وهي :

- **أحادية البعد** : أي أن هناك قدرة واحدة تفسر أداء الفرد على الأختبار (في النماذج احادية البعد)، أما في النماذج متعددة الأبعاد فيفترض وجود أكثر من قدرة تفسر أداء الفرد عليها .

- **افتراض منحنى خصائص الفقرة (ItemCharacteristics Curve)**

ويمكن وصف العلاقة بين تحصيل المفحوص على الفقرة وقدرته المقاسةبالاختبار من خلال الاقتران التراكمي الصاعد(دالة الاستجابه للفقرة)، إذ يوفر المنحنى احتمالات استجابة المفحوص عن الفقرة في مستويات القدرة المختلفة إجابه صحيحة، وأن صفة التراكمية تشير إلى زيادة احتمال الاجابهالصحيحه على الفقرة بازدياد قدرة المفحوص وتبدو في نماذج الاختبارات المعدة لقياس سمة واحدة، بدلالة معلم واحد هو معلم الصعوبة للفقرة، أو معلمتيالصعوبة للفقرة وتمييزها، أو ثلاثة معالم الصعوبة والتمييز والتخمين .

السرعة في الأداء : ليس لعامل السرعة دور في الإجابة عن فقرات الاختبار وان الإحفاق في الإجابه عن فقرات الإختبار يعود لأنخفاض قدرة الافراد وليس لتأثير عامل السرعة في الإجابة.

واستخدم في تحليل استجابات عينة الدراسة النموذج اللوحستيي احادي المعلم نموذج راش (RaschModel)

ويعد هذا النموذج من أبسط هذه النماذج واكثرها شيوعا ويعبر عنه بالدالة الرياضية التي تربط بين احتمال الاجابة الصحيحة للمفحوص على الفقرة، ومعلمة صعوبة الفقرة ومنحنيات خصائص الفقرة التي تختلف فقط في موقعها على متصل السمة، كما يفترض هذا النموذج تساوي القوة التمييزية لجميع الفقرات وتساوي (1) الصحيح، وأن معلمة التخمين تساوي صفراً لجميع الفقرات، والدالة الرياضية التي تمثل الاحتمالية يعبر عنها بالمعادلة التالية :

$$p_i(\theta) = \frac{e^{D(\theta - b_i)}}{1 + e^{D(\theta - b_i)}} \quad , (i = 1, 2, \ldots, n) \quad \ldots\ldots\ldots\ldots (11)$$

حيث :

$p_i(\theta)$ احتمال الاجابة الصحيحة لفرد قدرته(θ) عن الفقرة i

Dعامل تدريج (scalingfactor)ويساوي (1.7) ، θ = قدرة فرد ما ، b_i = معلم صعوبة الفقرة i ، a_i معلم تمييز الفقرة (Hambleton & Swaminathan, 1985).

ويعتمد هذا النموذج على الفرق بين القدرة (θ) التي يمثلها المفحوص حرف (S) في السمة التي يراد تقديرها (القدرة الكامنة وراء اجابة المفحوص ،ودرجة الفقرة (i)التي يرغب المفحوص بالاجابة عنها ،والتي يمثلها الرمز (bi)(التقي 2009).

a:ميل منحنى خصائص الفقرة عند نقطة انعطاف المنحنى (معلم تميز الفقرة) ويفترض هذاالنموذج ان تكون قيمة معامل التمييز في جميع الفقرات ثابته، وفي الغالب مساوية للواحد الصحيح .

bi:النقطة التي تقع على متصل القدرة عندما يكون ميل منحنى خصائص الفقرة اكبر ما يمكن (معامل الصعوبه).(حمادنة ،2009).

وتؤدي دالة معلومات الاختبار دورا مهما في نظريه الإستجابه للفقرة، إذ يمكن من خلالها تحديد الخطاالمعياري في التقدير الذي يرتبط عكسيا مع الجذر التربيعي لدالة المعلومات ، كما يتيح استخدام دالة معلومات الاختبار انتاج اختبار يتصف بالدقة اعتمادا على التقديرات الدقيقة لمعلم القدرة وفي الاختبارات المحكية خاصة (Gruijter & Kamb,2005).

ويتم تقييم الفقرات وفق نظرية الاستجابه للفقرة حسب مطابقتها النموذج الذي يصف العلاقة بين الاستجابه والقدرة باستخدام اختبار احصائي او تحليل البواقي(Residuals)، ويتم التعرف على الفقرات الضعيفة من خلال مؤشرات التمييز والتي اما تكون سالبة او موجبة ضعيفة، وتقديرات مؤشرات الصعوبة والا تكون الفقرات سهلة جدا او صعبة جدا لدى مجموعة من المفحوصين (Chernyshenko, Stark, Chan Drasgow & Williams, 2001)

وتم تقدير القدرة للمفحوصين من خلال نمط استجابة المفحوص على فقرات الاختبار باستخدام طريقة الأرجحية العظمى بايجاد قيمة (θ)، التي تعطي أرجحية عظمى للحصول على نمط الإجابة عن فقرات الاختبار، ويقوم هذا الأساس على إيجاد الأرجحية العظمى للقدرة (θ)، التي تتراوح بين (−∞، + ∞) من خلال الاعتماد على طرق رياضية وبتقريب هذه القيمة من خلال تحسين عملية التقدير بدءاً من قيمة تحدد مسبقاً (التقي، 2009).

وقد توصل فوربس (Forbes,1986). في دراسته لفاعلية استخدام نموذج راش في بناء الاختبارات التحصيلية في اللغة الانجليزية والعلوم والرياضيات انه من الممكن اعداد اختبارات تحصيلية اكثر جودة باستخدام هذا النموذج

الخصائص السكومترية في نظرية الاستجابة للفقرة

تفترض هذه النظرية وجود دالة لكل فقرة من فقرات الاختبار والتي تمثل منحنى خصائص لكل فقرة يعبر عن معالم الفقرة وهي:

-**معلم صعوبه الفقرة (bi)**:ويعرف بأنه تقدير القدرة المناظر لاحتمال الاجابه الصحيحة (0،50) عندما يكون نقطة تقاطع منحنى خصائص الفقرة مع المحور الصادي يساوي صفرا تقريبا، ويكون معلم التخمين للفقرة(C) يساوي صفرا .

اما اذا كانت قيمة المقطع الصادي لمنحنى الفقرة اعلى من الصفر فان صعوبه الفقرة هي القدرة الممثلة على محور السينات التي تقابل احتماليه الاجابه الصحيحة في منتصف المسافة بين تقاطع المنحنى مع المحور الصادي والقيمة (1) والاحتماليه تساوي (1+C/2) لان القيمة (1)هي اكبر قيمة احتمال الاجابهالصحيحةالتي تعبر عن مستوى صعوبه الفقرة .

وعند تحويل علامات القدرة للمفحوصين إلى مقياس معين وبمتوسط قدرة يساوي صفر (Θ=α) وانحراف معياري يساوي واحد فان مدى قيم معلم الصعوبهللفقراتيترواح من (−2إلى2+)،فقيم(b) القريبة من (−2)تكون الفقرة سهلة جدا وقيم الفقرات القريبة من (+2)تكون صعبه جداً (Hambleton&swamnathan,1985).

ويشير هارس (Harris,1989) إلىأن مدى قيمة معلم صعوبة نظريا يترواح ما بين (−∞إلى+∞) أو يقع المدى من ناحية عملية بين (−3إلى +3) عند تدرج القدرة على مقياس متوسط القدرة يساوي صفر وانحراف معياري يساوي (1).

كما يكون مدى القدرة للمفحوصين من(−∞إلى+∞) ونادرا ما يتجاوزمن(−3إلى +3)، ويرى وارم (Warm,1978)انه اذا كان مدى قيم صعوبة الفقرة(b) بين (−2.5 إلى +2.5) فالفقرة التي معلم صعوبتها يساوي (−2.5) فأقل سهلة جدا، والفقرة التي معلم صعوبتها(+2.5) فأعلى صعبة جدا، وإن الفقرة التي معامل صعوبتها يساوي صفرا متوسطة الصعوبة

معلم التمييز للفقرة (αi):

ويعبر عن قدرة الفقرة على التمييز بين مستويات القدرة المختلفة،ويتناسب معلم التمييز طرديا مع ميل منحنى خصائص الفقرة عند نقطة الانعطاف،وتزداد قيمة معامل تمييزالفقرة بزيادة شدة ميل الانحدار، ويعرف معلم تمييز الفقرة (αi) نظريا على مقياس مداه(−∞إلى+∞)، مما يشير إلى وجود فقرات ذات تمييز سالب وهذه يتم حذفها من الاختبار، وفقرات ذات تمييزموجب، وأن مدى معالم تمييز الفقرات يتراوح بين (صفر إلى 2)، وإذا كان معامل التمييز للفقرة أكثر من صفر دل ذلك على وجود علاقة موجبة بين

الاداء على الفقرة والقدرة المقاسة بالاختبار (Hambleton & swamnathan, 1985)، فالفقرات عالية التمييز يمكن أن تكون فقرات الاختبارجيدة في مدى ضيق من القدرة، والفقرات منخفضة التمييز تكون ضعيفة في مدى واسع من القدرة (Warm,1978) .

معلم تخمين الفقرة (Ci):هو الاستواء السفلي لمنحنى خصائص الفقرة ،ويتمثل في نقطة تقاطع المنحنى مع المحور الصادي ويعبر عن احتمال الاجابهالصحيحه عن الفقرة من المفحوصين ذوي القدرات المخفضه على الفقرات الصعبه جدا في التخميين او هو احتمال الاستجابهالصحيحه عند مستوى قدرة متدنية جدا، وهذا المعلم خاص بالفقرات الموصوعية ثنائية التدرج من نوع الاختبار من متعدد فقط.

ويشير وارم (Warm,1978)إلى ان معظم قيم التخمين(Ci) ينحسر مداها(صفر إلى 0.40) ،وتعد الفقرات غير مناسبة اذا كانت قيم معاملات تخمينها (0.30) فاكثر وكلما كان معلم التخميين منخفضا كان افضل، والفقرة التي قيمة تخمينها يقترب من الصفر فقرة مثالية ويظهر مراجعة الادب التربوي شيوع استخدام نظرية الاستجابه للمفردة في الدراسات والبحوث في اعداد الاختبارات والمقاييس التربوية والنفيسية ومن أهم هذه الدراسات:

وفي دراسة أجرتها عبد العزيز (1996) بعنوان استخدام نموذج راش في بناء بنك أسئلة لمادة علم النفس التعليمي على عينة بلغت(3236) طالبة، طبق عليهن اختبار مكون من (44) فقرة من الاختبار من متعدد ثلاثي التدريج ، واسفرت الناتئج إلى تدريج هذه الفقرات (بعد استبعاد بعض الفقرات والافراد) على متصل يحقق استقلالية القياس.

وأجرت الجمحاوي (2000) دراسة هدفت إلى المقارنة بين القدرة الرياضية ، ويكون الاختبار من (39) طبق على عينة بلغت (1061) طالباً وطالبة من الصف التاسع في مديريتي اربد الأولى والثانية ، أظهرتالنتائج اتفاقا عالياً في تقدير صعوبة الفقرات وتميزها، وفي توزيع قدرة المفحوصين في النظريتين الكلاسيكية والحديثة، وبينت النتائج ان عدد الفقرات التي تم اختبارها وفق النظرية الكلاسيكية بلغت (33) فقرة في حين طابقت (20) فقرةالنموذج الاحادي المعلم بينما طابقت (35) فقرة نموذج المعملين و(38) فقرة النموذج الثلاثي المعالم.

وفي دراسة أجرتها دينا (Dena ,2003) هدفت إلى استخدام اختبار متعدد المستويات وفق نماذج استجابة الفقرة في البحوث التطبيقية، وتم فيها تطبيق اختبار متعدد المستويات على أطفال من مراحل عمرية مختلفة (12– 18) عاما،وتم حساب الخصائص السيكومترية

للاختبارات المتعددة حسب نماذج نظرية استجابة الفقرة، والمعادلة الرأسية للتنبؤ بالسمة التي يمتلكها المفحوص من خلال الاختبار المشترك، وأظهرت النتائج أن القدرة تزداد بزيادة الفئة العمرية مع ازدياد صعوبة الفقرة وتميزها .

واستخدم دودين(2004)نظرية الاستجابة على الفقرة في تحليل الاستجابات على اختبار المهارات الكمية لطلبة الجامعة وقد اعتمدت هذه الدراسة على المنهجية التعليمية بجامعة الامارات العربية المتحدة في اعداد وبناء فقرات الاختبار التي يتكون من (40) فقرة من نوع الاختبار من متعدد وفق مجالات هي :الحساب والجبر ،والهندسة ،والاشكال البيانية ، والجداول، والاحصاء ، والاحتمالات ،على عينة بلغت (1005) من طلبة جامعة الامارات العربية المتحدة وتم الاعتماد على آليةخاصةبنظرية الاستجابة على مستوى الفقرة تعرف بتصميم الارتكاز (Anchor design) لاضافة فقرات بشكل مستمر تمكن من اعداد اختبارات متكافئة ويمكن لهذه المنهجة ان تطبق لقياس قدرات واستعدادات او مهارات لدى طلبة الجامعة، وسهولة تطبيقها في الاختبارات التحصيلية.

وفي دراسة اجرتها وابيرج (Wiberg,2004) وهدفت الى تقييم الاختبار النظري لرخصة القيادة السويدي وفق النظرية الكلاسيكية ونظرية الاستجابة للفقرة وقارنت نماذج نظرية الاستجابة للفقرة الاحادي والثنائي والثلاثي لتحديد النموذج الاكثر ملائمة لنمذجة الفقرات وتكون الاختبار من (65) فقرة من نوع الاختيار من متعدد محكي المرجع ولكل فقرة من (2–6) بدائل منها بديل واحد صحيح اعطيت الاجابة الصحية واحد علامة واحدة وصفر للاجابة الخاطئة على عينة بلغت (5404) مفحوصا واشارت النتائج وفق النظرية الكلاسيكية ان متوسط الحساب لعلامات الاختبار بلغ (51.44). بانحراف معياري (6.9) وبمدى من (13–65)، وان (65%) من المتقدمين اجتازوا الاختبار النظري وتم تقدير الثبات لمعامل (الفا) فبلغ(0.82)وتراوحت قيم معاملات الصعوبة بين(0.13 الى 0.97) وقيم التمييز للفقرة بمعامل الارتباط الثنائي من(0.03الى 0.48) كما اشارت النتائج الى اختلاف بين معاملات الارتباط الثنائي النقطي مما يشير الى الاختلاف في مدى تمييز الفقرات. وتمنمذجة الفقرات وفق نظرية الاستجابة للفقرة واشارت النتائج الى ان (53) لم تطابق النموذج احادي المعلم ، و(25) لمتطابق النموذج ثنائي المعلم ، وان(16) فقرة لم تطابق النموذج ثلاثي المعلم وعليه فان النموذج الثلاثي هو النموذج المفضل .

دراسة الشريفين (2006) الخصائص السيكومتريةلاختبار محكي المرجع في القياس والتقويم وفق النظرية الحديثة في القياس، وتم بناء اختبار تحصيلي في القياس والتقويم

التربوي من (50) فقرة من نوع الاختبار من متعدد باربعة بدائل على عينة بلغت (222) طالباً وطالبه واشارت النتائج مطابقة استجابات (37)فقرةمن الاختبار لافتراضات نموذج راش بعد حذف (13) فقرة لم تطابق افتراضات النموذج، وبلغ معامل الثبات للافراد(0.90) ،ومعامل ثبات الاختبار (0.95)، كما بينت الناتئج أن قيمة دالة المعلومات كانت اقصى ما يمكن عندما كانت (θ=b) كما يتوقع من نموذج راش وان الاختبار يقدم اكبر كمية من المعلومات عند مستويات القدرة المتوسطة ،وتقدم معلومات أقل عند مستويات القدرة العالية والمتدنية.

كما أجرى أحمد (2007) دراسة لبناء مقاييس للقبول في بالجامعات المصرية وبلغت العينة (553) طالباً وطالبة طبق عليهم مقاييس الاستعداد اللغوي والرياضي والاستعداد للغة الانجليزية ، واستخدم نموذج راش في تدريج المقاييس، وتم حذف (36) فرداً و ثلاث فقرات من مقاييس الاستعداد اللغوي، وحذف (30) فردا و(6) فقرات من مقياس الاستعداد الرياضي ، وحذف (46) فردا و (8) فقرات من مقياس الاستعداد للغة الانجليزية ، وأظهرت النتائج ارتفاع معاملات الثبات من (0.63،0.63،0.69) إلى (0.97،0.97،0.99).

وهدفت دراسة (الحواري، وعودة 2008) التعرف على الخصائص السيكومترية لصورة مختارة من اختبارات الرخصة الدولية لقيادة الحاسوب في الاردن ومعادلة درجاتها، تم تعريف معادلة القياس ، وتوضيح كيفية جمع البيانات اللازمة لها والطرق في عملية المعادلة (المعادلة الخطية الافقية، المعادلة المئينية الافقية) وفق النظرية الكلاسيكية، وباستخدام نموذج راش احادي المعلم من نماذج نظرية استجابة الفقرة،على عينة بلغت (1200) مفحوص اختيروا بالطريقة العشوائية وزعت إلى ثلاث مجموعات متكافئة ، وتم تطبيق النماذج الثلاثة للمقاييس الفرعية وهي : المفاهيم الاساسية ، واستخدام الحاسوب ، وادارة الملفات ، ومعالجة الكلمات ، وجداول البيانات ، وقواعد اليانات ، والعروض التقديمية ، والمعلومات والاتصالات ، واظهرت النتائج ان المعادلة وفق نموذج راش كانت الافضل والاكثر فاعلية في معادلة درجات المفحوصين من خلال وضعهم على تدريج مشترك بالاعتماد على المعايير السابقة.

وفي دراسة حمادنة(2009) باستخدام نظرية الاستجابة للمفردة في بناء اختبار محكي المرجع في الرياضيات وفق النموذج اللوجستي ثلاثي المعلم في الاردن ، وتم بناء اختبار في الرياضيات يقيس تحصيل الطالب في الاحصاء مكون من (28) فقرة من نوع

الاختبار من متعدد من اربعة بدائل على عينة بلغت (411) طالباً وطالبة بالصف الأول الثانوي العلمي ، واظهرت النتائج تحقق افتراضات نظرية الاستجابه للفقرة ، ومطابقة (24) فقرة من فقرات الاختبار لتوقعات النموذج اللوجستي ثلاثي المعلم وحذف (4) فقرات لم تطابق النموذج ، وان تقديرات معالم الفقرت (الصعوبة ، والتمييز ، والتخمين) لها كانت مقبولة ضمن المحكات الواردة في ادبيات القياس التربوي .

وفي دراسة أجرتها هادي، ومراد (2013) استخدام نموذج راش في اعداد وتقنين اختبار في اللغة العربية لقبول الطلبة في الدراسات العليا بكليات العلوم الانسانية بجامعة الكويت وبعد مراجعة الادبيات والاختبارات المتوافرة في هذا المجال تم تحديد ثلاثة محاور اساسية للاختبار وهي: القدرة على الفهم والتعبير، والتفكير، وتكونت الصورة الاولية للاختبار من (130) بنداً تم تطبيقها على عينة بلغت (86)طالباً وطالبة ، وأسفر التحليل عن حذف (30) بنداً وتعديل (16) بنداً ، تم تطبيق الصورة النهائية على عينة بلغت (335) طالباً وطالبة، وتم تحليل البيانات بالطريقة العادية وباستخدام نموذج راش ، وأسفرت نتائج الطريقة العادية عن حذف (17) بنداً، وبلغ معامل كرونباخ ألفا (0.86)، وتراوحت معاملات الاتساق الداخلي بين (0.88،0.46)، كما ارتبطت درجات الاختبار بالمعدل التراكمي ارتباطاً دالاً (0.01) .

أما نتائج التحليل باستخدام نموذج راش في تدريج بنود الاختبار (100)بند قفد تم حذف (26) فردا، و(19) بندا غير ملائمين للنموذج واظهر التحليل عن معامل ثبات جيد للاختبار بلغ (0.81و0.99) لقدرات الافراد، وتراوح مدى القدرة بين (+،−2)لوجيت بمتوسط = 0.5 وهو مستوى جيد .

تبين من عرض الدراسات ذات الصلة الاستخدام الواسع للنظرية الحديثة للقياس في بناء الاختبارات والمقاييس النفسية والتربوية،التي أكدت نتائجها دقة وموضوعية نظرية السمات الكامنة في اعداد الاختبارت وقياس مستوى أداء الافراد .

ونظراً لندرة الدراسات التي تناولت كيفية إعداد الاختبارات وفق نظرية الاستجابة للفقرة، لذا تميزت الدراسة الحالية في محاولتها لكيفية بناء اختبار محكي المرجع في البحث التربوي لدى عينة من طلبة المرحلة الجامعية في اليمن وحساب الخصائص السيكومترية وفق النموذج اللوجستي أحادي المعلم والذي لم يتم تناوله في البئية اليمنية خاصة.

مشكلة الدراسة :

إن الهدف من هذه الدراسة هو بناء اختبار في مادة البحث التربوي لطلبة المرحلة الجامعية والتحقق من خصائصه السيكومترية وفق نظريتي القياس الكلاسيكية والحديثة .

أسئلة الدراسة:

لتحقيق أهداف الدراسة حاولت الدراسة الاجابة عن الاسئلة الاتية :

1-ما دلالات الصدق المحققة لصورة الاختبار ؟

2-ما دلالات الثبات المتحققة لصورة الاختبار؟

3-ما القيم المتحققة لاحصائيات الفقرة وفق النظرية الكلاسيكية في القياس ؟

4-ما القيم المتحققة لمعالم الفقرة وقدرة الافراد وفق النظرية الحديثة للقياس؟

أهمية الدراسة:

تبدو أهمية هذه الدراسة بجانبين :

الاول:نظري من خلال عرضها للادب التربوي المتعلق ببناء الاختبارات في النظرتين الكلاسيكية والحديثة وكيفية التحقق من الخصائص السيكومترية حسب نماذج نظرية الاستجابه للفقرة ، وتقدير معالم الفقرات وتقدير قدرات الافراد.

أما الجانب التطبيقي فيظهر في عرضه لكيفية بناء أداة قياس موضوعية يتحقق فيها خصائص سيكومترية مناسبة لقياس تحصيل الطلبة في مساق البحث التربوي ، وبناء الاختبارات نظراً لاهمية اتقان طلبة الكليات التربوية لمهارات البحث العلمي.

تعريف المصطلحات:

نموذج راش(Rash,model) أحادي المعلم ،هو احد نماذج نظرية الاستجابة للفقرة ، ويفترض وجود قدرة واحدة تفسر اداء الفرد على الاختبار ، كما تفترض تساوي جميع فقرات الاختبار في التمييز بين مستويات القدرة التي يقيسها الاختبار، وأن معلم التخمين لجميع الفقرات يساوي صفراً(Wright&stone,1979)

صعوبة الفقرة : القيمة التي تندرج على متصل السمة يكون احتمال اجابة الفقرة عندها اجابه صحيحة مساويال(50%)، واتأخذفي نظرية الاستجابة للفقرة قيما بين (-∞إلى+∞) وبشكل عام من (-3 إلى +3).

قدرة الفرد: يتم تقديرها بطريقة الارجحية العظمى لاستجابات الفرد على فقرات الاختبار،

الاختبار التحصيلي :عرفه عودة (2001)، بأنه طريقةمنظمةلقياسمستوى تحصيل الطلبة لمعلومات ومهارات مادة دراسية بالاجابة عن مجموعة من الفقرات التي تمثل محتوى المادة الدراسية .

حدود الدراسة ومحدداتها:

-اقتصرت الدراسة على عينة من طلبة كلية التربية بالمحويت الفصل الدراسي الاول(2009-2010)

-اقتصار الدراسة علىاستخدامنموذج راش احادي المعلم ، وعلى دقة البرمجيات الاحصائية المستخدمة للاجابة عن اسئلة الدراسة .

منهجية الدراسة وإجراءاتها:

اعتمد الباحث في بناء الاختبار على خطوات اعداد الاختبارات محكية المرجع وفقا لافتراضات النموذج اللوجستي أحادي المعلم لنظرية الاستجابة للفقرة.

مجتمع الدراسة :

تكون مجتمع الدراسة من جميع طلبة السنة الرابعة المسجلين في كلية التربية بالمحويت، ونظرا لغياب عدد من الطلبة ، قد تم اختيار العينة من الطلبة المنتظمين في الدراسة.

عينة الدراسة:

بلغت عينة الدراسة(240) طالبا وطالبة من المنتظمين في الفصل الدراسي الاول (2009-2010) من طلبة السنة الرابعة لكلية التربية بالمحويت، اليمن.

أداة الدراسة :

تم بناء اختبار تحصيلي في مادة البحث التربوي من نوع الاختيار من متعدد من أربعة بدائل وفق الاسس المتبعة في بناء الاختبارات التحصيلية ، التي حددها غرونلند ولين (Gronlund & Linn, 1990)، وقد استعان الباحث ببعض الدراسات السابقة التي تناولت بناء ادوات لقياس موضوعات التحصيل في مساق البحث التربوي (علام، 2001)، ومراجع البحث التربوي، وتدريسه للمساق مدة طويلة، وتضمن المساق الموضوعات التالية : المفاهيم الرئيسة للبحث العلمي وخصائصة ، خطوات البحث العلمي: المشكلة وتحديدها، خطة البحث ، العينات وطرق اختيارها ، ادوات البحث ، اساليب البحث (التاريخي، الوصفي،التجريبي، الاجرائي) كتابه البحث ،وكيفية توثيق المراجع. حيث تم إعداد جدول لمواصفات الاختبار بعد تحليل محتوى المساق، وصياغة الأهداف السلوكية التي تضمنتها موضوعات المساق والتي يتوقع أن تتحقق عند الطلبة .

تم صياغة (54) فقرة من نوع الاختيار من متعدد بأربعة بدائل أحدها يمثل الاجابة الصحيحة لقياس الأهداف السلوكية مع مراعاة الفقرة للهدف السلوكي الذي تقيسه.

صدق الاختبار:

تم التحقق من صدق الاختبار بعرض الاختبار وجدول المواصفات على مجموعة من المتخصصين للحكم على مدى تمثيل الفقرات للمحتوى المراد قياسه، ودقة صياغة الفقرات،وأسفر التحكيم عن استبعاد (6)فقرات اتفق 80% من المحكمين على استبعادها وتعديل صياغة بعض الفقرات ، وتعديل صياغة بعض البدائل لبعض الفقرات ، وتكون الاختبار من(48) فقرة شكلت الصورة الأولية للاختبار بعد التحكيم.

التجريب الاولي للاختبار: تم تطبيق الاختبار على عينة بلغت (50) طالباً وطالبة من مجتمع الدراسة للتعرف على صعوبة الفقرات، وتحديد الزمن اللازم للتطبيق، والاطمئنان على وضوح صياغة الفقرات والاجابة عن تساؤلات الطلبة، واجراء التحليل لبيانات العينة الاستطلاعية باستخدام الحزمة الاحصائية (Spss)

وقد تم حساب معاملات الصعوبة بايجاد نسبة الطلبة الذين أجابوا إجابة صحيحة على الفقرة من الطلبة الذين حاولوا الاجابة عن الفقرة ، وحساب القدرة التميزية للفقرات بحساب معامل الارتباط الثنائي بوينت بايسيريال بين الدرجة على الفقرة والدرجة على الاختبار ، والجدول (1) يوضح ذلك

الجدول رقم (1)

معاملات الصعوبة مقدرة بنسبة النجاح على الفقرة ، ومعاملات التميز مقدرة بمعامل الارتباط الثنائي لفقرات الاختبار في العينة الاستطلاعية

رقم الفقرة	معامل الصعوبه	معامل التميز	رقم الفقره	معامل الصعوبه	معامل التميز	رقم الفقره	معامل الصعوبه	معامل التميز
1	0.35	0.443	17	0.32	0.532	33	0.28	0.336
2	0.33	0.465	18	0.41	0.486	34	0.32	0.483
3	0.35	0.454	19	0.44	0.515	35	0.34	0.412
4	0.19	0.318	20	0.45	0.582	36	0.45	0.418
5	0.41	0.438	21	0.37	0.546	37	0.59	0.437
6	0.42	0.516	22	0.41	0.532	38	0.42	0.501
7	0.37	0.532	23	0.39	0.343	39	0.45	0.442
8	0.42	0.503	24	0.38	0.501	40	0.52	0.435
9	0.43	0.448	25	0.36	0.492	41	0.31	0.386
10	0.47	0.376	26	0.37	0.426	42	0.30	0.347
11	0.41	0.371	27	0.35	0.347	43	0.34	0.343
12	0.42	0.358	28	0.25	0.384	44	0.37	0.382
13	0.45	0.402	29	0.31	0.563	45	0.35	0.421
14	0.39	0.381	30	0.52	0.432	46	0.26	0.387
15	0.49	0.435	31	0.25	0.415	47	0.18	0.21
16	0.39	0.527	32	0.16	0.326	48	0.41	0.375

يلاحظ من الجدول (1) أن قيم معاملات الصعوبه لفقرات الاختبار تراوحت بين (0.16) للفقرة (32) و (0.59) للفقره (37). وبمتوسط الصعوبة لجميع الفقرات بلغ

(0.374)، اما قيم معاملات التمييز فقد تراوحت بين (0.21) للفقره (46) و(0.59) للفقره (37) وبمتوسط كلي لجميع معاملات التمييز بلغ (0.431)

وبناء على هذه النتائج فقد تم حذف الفقرة (4) والفقرة (47) لان قيم معامل صعوبتيهما اقل من (0.20)

ونتيجة لذلك تشكلت الصورة النهائية للاختبار الذي تم تطبيقه على عينة الدراسة من (46) فقرة .

دلالات صدق الاختبار وثباته

تم التحقق من صدق الاختبار بطريقتين:

الطريقة الاولى:تم التحقق من صدق المحتوى بالاعتماد على تحليل المحتوى وبناء جدول المواصفات، وتأكد ذلك من خلال أحكام المحكمين من أعضاء الهيئة التدريسية كما سبق توضيحه في إجراءات إعداد الصورة الأولية للاختبار.

الطريقه الثانية: الصدق البنائي : تم تقديره بإجراء التحليل العاملي لاستجابات العينة الاستطلاعية على فقرات الاختبار بتحديد المكونات الرئيسة بإيجاد قيم الجذور الكامنة للعوامل المستخلصه(Eigen Valve) ونسبة التباين المفسرلكل عامل والنسبة التراكمية للتباين المفسر والجدول (2) يوضح نتائج العينة الاستطلاعية لفقرات الاختبار .

الجدول (2)

العوامل المستخلصة من التحليل العاملي لفقرات الاختبار ، وقيم الجذر الكامن ونسبة التباين المفسر لكل عامل والنسبة التراكمية للتباين المفسر من استجابات أفراد العينة الاستطلاعية

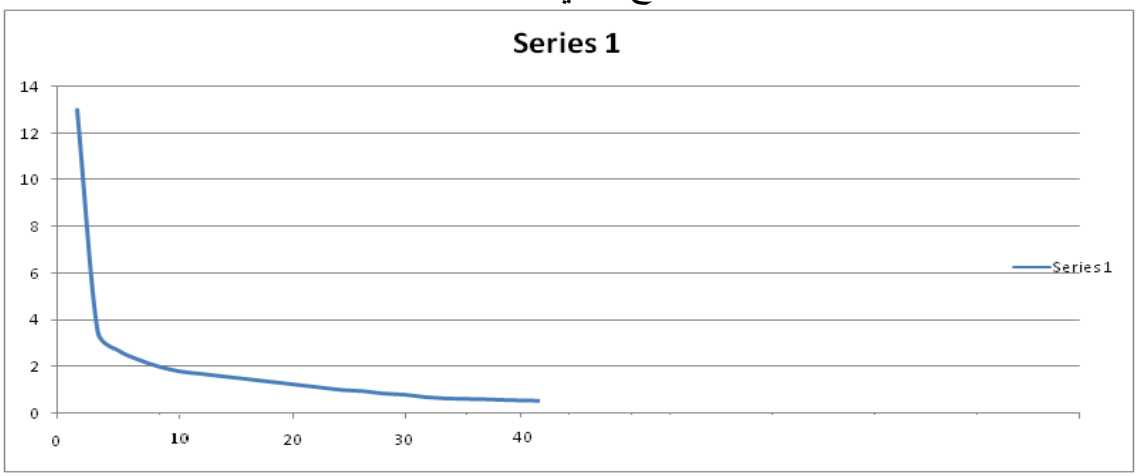
العامل	الجذر الكامن	نسبة التباين المفسر	النسة التراكمية للتابين المفسر
1	10.103	22.200	22.204
2	2.812	5.624	27.824
3	2.014	4.005	31.833
4	1.813	3.627	35.46
5	1.648	3.296	38.756
6	1.527	2.972	41.728
7	1.482	2.845	44.573
8	1.439	2.706	47.279
9	1.377	2.653	49.932
10	1.342	2.585	52.517
11	1.312	2.521	55.038
12	1.254	2.462	57.5
13	1.196	2.405	59.905

تبين من الجدول (2) أن هنالك(13) عاملا كانت قيم الجذر الكامن لها اكبر من (1) وفسرت ما نسبته (59.905) كما تبينت أن هناك ثلاث عوامل زادت قيم الجذر

الكامنلهما عن (2)، وفسرت ما نسبته (27.824) من التباين الكلي ومن الملاحظ أن قيمة الجذر الكامن للعامل الاول بلغت(10.103) وهي قيمة عالية اذا ما تم مقارنتها بقيم جذور العوامل الاخرى، وفسر العامل الاول وحده (22.204) من التباين الكلي على الاختبار وان جذره الكامن اكبر من (3.948) من قيمة الجذر الكامن للعامل الثاني والذي فسر فقط (5.624) من التباين الكلي.

واعتماداً على قيمة محك الجذر الكامن كمؤشر على أحادية البعد إذا كان حاصل قسمة الجذر الكامن للعامل الأول إلى العامل الثاني كبيرة وتزيد عن (2)، والمحك الآخر هو نسبة التباين المفسر، التي بلغت للعامل الأول (22.204) من التباين الكلي، مما يؤكد وجود عامل سائد في الاختبار ويتوافق ذلك مع اقتراح ريكاس (Rechase, 1985) من أنه إذا فسر العامل الأول (20%) من التباين المفسر على الأقل فإن ذلك يعد مؤشراً كافياً لتحقق أحادية البعد (Hambleton&swamnathan,1985)،(Hattie, 1985)، (الشريفين ، 2012).

كما تم استخدام الرسم البياني لقيم الجذر الكامنة للعوامل المكونة لصورة الاختبار الاعتيادية كما في الشكل (1) التمثيل البياني لقيم الجذور الكامنة للعوامل المختلفة التي اسفر عنها التحليل، ونلاحظ من الشكل (1) ان العامل الاول يمثل عاملا سائدا وان هناك تحولا في المنحنى يعد العامل الثالث، وبعد يصبح ميل المنحني طفيفا، كما يرجح افتراض احادية البعد لاغراض التحليل وفق نماذج نظرية استجابة الفقرة.



الشكل (1) التمثيل البياني لقيم الجذور الكامنة للعوامل المستخلصة من التحليل العاملي للاختبار

* أما دلالات الثبات الاختبار تم تقديرها من استجابات افراد العينة الاستطلاعية بالاتي :

-حيث بلغ معامل ثبات الاختبار كرونباخ الفا (0.87)،وبالطريقة النصفية بعد التصحيح بمعادلة سبيرمان براون (0.89) .

وقد أسفرت جميع الاجراءات السابقة على استجابات افراد العينة الاستطلاعية عن الصورة النهائية للاختبار مع توضيح التعليمات ،وقد تم تطبيق الاختبار على افراد عينة الدراسة في نهاية الفصل الدراسي الاول (2009-2010) ،و البالغ عددها (240) طالب وطالبة.

ادخلت البيانات إلى ذاكرة الحاسوب وقد تم استبعاد (12) ورقة لعدم اكتمال الاجابه عليها.

استخدم في تحليل البيانات برنامج الحاسوب (Rash measurement analysis software directory)

ويتميز نموذج راش الأحادي المعلم بحذف كل فرد فشل في الأجابة عن جميع الفقرات (اي حصل على صفر)، كما يحذف كل فرد نجح في الاجابه عن جميع فقرات الاختبار اجابه صحيحة ، اذ يعد هؤلاء الأفراد غير مطابقين للنموذج باستخدام الارجحية العظمى في تقدير المعالم .

كما يحذف البرنامج كل فقرة اخفق في الاجابه عنها جميع افراد العينة، وكل فقرة نجح في الاجابه عليها جميع افراد العينة اجابه صحيحة وتعد هذه الفقرات غير مطابقة لشروط النموذج .

المعالجة الاحصائية :

تم تحليل البينات وفق نظريتي القياس الكلاسيكية والحديثة باستخدام نموذج نظرية استجابة الفقرة احادي المعلم (نموذج راش) لفقرات نموذج الاختبار وتم استخدام الرزمة الاحصائية (Spss)والبرمجيات الاحصائية(ACER 301) لبيانات العينة النهائية وبتطبيق شروط نموذج راش احادي المعلم فقد تم استبعاد بيانات (8) من افراد العينة لتصبح عينة التحليل (220) طالباً وطالبة ولم تستبعد اي فقرةمن فقرات الاختبار(46) .

تقدير الخصائص السيكومترية للاختبار وفق النظرية الكلاسيكية للقياس

اولاً: تم استخدام الحزمة الاحصائية(Spss) لتقدير مايلي :

1-معاملات الصعوبه لفقرات الاختبار

2-معاملات التميز لفقرات الاختبار

3-معامل ثبات كرونباخ الفا للاختبار

4-معمل الثبات النصفي لفقرات الاختبار وتصحيحه

5-التحليل العملي لفقرات الاختبار باستخدام تحليل المكونات الرئيسة والجذر الكامن ونسبة التباين المفسر لكل عامل جذره الكامن اكبر من (1) (Hambleton and swamnathan,1985)

كما تم التحقق من احادية البعد من التمثيل البياني للجذور الكامنة للعوامل (Scree plot)

ثانياً:استخدام برنامج (EXCEL) وايجاد :

-حساب المتوسط الجسابي والانحراف المعياري لمعاملات الصعوبه لفقرات الاختبار

-حساب المتوسط الحسابي والانحراف المعياري لمعاملات التميز لفقرات الاختبار

تقدير الخصائص السيكومترية وفق النظرية الحديثة في القياس تم استخدام برمجية (Rumm 2020 , ACER 301,Excel) في تقدير الخصائص الاتية:

-قدرات الافراد على فقرات الاختبار

-الخطأ المعياري في تقدير قدرات الافراد على فقرات الاختبار

-المتوسط الحسابي والانحراف المعياري لتقديرات الافراد

-المتوسط الحساب والانحراف المعياري للاخطاء المعيارية للقياس

-احصائي المطابقة للافراد

-قيم معاملات الصعوبه لفقرات الاختبار

-المتوسط الحسابي والانحراف المعياري لتقديرات قيم معاملات الصعوبة لفقرات الاختبار

-الخطأ المعياري في تقدير قيم معاملات الصعوبة لفقرات الاختبار

-احصائي المطابقة لفقرات الاختبار

-قيم معاملات الصعوبة لفقرات الاختبار

-المتوسط الحسابي والانحراف المعياري لتقديرات قيم معاملات الصعوبة لفقرات الاختبار

-الخطأ المعياري في تقدير قيم معاملات الصعوبة لفقرات الاختبار

-احصاء المطابقة لفقرات الاختبار

نتائج الدراسة:

اولاً الاجابة عن السؤال الاول ونصه

ما دلالات الصدق المتحققة لنموذج الاختبار؟

تم تحقق الصدق بنوعين هما صدق المحتوى من خلال تحليل اهداف المحتوى واعداد جدول المواصفات وتاكد ذلك باحكام الاساتذة المحكمين اللذين اكدوا ملائمة نتائج تحليل اهداف المحتوى ،ومدى تمثيل الفقرات لاهداف المحتوى وقدرتها على قياس هذه الاهداف ،وقد بلغت النسبة المئوية لاتفاق المحكمين (80%) .

أما النوع الثاني فقد تحقق صدق البناء باستخدام برنامج الحاسوبٍ(Spss)(ACER,301) باجراء التحليل العاملي(Factor analysis) للمكونات الرئيسة ،وحساب قيم الجذور الكامنة (Eigen value) للعوامل المستخلصة ونسبة التباين التي يفسرها كل عامل والنسبة التراكمية للتباين المفسر للعوامل التي تزيد قيم جذورها الكامنة عن (الواحد) والجدول (3) يبينالتحليل العاملي للاختبار من استجابات افراد العينة النهائية

لجدول (3)

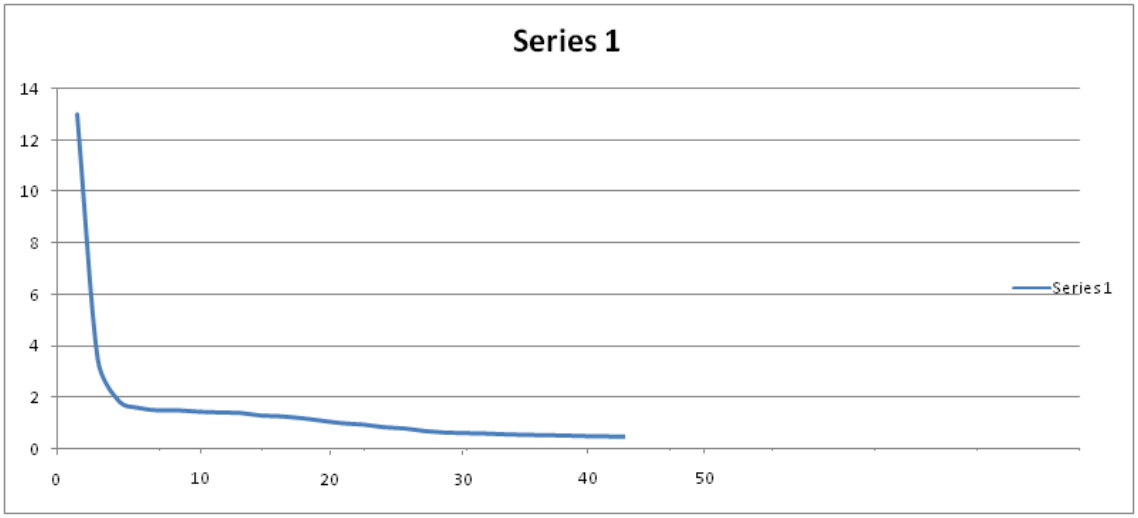
العوامل المستخلصة من التحليل العاملي لفقرات الاختبار وقيم الجذر الكامن ونسبة التباين المفسر لكل عامل،ونسبة التراكمية للتباين المفسر للعينة النهائية

العامل	الجذر الكامن	نسبة التباين المفسر %	النسبة التراكمية للتباين المفسر %
1	8.935	21.351	21.351
2	2.176	5.573	27.081
3	1.962	3.028	30.109
4	1.804	2.745	32.854
5	1.685	2.452	35.306
6	1.427	1.969	37.275
7	1.357	1.843	39.118
8	1.248	1.804	40.922
9	1.162	1.786	42.708
10	1.134	1.656	44.364
11	1.105	1.627	45.991
12	1.076	1.572	49.563
13	1.027	1.565	49.128

تبين من الجدول(3) ان هناك (13) عاملا زادت قيم جذورها الكامنة عن الواحد وفسرت ما نسبته (49.128) من تباين علامات افراد العينة النهائية ،وقد بلغت قيمة الجذر الكامن للعامل الاول (8.935) وهي قيمة مرتفعة عن قيم الجذور الكامنة لبقية العوامل، وان نسبة الجذر الكامن للعامل الاول إلى الجذر الكامن للعامل الثاني (3.227) ،وفسر العامل الاول(21.351) من تباين علامات الاختبار ، اما العامل الثاني فلم يفسر سوى (5.573)

،وهذان المؤشران يؤكدان تحقق افتراض احادية البعد للاختبار بصورته النهائية وفق المعيار الذي حدده (Reckase,1979) إن افتراض احادية البعد للاختبار يتحقق اذا فسر العامل الاول (20%) فاعلى من التباين المفسر لعلامات الاختبار .

كما تم استخدام الرسم البياني لقيم الجذور الكامنة للعوامل المكونة للاختبار كما في الشكل (2)



الشكل (2) التمثيل البياني لقيم الجذور الكامنة للعوامل المكونة للاختبار

يلاحظ من الشكل (2) ان قيمة الجذر الكامن للعامل الاول كبيرة مقارنة بقيمة الجذر الكامن للعامل الثاني كما يلاحظ ان العامل الاول يمثل عاملا سائدا وان هناك تغيراً في ميل خط المنحنى بعد العامل الثاني مما يرجح تحقق افتراض احادية البعد

ثانياً:نتائج الإجابة عن السؤال الثاني ونصه.

ما دلالات ثبات المتحققة للاختبار؟

تم تقدير دلالات الثبات كما يلي :

-تم ايجاد معامل الثبات بالطريقة النصفية وتصحيحه بمعادلة سبيرمان براون فبلغ (0.90) ،وبلغ بطريقة كرونباخ الفا(0.88) .

ثالثا:الاجابة عن السؤال الثالث ونصه:

ما القيم المتحققة لاحصائيات الفقرة للاختبار للعينة النهائية وفق النظرية الكلاسيكية ؟

للاجابة عن هذا السؤال تم اجراء عملية التحليل لبيانات العينة النهائية باستخدام برنامجين للحاسوب هما (Spss.ACER.301) لحساب معاملات الصعوبة والتميز لفقرات الاختبار ،وتم حساب نسبة الطلبة اللذين اجابوا اجابة صحيحة كمؤشر على صعوبة الفقرة ،

وحساب معامل الارتباط الثنائي بين العلامات على الفقرة وعلامات الاختبار كمؤشر على تمييز الفقرة،والجدول (4) يوضح ذلك.

الجدول (4)

معاملات الصعوبة مقدرة بنسبة النجاح على الفقرة، والتمييز مقدرة بمعامل الارتباط الثنائي للعينة النهائية

رقم الفقرة	معامل الصعوبة	معامل التمييز	رقم الفقرة	معامل الصعوبة	معامل التمييز
1	0.30	0.391	24	0.42	0.35
2	0.39	0.363	25	0.42	0.38
3	0.49	0.40	26	0.43	0.40
4	0.40	0.282	27	0.55	0.35
5	0.56	0.36	28	0.59	0.41
6	0.48	0.34	29	0.43	0.43
7	0.44	0.42	30	0.46	0.42
8	0.41	0.40	31	0.43	0.35
9	0.45	0.36	32	0.64	0.40
10	0.66	0.46	33	0.40	0.37
11	0.48	0.41	34	0.39	0.335
12	0.46	0.42	35	0.35	0.41
13	0.41	0.381	36	0.42	0.38
14	0.38	0.37	37	0.57	0.49
15	0.41	0.56	38	0.44	0.31
16	0.34	0.48	39	0.42	0.37
17	0.44	0.39	40	0.39	0.38
18	0.45	0.41	41	0.68	0.54
19	0.46	0.39	42	0.41	0.28
20	0.41	0.47	43	0.54	0.39
21	0.46	0.42	44	0.27.	0.42
22	0.45	0.45	45	0.40	0.44
23	0.43	0.29	46	0.63	
المتوسط الحسابي= 0.423			0.419		

تبين من الجدول (4) ان قيم معاملات الصعوبة للفقرات تتراوح بين(0.26) و(0.68) وبلغ المتوسط الحسابي لمعاملات صعوبة فقرات الاختبار (0.423) وعند اخذ التخمين بعين الاعتبار فان افضل قيمة لمعامل الصعوبه هي (1+فرصة التخمين /2) (عودة ،2001) ، وتساوي هنا (1+0.25 مقسوما على 2) وتساوي(0.625)لان بدائل الاجابه على الفقرات اربعة بدائل، وعليه فان المدى المقبول لمعامل صعوبة الفقرات يمتد من (0.355 إلى 0.895)اضافة او حذف(+،−0.27) للقيمة ، وعليه فان الفقرات التي يقل معامل صعوبتها عن (0.355) تعد فقرات صعبه ، وهي الفقرات (1,16,44) ولم تكن هنالك فقرة يزيد معامل صعوبتها عن (0.90) .

اما قيم معاملات التميز و تراوحت ما بين (0.27,0.56) وبمتوسط حسابي لجميع معاملات التميز بلغ (0.419) وحسب ما اعتمده ايبل (Ebel) المشار اليه في (عودة ،2001) بان معامل التميز المقبول للفقرة يجب ان يزيد عن (0.25)، وعليه فان جميع فقرات الاختبار لها قدرة تميزية مقبوله .

رابعا:الاجابة عن السؤال الرابع ونصه:

ما القيم المتحققة لمعالم الفقرات،والقدرة للاختبار وفق النظرية الحديثة في القياس ؟

تم تقدير قيم معالم فقرات الاختبار من تحليل استجابات افراد عينة الدراسة (220) عن فقرات الاختبار (46) فقرة . وهو عدد مناسب وفق ما اشار اليه فان (Fan,1998) من ان عدد الافراد وفق النموذج الاحادي المعلم (نموذج راش) يجب ان الايقل عن (200) باستخدام البرنامج (ACER,301) ويحذف البرنامج استجابات الافراد اللذين يحصلون على العلامة صفر على الاختبار أو الحاصلين على العلامة التامة (46) على الاختبار، وكذلك الفقرات الذي يفشل جميع افراد العينة في الاجابة عليها جميعا ، وكذلك الفقرات التي ينجح جميع افراد العينة في الاجاية عليها اجابة صحيحة ، ولم يحصل اي فرد على العلامة الكاملة (46) ولا أي فرد حصل على العلامة صفر، وكذلك لم تكون هناك فقرة اجاب عليها جميع افراد العينة اجابة صحية او اجابة خاطئة، ونتيجة لذلك لم تستبعد اجابات اي فرد من العينة وكذلك لم تستبعد اي فقرة من فقرات الاختبار.

ويبين الجدول (5) معاملات صعوبة الفقرات الاخطاء المعيارية في تقدير معلم الصعوبة لكل فقرة من فقرات الاختبار وفق نموذج راش.

الجدول (5)

صعوبة الفقرات والاخطاء المعيارية لها مقدرة باللوجيت و مرتبة تنازليا حسب صعوبة الفقرة للاختبار بصورته النهائية.

رقم الفقرة	معامل الصعوبة	الخطأ المعياري	رقم الفقرة	معامل الصعوبة	الخطأ المعياري
37	1.674	0.109	34	0.209	0.09
8	1.281	0.107	11	0.9151	0.09
46	1.128	0.099	14	0.123	0.09
23	1.007	0.097	36	0.082	0.09
43	1.005	0.097	32	0.051	0.09
10	1.003	0.096	29	0.038	0.09
38	1.00	0.095	31	0.024	0.09
41	0.998	0.094	40	0.016	0.09
20	0.813	0.794	30	0.009	0.09
26	0.819	0.793	9	0.208-	0.09
39	0.783	0.093	35	0.334-	0.09
28	0.776	0.793	7	0.343-	0.091
18	0.711	0.791	1	0.428-	0.091
22	0.706	0.09	13	0.534-	0.091
21	0.572	0.09	16	0.678-	0.091
33	0.534	0.09	2	1.174-	0.092
45	0.520	0.09	4	1.225-	0.092
17	0.425	0.09	44	1.356-	0.092
5	0.325	0.09	25	1.418-	0.093
6	0.298	0.09	19	1.462-	0.099
3	0.251	0.09	12	1.491-	0.112
24	0.231	0.09	15	1.862-	0.114
27	0.223	0.09	42	2.648-	0.172
المتوسط الحسابي		صفر		0.10	
الانحراف المعياري		1.163		0.032	

يلاحظ من الجدول (5) ان مدى قيم صعوبة الفقرات امتد من (−2.648) إلى(1.674) لوجيت وبخطأ معياري يتراوح بين(0.09 إلى0.172)وبمتوسط حسابي يساوي صفر بخطأ معياري (0.1)

وبتطبيق المعيار التي اشارت اليه جمحاوي (2000) من ان الفقرات التي يزيد معامل صعوبتها عن(+1.50) لوجيت تعد ضمن الفقرات الصعبة وان الفقرات التي يقل معامل صعوبتها عن (−1.50) لوجيت تعد من الفقرات السهلة ، وان الفقرات التي تتراوح قيم معامل صعوبتها بين (−1.50،+1.50) هي ضمن المدى المقبول لمعاملات الصعوبة

،وعليه فان الفقرة (37) تعد صعبه لان معامل صعوبتها زادعن(+1.50) وان الفقرتان (15,42) فقرات سهلة لان معامل صعوبه كل منهما زاد عن(−1.50) ،واما بقية الفقرات التي تقع قيم معامل صعوبتها بين (−1.50إلى +1.50) فهي ضمن المدى المتوسط والمقبول بمعاملات صعوبة الفقرات .

وللتحقق من مؤشرات حسن المطابقة لكل فقرة من فقرات الاختبار ومقارنتها بالمدى اللذي يمكن ان يقع ضمنه احصائي المطابقة بتحليل البواقي (FitResid) هو (−2.5 إلى+2.5) www.RUMMlab.com والجدول (6) يوضح ذلك .

الجدول (6)

قيمة احصائي المطابقة لقيم صعوبة فقرات الاختبار

رقم الفقرة	احصائي المطابقة لصعوبة الفقرات	رقم الفقرة	احصائي المطابقة لصعوبة الفقرات	رقم الفقرة	احصائي المطابقة لصعوبة الفقرات
1	0.254−	17	0.752	33	0.156
2	0.325−	18	1.217−	34	1.079−
3	1.65−	19	0.670−	35	0.806−
4	1.18−	20	3.62	36	−1.157
5	2.306	21	2.436	37	0.365
6	1.139	22	1.431	38	0.569
7	2.054	23	2.283	39	1.748−
8	2.108	24	0.376	40	2.891
9	0.913	25	0.734	41	0.987
10	1.954−	26	0.724	42	3.343
11	0.256−	27	1.417	43	1.634
12	0.325	28	2.417−	44	0.401
13	1.337	29	3.424−	45	0.623−
14	0.98	30	1.545	46	0.408
15	0.814	31	0.950−		
16	0.71−	32	0.950−		

يلاحظ من الجدول (6) ان الفقرات(20,29,40,42) تقع قيم مؤشرات المطابقة لها خارج المدى الذي يمكن ان يقع ضمنه إحصائي المطابقة ، لذا تعد هذه الفقرات غير مطابقة لنموذج راش احادي المعلم واما بقية الفقرات فهي مطابقة للنموذج .

كما يعطي النموذج ايضا تقديرات لقدرات الافراد، اللذين طبق عليهم فقرات الاختبار المبنية وفق هذا النموذج ،ويوضح الجدول (7) تقديرات الافراد والمناظرة للعلامة الخام في قياسها للاختبار في صورته النهائية مرتبة تنازليا.

الجدول (7)

قدرات الافراد في قياسها للاختبار بصورته النهائية مرتبة تنازليا

العلامة الخام	تكرار الافراد	قيم القدرة المناظرة للعلامة الخام	العلامة الخام	تكرار الافراد	قيم القدرة المناظرة للعلامة الخام
45	0	3.923	22	8	1.528
44	0	3.224	21	11	0.081
43	1	2.812	20	13	0.144
42	3	2.575	19	9	0.133
41	1	2.240	18	8	0.123
40	1	2.03	17	7	0.096
39	2	1.793	16	4	0.027
38	1	1.579	15	5	0.026-
37	1	1.512	14	6	0.148-
36	1	1.385	13	2	0.186-
35	1	1.267	12	1	0.262-
34	0	1.07	11	5	0.482-
33	2	0.956	10	3	0.674-
32	3	0.86	9	2	0.785-
31	0	0.721	8	1	1.154-
30	15	0.634	7	3	1.836-
29	22	0.567	6	1	1.969-
28	16	0.482	5	2	2.034-
27	18	0.417	4	4	2.096-
26	15	0.364	3	2	2.385-
25	9	0.245	2	3	2.659-
24	7	0.163	1	0	
23	3	0.091			

يتضح من الجدول (7) أن معالم القدرة الفعلية لافراد عينة الدراسة، قد تراوحت بين (2.659-) لوجيت التي يقابل العلاقة (ح)و(2.812) لوجيت التي تقابل العلامة الخام (43)، وكانت أدنى قدرة (2.659-). وتقابل العلامة الخام (2)

كما اظهرت النتائج وجود (7) افراد من العينة غير مطابقين لنموذج راش احادي المعلم لان قيم التقديرات لهم خارج مدى المطابقة (2.5- إلى+2.5) لوجيت.

مناقشة النتائج:

تم التحقق من صدق الاختبار من خلال صدق المحتوى بتحليل الاهداف واعداد جدول مواصفات وصياغة فقرات تمثل المجال السلوكي المحدد في مساق البحث التربوي وتحقق النوع الثاني من الصدق من خلال اجراء التحليل العاملي واستخراج قيم الجذور الكامنة للعوامل المستخلصة ونسبة التباين المفسر لكل عامل والنسبة التراكمية للتباين المفسر ، وتبين ان قيمة الجذر الكامن للعامل الاول بلغ (10.103) وهي اكبر من قيمة الجذر الكامن للعامل الثاني (2.812) ، وقد فسر العامل الاول وحده ما نسبته (22.204) من التباين الكلي للاختبار ، كما بلغت نسبة ما يفسره العامل الاول الى نسبة ما يفسره العامل الثاني (3.948) وهي قيمة قريبة من اربعة اضعاف تقريبا ويستنتج من ذلك ان الاختبار يقيس سمة واحدة و بالاعتماد على ما اشار اليه ريكاسي(Reckase, 1979) بانه اذا فسر العامل الاول اكبر التباين وان تكون القيمة على الاقل 20% فان هذا مؤشر على احادية البعد وبالاعتماد ايضا على مؤشر لورد (Lord,1980) الذي ينص على ان الفقرات تحقق احادية البعد اذا كانت قيمة الجذر الكامن للعامل الاول كبيرة مقارنة بقيمة الجذر الكامن للعامل الثاني وهذا ما أكدته النتائج .

مناقشة النتائج المتعلقة بدلالات الثبات للصورة النهائية للاختبار

فقد تم تقدير دلالات الثبات بطريقتين الطريقة النصفية والتصحيح بمعادلة سبيرمان براون فبلغ (0.89) وباستخدام معامل(الفا كرونباخ فبلغ 0.87) وان القيمة العالية لمعامل الثبات المحسوب ب (KR-20) تؤكد ايضا تحقق احادية البعد حسب رأي كرونباخ والمشار اليه في (Hattie,1985) و(حمادنة ،2009) اذ يرى ان معامل الثبات يعد مؤشرا جيدا لتحقق احادية البعد لانهيمثل متوسط كل من المعاملات النصفية الممكنة اضافة الى ان (KR-20) تمثل القيمة المتوقعة بالنسبة للتباين المفسر من العوامل المشتركة بين الفقرات عند ارتباط عينتين عشوائيتين من تجمع الفقرات الكلي .

اما مناقشة النتائج المتعلقة بمطابقة استجابات العينة النهائية عن فقرات الاختبار لتوقعات النموذج اللوجستي احادي المعلم وفق نظرية الاستجابة للفقرة بان نتائج تحليل استجابات (220) فردا عن (46) فقرة باستخدام برمجية (ACER,301) فقد تم حذف استجابات (7) افراد من العينة النهائية لان قيم مؤشرات المطابقة بقع خارج المدى التي يمكن ان يقع ضمنه احصائي المطابقة ويمكن تفسير ذلك من تنوع تخصصات الطلبة في

الاقسام العلمية والانسانية ، والى غياب الجدية من بعض الطلبة في الاجابة عن بعض الفقرات ولجوء بعض الطلبة من ذوي المستوى المتدني الى التخمين في الاجابة عن بعض فقرات الاختبار .

كما اشارت النتائج ان تقديرات قدرات الافراد تراوحت بين (−2.659) لوجيت والتي تقابل العلامة الخام(2) و (2.812) لوجيت والتي تقابل العلامة (43) وان معظم افراد عينة الدراسة كانت استجاباتهم طابقة لتوقعات النموذج احادي المعلم بنسبة (96.8) اما بنسبة لتقديرات معلم صعوبة الفقرة بالاعتماد على نموذج راش احادي المعلم اشارت النتائج الى ان قيم معلم صعوبة الفقرات تراوحت بين (−2.648)الى (1.674) لوجيت وبمتوسط حسابي قدره صفر لوجيت وبانحراف معياري قدره (1.163) لوجيت كما بلغ متوسط الاخطاء المعيارية(0.1)لوجيت وبانحراف معيار يقدره(0.032)مما يشير الى دقة تقدير صعوبة فقرات الاختبار

التوصيات:

على ضوء النتائج التي توصلت اليها هذه الدراسة يوصي الباحث بالاتي:

1. مناسبة الاختبار المعد وفق نظرية الاستجابة للفقرة لقياس تحصيل طلبة كليات التربية في البحث التربوي
2. −اذ تحقق للاختبار دلالات صدق وثبات مناسبة وفق معايير النموذج اللوجستي احادي المعلم
3. −اجراء دراسة للخصائص السيكومترية للاختبار وفق النموذج اللوجستي ثنائي المعلم وثلاثي المعالم
4. −الاستناد على المنهجية المستخدمة في اعداد الاختبار في اعداد اختبارات لمساقات اخرى والتأكد من خصائصها السيكومترية

المراجع

* المراجع العربية

1. أحمد، عبد الناصر عبد الحليم (2007). **دراسة سيكومترية لبناء مقياس للقبول بالجامعات المصرية باستخدام نظرية الاستجابة للمفردة وعلاقته بالنجاح الأكاديمي**.أطروحة دكتوراه غير منشورة كلية التربية : جامعة الفيوم.
2. التقي، أحمد(2009). **النظرية الحديثة في القياس**، دار المسيرة للنشر والتوزيع والطباعة: عمان، الأردن
3. −جمحاوي،إيناس(2000) .**مقارنة خصائص الفقرات وفق النظريةالتقليدية ونظرية الاستجابة للمفردة في مقياس للقدرة الرياضية**.رسالة ماجستير غير منشورة، جامعة اليرموك ، اربد : الاردن.

4. حمادنة، إياد محمد (2009). استخدام نظرية الاستجابة للفقرة في بناء اختبار محكي المرجع في الرياضيات وفق النموذج اللوجستي ثلاثي المعلم، **مجلة العلوم التربوية والنفسية**، المجلد (10)، العدد (2) ص 216 إلى ص 230
5. الحواري،أروى. وعودة، أحمد (2008). الخصائص السيكومترية لصور مختارة من اختبارات الرخصة الدولية لقيادة الحاسوب في الأردن، **المجلة الأردنية في العلوم التربوية**. مجلد (4)، عدد (4) ص 297 إلى ص 319 .
6. دودين، محمد (2004). استخدام نظرية الاستجابة على الفقرة في تحليل الاستجابات على اختبار المهارات الكمية لدى طلبة الجامعة، **المجلة التربوية**، العدد (72) ص (105) إلى (125)
7. –الشريفين، نضال كمال (2006). الخصائص السيكومترية لاختبار محكي المرجع في القياس والتقويم وفق النظرية الحديثة في القياس التربوي والنفسي،**مجلة العلوم التربوية والنفسية**، المجلد (7) العدد (4) ص 81 الى ص109 .
8. عبابنة، عماد غصاب (2009). **الاختبارات محكية المرجع فلسفتها وأسس تطويرها:** دار المسيرة، عمان، الأردن .
9. عبد العزيز ،شادية (1996).**استخدام نموذج راش في بناء بنك للاسئلة لمقرر في علم النفس التعليمي وتحديد الدرجات الفاصلة المقابلة للتقديرات الجامعية** . رسالة ماجستير غير منشورة ،كلية البنات، جامعة عين شمس،جمهورية مصر العربية.
10. علام، صلاح الدين محمود (1986). **تطورات معاصرة في القياس النفسي والتربوي**، الكويت، جامعة الكويت .
11. علام ، صلاح الدين(2000).**القياس والتقويم التربوي والنفسي** .القاهرة: دار الفكر العربي.
12. –علام ،صلاح الدين(2001) .**الاختبارات التشخيصية مرجعية المحك في المجالات التربوية والنفسية والتدريبية**.القاهرة، دار الفكر العربي .
13. –علام،أحمد (2005).القياس والتقويم في العملية التدريسية. عمان : دار الامل .
14. عودة، أحمد(2001). **القياس والتقويم في العملية التدريسية**، دار الأمل للنشر والتوزيع: اربد، الأردن .
15. هادي، فوزية عباس ومراد،صلاح أحمد (2013). استخدام نموذج راش في إعداد وتقنين اختبار في اللغة العربية لقبول طلاب الدراسات العليا لكليات العلوم الانسانية بجامعة الكويت، **المجلة التربوية العدد** 106 الجزء الاول ص15 الى ص53.

*** المراجع الأجنبية**

1. -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 Rresearch, 36** (4), 523-562.
2. -Crocker,L., &Algina, J.(1986). **Introduction to classical and modern test theory.** New York : CBS college publishing
3. Dena. A. paster (2003). The use of Multilevel Item response Theory Modeling in applied Research:**Applied Measurement in Education**, 16 (3) 223 – 243 .
4. -Embretson,S.,&Reise,S.(2000) **.Item response theory for psychologists.** New Jersey :Lawrence Erlbaum Associates.
5. -Fan,X(1998).Item response theory and classical theory:An empirical comparsion of their item /person statistics**.Educational and psychological Measurement**, 58 357 – 381 .
6. Forbes, D.W.(1986) **The Rasch model asa practical and effective procedure for Educational measurement** :paper presented at Their Annual Meeting of NCME. San Francisco.
7. Gruijter, D., &kamp, L. (2005**). Statistical test theory for education and psychology.**Retrieved December 30, 2005 from: www.leidenunive.nl / gruijterdnmde.

8. Hambleton, R., & Jones, R. (1994). Comparison of classical test theory and item response theory and their applications to test development. **Educational Measurement, 31**(6), 38- 47 .
9. Hambleton, K &Swaminathan, H.,Rogers, H. (1999). Fundamentals **of item response theory.** Newbury Park, CA: sage publication.
10. Hambletone, R. K., &Swaminathan, H. (1985**). Item response theory: principles and applications.** Boston , MA:
11. Hambletons,R.K&Jonse, R.W. (1993). Comparison of classical test theory and Item response theory and their applications to test devel-opment**. Educational Measurement, Issues and practice,4, 38-47 .**
12. Harris, D. (1989). Comarison of 1- , 2-, and 3- parameter IRT models. **Educational measurement: Issue and practice** . 8: 35-41 .
13. Lord, F.M.(1980). **Application of Item Response Theory to Practical Testing Proplems**. New Jersey: Lawrence Eelbaum.
14. Reckase, M. D. (1979). UnifactorIatent trait models applied to multi – Factor tests: Results and implications. **Journal of Educational statistics,** 4. 207-30 .
15. sheng, Y. (2005). **Bayesian analysis of hierarchical IRT models: comparing and combing the unidimensional& multi- unidimensional IRT models.** Unpublished doctoral dissertation, university of Missouri, Columbia.
16. Warm,T.A.(1978) **. Aprimer of Item Response theory Oktahoma.** Oklahoma: U.S. cost Guard Institute 73/ 69.
17. Wiberg, narie (2004). Classical Test Theory VS. **Item Response Theory an Evaluation of The Theory Testin The Scwedishdrivmglicence test.** EM No .50 UmedUniverstiy. www.rummlab .com
18. Wright. D& stone. M (1979). **Best Test Design: A handbook For Research Measurement.** Chicago. MESA Press .

Psychometric Properties of a Criterion-referenced Test in Educational Research According to the Item-Response Theory

Dr. Mohammed AbdulAlmajeed Al-masri

ABSTRACT:

The aim of this study was to construct a criterion-referenced test in educational research, using item-response theory (IRT), applying One-Parameter Logistic Model. The test consisted of (48) MCQ with four alternative items. The subjects of study were (220) students in the Faculty of Education, in Mahweet-Yamen. The results showed that all test items proved the assumption of the model except (4) items, that were deleted. The estimates of difficulty and discrimination parameters were acceptable. Results also showed that (7) subjects were excluded because they were misfit the model. Parameters estimates of subjects ranged between (-2.659) and (2.812) logit.