استخدام مؤشرات ملاءمة المفردة (الداخلية – الخارجية) في تدريج مفردات اختبار تحصيلي في سيكولوجية التعلم وفق نموذج راش على طلبة الفرقة الأولى بكلية التربية جامعة الإسكندرية

د. خالد حسن بكر الشريف*

المستخلص

هدف البحث إلى استخدام نظرية الاستجابة للمفردة في تدريج مفردات اختبار محكي المرجع في مقرر سيكولوجية التعلم وفق نموذج راش ولتحقيق هذا الهدف تم صياغة (٢٠) من المفردات ثنائية الاستجابة ، وروعي في صياغة هذه المفردات أن تقيس تحصيل الطالب في موضوعات المقرر التي تم تدريسها .

وقد تكونت عينة الدراسة من (١٤٢) طالب وطالبة من طلبة الفرقة الأولى شعبة علم النفس بكلية التربية جامعة الإسكندرية وقد أجابوا عن جميع المفردات .

وقد تم تقدير الخصائص السيكومترية للاختبار ببرنامج Winsteps وفق النموذج أحادي البارامتر والذي كان الأنسب لبيانات البحث وأشارت النتائج إلى تمتع الاختبار بمستوى مناسب من الصعوبة وتم استبعاد ٥ مفردات لانخفاض بارامترات صعوبتها حيث تراوحت قيم بارامترات الصعوبة بين (١.٨- ، ٢.٨٧) لوجيت بمتوسط قدره (١.٠١٨٧) وأعطى الاختبار أكبر كمية معلومات عند مستويات القدرة فوق المتوسط وقد تراوحت قيم بارامترات القدرة بين (-٤.٨٧ ، ٥.١٤) لوجيت بمتوسط قدره (١.٩٥٢٠) كما توصل البحث إلى عدة نتائج مهمة في مجال القياس التربوي

وفي ضوء نتائج الدراسة تمت التوصية بمجموعة من التوصيات

الكلمات المفتاحية : ملاءمة المفردة – نموذج راش – اختبار تحصيلي

Using (In-Out)Fit Parameters for calibrating Psychology of Learning Exam Items for Psychology Undergraduate Students in College of Education Alexandria University

Dr.Khaled Hassan Bakr Elsherif

Abstract

The aim of this paper is to use the item response theory in grading the psychology of learning test according to the Rush model.

The study sample consisted of (142) male and female students of the first year, psychology department, Faculty of Education, Alexandria University, and they answered all the vocabulary.

The psychometric characteristics of the test were estimated using Winsteps program according to the one-parameter model, which was the most appropriate for the research data. The results indicated that the test had an appropriate level of difficulty, and 5 items were excluded due to the low parameters of their difficulty, where the values of the difficulty parameters ranged between (1.8-, 2.87) logit with an average of (1.0187). The test gave the largest amount of information at power levels above the average. The values of the power parameters ranged between (-4.87 ،5.14) logit with an average of (1.9520)

In light of the results of the study, a set of recommendations were recommended

Keywords: Item-Fit - Rasch Model – Psychology of learning achievement test

◆ أستاذ مساعد بقسم علم النفس التربوي- كلية التربية- جامعة الإسكندرية.

مقدمة البحث

تعاني النظرية التقليدية في القياس Classical Test Theory (CTT) من انتقادات عديدة في الأوساط التربوية تجعل مصممي الاختبارات التحصيلية يلجأون إلى نظرية السمات الكامنة Latent Traits Theory من أجل تقدير أدق للقدرات التحصيلية موضع القياس ؛ فنظرية السمات الكامنة نظرية حديثة نسبيا مقارنة مع النظرية التقليدية ونماذجها المتعددة تتيح أمام مصممي الاختبارات خيارات أفضل أثناء عملية تدريج الأسئلة سواء في برامج بنوك الأسئلة أو في الاختبارات المباشرة حيث لم تعد الدرجة الكلية للاختبار هي المحك الرئيسي في الحكم واتخاذ القرارات بل تتمتع المفردات Items باستقلالية تجعل الإجابة عليها بطريقة صحيحة معبرة عن قدرات أو سمات موجودة بالفعل ؛ لذلك تهتم نظرية الاستجابة للمفردة Item Response Theory (IRT) بزيادة احتمالات تقديم الاستجابات الصحيحة فنجد مثلا نموذج راش أحادي البارامتر ؛ معادلته الرياضية دالة في متغيرين : قدرة الفرد وصعوبة المفردة وكلما زاد الفرق بينهما كلما زاد احتمال ظهور الاستجابة الصحيحة ؛ وبالتالي فإن التوسع في تصميم الأسئلة وتدريجها باستخدام نظرية الاستجابة للمفردة ونماذجها المتعددة يقلل من ظهور أخطاء القياس ويعطي المزيد من الفرص نحو تحسين أداء الاختبارات لوظائفها حيث هناك شروط للقياس ومحكات موضوعية لاستبعاد الأفراد أو المفردات غير الملائمة أثناء عملية التدريج .

ويشير جاد الرب (٢٠١٦) إلى أن هناك إتجاهات حديثة في تطوير الاختبارات النفسية ؛ منها الاتجاه الذي يعمل على التكامل بين علم النفس المعرفي ، والقياس النفسي معا ويتضح هذا الاتجاه في أبحاث امبرستون (Embretson ، 1996 ، 1998) و أبحاث فريدركسون وميسليفي وبيجار (Frederiksen, Mislevy, and Bejar ، 1993) ، وإتجاه آخر في القياس النفسي يوافق هذا التوجه في التركيز على المفردة بدلا من التركيز على الاختبار ككل ، هذا الاتجاه أصطلح على تسميته نظرية الاستجابة على المفردة (IRT) للتأكيد على أهمية المفردة الواحدة وأن استجابة الفرد على المفردة هي أساس نمذجة العلاقة بين احتمالات استجابة الفرد بشكل معين على المفردة وبارامترات المفردة مثل صعوبتها وقدرتها على التمييز. (جاد الرب ، ٢٠١٦ ، ٤٨٤-٤٨٥)

ويشير "صلاح الدين علام" (٢٠٠٥) إلى أن نظرية الاستجابة للمفردة تحاول نمذجة العلاقة القائمة بين مستوى سمة معينة لدى الفرد Trait Level التي يقيسها اختبار معين ، واستجابته لمفردة من مفردات الاختبار ونظرا إلى أن مستويات السمة غير منظورة بطبيعتها وتؤثر في المتغيرات المنظورة فإنه يتم تقديرها من هذه المتغيرات ، أي من الاستجابات الملاحظة . وعادة يفترض أن السمة التي تنطوي عليها الاستجابات بمثابة قدرة Ability من نوع معين ، أي خاصة تميز الفرد بحيث توجد علاقة منتظمة بين مستويات القدرة لمختلف الأفراد ، واحتمالات إجابتهم على مفردات اختبارية مختلفة إجابة صحيحة.

وعلى الرغم من أن هذه العلاقة كانت توصف استنادا إلى النموذج الاعتدالي التراكمي Normal Ogive Model وذلك في العقدين الثالث والرابع من القرن العشرين ، إلا أنه في الإطار المعاصر لنظرية الاستجابة للمفردة أصبح وصف هذه العلاقة يستند إلى نموذج دالة الترجيح اللوغاريتمي Logistic Function وعلى العكس من نظرية القياس التقليدية فإن نظرية الاستجابة للفردة تستند إلى افتراضات قوية فيما يتعلق بما يحدث عند استجابة الفرد لمفردات اختبارية. وتؤدي هذه الافتراضات إلى ميزات متعددة. فمثلا يمكن وصف المفردة الاختبارية بطريقة مستقلة عن أي عينة من المفردات التي اختبر بها ، وكذلك يمكن التنبؤ بخصائص الاختبار قبل تطبيقه . (علام ، ٢٠٠٥ ، ٥٣-٥٤)

وتعد مسألة التحقق من الثبات والصدق ضرورية في مجال مقاييس التربية و علم النفس ، بهذا المعنى ، كانت نظرية الاختبار الكلاسيكي (CTT) تخدم مجال الاختبار على نطاق

واسع لنحو ١٠٠ عام. ومع ذلك ، فإن تطبيق نظرية الاستجابة للمفردة (IRT) للتقييم النفسي والتربوي قد تسبب في تغييرات كبيرة وإيجابية في تطوير الاختبارات النفسية.

وتتميز نظرية الاستجابة للمفردة بمميزات عديدة مقارنة مع نظرية الاختبار الكلاسيكي حيث توفر بارامتر لقياس قدرات الأفراد ؛ وتوجد استقلالية في القياس فلا تعتمد درجة مفردة على درجة مفردة أخرى وهناك دقة في القياس وتدريج للمفردات من حيث الصعوبة وهناك طرق للتحقق من ملاءمة النموذج المستخدم لبيانات الأفراد بموضوعية Zanon et al(2016)

نظرية الاختبار الكلاسيكي

إن بيانات الاختبارات المستمدة من نظرية الاختبار الكلاسيكي تعتمد بشكل كبير على خصائص العينة التي تم تطبيق الاختبار عليها ، وتعتمد درجات المتحنين على اختيار مفردات الاختبار ؛ ولكن يمكن أن ينتج اختبار أفضل إذا كانت مفردات الاختبار مطابقة لمستويات قدرة المتحنين ؛ وبالتالي فإن المتحنين الذين يملكون قدرات أكبر يجيبون مفردات أكثر صعوبة من أقرانهم كذلك تفسير درجات المتحنين سيعتمد على مدى صعوبة المفردات فلو كان اختبار ذكاء سهل سيبدو المستجيبون أكثر ذكاء في حين لو كان الاختبار أكثر صعوبة سيبدو المستجيبون أقل ذكاءا

المشكلة الثانية في نظرية الاختبار الكلاسيكي هي أن خصائص الاختبار تعتمد على العينة فنفس الاختبار لو تم تطبيقه على ذوي مستويات قدرات مرتفعة لو أعيد تطبيقه على ذوي مستويات قدرات أقل سينتج مستويات صعوبة مختلفة . حيث ستظهر صعوبة المفردات في العينة الأولى أقل من صعوبة العينة الثانية . ومن ثم فإنه لا يمكن دراسة خصائص الاختبار السيكومترية إلا على نفس العينة ؛ في حين أن نظرية الاستجابة للمفردة تعتمد فيها بارامترات الاختبارعلى السمات الكامنة للأفراد وبالتالي ستتغير خصائص المفردات والاختبار عندما يجيب عليه أشخاص آخرون (عينات بمستويات مختلفة من السمات الكامنة) على الاختبار.

والمشكلة الثالثة في نظرية الاختبار الكلاسيكي هو أن النظرية تفترض أن أخطاء القياس متساوية لكل الأفراد ؛ وهذا الأمر يمثل مشكلة لأن الأفراد الذين لديهم مستويات مختلفة من القدرة سيظهرون كذلك مستويات مختلفة من الأخطاء (كالتخمين والغش و.....إلخ) والمشكلة الرابعة في نظرية الاختبار الكلاسيكي هو أنها لا تسمح بالتنبؤات الدقيقة حول النتائج المحتملة للمستجيب أو لعينة على مفردة ما بدلالة درجات قدراتهم فقط في حين أن هذه المعلومات التي توفرها نظرية الاستجابة للمفردة مهمة لمصمم الاختبار كي يتخذ قرارات صحيحة بشأن درجات الاختبار.

ويرى معظم خبراء القياس من أمثال لورد Lord ونوفيك Novick أن النظرية الحديثة في القياس تعد امتدادا وتوسيعا لمفاهيم ومبادىء النظرية الكلاسيكية ؛ لذلك اشتقا كثيرا من مبادىء النظرية المعاصرة من مبادىء النظرية الكلاسيكية ، غير أنه تبين بعض مبادىء القياس التي كان متعارفا عليها لم تعد ملائمة أو ضرورية ، وتحتاج إما إلى مراجعة أو إلى استبدال . فكثير من مبادىء النظرية المعاصرة تختلف اختلاف جوهري عن المبادىء الكلاسيكية ومن بين الاختلافات ما يلي : (Lord and Novick,1978; Douglas, 1990; Embertson and Reice,2000)

-أساليب تقييم خصائص المفردات الاختبارية :

تعتمد النظرية الكلاسيكية في تقييم خصائص مفردات الاختبار على قيم معامل الصعوبة ، أي نسبة عدد الأفراد في جماعة مرجعية محددة الذين يجيبون على المفردة إجابة صحيحة . وكلما زادت هذه النسبة دل ذلك على سهولة المفردة ، وكلما قلت دل ذلك على صعوبة المفردة وكذلك تعتمد على قيم معامل التمييز الذي يتم تقدير قيمته باستخدام معامل الارتباط الثنائي المتسلسل أو غيره من الأساليب الإحصائية ، اعتمادا على درجات أفراد الجماعة المرجعية المعينة. وتختلف قيم معامل الصعوبة ومعامل التمييز اختلافا ملحوظا باختلاف عينة

الأفراد ، وبخاصة إذا لم تكن هذه العينة عشوائية أو ممثلة لمجتمع هؤلاء الأفراد . بينما تكون قيم هذين المعاملين مستقلة عن عينة الأفراد التي استخدمت في تقديرها في إطار بعض نماذج نظرية الاستجابة للمفردة ، إذا افترضنا وجود مجتمع كبير من الأفراد.

-تفسير درجات الاختبار : تعتمد النظرية الكلاسيكية في تفسيرها لدرجات الاختبار على معيار الجماعة Norm Group، وذلك بتحويل الدرجات الخام في الاختبار إلى درجات معيارية أو بأي تحويل خطي استنادا إلى التوزيع الاعتدالي. بينما تعتمد نظرية الاستجابة للمفردة على المقارنة بين مستويات قدرة الأفراد ، وصعوبة المفردات على ميزان مشترك Common Scale. فصعوبة المفردات ، ومستويات قدرة الأفراد يتحدد موقعها على متصل واحد . وبذلك يمكن تفسير الفرق بين مستوى قدرة الفرد وصعوبة المفردة تفسيرا مباشرا . فإذا كانت درجة قدرة الفرد مساوية لصعوبة المفردة فإن احتمال الإجابة الصحيحة على المفردة = (٠.٥) وهذا يعني أن تفسير الدرجات يكون بالإشارة إلى المفردات الاختبارية ، وليس إلى الجماعة المرجعية كما هو الحال في النظرية الكلاسيكية . ومع هذا يمكن أيضا إجراء تحويلات خطية معينة على الدرجات بحيث يتيسر تفسير الدرجة في إطار الجماعة المرجعية.

-تحقق خصائص الميزان الفتري

تعتمد النظرية الكلاسيكية في انتقاء مفردات الاختبارات على فرضية أن درجاتها تتخذ شكل التوزيع الاعتدالي في المجتمع المستهدف. ويمكن أن يتحقق ذلك بانتقاء المفردات التي تكون معاملات صعوبتها قريبة من (٠.٥) وإذا لم يتحقق ذلك تستخدم أساليب إحصائية لتحويل التوزيعات المختلفة إلى توزيعات اعتدالية ، وبذلك يمكن تبرير تحقق خصائص الميزان الفتري في إطار افتراضات معينة فيما يتعلق بتوزيع درجات الاختبار الحقيقية في المجتمع المستهدف. فهذا التوزيع يفترض أنه اعتدالي ، وأن هذه الدرجات تفي بخصائص الميزان الفتري. غير أن هذا يجعل خصائص الميزان متعلقة بمجتمع محدد من الأفراد ، فإذا ما طبق الاختبار على أحد أفراد مجتمع آخر ربما يصعب تبرير تحقق خصائص الميزان الفتري في هذه الحالة ، وهذا يعد أحد أوجه القصور في النظرية الكلاسيكية في القياس ، لكن بعض نماذج نظرية الاستجابة للمفردة مثل نموذج راش Rasch Model، يمكن أن يؤدي إلى تحقق خصائص الميزان الفتري وربما الميزان النسبي دون ضرورة أن يكون توزيع مستويات القدرة في المجتمع المستهدف اعتداليا حيث أن النموذج يستند إلى مبادىء أساسية في القياس Fundamental Measurement ومقارانات تتصف بعدم التغير Invariance ، مما يعد أحد ميزات هذه النماذج . (علام ، ٢٠٠٥ ، ٥٤-٥٥)

ومنها مايلي :-

١-**خصائص الخطأ المعياري للقياس** تعتبر النظرية الكلاسيكية في القياس أن الخطأ المعياري Standard Error ثابت بالنسبة لمجتمع معين من الأفراد المختبرين . ولكن نظرية الاستجابة للمفردة لا تفترض تساوي تباين أخطاء القياس لجميع الأفراد الذين يطبق عليهم الاختبار. فبعض الأفراد يكون أداؤهم في الاختبار أكثر اتساقا من غيرهم من الأفراد ، ويختلف هذا الاتساق باختلاف مستوى قدرتهم فبدلا من تقديم خطا معياري واحد للقياس ينطبق على جميع المختبرين بغض النظر عن درجاتهم في الاختبار ، تجعل نظرية القياس المعاصرة من الممكن تقديم تقديرات منفصلة للخطأ المعياري لكل فرد ، ولكل مستوى قدرة معين ؛ لذلك فإن الخطأ المعياري للقياس في هذه الحالة يتميز بإمكانية التعميم على مجتمعات من الأفراد مختلفي القدرة. ويمكن إيجاد متوسط الأخطاء المعيارية الذي يعد مؤشرا لهذه الأخطاء في مجتمع هؤلاء الأفراد.

٢-**تأثر الثبات بعدد مفردات الاختبار** فالنظرية الكلاسيكية للقياس تعتبر أنه كلما زاد عدد مفردات الاختبار ارتفع ثبات درجاته . ويمكن التنبؤ بقيمة الثبات عندمايزداد أو يقل عدد مفردات الاختبار باستخدام الصيغة الرياضية التي اقترحها سبيرمان وبروان Spearman-Brown أما نظرية الاستجابة للمفردة فإنها تؤكد أن الاختبارات القصيرة التي تشتمل على نوعية جيدة

من المفردات يمكن أن تكون أكثر ثباتاً إذا افترضنا أن تمييز مفرداتها ثابت ، كما تؤكد أن الدرجات المتطرفة تكون أخطاؤها المعيارية أكبر (علام ، ٢٠٠٥ ، ٥٦)

مشكلة البحث

يعتمد معظم الباحثين ومصممو الاختبارات النفسية والتربوية على نظرية الاختبار الكلاسيكي ؛ حيث يتم اختيار المفردات ذات معاملات الصعوبة المناسبة والدلالة التمييزية العالية ؛ بالتالي هناك قلة في الاختبارات النفسية والتحصيلية المصممة وفق النظرية الحديثة في القياس والتي تعتمد على قدرة الفرد عند تقييم مفردات الاختبار كمحك أساسي للحكم على ملاءمة هذه المفردات فقياسها لقدرات الأفراد وتحقيقها للهدف الذي صممت من أجله هو الأساس الذي يتم تقييمها عليه وبالتالي أخذ قرار استبعادها أو الإبقاء عليها في الاختبار

وفي هذا الإطار تشير نتائج دراسة الزبيدي (٢٠١٨) إلى أهمية قيام المتخصصين في علم النفس والقياس النفسي والتربوي بتصميم اختباراتهم وفق نماذج نظرية الاستجابة للمفردة حيث يمكن الارتكان لعدة مؤشرات للثبات مثل معامل الفصل ومعامل ثبات تقدير القدرة (الأفراد) و معامل ثبات تقدير صعوبة الاختبار (المفردات) وفي نفس الإطار يشير الخياط (٢٠١٢) إلى تفوق نماذج نظرية الاستجابة للمفردة على نظرية الاختبار الكلاسيكي فالنماذج الكلاسيكية تستبعد المفردات إذا كان معامل تمييزها منخفض أما نماذج نظرية الاستجابة للمفردة فتستبعد المفردات ذات معامل التمييز المتطرف (المرتفع أو المنخفض) ؛ ولذلك كانت المحكات التي يتم على أساسها استبعاد المفردات من الصورة النهائية للاختبار مثار جدل بين المتخصصين في القياس التربوي ولحسم هذا الجدل يفضل المتخصصون استخدام أكثر من محك واكثر من مؤشر كي يكون الباحث مطمئنا لقراره.

ويؤكد كانج وشين Kang and Chen (2008) إلى وجود قصور في المؤشرات المستخدمة لقياس مدى ملاءمة المفردات في الاختبار المصمم لنموذج راش ؛ أن بعض المؤشرات مازالت حساسة لحجم العينة وطول الاختبار وهي أمور تحررنا منها في ضوء نماذج نظرية الاستجابة للمفردة IRT حيث أنه وجد اختبار مكون من عشر مفردات ثنائية الاستجابة أظهرت هذه الإحصائيات التقليدية معدلات خطأ تجريبية من النوع الأول متضخمة تصل إلى ٠.٩٦ بينما اختيار واستخدام اختبارات جودة ملاءمة مناسبة لطرق تصميم المفردات ونماذج تحليلها من شانه أن ينتج معدلات خطأ تجريبية من النوع الأول في أدنى مستوياتها (ما بين ٠.١٤٢ - ٠.٣٠٤) في حالة المفردات ثنائية الاستجابة. وتؤكد ذلك نتائج دراسة لاهويس وأخرين LaHuis et al (2011) حيث أشارت إلى أهمية استخدام أكثر من مؤشر لملاءمة المفردات الاختبارية ؛ وأن مربع كاي يصلح كاختبار لجودة الملاءمة وكذلك مربع كاي مطروحا منه الانحراف المعياري و مربع كاي المعدل بدرجات الحرية ، وأن خطأ النوع الأول يتعلق باستخدام نموذج تدريج غير مناسب من نماذج الاستجابة للمفردة وهذا يوضح كثرة البدائل التي توفرها بيانات التحليل تبعا لنظرية الاستجابة للمفردة فبدلا من الارتكان إلى مؤشر واحد للخصائص السيكومترية لكل العينة توفره النظرية الكلاسيكية في القياس ؛

ويمكن تلخيص مشكلة البحث في الأسئلة التالية

١-ما درجة ملائمة البيانات المستمدة من تطبيق اختبار سيكولوجية التعلم الفرقة الأولى للتدريج بنموذج راش أحادي البارامتر؟

٢- ما مناسبة مدى مؤشرات الملاءمة (الداخلية - الخارجية) لكل مفردة في اختبار سيكولوجية التعلم المدرج وفق نموذج راش ؟

أهداف البحث

١-الكشف عن درجة ملاءمة بيانات اختبار تحصيلي في سيكولوجية التعلم لافتراضات نموذج راش

٢- تطوير اختبار تحصيلي في مقرر سيكولوجية التعلم لطلبة الفرق الأولى علم النفس ومدرج وفق نموذج راش

٣-تفسير مؤشرات ملاءمة مفردات الاختبار المصمم وفق نموذج راش باستخدام أكثر من مؤشر -ZSTD MNSQ-

أهمية البحث

في الجانب النظري ؛ ان استخدام نموذج راش أحادي البارامتر يحقق موضوعية القياس ؛ وذلك لأن تدريج بارامترات المفردات التي تعبر عن أداء الأفراد على مقياس السمة الكامنة يتم بشكل مستقل عن خصائص العينة وهي نقطة خلاف رئيسية مع النظرية الكلاسيكية في القياس CTT وذلك لأن تحليل مفردات أي مقياس وفق النظرية الكلاسيكية يعطي قيم غير ثابتة لمعاملات الصعوبة والتمييز ؛ وبالتالي فإن هذه القيم تختلف باختلاف العينات ، ناهيك عن التقديرات غير الثابتة لأداء الأفراد .

وأساتذة ومعلمي مقررات علم النفس في حاجة إلى تصميم اختبارات محكية سليمة التدريج ودقيقة في قياس القدرات التحصيلية فهو التخصص الذي تنتمي إليه نظريات القياس ولا يصح أن يصمم أساتذة ومعلمو علم النفس العام بصفة عامة وعلم النفس التربوي والقياس النفسي بصفة خاصة اختبارات وفق النظرية الكلاسيكية في القياس دون التحقق من ملاءمة مفرداتها وسلامة تدريجها ومدى تحقيقها للأهداف التي صممت من أجل قياس تحققها وهو ما توفره لهم نظرية الاستجابة للمفردة

تنبع أهمية دراسة درجة ملاءمة مفردات الاختبارات النفسية والتربوية وفق نظرية الاستجابة للمفردة من منطلق أن نجاح عملية تدريج هذه المفردات ينبىء بمدى النجاح في قياس مستويات الأداء المستهدفة ؛ ويشير إلى مدى تمكن أساتذة علم النفس من مهارات القياس النفسي ، ومن ناحية أخرى فإن انخفاض مستويات الطلبة التحصيلية. يمكن تفسيره بثقة إلى ضعف امتلاكهم للقدرات المقاسة وليس لأي عوامل أخرى مثل أخطاء القياس

ويشير الشريف (٢٠١٨) إلى أن عدم اعتماد بعض مدرسي علم النفس في اختباراتهم الجامعية على نظرية الاستجابة للمفردة يعود إلى أسباب تعزى إلى ضعف مهارات تدريج الاختبارات وفق نماذج هذه النظرية ، وغياب ثقافة تصميم بنوك الأسئلة وتدريجها وهذا الأمر يجب أن ينتهي خاصة في ظل انتشار وحدات القياس والتقويم في مختلف الكليات الجامعية ضمن مشروعات التعليم العالي الأمر الذي يجعل التطبيق والتصحيح على عينات كبيرة سهل نسبياً مقارنة بالماضي .كما تكمن أهمية البحث في كونه سيكشف أسباب استبعاد بعض المفردات الاختبارية لإعادة النظر في طرق تصميم المفردات الاختبارية في اختبارات علم النفس بصفة عامة وعلم النفس التربوي بصفة خاصة من أجل تطوير الواقع وتحسينه وحل ما قد يوجد من مشكلات قد تؤثر على تصنيف الطلبة وتوجيههم إلى التخصصات المناسبة لقدراتهم المقاسة.

والجامعات المصرية في حاجة إلى تصميم المزيد من المقاييس والاختبارات التحصيلية وفق نماذج نظريات القياس الحديثة وفي مقدمتها نماذج نظرية الاستجابة للمفردة الأحادية والمتعددة التي تحدد المستويات التحصيلية المقاسة بدقة أكبر من نظرية الاختبار الكلاسيكي CTT في ظل جهود هذه الجامعات نحو الحصول على الاعتماد المؤسسي والاعتماد البرامجي .

وتكمن أهمية البحث كذلك في تناوله متغير ملاءمة المفردات الاختبارية Item Fit، الداخلية والخارجية بأكثر من مؤشر حيث يعد دراسة سلوك (المفردات / الأفراد) بدقة سبيل للاطمئنان إلى سلامة التدريج وتحقيق افتراضات النموذج المستخدم وهو في هذا البحث نموذج

راش ، وتحديد العلاقة الاحتمالية بين بارامترات (القدرة / الصعوبة) يسهم في تحديد دقيق للقدرات المقاسة

تطبيق افتراضات نظرية الاستجابة للمفردة وتحديداً نموذج راش فيما يتعلق بتوضيح أثر خصائص (المفردة / الفرد) على دقة تقدير البارامترات في اختبار تحصيلي في مقرر سيكولوجية التعلم.

-زيادة قدرة مصممي اختبارات علم النفس على تحديد الطريقة التي بواسطتها يمكن اختيار المفردات المناسبة في الاختبارات السيكولوجية ، وذلك للحصول على دقة أكبر في تقدير قدرات الأفراد (وصولا إلى نشر ثقافة بنوك الأسئلة التي تتيح سهولة استخدام عدد كبير من المفردات المدرجة والممثلة للأهداف التي وضعت من أجل قياسها)

-توفير اختبار تحصيلي في علم النفس التربوي بعد تدريجه وفق نموذج راش يصلح للاستخدام في البيئة المصرية.

الإطار النظري

مع ظهور مشكلات نظرية الاختبار الكلاسيكي تطورت عدة نماذج لنظرية الاستجابة للمفردة منها نموذج راش وهو نموذج أحادي البارامتر أو المعلم (بارامتر صعوبة المفردة) ، ويفترض توازي المنحنيات المميزة لجميع مفردات الاختبار ، ويعد نموذج راش أبسط نماذج نظرية الاستجابة للمفردة وأكثرها استخداما في التطبيقات التربوية ، وفي مجال بناء الاختبارات التحصيلية والعقلية وتحليل مفرداتها . كما يعد أكثر النماذج التي أجريت حولها دراسات متعددة للتحقق إمبريقيا من خصائصه وجدواه في بناء الاختبارات ، ومواجهة أوجه قصور المدخل الكلاسيكي في القياس التربوي والنفسي.

والاختبارات السيكولوجية والتربوية التي يعدها أساتذة ومعلمو علم النفس تفترض وجود سمات أو خصائص معينة يشترك فيها جميع الأفراد، ولكنهم يختلفون في مقدارها، وعلى الرغم من أن هذه السمات غير منظورة ، إلا أنه يمكن الاستدلال على مقدارها من السلوك الملاحظ للفرد الذي يتمثل في استجابته لمفردات الاختبار ، وهذا ما قد يفسر تسميتها بنظرية السمات الكامنة Latent Traits Theory . فالسمة التي تكمن وراء استجابة الفرد لمفردات اختبار لفظي تختلف عن السمة التي تكمن وراء استجابته لمفردات اختبار عددي أو مكاني. ولكن يمكن أن تكمن سمة واحدة وراء استجابته لمفردات اختبارين مختلفين متعلقين بالمحتوى نفسه. لذلك فإن الهدف الأساسي لكلا النوعين من النماذج السيكومترية تحديد العلاقة بين استجابات الأفراد لاختبار معين ، والسمة الكامنة وراء هذه الاستجابات.

وتتفق كل من النماذج الكلاسيكية ونماذج الاستجابة للمفردة الاختبارية في افتراض وجود متصل للسمة ، وأنه يمكن تقدير احتمال إجابة فرد إجابة صحيحة عن مفردة اختبار إذا علمنا موقعه على هذا المتصل. وبتعبير رياضي فإن الاختبار يعد بمثابة راسم فوقي (دالة) Mapping من عينة الأفراد المختبرين إلى متصل السمة الكامنة. كما تتفق هذه النماذج في أن مقدار هذا الاحتمال يكون دالة متزايدة مطردة لمواقع الأفراد على متصل السمة. وهذا يعني أن احتمال توصل الفرد إلى الإجابة الصحيحة على مفردة اختبار يزداد بازدياد مقدار السمة لديه. لكن نماذج الاستجابة للمفردة تختلف عن النظرية الكلاسيكية في القياس في تصورها لشكل هذه الدالة. فبينما تفترض النماذج الكلاسيكية أن هذه الدالة تتخذ شكل المنحنى الاعتدالي في الاختبارات مرجعية المعيار ، وتتخذ شكلا ملتويا التواءً سالبا في الاختبارات مرجعية المحك ، تفترض نماذج السمات الكامنة أنها تتخذ شكل المنحنى اللوغاريتمي ، وتسمى هذه الدالة في الحالتين بالدالة المميزة للمفردة Item Characteristic Curve . (علام ، ٢٠٠٧ ، ٢١٤)

نماذج الاستجابة للمفردات الاختبارية

تهدف نماذج السمات الكامنة غلى تحديد علاقة بين أداء الفرد في اختبار معين وبين السمات أو القدرات التي تكمن وراء هذا الأداء وتفسره ، لذلك فإن هذه النماذج هي دوال رياضية

احتمالية.وتنقسم هذه النماذج إلى مجموعتين إحداهما : النماذج الاستاتيكية ، والأخرى النماذج الديناميكية. وتهتم النماذج الاستاتيكية بالقياس في مدة زمنية واحدة ، كما تهتم بتحديد العمليات التي ينطوي عليها الأداء في الاختبارات السيكلوجية والتربوية. اما النماذج الديناميكية فتهتم بمشكلة قياس التغير في السمات النامية المختلفة في فترات متباعدة. غير أن النماذج الاستاتيكية هي التي نالت كثيراً من الاهتمام في مجال الاختبارات مرجعية المحك . ومن هذه النماذج :

نموذج راش : يتضمن نموذج راش واقعاً سيكولوجياً في صورة رياضية . فالمعادلة تفسر احتمال أن يجيب فرد إجابة صحيحة على مفردة اختبار. فكلما زادت قدرة الفرد زاد احتمال إجابته إجابة صحيحة على مفردة اختبار وأن هناك احتمالا كبيرا أن يجيب أي فرد إجابة صحيحة على مفردة تتميز بالسهولة بالنسبة له . فإذا كانت هاتان القضيتان اللتان تقرران ديناميكية التفاعل بين الأفراد ومفردات الاختبارات صحيحتين ، فإنه يصبح نموذج راش الذي تمثله المعادلة الرياضية تعبير واقعي عن هذا التفاعل ويفسر النموذج بذلك أداء الأفراد في الاختبارات والمقاييس . فإذا حققت مجموعة من البيانات المستمدة من هذه الاختبارات متطلبات نموذج راش ، فإن هاتين القضيتين يجب أن تكونا صحيحتين . أما إذا لم تحقق البيانات هذه المتطلبات ، يكون معنى هذا أن البيانات تناقض الأساس المنطقي الذي بني عليه النموذج.

وهذا وتعني المطابقة أو الملاءمة Fit لأي مفردة من مفردات اختبار ما أنه لم تطابق النموذج حينئذ لا يكون صحيحا أن الفرد الأكثر قدرة يحتمل بدرجة أكبر أن يجيب على المفردة إجابة صحيحة من الفرد الأقل قدرة . لذلك يجب التساؤل عن سبب ذلك أي يجب أن يكون السؤال: ماالذي تقيسه هذه المفردة؟ وبالمثل إذا لم يطابق الفرد نموذج راش فإنه لا يكون صحيحا أيضا أن احتمال إجابته إجابة صحيحة على مفردة تتميز بالسهولة أكبر من احتمال إجابته إجابة صحيحة على مفردة تتميز بالصعوبة. وهنا التساؤل يكون عن سبب حدوث ذلك أي عما حدث لهذا الفرد أثناء إجابته على المفردة. إذ ربما لم تميز المفردة بين مستويات قدرة الأفراد ، وبذلك لا تصلح لقياس هذه القدرة (علام ، ٢٠٠٧ ، ٢٤٢) مماسبق يتضح أن النموذج الذي لا يساعد في التفسير السيكولوجي للسمات المراد قياسها لا يصلح للقياس النفسي والتربوي وهذا يوضح أيضا أهمية نموذج راش فهو نموذج رياضي يشتمل على أقل عدد ممكن من البارامترات ، وفي الوقت نفسه له أساس سيكولوجي

مفهوم ملاءمة المفردات (الداخلية-الخارجية) لنموذج راش Items Fit (In-Out)

يقصد بالملاءمة (أو المطابقة) لأي مفردة من المفردات نوعين: التقاربية والتباعدية أو الداخلية والخارجية وهي التي تقاس بمؤشرات معينة من البرامج المستخدمة فمثلا برنامج Winsteps يوفر مؤشرين مؤشر MNSQ ومؤشر ZSTD والمفردة الملاءمة داخلياً هي المفردة التي يجيب عليها إجابة صحيحة ذوي القدرة المرتفعة ، ويجيب عليها إجابة خاطئة ذوي القدرة المنخفضة ، ويشير خضر (٢٠٠٧) إلى أن مؤشر الملاءمة الداخلية للمفردات يشير إلى مدى تقارب البيانات من النموذج Infit Mean Square بينما مؤشر الملاءمة الخارجية يشير إلى مدى تباعد البيانات من النموذج Outfit Mean Square. (خضر ،٢٠٠٧، ١٥٥)

والنموذج الذي يعتمد على الدرجة الثنائية هو واحد من مجموعة نماذج إحصائية أطلق عليها راش اسم نماذج القياس Measurement Models وقد طور هذا النموذج بطريقة مستقلة عن غيره من نماذج الاستجابة للمفردة ويمكن اعتباره نموذجا فيه المنحنى المميز للمفردة تمثله دالة ترجيح لوغاريتمي أحادية البارامترات One-Parameter Logistic Function

ويشير كانج وشين (2008) Kang and Chen إلى أن استخدام نموذج راش في تدريج الاختبارات التحصيلية ذات المفردات الثنائية (الإجابة لها احتمالان ١ إجابة صحيحة ، و٠ إجابة خاطئة) قد سمح باستخدام عدة مؤشرات للتحقق من جودة ملاءمة المفردات للنموذج حيث ؛ استخدم اختبار مربع كاي وهو اختبار تقليدي يصلح في حالة المفردات الثنائية واستخدم مؤشر

آخر وهو الانحراف المعياري مطروحاً منه مربع الدرجة s-x2 وقد أظهرت الاختبارات التي أجريت عليه إلى زيادة في احتمالية ظهور خطأ من النوع الأول ؛ لذلك أكدت النتائج أن هذا المؤشر له نتائج واعدة ولكن في حالة المفردات المتعددة (الإجابات لها أكثر من احتمالين)

ملاءمة المفردات الاختبارية للنموذج

يتم التحقق من ملاءمة مجموعة المفردات التي تشكل الاختبار كوحدة واحدة لنموذج معين ، وكذلك التحقق من ملاءمة كل مفردة على حدة للنموذج ، وبخاصة إذا كان الاختبار يشتمل على أنواع مختلفة من المفردات التي تتطلب نماذج مختلفة لتقدير بارامتراتها ويقصد بملاءمة المفردات لنموذج معين إمكانية النموذج تفسير كيفية استجابات الأفراد لإحدى المفردات ، أو التنبؤ بهذه الاستجابات.

ملاءمة المفردات الاختبارية لنموذج راش : يعد نموذج راش Rasch Model أحد نماذج تحليل المفردات الاختبارية الذي تكون فيه المنحنيات المميزة للمفردات متساوية الميل ، أي أن القوة التمييزية للمفردات ثابتة .

وقد تم اقتراح عدة مقاييس ومؤشرات لتقييم ملاءمة المفردات لنموذج راش ، وقد اقترح أندرسون Anderson (1973) مقياسا لحسن الملاءمة لنموذج راش اعتمادا على المقارنة بين قيم بارامترات الصعوبة التي يتم تقديرها من مجموعات فرعية مختلفة ، وقيم هذه البارامترات المستمدة من المجموعة ككل. ومن ثم التوصل إلى نسبة الأرجحية المشروطة Conditional Likelihood Ratio وقد وجد أن ضعف لوغاريتم هذه النسبة يتخذ شكل توزيع مربع كاي عندما يكون نموذج راش صحيحا.

والقيم المرتفعة لمربع كاي تساعد في تشخيص المفردات التي تسلك بطريقة متمايزة Differential Functioning بالنسبة للنموذج (أي أن المفردات لا تلائم النموذج في هذه الحالة) ، ويمكن عندئذ تحليل أي مفردة اختبارية غير ملائمة للنموذج للتعرف على تحيزها فيما يتعلق ببعض خصائص الأفراد ، مثل : النوع والخلفية الثقافية ، وذلك بإيجاد انحدار البواقي الخاص بالمفردة على هذه المتغيرات وإجراء التعديلات اللازمة على هذه المفردة

وقد أجرى سميث Smith (1988) مقارنة بين فاعلية مقياس إحصائي لإجراء ملاءمة بين المفردات الاختبارية Between-Item Fit Statistics لنموذج راش ، من خلال ضم الدرجات الخام لكي تشكل مجموعات مختلفة من مستويات القدرة المقاسة ، وكذلك مقياس إحصائي لإجراء ملاءمة عامة Total Fit Statistics للنموذج. وقام بدراسة خصائص هذين المقياسين وقدرتهما على كشف مختلف أنواع أخطاء القياس المنتظمة والعشوائية.

وقد اقترح روست ، ودافير (Rost and Davier,1994) مؤشرا للتحقق من ملاءمة المفردات المشروطة Conditional Item- Fit Index لنماذج راش للمفردات ثنائية الاستجابة ، ومتعددة الاستجابات المرتبة . واعتمد هذا المؤشر على الأرجحية المشروطة لنمط المفردة ، ولا يعتمد ذلك على بارامترات المفردة ، او على بواقي الدرجات.

وقد وجد البعض أنه من المفيد حساب البواقي المعيارية Standardized Residuals للتحقق من ملاءمة المفردات لنموذج راش . فهذه البواقي يمكن حسابها لكل خلية من خلايا مصفوفات البيانات الاختبارية التي تمثل صفوفها المفردات وأعمدتها الأفراد . وناتج جميع هذه البواقي لجميع الخلايا يمكن استخدامه كمؤشر Index لملاءمة المفردات أو الأفراد للنموذج. (علام ، ٢٠٠٥ ، ١٠٩-١١٠)

والخلاصة أن تقييم حسن ملاءمة البيانات الاختبارية لأحد نماذج الاستجابة للمفردة - وفي هذا البحث نموذج راش- يتطلب التحقق من افتراضات النموذج المستخدم ، وبخاصة افتراض البعد الأحادي ، والاستقلال الموضعي ، والتحقق من تنبؤات النموذج من خلال فحص البواقي المعيارية لملاءمة البيانات للنموذج ، وتحديد طبيعة عدم الملاءمة Misfit

وعلى الرغم من تعدد أساليب هذا التقييم إلا أنه يمكن تصنيفها في ثلاث مجموعات- وبخاصة فيما يتعلق بنماذج راش- كالتالي :

إيجاد قيمة مربع كاي وذلك بمقارنة تكرارات الإستجابات الملاحظة بالاستجابات المتوقعة لمجموعات فرعية من الأفراد المختبرين المحددة مسبقا

-إيجاد القيم المعيارية (Z) استنادا إلى دالة الأرجحية للاستجابات أو لنمط المفردات.

-إيجاد بواقى الدرجات Score Residuals استنادا إلى متوسطات انحرافات الدرجات الملاحظة عن الدرجات المتوقعة لاستجابات الأفراد المختبرين.

ويعتمد اختيار إحدى الطرق واستخدامها في تطبيق معين على معرفة كيفية ملاءمتها لافتراضات النموذج المستخدم (علام ،٢٠٠٥، ١١٢)

افتراضات نماذج الاستجابة للمفردة الاختبارية أحادية البعد :

تستند هذه النماذج إلى افتراضات قوية ينبغي تحققها في البيانات لكي تؤدي إلى نتائج يمكن الوثوق بها :

(١)**المنحنيات المميزة للمفردة** : احتمال الإجابة الصحيحة يكون دالة متزايدة مطردة لمواقع الأفراد على متصل السمة ، وبذلك يزداد احتمال توصل الفرد إلى الإجابة الصحيحة على مفردة الاختبار بازدياد مقدار السمة لديه.

وتفترض هذه النماذج وجود دالة مميزة خاصة بكل مفردة يتخذ كل مها شكل منحنى الترجيح اللوغاريتمي الاحتمالي Logistic Curve الذي يسمى المنحنى المميز للمفردة Item Characteristic Curve.

وشكل هذا المنحنى يوضح كيفية تغير مستوى السمة في علاقتها بالتغيرات في احتمالات إجابة معينة. فإذا كانت المفردة ثنائية الدرجة Dichotomous Item كما في مفردات الاختيار من متعدد ومفردات الصواب أو الخطأ ، فإن المنحنى المميز للمفردة يعبر عن انحدار "احتمال الإجابة الصحيحة" على "مستوى السمة". (علام ، ٢٠٠٥ ، ٥٩)

(٢)أبعاد الفضاء الكامن : ويقصد بالفضاء الكامن Latent Space مجموعة من الأبعاد غير المرتبطة مثنى مثنى ، وهذه الأبعاد تناظر السمات الكامنة. وموقع الفرد في الفضاء الكامن يتحدد من درجات السمة الكامنة لديه . وهذه الدرجات تعد بمثابة إحداثيات لتحديد موقع الأفراد تحديدا مناسبا في الفضاء الكامن . وتعتمد أبعاد هذا الفضاء على عدد السمات التي ينطوي عليها أداء الأفراد في مفردات الاختبار. ويمكن أن تتعدد هذه الأبعاد كما يمكن افتراض أن الفضاء الكامن أحادي البعد ، أي افتراض أن مفردات الاختبار متجانسة وتقيس قدرة أو سمة كامنة واحدة . (علام ، ٢٠٠٥ ، ٦٠)

(٣) الاستقلال الموضعي ويقصد به أن احتمال الإجابة الصحيحة للفرد على مفردة اختبارية يكون مستقلا عن ناتج إجابته على أي مفردة أخرى في الاختبار ، عند ضبط كل من القيمة التقديرية لقدرته ، والقيمة التقديرية لصعوبة المفردة ، لذلك فإن الاستقلال الموضعي Local Independence يشير إلى أن مفردات الاختبار تكون غير مرتبطة.وربما يبدو هذا متعارضا مع أحد مبادىء نظرية القياس الكلاسيكية التي تؤكد أهمية الاتساق الداخلي للاختبارات ، أي أنه كلما كانت درجات المفردات مرتبطة زاد ثبات هذه الدرجات.

ويشير علام (٢٠٠٥) إلى أن هذا التعارض يزيله شرط أحادية البعد في نظرية الاستجابة للمفردة حيث تكون إجابات الأفراد على المفردات مشروطة بمستوى قدرة كل منهم ،وصعوبة المفردات. أي أن الاستقلال الموضعي يدل على أنه لا توجد ارتباطات باقية بين المفردات عندما يتم إشراط Conditioned القيم التقديرية لقدرة الفرد ، وصعوبة المفردة. ولعل هذا يشير إلى أن مفردات الاختبار تقيس بعدا واحدا. ولكي يتحقق هذا الافتراض ينبغي أن يكون نمط الاستجابة الفرد (أي واحد صحيح ، أو صفر) مساويا حاصل ضرب احتمالات حصوله على درجة معينة في كل مفردة من مفردات الاختبار . وهذا يتحقق غذا كانت جميع المفردات تقيس سمة كامنة

واحدة أو قدرة واحدة. فأحادية البعد تتطلب إذن وجود سمة كامنة واحدة بحيث يكون جميع أزواج المفردات تحقق افتراض الاستقلال ، وهذا يعني أن المفردات يجب أن تكون مستقلة إحصائيا لأي مجتمع فرعي متجانس من الأفراد فيما يتعلق بالسمة الكامنة لذلك فإن الاستقلال الموضعي يتعلق أيضا بعدد السمات او القدرات التي ينطوي عليها الأداء على المفردات ، أي أن الاستقلال الموضعي يعد مؤشرا لأحادية البعد إذا كان النموذج المستخدم يحدد قيمة تقديرية لقدرة الفرد على بعد أحادي Embertson and Reise(2000)

محكات اختيار النموذج المناسب

إن اختيار أحد النماذج الأحادية لاستخدامه في موقف تطبيقي معين ، ومطابقته لمجموعة بيانات واقعية لا يعد أمرا يسيرا ؛ وهناك مجموعة من المحكات التي يمكن الاستناد إليها في هذا الشأن :

(١) واقعية افتراضات النموذج : فإذا كانت المفردات الاختبارية تتباين فقط في صعوبتها وليس في تمييزها ، فإن النموذج الأحادي البارامتر يكون أكثر ملاءمة. ويرى رايت (1977) Wright أنه ينبغي عند بناء الاختبار العمل على تجنب حدوث التخمين في الإجابة ، وذلك بالاهتمام بتصميم مشتتات Distractors ملائمة لكل مفردة ، وبذلك لا تكون هناك حاجة إلى تقدير بارامتر التخمين ويقترح كروكر Crocker ، وألجانيا Algina أن هذا النموذج يناسب بدرجة أكبر مفردات الاختيار من متعدد ، والصواب أو الخطأ .

(٢) تحمل النموذج لمخالفة افتراضاته : فاختيار نموذج بسيط مناسب يعتمد على مدى تحمله لمخالفة افتراضاته في تطبيق معين . فعلى الرغم من أن النموذج أحادي البارامتر يستند إلى عدة افتراضات تتعلق بعدم التخمين ، وتساوي جميع المفردات في التمييز ، وقياسها سمة أحادية البعد ، واتخاذ المنحنى المميز للمفردة شكل منحنى الترجيح اللوغاريتمي ، إلا أن مخالفة بعض هذه الافتراضات في البيانات الاختبارية ربما لا تؤثر كثيرا في دقة التقديرات الناتجة عن استخدام هذا النموذج في مواقف تطبيقية معينة .

ويعد نموذج راش أكثر نماذج نظرية الاستجابة للمفردة استخداماً في تصميم الاختبارات لأنه أقل نماذج الاستجابة للمفردة في عدد الفروض اللازم توافرها في البيانات لكي يعطي تقديرات دقيقة متسقة.

وتمكن علماء القياس من إيجاد حلول مناسبة لمشكلة تقدير بارامتر الصعوبة، وبارامتر قدرة الفرد لهذا النموذج ، في حين أنهم يواجهون حتى الآن مشكلات سيكومترية وإحصائية في تقدير بارامترات النموذجين الآخرين ، وبخاصة النموذج ثلاثي البارامتر.

القيم التقديرية لقدرة فرد حصل على نفس الدرجة في الاختبار الذي بني على أساس نموذج راش تكون متساوية بغض النظر عن عينة المفردات التي اختبر بها، في حين أنه ليس من الضروري أن يحدث ذلك في حالة النموذجين الآخرين ؛ لذلك يطلق علي نموذج راش " نموذج البارامتر الحر في تحليل المفردات Sample-Freed Model

(٣)أنواع البيانات الاختبارية المراد تحليلها : يعتمد اختيار النموذج المناسب أيضاً على نوع المفردات الاختبارية وكيفية تصحيحها أو أساليب تقدير درجاتها. كما ينبغي عند اختيار النموذج المناسب مراعاة كمية البيانات المتوافرة التي تستخدم في تقدير بارامترات النموذج.

فإذا لم تكن هناك بيانات كافية كما في بعض الحالات فإن أخطاء تقدير بارامترات نموذج معين سوف تنخفض إذا أختير نموذج عدد بارامتراته أقل ، حتى وإن كان أقل ملاءمة Lord (1983) (علام ، ٢٠٠٥ ، ٨٥-٨٧).

الدراسات السابقة

دراسة أبو جراد (٢٠٠٨)

هدفت إلى إعادة تدريج اختبار الذكاء الثالث كاتل الصورة (أ) باستخدام نموذج راش الأحادى المعلم، والى عمل معايير مختلفة تفسر من خلالها مستويات القدرة للأفراد، وقد تم تطبيق الاختبار على عينة من طلبة كلية التربية بجامعة القدس المفتوحة بلغ عددها (٢٤٠) طالبا وطالبة، واستخدم برنامج (spss (version 13 ، وبرنامج Winsteps (version 2.88) ؛ لإجراء المعالجات الإحصائية اللازمة وفقا لنموذج راش الأحادى المعلم، وتوصلت نتائج الدراسة إلى حذف فقرتين من فقرات اختبار التصنيف؛ لعدم ملاءمتها لنموذج راش، وعليه بلغ عدد فقرات الاختبار بعد تدريجه باستخدام نموذج راش (٤٨) فقرة، كما توصلت الدراسة الحالية إلى معايير للاختبار عن طريق إيجاد الرتب المئينية، والدرجات التائية، ونسب الذكاء الانحرافية المناظرة للتقديرات المختلفة لقدرة الأفراد لكل من الذكور والإناث.

دراسة الحموري (٢٠١١)

هدفت إلى استقصاء إمكان استخدام نماذج نظرية الاستجابة للمفردة في بناء اختبار تحصيلي من إعداد المعلم. تم بناء اختبار تحصيلي مكون من (٢٣) فقرة من نمط الاختيار من أربعة بدائلتقيس العدد نفسه من الأهداف المقصودة. طبق الاختبار على (٢٨٤) طالبا وطالبة . وحللت البيانات الناتجة من الاستجابة للفقرات وفق نظرية الاستجابة للمفردة وكشفت نتائج الدراسة عدم إمكان استخدام هذه النظرية لبناء الاختبارات ذات العدد القليل من الفقرات.

دراسة فراج والشريف (٢٠١٣)

إن دقة النتائج الواردة من الاختبارات التحصيلية المبنية على أساس النظرية الكلاسيكية في القياس كانت مثار جدل بين الباحثين التربويين والمتخصصين في مجال القياس والتقويم في ضوء العوامل التي تؤثر في تقدير الدرجة الحقيقية للمتعلم باستخدام هذا النموذج الكلاسيكي ، وهو ما يدعو إلى استخدام أحد نماذج نظرية القياس الحديثة(السمات الكامنة) وذلك للتحقق من فاعلية كلا النموذجين في تقدير المستوى التحصيلي للمتعلمين بدقة ، وهوهدف أساسي للعديد من الدراسات المقارنة بين النظرية الكلاسيكية ونظرية السمات الكامنة ، وقد قام الباحثان بتصميم اختبار تحصيلي مرجعي الهدف مكون من ٢٥ مفردة في مادة الجبر وقد تم تطبيقه على عينة مكونة من ٨٠ تلميذ من تلاميذ الصف الأول الإعدادي بمحافظة الإسكندرية ، وقد تم تحليل البيانات بطريقتين مختلفتين : الأولى في ضوء النظرية الكلاسيكية واعتمدت الدرجة الكلية كمحك ، والثانية في ضوء نظرية الاستجابة للمفردة الاختبارية وقد اعتمدت قدرة التلميذ على تحقيق الهدف الذي تقيسه المفردة كمحك.

وقد أشارت النتائج إلى تمتع الاختبار التحصيلي مرجعي الهدف بمؤشرات صدق وثبات مناسبة كذلك أكدت النتائج قابلية البيانات للتحليل باستخدام نظرية الاستجابة للمفردة ، وقد كانت البيانات المستمدة من النموذجين قابلة للمقارنة ، وأن نظرية السمات الكامنة كانت أكثر دقة في تحديد قدرة المفحوصين على الأداء على مفردات الاختبار وبأخطاء معيارية أقل في القياس من نظيراتها في النظرية الكلاسيكية، وكذلك اختلف النموذجان في تحديد مؤشرات صعوبة مفردات الاختبار.

دراسة حمادنة وبني خالد (٢٠١٣)

هدفت إلى بناء مقياس اتجاهات نحو العنف الالكتروني لدى عينة من مستخدمي مواقع التواصل الاجتماعي بجامعة آل البيت، وفق النظرية الحديثة في القياس. ولتحقيق هدف الدراسة تم تحديد الأبعاد الرئيسة للمقياس وصياغة ٤٤ فقرة موزعة على الأبعاد بشكل غير متساوٍ. وقد تم تطبيق الأداة على عينة مؤلفة من ٥٥٠ طالبا وطالبة من طلبة جامعة آل البيت ممن يستخدمون مواقع التواصل الاجتماعي.

وعند تحليل استجابات أفراد العينة- وفق نظرية استجابة الفقرة (IRT)- لم يتبق من فقرات المقياس سوى ٢٩ فقرة، موزعة في ثلاثة مجالات رئيسة. وتمتع المقياس بصورته النهائية ٢٩ فقرة بخصائص سيكومترية مناسبة، حيث بلغت قيمة معامل الثبات للمقياس ٠.٩٤، كما تمتع المقياس بدلالات متعددة للصدق.

دراسة زانون وآخرين (٢٠١٦).Zanon, et al

أصبحت نظرية الاستجابة للمفردة (IRT) إطارًا منهجيًا شائعًا لنمذجة بيانات الاستجابة في العديد من الاختبارات في مجالي التعليم والصحة ؛ ومع ذلك ، لا ينتشر استخدام النظرية على نطاق واسع بين علماء النفس. تهدف هذه الورقة إلى تقديم تطبيق تعليمي لنظرية الاستجابة للمفردة IRT وتسليط الضوء على بعض هذه المزايا لتطوير الاختبار النفسي. تم استخدام منهجية نظرية الاستجابة للمفردة IRT في تحليل نتائج مقياسين (مقياس التأثير الإيجابي والسلبي) وتم التطبيق بطريقة التقرير الذاتي. كان المشاركون ٨٥٣ طالبًا جامعيًا (٥٧٪ نساء) تتراوح أعمارهم بين ١٧ و ٣٥ عامًا. كشفت تحليلات IRT أن مقياس التأثير الإيجابي يحتوي على مفردات ذات تمييز معتدل وتقيس قدرات المستجيبين على مستوى أقل من المتوسط. قدم مقياس التأثير السلبي أيضًا مفردات ذات تمييز معتدل وتقيم المستجيبين عبر متصل السمة ؛ وتم استغلال مزايا نظرية الاستجابة للمفردة في تحسين قياس المقاييس للسمة موضع القياس. و تسمح لصانعي الاختبار بتحسين وزيادة مؤشرات الثبات والصدق لهذه المقاييس.

دراسة دبوس (٢٠١٦)

وهدفت إلى استخدام نظرية الاستجابة للفقرة في بناء تجمع فقرات اختبار محكي المرجع في الرياضيات وفق النموذج اللوجستي ثنائي المعلم ولتحقيق هذا الهدف تم إعداد تجمع من الفقرات ثنائية التدريج مكون من (٥٠) فقرة ، وتجمع من الفقرات متعددة التدريج مكون من (١٠) فقرات وروعي في بناء هذه الفقرات أن تقيس تحصيل الطالب في الرياضيات في المعرفة المفاهيمية والمعرفة الإجرائية وحل المشكلات .

وقد تكونت عينة الدراسة من (٥٠٢) طالب وطالبة أجابوا عن جميع الفقرات . وقد أظهرت النتائج تحقق افتراضات نظرية الاستجابة للفقرة كما بينت نتائج التحليل فيما يتعلق بمطابقة الفقرات الثنائية التدريج للنموذج الثنائي المعلم أن جميع الفقرات كانت مطابقة للنموذج الثنائي المعلم ، وأن الفقرات المتعددة التدريج كانت مطابقة لنماذج الاستجابة المتدرجة GRM وفي ضوء نتائج الدراسة تمت التوصية بمجموعة من التوصيات.

دراسة الموسوي (٢٠١٧)

استهدفت تطوير مقياس لجودة حياة طالب المرحلة الثانوية ، وذلك باستخدام نموذج راش، ولتحقيق هذا الهدف ، قام الباحث بإعداد الصورة الأولية للمقياس من (٣٠) مفردة وفق تدريج ليكرت ، حيث تم تطبيق المقياس على (٤٤٢) طالبا بالمرحلة الثانوية تتراوح أعمارهم من (١٥ ، ١٧) عاما وأشارت النتائج إلى مطابقة (٢٤) مفردة لافتراضات سلم التقدير المنبثق من نموذج راش أحادي البارامتر كأحد نماذج النظرية الحديثة الملائمة للمقياس المستخدم.

تكون المقياس في صيغته النهائية من (٢٤) مفردة ذات خصائص سيكومترية مناسبة ، حيث بينت النتائج أن المقياس المطور يقدم أكبر مقدار من المعلومات للأفراد عن جودة حياةٍ المراهق ذات المستوى المتوسط ، إذا كان متوسط قيم القدرة (٠.٥) ، وهذه القيمة مساوية تقريبا لمتوسط قيم الصعوبة لمفردات المقياس (٠.٥٧) وبذلك تكون مماثلة للقيمة المتوقعة في نموذج القياس.

دراسة الزبيدي (٢٠١٨)

هدفت إلى استخدام النموذج الأحادي البارامتر (راش) في بناء اختبار تحصيلي محكي المرجع في مقرر الاختبارات والمقاييس لطلبة الدبلوم التربوي في كلية التربية بجامعة الطائف ولتحقيق هذا الهدف تم بناء اختبار محكي المرجع في مقرر الاختبارات والمقاييس يقيس تحصيل

الطلبة ويقدم تقويماً موضوعياً للمهارات المطلوب إتقانها مكون من (٣٧) فقرة من نوع الاختيار من متعدد بأربع بدائل، وطبق الاختبار على عينة مكونة من (١٤٧) طالبا أشارت النتائج على ملائمة الاستجابات عن (٣٧) فقرة وهي جميع فقرات الاختبار بصورته النهائية لافتراضات نموذج راش وقد بلغ معامل الثبات للأفراد (٠.٩) أما معامل الثبات للاختبار (٠.٧٩) ، وخلصت الدراسة إلى مجموعة من التوصيات.

دراسة الشريف (٢٠١٨)

هدفت إلى التعرف على الخصائص السيكومترية لاختبار مقرر مهارات التعلم والتفكير على عينة (١٦٢) من طلاب كلية التربية جامعة الملك فيصل باستخدام النظرية الكلاسيكية وكذلك تحديد الخصائص السيكومترية لنفس الاختبار باستخدام نموذج راش والتحقق من إمكانية مقارنة درجات الطلاب على الاختبار بكلا الإطارين القياسيين .

أشارت النتائج إلى اختلاف المفردات المكونة لاختبار مقرر مهارات التعلم والتفكير باختلاف النموذج الإحصائى المستخدم ؛ حيث أفرزت مؤشرات نظرية الكلاسيكية للصدق والثبات اختبار مكون من (٣٠) مفردة، بينما أفرز نموذج راش اختبار مكون من (٢٣) مفردة ، وانطبقت البيانات المستمدة من الاختبار على نموذج راش ، كما أنه لم يتغير تدريج مفردات اختبار مقرر مهارات التعلم والتفكير بعد تدريجه وفق نموذج راش بتغير مستوى قدرات أفراد العينة المستخدمة فى الحصول على هذا التدريج ، وتم حذف (٧) مفردات غير مطابقة لنموذج راش كما بينت النتائج أن تقديرات بارامترات الصعوبة لمفردات الاختبار ضمن المحكات المقبولة في أدبيات القياس التربوي والنفسي. وتراوحت بارامترات القدرة لأفراد العينة بين (- ٤.٠٣) لوجيت و (٤.٠٤) لوجيت .

دراسة رشيد ومنصور (٢٠١٨)

هدفت الدراسة إلى استخدام نموذج راش في تدريج اختبار محكي المرجع لقياس مستويات التفكير الهندسي مبني وفق نظرية فان هيل ، تكون الاختبار من (٢٨) مفردة من نوع الاختيار من متعدد طبق الاختبار على عينتين استطلاعية وأساسية مكونتين من (٦٨١) تلميذ وتلميذة من تلاميذ المرحلة الثانوية . أظهرت النتائج تدريج الاختبار بعد استبعاد (٦) مفردات و(١١) فرد ، بأن الاختبار ثابت ، حيث بلغ معامل الثبات للمفردات (٠.٩٩) ، ومعامل الثبات للأفراد (٠.٦٥) ، كما أن مفردات الاختبار غطت مدى الصعوبة بشكل جيد ، فتراوحت قيم الصعوبة للمفردات مابين (١.٦١- ، ١.٤٦) لوجيت وقيم قدرات الأفراد مابين (٣.٤٤- ، ٤.٦٣) لوجيت ، ووقعت جميع المفردات وعددها (٢٢) مفردة ضمن قيم بارامترات المطابقة لنموذج راش (MNSQ) في المدى المقبول (٠.٧- ، ١.٣) ، وقيم بارامترات الملاءمة (ZSTD) في المدى المقبول بين (٠.٢- ، ٠.٢)

دراسة علي (٢٠١٩)

هدفت إلى استخدام أحد نماذج النظرية الحديثة في القياس النفسي والتربوي (نظرية الاستجابة للمفردة الاختبارية) وهو نموذج راش لميزان التقدير في بناء وتقنين مفردات مقياس الاتجاه نحو مهنة التدريس. وتكونت عينة البحث من ١١٢٠ طالب وطالبة بكلية التربية بجامعة المنيا. بلغ عدد الذكور ٢٢٥ والاناث ٨٩٥ . بلغ عدد طلبة الفرقة الثالثة ٧٣٣ ، وعدد طلبة الفرقة الرابعة ٣٨٧. وتم تحليل البيانات باستخدام برنامجWinsteps ، وأسفر التحليل الأول عن استبعاد ٤ مفردات من أصل ٤٣ مفردة وحذف بيانات ٣٤٦ فرد لم تكن مطابقة لحدود الملاءمة المقبولة في هذا النموذج، ثم تمت اعادة التحليل واشتمل المقياس في صورته النهائية على ٣٩ مفردة باستخدام بيانات ٧٧٤ فرد. وباستخدام التحليل العاملي الاستكشافي بطريقة المكونات الأساسية تم التحقق من افتراض أحادية البعد، كما تم فحص أعلى معاملات ارتباط بين المفردات ولم تصل قيمتها لحدود الاعتماد الموضعي، وبذلك تم التأكد من توافر الشرط الثاني لاستخدام نموذج راش وهو الاستقلال الموضعي أو استقلالية القياس، ولم يتم تحديد زمن معين للإجابة على المقياس وبذلك كان المقياس متحررًا من عامل السرعة. وقد تراوحت قيم متوسطات احصاءات المطابقة الداخلية والخارجية فيما بين (٠.٦ : ١.٤) ماعدا مفردتين فقط كان

متوسط احصائي المطابقة الداخلية والخارجية لهما أكبر قليلاً عن المعيار ١.٤، لذلك أوصى الباحث بحذف هاتين المفردتين عند استخدام المقياس في دراسات مستقبلية. أما عن معاملات الثبات، فقد كان معامل ثبات المفردات ٠.٩٩ وكان معامل الفصل بين المفردات ١٢.٦٤، وقد كان معامل ثبات الأفراد ٠.٩٣ ومعامل الفصل بين الأفراد ٣.٥٩ وهي قيم ثبات مرتفعة ومعاملي الفصل يزيدان عن القيمة المقبولة وهي ٢. مما يشير إلى تمتع المقياس بخصائص سيكومترية مقبولة في ضوء تطويره وفقا لافتراضات النظرية الحديثة في القياس وبالتحديد نموذج راش لميزان التقدير .

التعقيب على الدراسات السابقة

-يلاحظ قلة عدد الدراسات التي صممت اختبارات تحصيلية في علم النفس وتدريجها وفق نموذج راش فباستثناء البحث الحالي ودراسة سابقة للباحث الحالي الشريف (٢٠١٨) ، ودراسة Zanon et al(2016) لم يجد الباحث -في حدود علمه - دراسات أخرى يجريها خبراء في علم النفس وتهتم بتدريج الاختبارات التحصيلية في مقررات علم النفس وفق نماذج النظرية الحديثة في القياس

-بعض الدراسات استخدمت المنهج المقارن مثل دراسة فراج والشريف (٢٠١٣) ؛ حيث كان هناك تفوق ملحوظ لنموذج من نماذج نظرية الاستجابة للمفردة وهو نموذج راش في تدريج مفردات الاختبارات المصممة في هذه الدراسات مقارنة بالنظرية الكلاسيكية في القياس.

-العديد من الدراسات استخدمت نموذج راش أحادي البارامتر في تدريج المفردات ؛ حيث كان هو الأنسب لتحليل البيانات وتصميم المفردات ثنائية التقدير (صفر ، ١) مثل دراسات أبوجراد (٢٠٠٨) ، والشريف (٢٠١٨) ، والزبيدي (٢٠١٨) ،ودراسة الموسوي(٢٠١٧) ودراسة علي (٢٠١٩).

-لم يكن الاهتمام بعدد المفردات في الاختبار بقدر ماكان الاهتمام بمؤشرات مطابقة المفردات (الداخلية والخارجية) لنموذج راشوبأكثر من مؤشر مثل البحث الحالي ودراسة الحموري (٢٠١١) ، ودراسة فراج والشريف (٢٠١٣) ، ودراسة الشريف (٢٠١٨) وأشارت دراسة علي (٢٠١٩) إلى مؤشرين : للمطابقة مؤشر MNSQ ومؤشر ZSTD

مصطلحات البحث

نموذج راش : يطلق عليه النموذج أحادي البارامتر One Parameter Model ، وقد اقترحه جورج راش ويهتم بتحديد موقع المفردة الاختبارية على سلم صعوبة لجميع المفردات التي تشكل الاختبار ، كمايهتم بتدريج مستويات قدرة الفرد باختبار معين على نفس متصل المفردات (علام ، ٢٠٠٧)

اللوجيت : القيم التقديرية لقدرة الأفراد ، وكذلك القيم التقديرية لصعوبة المفردات تقاس بوحدة مشتركة تسمى " وحدة الترجيح اللوغاريتمي" Logit . فقدرة الفرد كما تقدر بهذه الوحدة هي اللوغاريتم الطبيعي لترجيح الإجابات الصحيحة على المفردات التي من النوع الذي استخدم في تعريف نقطة صفر ميزان صعوبة المفردات.

وصعوبة المفردة كما تقدر بهذه الوحدة هي اللوغاريتم الطبيعي لترجيح الإجابة الخطأ على هذه المفردة للأفراد الذين تقع قدرتهم عند نقطة صفر ميزان القدرة (علام ، ٢٠٠٥ ، ٦٠)

ملاءمة المفردة Item-Fit : المفردة الملاءمة داخلياً هي المفردة التي يجيب عليها إجابة صحيحة ذوي القدرة التحصيلية المرتفعة ، ويجيب عليها إجابة خاطئة ذوي القدرة التحصيلية المنخفضة.

طلبة الفرقة الأولى : شعبة علم النفس بكلية التربية جامعة الإسكندرية لائحة الساعات المعتمدة ؛ يبلغ متوسط أعمارهم (١٨) سنة. (٢٠٠٨).

منهج البحث

اتبع الباحث في البحث المنهج الوصفي التحليلي لشرح نماذج نظرية الاستجابة للمفردة تحديداص نموذج راش وذلك بهدف تحديد الافتراضات والشروط الواجب توافرها في مفردات اختبار سيكولوجية التعلم لطلبة الفرقة الأولى علم النفس كي يتم تدريجه على نموذج راش.

عينة البحث

تكونت عينة البحث من ١٤٢ طالب وطالبة من طلبة كلية التربية جامعة الإسكندرية الفرقة الأولى شعبة علم النفس وذلك خلال الفصل الدراسي الأول من العام الجامعي ٢٠٢٠- ٢٠٢١ م الذين كانوا يدرسون مع الباحث مقرر سيكولوجية التعلم في لائحة الساعات المعتمدة.

إجراءات البحث

بعد الانتهاء من تطبيق الاختبار، استخدم الباحث البرنامج الإحصائي (SPSS v.21) ، والبرنامج الإحصائي(WINSTEPS v.3730) في إجراء التحليلات الإحصائية اللازمة للإجابة عن أسئلة البحث، والتي تتضمن:

-حساب معامل ألفا .

-تم حساب مربع كاي؛ للتأكد من أحادية البعد للاختبار.

-حساب قيم بارامترات قدرات الأفراد.

-حساب الخطأ المعياري في تقدير قدرات الأفراد .

-حساب المتوسط الحسابي والانحراف المعياري لتقديرات قدرات الأفراد .

-حساب المتوسط الحسابي والانحراف المعياري للأخطاء المعيارية في القياس لتقديرات قدرات الأفراد.

-حساب قيم بارامترات الصعوبة للمفردات.

- حساب المتوسط الحسابي والانحراف المعياري لتقديرات قيم بارامترات صعوبة المفردات.

-حساب المتوسط الحسابي والانحراف المعياري للأخطاء المعيارية في القياس لتقديرات قيم بارامترات صعوبة المفردات.

أداة البحث

تصميم اختبار سيكولوجية التعلم للفرقة الأولى شعبة علم النفس إعداد الباحث

هو اختبار تحصيلي مرجعي الهدف في ضوء نظرية الاستجابة للمفردة الاختبارية

يعاني الطلبة في بداية المرحلة الجامعية من صعوبات عديدة في تعلم مقررات علم النفس بالجامعة خاصة إذا كانوا من خلفية علمية ولم يدرسوا مقرر علم النفس العام في المرحلة الثانوية ؛ وهذه الصعوبات تخفق الاختبارات التحصيلية في تشخيصها ولذلك فإن تصميم الاختبارات التحصيلية مرجعية الهدف في مقررات علم النفس تسهم في تحديد مستويات القدرات التحصيلية التي يمتلكها الطلبة بدقة والتغلب على مثل هذه الصعوبات

وقد مر تصميم الاختبار بعدة مراحل توضحها الخطوات التالية:-

مرحلة التحليل :-

تحديد الكفايات الأساسية في الجزء الأول من الفصل الدراسي بمقرر سيكولوجية التعلم.

تحديد النطاق السلوكي للكفايات.

صياغة الأهداف السلوكية وترتيبها في نسق هرمي.

مرحلة البناء :- وتضمنت كتابة مفردات الاختبار ، وقام الباحث بصياغة مفردة واحدة لقياس كل هدف من أهداف الاختبار السلوكية ، وقد روعى أن تكون المفردات من نوع الاختيار من متعدد (أربع بدائل). ومن نوع الصواب أو الخطأ

مرحلة التجريب :-

أ- التحقق من صدق محتوى الاختبار:

- قام الباحث بتحديد موضوعات الدروس المختارة؛ وقد تم تحليل المحتوى وفق الطريقة التي حددها صلاح علام (٢٠٠٧) ؛ والطريقة جمعت بين تحليل الإجراءات والتحليل البنائي الهرمي ، وقد تمت صياغة الأهداف السلوكية في ضوء هذا التحليل بطريقة موضوعية قابلة للملاحظة والقياس وعددها (٢٠) هدف سلوكي وكل هدف تم تحويله إلى مفردة في الاختبار .

- قام الباحث بتحكيم المفردات بواسطة مجموعة من الخبراء في مجال القياس النفسي وتدريس الرياضيات* ، وطلب من كل محكم توضيح مدى مناسبة كل مفردة لقياس المجال الذي تقيسه وتم في هذه الخطوة حساب معامل صدق المحتوى لكل مفردة بمعادلة لوشي Lawshe وقد أكد كل من روفينلي وهامبلتون(1971) Rovinelli and Hambleton أن طريقة حساب معامل اتفاق المحكمين تتطلب اختيار عدد من المتخصصين للتحقق من أن كل مفردة تقيس الهدف الإجرائي الذي وضعت لقياسه في الاختبار . (منسي ، ٢٠٠٣: ٢٠٨) والجدول (١) يوضح نتيجة ذلك

جدول (١) مواصفات اختبار سيكولوجية التعلم ومعامل صدق المحتوى بصيغة لوشي ن =٥

المفردة	الهدف الذي تقيسه	تقيس	لا تقيس	معامل Lawshe
١	أن يتعرف الطالب مفهوم عملية التعلم	٤	١	٠.٦
٢	أن يحدد الطالب مجال تطبيقات علم النفس التربوي	٥	٠	١
٣	أن يحدد الطالب خصائص الدوافع في عملية التعلم	٥	٠	١
٤	أن يحدد الطالب تصنيفا علميا للدوافع	٤	١	٠.٦
٥	أن يحدد الطالب خصائص الدوافع الثانوية	٥	٠	١
٦	أن يبرز الطالب أهمية دراسة سيكولوجية التعلم بالنسبة له	٤	١	٠.٦
٧	أن يصف الطالب النظرة التقليدية لعملية التعلم	٥	٠	١
٨	أن يحدد الطالب دور الخبرة في عملية التعلم	٥	٠	١
٩	أن يتعرف الطالب شرطا من شروط التعلم	٥	٠	١
١٠	أن يحدد الطالب خصائص الدوافع الأولية	٤	١	٠.٦
١١	أن يتعرف الطالب خصائص عملية التعلم	٥	٠	١
١٢	أن يميز الطالب بين : الدوافع الأولية والدوافع الثانوية	٤	١	٠.٦
١٣	أن يميز الطالب بين عمليتي : التعليم والتعلم	٥	٠	١
١٤	أن يتعرف الطالب استمرارية عملية التعلم مدى الحياة	٤	١	٠.٦
١٥	أن يحدد الطالب العلاقة بين : الدوافع الأولية والدوافع الثانوية	٥	٠	١
١٦	أن يميز الطالب بين التغيرات السلوكية الناتجة عن التعلم والناتجة من عوامل أخرى	٥	٠	١
١٧	أن يتعرف الطالب دور المعلم في تنمية الدافعية نحو تعلم مادة تخصصه	٥	٠	١
١٨	أن يميز الطالب مفهوم الدافعية كشرط من شروط التعلم	٥	٠	١
١٩	أن يتعرف الطالب دور الدوافع في عملية التعلم	٤	١	٠.٦
٢٠	أن يصنف الطالب دافعا من الدوافع الاجتماعية	٥	٠	١

وقد تراوحت نسب لوشي لل (٢٠ مفردة) بين (٠.٦، ١) وهي قيم مقبولة

ثبات تدريج القدرة ومعاملات الصعوبة لمفردات الاختبار بطريقة معامل الفصل:

قام الباحث باستخدام طريقة معامل الفصل ببرنامج وينستبس Winsteps 3.73 ؛ في تقدير قيمة ثبات تقدير الاختبار التحصيلي مرجعي الهدف للقدرات المقاسة وذلك على عينة مكونة من (١٤٢) طالب وطالبة من الفرقة الأولى شعبة علم النفس بكلية التربية جامعة

* شارك في التحكيم : أ.د/ أنس الطيب ، و د / حسن بدري ، ود/آدم بشير ، وأ.د.أحمد رجب ، ود .شهاب بن شواشة

الإسكندرية خلال الفصل الدراسي الأول من العام الجامعي ٢٠٢٠ /٢٠٢١ م والجدول (٢) يوضح نتيجة ذلك :

جدول (٢) معاملات الثبات بطريقة معامل الفصل

		العدد	القدرة	الخطأ النموذجي	الملاءمة الداخلية		الملاءمة الخارجية	
					MNSQ	ZSTD	MNSQ	ZSTD
المتوسط	١٣٢.٨	١٤٢		٠.٥٨	٠.٩٩	٠.٢	٠.٨٨	٠.١
الانحراف المعياري	١١		١.٤٣	٠.٢٨	٠.١٦	٠.٤	٠.٤	٠.٩
أكبر درجة	١٤١	١٤٢	٢.٨٧	١.٠٣	١.٢٤	٠.٩	١.٥٣	٢
أقل درجة	١٠٢	١٤٢	-١.٨	٠.٢٢	٠.٦٦	-٠.٧	٠.٠٥	-١.٦
RMSE	الحقيقي	٠.٦٦	الانحراف المعياري الحقيقي	١.٢٦	معامل الفصل	١.٩١	ثبات المفردات الحقيقي	٠.٧٨
RMSE	النموذجي	٠.٦٤	الانحراف المعياري النموذجي	١.٢٨	معامل الفصل	١.٩٨	ثبات المفردات النموذجي	٠.٨

ويتضح من الجدول (٢) أن :

قيمة معامل الفصل (١.٩١) والنموذجية (١.٩٨) تقترب من القيمة المثالية وهي ٢ وتدل على أن الدرجات المقدرة مقسمة إلى فئتين : متقنين وغير متقنين والمعنى يستقيم مع فكر نظرية الاستجابة للمفردة التي تسعى إلى تقسيم المفحوصين إلى فئتين (متقنين / غير متقنين)

والمفردة الملاءمة داخليا هي المفردة التي يجيب عليها إجابة صحيحة ذوي القدرة المرتفعة ، ويجيب عليها إجابة خاطئة ذوي القدرة المنخفضة ، ويشير خضر (٢٠٠٧) إلى أن مؤشر الملاءمة الداخلية للمفردات يشير إلى مدى تقارب البيانات من النموذج Infit Mean Square بينما مؤشر الملاءمة الخارجية يشير إلى مدى تباعد البيانات من النموذج Outfit Mean Square. (خضر ،٢٠٠٧ ،١٥٥)

ويشير خضر (٢٠٠٧) إلى أن المدى المقبول لمؤشرZSTD الملاءمة الداخلية (من -٢ :إلى ٢) ، والمدى المقبول لمؤشرZSTD الملاءمة الخارجية (من -٢ :إلى ٢) (خضر ،٢٠٠٧ ،١٥٦)

ويتضح من جدول (٢) أن قيمة الملائمة الداخلية للمتوسط بمؤشر MNSQ ٠.٩٩في المدى المقبول (من ٠.٥ إلى ١.٥) وقيمة الملاءمة الداخلية للمتوسط بمؤشر ZSTD ٠.٢ في المدى المقبول (من ٢- إلى ٢) كذلك قيمة الملائمة الخارجية للمتوسط بمؤشر MNSQ ٠.٨٨في المدى المقبول (من ٠.٥ إلى ١.٥) وقيمة الملاءمة الخارجية للمتوسط بمؤشر ZSTD ٠.١ في المدى المقبول (من ٢- إلى ٢)ويمكن حساب خارج قسمة معامل ثبات تدريج المفردات الحقيقي على معامل ثبات تدريج المفردات النموذجي إذا كانت القيمة أكبر من ٠.٩ فإنها تدل على أن المفردات المدرجة تنتمي إلى نفس البعد ويتضح من الجدول السابق كذلك أن قيمة ثبات تدريج المفردات ٠.٧٨ وهي قيمة مقبولة تدل على ثبات عملية التدريج .

أما عن معاملات الثبات، فقد كان معامل ثبات المفردات 0.78 وكان معامل الفصل بين المفردات ١.٩١، وقد كان معامل ثبات الأفراد 0.7

هي قيم ثبات مرتفعة ومعاملي الفصل يزيدان عن القيمة المقبولةٍ وهي(.02)مما يشير إلى تمتع المقياس بخصائص سيكومترية مقبولة في ضوء تطويره وفقا لافتراضات النظرية الحديثة في القياس وبالتحديد نموذج راش لميزان التقدير.

نتائج البحث وتحليلها الإحصائي

الإجابة على السؤال الأول

ما مدى ملائمة البيانات المستمدة من تطبيق اختبار سيكولوجية التعلم الفرقة الأولى للتدريج بنموذج راش أحادي البارامتر؟

لابد من بحث مدى تحقق افتراضات نظرية الاستجابة للمفردة في بيانات الاختبار (المفردات/ الأفراد)

أ-شرط أحادية البعد :

أن تقييم حسن مطابقة البيانات الاختبارية لأحد نماذج الاستجابة للمفردة -وفي هذا البحث نموذج راش- يتطلب التحقق من افتراضات النموذج المستخدم ، وبخاصة افتراض البعد الأحادي ، والاستقلال الموضعي ، والتحقق من تنبؤات النموذج من خلال فحص البواقي المعيارية لمطابقة البيانات للنموذج ، وتحديد طبيعة عدم المطابقة Misfit

ويشير علي (٢٠١٩) إلى إمكانية التحقق من شرط أحادية البعد بقسمة معامل ثبات تدريج المفردات الحقيقي (الواقعي) على معامل ثبات تدريج المفردات النموذجي إذا كانت القيمة أكبر من ٠.٩ فإنها تدل على أن المفردات المدرجة تنتمي إلى نفس البعد

وذلك فإن شرط أحادية البعد قد تحقق لأن ٠.٧٨ / ٠.٨ = ٠.٩٧

بالرجوع لبيانات الجدول السابق فإن قيمة معامل ثبات التدريج لو كانت أكبر من ٠.٩ فإنها تكون مؤشر على أحادية البعد

شرط جودة الملاءمة:

يشير علام (٢٠٠٥) إلى أن إيجاد قيمة مربع كاي من أساليب تقييم جودة البيانات الإختبارية وذلك بمقارنة تكرارات الإستجابات الملاحظة بالاستجابات المتوقعة لمجموعات فرعية من الأفراد المختبرين المحددة مسبقا (علام ،٢٠٠٥، ١١٢)

والجدول (٣) يلخص نتيجة اختبار مربع كاي

جدول (٣) اختبار مربع كاي لجودة المطابقة بين التكرارات المشاهدة والتكرارات المتوقعة ن = ١٤٢

رقم المفردة	المتوسط	الانحراف المعياري	التكرارات		قيمة مربع كاي	درجات الحرية	مستوى الدلالة
			إجابة صحيحة	إجابة خاطئة			
١	.990	.0840	141	1	138.028	1	0.00
٢	.760	.4280	108	34	38.563	1	0.00
٣	.980	.1440	139	3	130.254	1	0.00
٤	1.00	.0000	142	0	-	-	-
٥	.720	.4510	102	40	27.070	1	0.00
٦	.990	.0840	141	1	138.028	1	0.00
٧	.960	.2020	136	6	119.014	1	0.00
٨	.930	.2570	132	10	104.817	1	0.00
٩	.990	.1180	140	2	134.113	1	0.00
١٠	.990	.0840	141	1	138.028	1	0.00
١١	.910	.2890	129	13	94.761	1	0.00
١٢	.990	.1180	140	2	134.113	1	0.00
١٣	.960	.2020	136	6	119.014	1	0.00
١٤	.990	.0840	141	1	138.028	1	0.00
١٥	.960	.1850	137	5	122.704	1	0.00
١٦	.950	.2170	135	7	115.380	1	0.00
١٧	.970	.1660	138	4	126.451	1	0.00
١٨	.910	.2890	129	13	94.761	1	0.00
١٩	.990	.1180	140	2	134.113	1	0.00
٢٠	.840	.3700	119	23	64.901	1	0.00

يتضح من الجدول (٣) أن نتيجة اختبار مربع كاي لجودة المطابقة بين التكرارات المشاهدة والتكرارات المتوقعة كلها كانت دالة إحصائيا عند مستوى ٠.٠١ باستثناء المفردة رقم (٤) التي أجاب عنها جميع أفراد العينة (ن = ١٤٢) إجابة صحيحة وهذا يشير إلى أن الأداء على غالبية المفردات كان مرتفعا أكثر من المتوقع أي أنه هناك فروق ذات دلالة إحصائية بين التكرارات المشاهدة والتكرارات المتوقعة لصالح التكرارات المشاهدة لدرجات حرية ١ عند مستوى ٠.٠١

وهذا يشير إلى أن المفردات كلها تقيس نفس السمة (القدرة التحصيلية) مما يحقق شرط أحادية البعد

قام الباحث بتطبيق الاختبار على عينة مكونة من ١٤٢ طالب وطالبة من الفرقة الأولى شعبة علم النفس بكلية التربية ،وقد بلغت قيمة معامل ثبات ألفا كرونباخ الكلي (٠.٤٩٦) ، كمؤشر على الاتساق وليس الثبات تم حساب ثبات التدريج للمفردة وللأفراد جدول (٢) للصورة الأولية من الاختبار وعدد مفرداتها (٢٠) مفردة.

ويشير أبوهاشم (٢٠٠٧) إلى أن طريقة(ألفا كرونباك) تعد مؤشراً على الاتساق الداخلي لمفردات المقياس مع بعضها البعض ، ومع الدرجة الكلية ، ويؤكد أن الاتساق الداخلي أكثر قرباً لمعنى الثبات. قام الباحث بحساب معامل ألفا كرونباخ لعينة مكونة من ١٤٢ طالب ومن طلبة الفرقة الأولى علم النفس والجدول (٤) يوضح نتائج هذا الحساب

جدول (٤) معاملات ألفا في حالة حذف درجة المفردة ن = ١٤٢

رقم المفردة	معامل الثبات	معامل الارتباط بالكلية	رقم المفردة	معامل الثبات	معامل الارتباط بالكلية
١	0.500	-.013	١١	0.428	0.372
٢	0.516	.096	١٢	0.454	0.473
٣	0.493	.085	١٣	0.483	0.153
٤	0.497	.000	١٤	0.461	0.548
٥	0.479	.207	١٥	0.457	0.318
٦	0.496	.047	١٦	0.452	0.316
٧	0.508	-.002	١٧	0.464	0.293
٨	0.528	-.070	١٨	0.483	0.152
٩	0.498	.025	١٩	0.459	0.427
١٠	.5000	-.013	٢٠	0.483	0.172

يتضح من جدول (٤) أن معاملات ثبات المفردات أغلبها جاءت أقل من أو يساوى معامل ثبات ألفا الكلي، مما يدل على ثبات المفردات والاختبار ، حيث أن حذف أي مفردة يؤثر سلبا على الاختبار. وبعضها أدى حذفها إلى ارتفاع قيمة معامل ثبات ألفا الكلي وهي المفردات أرقام : ١، ٢، ٤ ، ٧ ،٨ ، ٩ ، ١٠ لكن معاملات ارتباط كل مفردة بالدرجة الكلية ضعيفة وغير دالة على قوة العلاقة وقد يكون ذلك أفضل لتحقيق شرط آخر هو الاستقلال الموضعي غير أن اتخاذ قرار بشأن استبعاد أي مفردات يكون بعد اكتمال التحقق من كل الشروط ومنها محكات الملاءمة الداخلية والخارجية كما يحسبها برنامج وينستبس لكل مفردة حيث تكون الصورة أوضح والقرار أصوب وهذا ما يميز النماذج الحديثة عن النظرية الكلاسيكية ؛ فكرة وجود عدة محكات يمكن الارتكان إليها قبل اتخاذ القرارات السابق توضيحها.

ومع استخدام نظرية الاستجابة للمفردة في التحليل الإحصائي لم تعد الدرجة الكلية للاختبار هي المحك الوحيد للحكم على دقة البيانات أو تحليل المفردات وأصبح هناك مؤشرات أخرى ؛

وبحساب نسبة التباين التي أمكن تفسيرها من درجات الاختبار يمكن الرجوع إلى إحصاءات dimensionality ٍItem اتضح أن أول عامل غير مفسر في التباين لم يزد عن ٢.١ وهذا يشير إلى أحادية البعد أيضا ، وثاني عامل غير مفسر لم يزد عن ١.٨ ، وبقية العاوامل غير المفسرة لم تزد عن ذلك وكان عددهم جميعا خمس عوامل .

شرط الاستقلال الموضعي ويقصد به أن احتمال الإجابة الصحيحة للفرد على مفردة اختبارية يكون مستقلا عن ناتج إجابته على أي مفردة أخرى في الاختبار ويمكن بحث مدى تحققه من خلال حساب مصفوفة معاملات ارتباط درجات أفراد العينة ن = ١٤٢ في كل مفردة بدرجات المفردات الاختبارية الأخرى

جدول (٥) مصفوفة معاملات الارتباط بين درجات مفردات الاختبار

م	2	3	5	7	8	9	11	12	13	15	16	17	18	19	20
2	1.000	.147	.162	.128	-.090	-.067	.394	-.067	-.118	-.018	-.052	-.096	.051	-.067	-.023
3	.147	1.000	-.092	.456	.151	-.018	-.047	-.018	.213	-.028	-.033	-.025	-.047	-.018	-.065
5	.162	-.092	1.000	-.054	-.050	.058	.344	.191	-.054	-.035	.002	.083	-.036	.191	.235
7	.128	.456	-.054	1.000	-.058	-.025	-.067	-.025	-.044	-.040	-.048	-.036	-.067	.272	-.092
8	-.090	.151	-.050	-.058	1.000	-.033	-.087	-.033	.079	.097	.064	-.047	.008	-.033	-.121
9	-.067	-.018	.058	-.025	-.033	1.000	-.038	-.014	-.025	-.023	.249	-.020	-.038	-.014	.110
11	.394	-.047	.344	-.067	-.087	-.038	1.000	.169	.055	.204	.266	.094	.069	.169	-.007
12	-.067	-.018	.191	-.025	-.033	-.014	.169	1.000	.272	.301	.249	.341	.377	.493	.272
13	-.118	.213	-.054	-.044	.079	-.025	.055	.272	1.000	.150	.114	.176	.055	.272	.003
15	-.018	-.028	-.035	-.040	.097	-.023	.204	.301	.150	1.000	.309	.660	.204	.301	.020
16	-.052	-.033	.002	-.048	.064	.249	.266	.249	.114	.309	1.000	.158	.153	.249	.165
17	-.096	-.025	.083	-.036	-.047	-.020	.094	.341	.176	.660	.158	1.000	.094	.341	.156
18	.051	-.047	-.036	-.067	.008	-.038	.069	.377	.055	.204	.153	.094	1.000	.169	.059
19	-.067	-.018	.191	.272	-.033	-.014	.169	.493	.272	.301	.249	.341	.169	1.000	.110
20	-.023	-.065	.235	-.092	-.121	.110	-.007	.272	.003	.020	.165	.156	.059	.110	1.000

يتضح من الجدول السابق انخفاض قيم معاملات ارتباط درجات كل مفردة ببقية المفردات الاختبارية ؛ بالتالي درجة الفرد على كل مفردة مستقلة عن درجته على أي مفردة اختبارية أخرى وأن ارتفاع أو انخفاض الأداء على اي مفردة اختبارية لا يرتبط بالضرورة بارتفاع أو انخفاض الأداء على أي مفردات أخرى في الاختبار .

الأمر الذي يحقق شرط الاستقلال الموضعي.

السؤال الثاني

- مامناسبة مدى مؤشرات الملاءمة (الداخلية – الخارجية) لكل مفردة في اختبار سيكولوجية التعلم المدرج وفق نموذج راش ؟

للإجابة عن هذا السؤال إستخدم الباحث برنامج ونستبس للحصول على مؤشرات جودة الملاءمة للمفردة داخليا وخارجيا بأكثر من مؤشر والجدول (٦) يوضح نتيجة هذا الاختبار :

جدول (٦) معامل الصعوبة ومؤشرات الملائمة الداخلية والخارجية لمفردات اختبار سيكولوجية التعلم ن= ١٤٢

رقم المفردة	معامل الصعوبة باللوجيت	الملاءمة الداخلية		الملاءمة الخارجية		معامل تمييز المفردات	
		MSQ	ZSTD	MSQ	ZSTD	الحقيقي	المتوقع
1	-1.8	1.1036	0.4211	1.2228	0.6112	0.1069	0.1352
2	2.57	1.0371	0.461	1.2239	1.9512	0.4652	0.513
3	-0.6	1.135	0.4211	0.6371	-0.3194	0.1822	0.2109
4	-3.04	1	0	1	0	0	0
5	2.87	0.997	-0.019	1.001	0.041	0.4682	0.5484
6	-1.8	1.088	0.4011	0.6745	0.1807	0.1061	0.1352
7	0.2	1.1837	0.6012	1.3213	0.7713	0.2572	0.2701
8	0.81	1.2351	0.9312	1.3825	1.1114	0.3307	0.323
9	-1.05	1.1419	0.4311	0.9114	0.2109	0.151	0.1805
10	-1.8	1.1036	0.4211	1.2228	0.6112	0.1069	0.1352
11	1.14	0.8408	-0.7192	0.5958	-1.5994	0.305	0.3548
12	-1.05	0.7354	-0.2193	0.2309	-0.9398	0.1212	0.1805
13	0.2	0.9952	0.111	1.529	1.1115	0.2358	0.2701
14	-1.8	0.6631	-0.0893	0.0485	-0.95	0.0828	0.1352
15	-0.02	0.8754	-0.1791	0.7113	-0.4093	0.2033	0.2534
16	0.38	0.8928	-0.2191	0.6643	-0.7293	0.2397	0.285
17	-0.27	0.8851	-0.1091	0.7185	-0.2993	0.1842	0.2342
18	1.14	1.0315	0.221	1.2453	0.9012	0.3378	0.3548
19	-1.05	0.7571	-0.1792	0.3507	-0.6396	0.123	0.1805
20	1.93	1.0219	0.211	1.0248	0.201	0.4159	0.4393

يتضح من الجدول أن مؤشرات الملائمة الداخلية للمفردات حسب مؤشر MSQ كلها في المدى المقبول (من ٠.٥ إلى ١.٥) ، ويتضح كذلك أن أن مؤشرات الملائمة الخارجية للمفردات حسب مؤشر MSQ كلها في المدى المقبول (من ٠.٥ إلى ١.٥)

كما يتضح أن مؤشرات الملائمة الداخلية للمفردات حسب مؤشر ZSTD كلها في المدى المقبول (من -٢ إلى ٢) ، ويتضح كذلك أن أن مؤشرات الملائمة الخارجية للمفردات حسب مؤشر ZSTD كلها في المدى المقبول (من -٢ إلى ١.٥٢)

بالنسبة لمعامل الصعوبة المقدر باللوجيت يلاحظ أن بعض المفردات كانت معاملاتها سالبة وتشير إلى أداء ضعيف لكن جودة المطابقة الواضحة من مؤشرات الملائمة تشير إلى أن الذين أجابوا عن هذه المفردات ذوي القدرة المنخفضة

بالنسبة ل PTMeasure فإنه يمكن تتبعها من خلال المقارنة بين المعامل الحقيقي لارتباط درجة المفردة بالدرجة الكلية للاختبار بعد حذف درجة المفردة منها والمعامل المتوقع لارتباط درجة المفردة بالدرجة الكلية للاختبار بعد حذف درجة المفردة منها ويجب ألا يكون هذا الفرق كبير

وفي هذا الإطار تحديداً يؤكد Magno(2009) على أهمية تقدير الاتساقات الداخلية لكل اختبار يتم تحليل نتائجه باستخدام النظرية الكلاسيكية وذلك على العينة الحالية ؛ وذلك لأن تقديرات الاتساق الداخلي مرتبطة بالعينة التي تم تطبيق الاختبار عليها.

بعد الاطمئنان لقيمة معامل الفصل من الإجابة على السؤال الأول في مشكلة البحث

تم بحث إحصاءات عدم الملائمة بالنسبة للأفراد person misfit وبناء عليه تم حذف الأفراد الذين مؤشرات الملائمة الداخلية والخارجية لقدراتهم غير مقبولة وكذلك التي هناك تفاوت بين معامل ارتباطها بالدرجة الكلية الواقعي والمتوقع أو قيمته سالبة

وهم الأفراد الذين أرقامهم كالتالي : (١٢٩-١٣-٨-٩٤-٩١-٩٢-١٤٠-١١٧-٤١-٧٠-٩٥-١٣٨-١٤٢-٨٧-٢-٥١-٥٢-٧-٤٩) وبذلك تبقى في العينة ١٢٣ فرد

وبإعادة التحليل مرة أخرى بعد حذف الأفراد ن = ١٢٣ والجدول (٧) يوضح نتيجة ذلك :

جدول (٧) معاملات صعوبة المفردات بعد حذف الأفراد ن = ١٢٣

رقم المفردة	معامل الصعوبة باللوجيت	الملاءمة الداخلية		الملاءمة الخارجية		معامل تمييز المفردات	
		MSQ	ZSTD	MSQ	ZSTD	الحقيقي	المتوقع
1	-1.8	1.1036	0.4211	1.2228	0.6112	0.0693	0.1352
2	2.57	1.0371	0.461	1.2239	1.9512	0.4745	0.513
3	-0.6	1.135	0.4211	0.6371	-0.3194	0.1986	0.2109
4	-3.04	1	0	1	0	0	0
5	2.87	0.997	-0.019	1.001	0.041	0.5485	0.5484
6	-1.8	1.088	0.4011	0.6745	0.1807	0.1139	0.1352
7	0.2	1.1837	0.6012	1.3213	0.7713	0.1736	0.2701
8	0.81	1.2351	0.9312	1.3825	1.1114	0.1764	0.323
9	-1.05	1.1419	0.4311	0.9114	0.2109	0.13	0.1805
10	-1.8	1.1036	0.4211	1.2228	0.6112	0.0693	0.1352
11	1.14	0.8408	-0.7192	0.5958	-1.5994	0.4801	0.3548
12	-1.05	0.7354	-0.2193	0.2309	-0.9398	0.3343	0.1805
13	0.2	0.9952	0.111	1.529	1.1115	0.214	0.2701
14	-1.8	0.6631	-0.0893	0.0485	-0.95	0.3217	0.1352
15	-0.02	0.8754	-0.1791	0.7113	-0.4093	0.3218	0.2534
16	0.38	0.8928	-0.2191	0.6643	-0.7293	0.362	0.285
17	-0.27	0.8851	-0.1091	0.7185	-0.2993	0.2902	0.2342
18	1.14	1.0315	0.221	1.2453	0.9012	0.3069	0.3548
19	-1.05	0.7571	-0.1792	0.3507	-0.6396	0.3091	0.1805
20	1.93	1.0219	0.211	1.0248	0.201	0.4263	0.4393

يتضح من الجدول (٧) أن بعض المفردات معاملات صعوبتها منخفضة (بالسالب) وهي المفردات أرقام : ١ ، ٤ ، ٦ ، ١٠ ، ١٤ ، وبالتالي يمكن حذفها

مؤشرات الملائمة الداخلية والخارجية للمفردات سواء بمؤشر ZSTD أو مؤشر MSQ ملائمة وفي المدى المناسب يلاحظ كذلك أن مؤشر ptmeasure كانت قيمه في المدى المناسب والفرق بين معامل الارتباط الحقيقي ومعامل الارتباط المتوقع لم يكن كبيرا في غالبية المفردات مع ملاحظة أن مؤشر ptmeasure يشير غلى تمييز المفردات وهو مؤشر على الصدق وليس الثبات والجدول (٨) يوضح نتيجة ذلك

جدول (٨) الدرجات الخام والقدرات المناظرة لها باللوجيت والنورميد لاختبارسيكولوجية التعلم
ن= ١٢٣

الدرجة الخام	القدرة باللوجيت	الخطأ المعياري	دالة المعلومات	القدرة بالنورميد	الخطأ المعياري بالنورميد
0	-4.87	1.86	0.29	-328	173
1	-3.59	1.06	0.9	-208	98
2	-2.78	0.79	1.61	-132	74
3	-2.25	0.68	2.16	-83	63
4	-1.83	0.62	2.58	-44	58
5	-1.46	0.59	2.89	-10	55
6	-1.13	0.57	3.12	21	53
7	-0.82	0.55	3.26	50	52
8	-0.51	0.55	3.34	79	51
9	-0.22	0.55	3.35	106	51
10	0.08	0.55	3.31	134	51
11	0.39	0.56	3.22	163	52
12	0.71	0.57	3.06	192	53
13	1.04	0.59	2.86	224	55
14	1.41	0.62	2.59	258	58
15	1.82	0.66	2.28	296	62
16	2.3	0.73	1.9	341	68
17	2.9	0.84	1.43	397	78
18	3.79	1.09	0.83	480	102
19	5.14	1.88	0.28	605	175

يتضح من الجدول (٨) أنه كلما زادت قيمة لقدرة قلت قيمة الخطأ المعياري ؛ وذلك حتى الدرجة الخام (١٠) بعد ذلك عاودت قيمة الخطأ المعياري في الازدياد بنفس المعدل تقريبا وصولا لنفس قيمة الخطأ المعياري المقابل لأقل قدرة عند الوصول لأعلى قدرة وهي (٥.١٤) كان الخطأ المعياري (١.٨٨) في حين كان الخطأ المعيارِي المقابل لأقل قدرة (٤.٨٧-) هو (١.٨٦) وذلك عند تقدير القدرة باللوجيت ؛ الأمر نفسه تقريبا تكرر عند تقدير القدرة بالنورميد وقد يرجع ذلك إلى أخطاء القياس التي هي في أقل معدلاتها في نماذج الاستجابة للمفردة مقارنة بالنظرية الكلاسيكية وهذه النتيجة تتفق مع ما توصل إليه Magno(2009) حيث أكد على انخفاض قيمة الخطأ المعياري في نظرية الاستجابة للمفردة مقارنة بالنظرية الكلاسيكية وأنه يصل إلى أدنى مستوياته في القيم القريبة من متوسط الدرجة الكلية وهو ما لم يتحقق في البحث الحالي حيث وصل إلى أقل مستوياته عند الدرجة الخام (١٠) في حين كان متوسط الدرجة الكلية للاختبار التحصيلي (١٨.٧٧) بانحراف معياري قدره (١.٤).

مناقشة النتائج والاستنتاجات

تم تدريج مفردات اختبار تحصيلي في سيكولوجية التعلم للفرقة الأولى علم النفس وفق نموذج راش تكون الإختبار في صورته الأولية من ٢٠ مفردة وتم تطبيقه كاختبار منتصف الفصل الدراسي إلكترونيا عبر منصة ميكروسوفت المعتمدة من جامعة الإسكندرية نظرا لظروف جائحة كورونا ، واستقرت الصورة النهائية للاختبار في ١٥ مفردة بعد الرجوع لعدة محكات للحكم على مدى جودة وملاءمة البيانات الاختبارية لنموذج راش كان أهمها في هذا البحث ملاءمة المفردة Item-Fit سواء داخليا أو خارجيا ؛ ولم تعد الدرجة الكلية للاختبار هي المحك الوحيد ولم تكن الأولوية لمؤشرات نظرية القياس الكلاسيكية ؛ بل كانت الأولوية لمؤشرات نظرية الاستجابة للمفردة رغم استخدام بعض المؤشرات الكلاسيكية مثل معامل صدق المحكمين ومعامل ألفا وتم تقييم كل مفردة حسب سلوكها الذي تبحثه نظرية الاستجابة للمفردة من خلال دالة احتمالية للإجابة الصحيحة تعتمد على بارامتري : قدرة الفرد ، وصعوبة المفردة وبالتالي استبعاد أي مفردة يكون إذا كانت مؤشرات الملاءمة الداخلية والخارجية لها غير مناسبة أو معامل صعوبتها منخفض بشكل واضح يجعلها لا تميز بين مستويات الأداء بشكل سليم في القدرة التحصيلية المقاسة :

والجدول (٩) يوضح المؤشرات التي تم في ضوؤها استبعاد خمس مفردات من الصورة النهائية :

جدول (٩) الصورة النهائية للاختبار بعد حذف المفردات (١ ، ٤ ، ٦ ، ١٠ ، ١٤) ن = ١٢٣

رقم المفردة	معامل الصعوبة باللوجيت	الملاءمة الداخلية		الملاءمة الخارجية		معامل تمييز المفردات	
		MSQ	ZSTD	MSQ	ZSTD	الحقيقي	المتوقع
2	2.57	1.0371	0.461	1.2239	1.9512	0.4745	0.513
3	-0.6	1.135	0.4211	0.6371	-0.3194	0.1986	0.2109
5	2.87	0.997	-0.019	1.001	0.041	0.5485	0.5484
7	0.2	1.1837	0.6012	1.3213	0.7713	0.1736	0.2701
8	0.81	1.2351	0.9312	1.3825	1.1114	0.1764	0.323
9	-1.05	1.1419	0.4311	0.9114	0.2109	0.13	0.1805
11	1.14	0.8408	-0.7192	0.5958	-1.5994	0.4801	0.3548
12	-1.05	0.7354	-0.2193	0.2309	-0.9398	0.3343	0.1805
13	0.2	0.9952	0.111	1.529	1.1115	0.214	0.2701
15	-0.02	0.8754	-0.1791	0.7113	-0.4093	0.3218	0.2534
16	0.38	0.8928	-0.2191	0.6643	-0.7293	0.362	0.285
17	-0.27	0.8851	-0.1091	0.7185	-0.2993	0.2902	0.2342
18	1.14	1.0315	0.221	1.2453	0.9012	0.3069	0.3548
19	-1.05	0.7571	-0.1792	0.3507	-0.6396	0.3091	0.1805
20	1.93	1.0219	0.211	1.0248	0.201	0.4263	0.4393

ومن جدول (٩) يتضح أن الصورة النهائية للاختبار تكونت من ١٥ مفردة ؛ كل مفردة معامل صعوبتها في الحدود المسموح بها ومؤشرات الملاءمة الداخلية والخارجية توضح ذلك فالمفردة السهلة لا يصح أن يجيب عليها الطالب المتفوق إجابة خاطئة والعكس المفردةٍ الصعبة لا يصح أن يجيب عنها الطالب المتأخر إجابة صحيحة ولذلك فقد تم بحث الأفراد أولا (بارامترات القدرة) ومن ثم حذف الأفراد غير الصادقين ومن ثم إجراء التحليل مرة أخرى وهو ما تم في جدول (٨) وبالتالي تم تحديد الدرجات الخام المباشرة والقدرة باللوجيت المقابلة لكل درجة حسب تدريج نموذج راش باستخدام برمجية Winsteps و بالتالي أعطى الاختبار أكبر كمية

معلومات عند مستويات القدرة فوق المتوسط وقد تراوحت قيم بارامترات القدرة بين (4.87- ، 5.14) لوجيت بمتوسط قدره (1.9520) وهذه النتيجة توضح أن مستويات القدرة للأفراد أعلى من مستويات صعوبة المفردات بنسب واضحة وهو ما يزيد احتمالية الإجابات الصحيحة على المفردات ويجعل الباحث مطمئن إلى أن سلوك كل مفردة يقترب إلى حد كبير من شكل الدالة اللوغاريتمية التي تصف العلاقة بين قدرة الفرد وصعوبة المفردة وهذه النتيجة تتفق مع دراسات : فراج والشريف (٢٠١٣) ، وحمادنة وبني خالد (٢٠١٣) و(2016) Zanon et al في استغلال مميزات نظرية الاستجابة للبند في تحسين الشروط السيكومترية للاختبارات المصممة في هذه الدراسات ؛ واختلفت مع نتائج دراسة الحموري (٢٠١١) في أن قلة عدد المفردات المستخدمة في تصميم الاختبارات لا تقف حائلا أمام جودة التدريج طالما نستخدم نماذج قياس حديثة مثل نموذج راش حيث تكون هناك عدة محكات للحكم على سلوك كل بند وليس محك واحد كما في النظرية الكلاسيكية في القياس -وقد سبق توضيح ذلك في الإطار النظري – أما دراسة دبوس (٢٠١٦) فقد استخدمت نموذج لوجيستي ثنائي البارامتر في تدريج مفردات اختبار محكي المرجع في الرياضيات في حين اتفق البحث الحالي مع دراسات : الموسوي (٢٠١٧) ، والزبيدي (٢٠١٨) ، والشريف (٢٠١٨) في استخدام نموذج راش أحادي البارامتر في بناء الاختبارات التحصيلية محكية المرجع ولم يستخدم البحث الحالي المنهج المقارن بل ركز على أحد مؤشرات نظرية الاستجابة للمفردة وهو مؤشر جودة الملاءمة (الداخلية - الخارجية) بأكثر من مقياس (مؤشر ZSTD أو مؤشر MSQ) وعدد قليل من الدراسات السابقة ركزت على هذا المحك منها البحث الحالي ودراسة علي (٢٠١٩) - في حدود علم الباحث- رغم أهميته ، واستخدمت برمجيات متعددة في التحليل منها برمجيات : Winsteps في دراسات علي (٢٠١٩) والشريف (٢٠١٨) والبحث الحالي والحجوري (٢٠٢٠) استخدم برنامج R واستخدم الخياط (٢٠١٢) برنامج (Rumm , 2010) وأثبتت نتائج الخياط (٢٠١٢) ملاءمة (٢٥) مفردة من الاختبار المصمم وفق نموذج راش للبيئة المحلية بحيث تمتع بخصائص إحصائية مناسبة من حيث معامل صعوبة كل مفردة وهو ما اتفق مع البحث الحالي في ١٥ مفردة

واختلف البحث الحالي مع دراسة .LaHuis et al (2011) التي استخدمت اختبار مربع كاي في حساب مؤشرات جودة المطابقة ؛ حيث أنه مؤشر لابارامتري يتبع النظرية الكلاسيكية في القياس في حين وفرت نظرية الاستجابة للمفردات عدة محكات لحساب جودة المطابقة منها المؤشرات المستخدمة في البحث الحالي وهي (مؤشر ZSTD أو مؤشر MSQ)

ولم توضح دراسات مثل الزبيدي (٢٠١٨) كيف تم التحقق من مؤشرات جودة المطابقة للمفردات و اكتفى بالتحقق من افتراضات نموذج راش أحادي البارامتر واستخراج معامل ثبات الأفراد ومعامل ثبات المفردات رغم وجود جداول جاهزة في المخرجات تشرح هذه المؤشرات تحتاج فقط لتفسير ويجب الارتكان إليها في استبعاد أو إبقاء أي مفردة في الصورة النهائية وتوضيح مراحل التحليل التي تم إجراؤها . حيث يبدأ التحليل عادة باستبعاد الأفراد غير الملائمين في المرحلة الأولى ثم يعاد التحليل مرة أخرى لاستبعاد المفردات غير الائمة وفق مؤشرات جودة المطابقة (الداخلية- الخارجية) ولم تهتم بعض الدراسات العربية التي تم الرجوع إليها بتوضيح المؤشرات التي ارتكنت إليها في حساب جودة مطابقة المفردات مثل دراسات : بركات (٢٠١١) ، والحموري (٢٠١٠)

وقد اتفق البحث الحالي مع رشيد ومنصور (٢٠١٨) في ضرورة توضيح عدد المفردات التي تم استبعادها وعدد الأفراد الذين تم استبعادهم أثناء التحليل ؛ وعلى العكس من النظرية الكلاسيكية فإن استبعاد الأفراد غير الملائمين يؤدي إلى ارتفاع قيمة معامل الثبات للأفراد ، واستبعاد المفردات غير الملائمة تؤدي إلى ارتفاع قيمة معامل الثبات للمفردات وفي البحث الحالي تم استبعاد ٥ مفردات ، و ١٩ فرد ، وعادة ما تكون قيمة معامل ثبات المفردات أعلى من قيمة معامل

ثبات الأفراد . واتفق البحث الحالي مع رشيد ومنصور (٢٠١٨) في استخدام مؤشري مؤشر ZSTD أو مؤشر MSQ في حساب جودة المطابقة للمفردات (الداخلية -الخارجية) ؛ فحسب مؤشر MSQ المفردات كلها كانت في المدى المقبول (من ٠.٥ إلى ١.٥)

كما اتضح أن مؤشرات الملائمة الداخلية للمفردات حسب مؤشر ZSTD كلها كانت في المدى المقبول (من -٢ إلى ٢)

البحوث والدراسات المقترحة والتوصيات

١-دراسة مقارنة بين مؤشري: (MSQ - ZSTD) في حساب جودة المطابقة الداخلية والخارجية المفردات في تدريج الاختبارات التحصيلية وفق نموذج راش أحادي البارامتر.

٢- بناء بنك أسئلة مدرج وفق نموذج راش في مقرر علم نفس التعلم كأساس لتصميم اختبارات المقرر التحصيلية في لائحة كلية التربية .

٣-تدريج الاختبارات التحصيلية في بقية مقررات شعبة علم النفس وفق نماذج نظرية الاستجابة للمفردة باستخدام منهجية وينستبس كأساس للتدريج مع تحديد محكات استبعاد المفردات

٤-دراسة مقارنة بين مؤشرات ملاءمة المفردات ومؤشرات ملاءمة الأفراد في تقييم مدى ملاءمة البيانات المستخدمة لنموذج من نماذج نظرية الاستجابة للمفردة

٥-دراسة مقارنة بين مؤشرات ملاءمة المفردات (الداخلية -الخارجية) لنموذج نظرية الاستجابة للمفردة المستخدم (راش -لورد -بيرنبوم) في حالة كل من المفردات الثنائية والمفردات المتعددة الاستجابة.

التوصيات

-اقتراح تصميم بنوك أسئلة في كل مقرر من مقررات علم النفس التربوي وتدريج عينات من أسئلتها وفق نموذج راش في دراسات وبحوث مماثلة.

-ضرورة التوسع في التدريب على تصميم وتدريج بنوك الأسئلة وفق نماذج نظرية الاستجابة للمفردة ضمن برامج التدريب في وحداتالقياس والتقويم بكليات الجامعة.

-ضرورة ربط أساليب وفنيات صياغة الأسئلة وتدريجها وفق النظرية الحديثة بطرق التدريس المستخدمة خاصة : الطريقة الاستقصائية ، وطريقة المشروعات ، والطريقة التعاونية ، والطريقة الحوارية.

-ضرورة تدريب المعلمين على استخدام نموذج راش في تصميم اختبارات مقررات باعتباره المدخل الصحيح لفهم واستخدام بقية نماذج نظرية الاستجابة للمفردة.

المراجع

١- أبو جراد ،حمدي يونس (٢٠٠٨) استخدام نموذج (راش) فى تطوير اختبار كاتل الثالث للذكاء الصورة (أ) ،مجلة الجامعة الإسلامية للبحوث الإنسانية ؛ ١٦ (٢) ،

٢- أبو هاشم ، السيد (٢٠٠٧). التوجهات المستقبلية للتقويم النفسي والتربوي وتطبيقاتها في مجال التربية الخاصة ، المجلة العربية للتربية الخاصة ، الأكاديمية العربية للتربية الخاصة ، (١١) ، ١٥٧- ١٨٥

٣- الحجوري ، عيسى بن سلامة بن صبيان (٢٠٢٠). استخدام نظرية الاستجابة للمفردة في تقدير الخصائص السيكومترية للاختبار المركزي لمادة العلوم للصف الثالث المتوسط بالمدينة المنورة ، محلة العلوم التربوية والنفسية ، ٤ (١٤) ، ص : ٧٥ -٩٩

Available at://www.ajsrp.com
http://doi.org/10.26389/AJSRP.E220719

٤- الحموري ، هند عبد المجيد (٢٠١١). دراسة استكشافية لملاءمة نماذج نظرية الاستجابة للمفردة في بناء اختبار تحصيلي من إعداد المعلم ، مجلة العلوم التربوية والنفسية ، جامعة البحرين ، ١٢(٢) ، ٤٧- ٨١

٥- الخياط ، ماجد محمد (٢٠١٢) .درجة مطابقة اختبار تحصيلي وفق نموذج راش أحادي المعلمة في الكشف عن مستوى المعرفة العلمية في المهارات الرياضية لدى طلبة الصف الثامن الأساسي ، مجلة جامعة الأقصى (سلسلة العلوم الإنسانية) ١٦(١) ، ٨٧- ١١١.

٦- الزبيدي ، محمد بن حسن يحيى (٢٠١٨). بناء اختبار تحصيلي محكي المرجع في مقرر الاختبارات والمقاييس باستخدام النموذج الأحادي المعلم (نموذج راش) لطلبة الدبلوم التربوي في كلية التربية بجامعة الطائف ، المجلة الدولية التربوية المتخصصة، ٧(٢) ، ٤٠ -٧٠

٧- الشريف ، خالد حسن بكر.(٢٠١٨). فاعلية نموذج راش في تدريج مفردات اختبار مقرر مهارات التعلم والتفكير على عينة من طلاب كلية التربية جامعة الملك فيصل ، مجلة العلوم الإنسانية والاجتماعية في جامعة الإمام محمد بن سعود الإسلامية ، العدد (٥٠) ، ٩١-١٧١.

٨- الموسوي ، نعمان محمد صالح (٢٠١٧) . توظيف نموذج راش في تطوير أداة لقياس جودة الحياة لدى طلبة المرحلة الثانوية في مملكة البحرين ، مجلة الطفولة العربية ، (٧٠) ، ٤٣- ٦١.

٩- جاد الرب ، هشام فتحي محمد (٢٠١٦). التحليل المعرفي لبنود اختبارات الاستدلال المجرد وعلاقته بصعوبة البند باستخدام نموذج راش ، المجلة المصرية للدراسات النفسية ، ٢٦(٩٢) ، ٤٨٤ -٥٢٠.

١٠- حسن ، عزت عبد الحميد (٢٠١١). الإحصاء النفسي والتربوي تطبيقات باستخدام برنامج SPSS18 ، دار الفكر العربي ، مصر ، القاهرة.

١١- حمادنة ، إياد وبني خالد ، محمد (٢٠١٣) . بناء مقياس اتجاهات نحو العنف الإلكتروني لدى عينة من مستخدمي مواقع التواصل الاجتماعي بجامعة آل البيت ، مجلة المنارة ، ١٩ (٣) ، ٧٧-١٠٢

١٢- خضر ، عادل سعد يوسف (٢٠٠٧). **بنوك الأسئلة بين النظرية والتطبيق** ، القاهرة ، دار السحاب للنشر والتوزيع.

١٣- دبوس ، محمد.(٢٠١٦). استخدام نظرية الاستجابة للفقرة في بناء فقرات اختبار محكي المرجع في الرياضيات بفقرات ثنائية التدريج ومتعددة التدريج وفق النموذج اللوجستي ثنائي المعلم ، *مجلة جامعة النجاح للأبحاث (العلوم الإنسانية)* ، ٣٠ (٧) ،

١٤- دودين ، حمزة (٢٠١٠). **التحليل الإحصائي المتقدم للبيانات باستخدام SPSS** ،المملكة الأردنية الهاشمية ،عمان، دار المسيرة.

١٥- رشيد ، زياد ومنصور ،بوقصاره (٢٠١٨). استخدام نموذج راش اللوغاريتمي أحادي البارامتر في تدريج اختبار لمستويات التفكير الهندسي مبني وفق نموذج فان هيل ، مجلة الباحث في العلوم الإنسانية والاجتماعية ، العدد (٣٣) ، ٦٢٣-٦٥٠.

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

١٧- علام ، صلاح الدين محمود (٢٠٠٧).**الاختبارات التشخيصية مرجعية المحك** ، في المجالات التربوية والنفسية والتدريبية ، القاهرة ، دار الفكر العربي.

١٨- علي ، مصطفى علي خلف (٢٠١٩) استخدام نظرية الاستجابة للمفردة الاختبارية في بناء مقياس الاتجاه نحو مهنة التدريس ، مجلة البحث في التربية وعلم النفس كلية التربية ، جامعة المنيا ، ٣٤(١) ، ١- ٤٧

١٩- فراج ، محمد أنور والشريف ، خالد حسن بكر (٢٠١٣). فاعلية استخدام كل من نظرية السمات الكامنة و النظرية الكلاسيكية في قياس التحصيل الدراسي في الجبر بالمرحلة الإعدادية - دراسة مقارنة ، *مجلة العلوم الإنسانية والاجتماعية في جامعة الإمام محمد بن سعود الإسلامية* ، العدد (٢٨) ، رجب ١٤٣٤ هجري.

٢٠- منسي ، محمود عبد الحليم (٢٠٠٣). **التقويم التربوي** ،الإسكندرية ، دار المعرفة الجامعية ، ط٢

21-Anderson, E.B. (1973). A goodness of fit test for the Rasch model. Psychometrika, 38, 123-140.

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

23-Embertson, S. and Reise.(2000). Item Response Theory for Psychologists. N.J.: Lawrence Erlbaum.

24-Hambleton, R. K.(1994). Item Response Theory: A broad Psychometric Framework For Measurement Advances 1, 2. Psicothema : 6(3), 535-556.

25-Kang,T. and Chen, T.(2008). Performance of the Generalized S-X 2 Item Fit Index for Polytomous IRT Models, *Journal of Educational Measurement, 45(4), 391-406.*

26-LaHuis, D.M. ; Clark, P. ; and O"Brien, E.(2011). An Examination of Item Response Theory Item Fit Induces for the Graded Response Model , Organizational Research Methods , 14(1) ,10-23.

27-Linacre ,J.M.(2012). A User's Guide to WINSTEPS MINISTEPS, Rasch-Model Computers Programs ,Program Manual 3.75.0

28-Lord, F.M. Applications of item response theory to practical testing problems. Hillsdale NJ: Erlbaum, 1980.

29-Lord, F. M. (1983). Small N justifies Rasch model. In D. J. Weiss (Ed.), New horizons ira testing (pp. 51-61). New York: Academic Press.

30-Lord, F. M. , & Novick, M. R. (1968). Statistical theories of mental test scores. Reading, MA: Addison-Wesley.

31-Magno, C.(2009). Demonstrating the Difference between Classical Test Theory and Item Response Theory Using Derived Test Data, The International Journal of Educational and Psychological Assessment April 2009, Vol. 1, Issue 1, pp. 1-11.

32-Rost, J., & von Davier, M. (1994). A conditional item-fit index for Rasch models. Applied Psychological Measurement, 18, 171-182

33-Rovinelli, R. J., Hambleton, R. K. (1977). On the use of content specialists in the assessment of criterion-referenced test item validity. Dutch Journal for Educational Research, 2, 49–60.

34-Smith, R. M. (1988). The distributional properties of Rasch standardized residuals. Educational and Psychological Measurement, 48, 657-667.

35-Wright, B. D. (1977). Solving measurement problems with the Rasch model. Journal of Educational Measurement, 14, 97–115.

36-Zanon, C., Hutz, C.S., Yoo, H. and Hambleton, R.K. (2016). An application of item response theory to psychological test development. Psicol. Refl. Crít. 29, 18 (2016).https://doi.org/10.1186/s41155-016-0040-