فاعلية نموذج راش في تدريج مفردات اختبار مقرر مهارات التعلم والتفكير على عينة من طلاب جامعة الملك فيصل

"دراسة مقارنة مع النظرية الكلاسيكية"

د. خالد حسن بكر الشريف

قسم علم النفس - كلية التربية - جامعة الملك فيصل

كلية التربية جامعة الإسكندرية

فاعلية نموذج راش في تدريج مفردات اختبار مقرر مهارات التعلم والتفكير على عينة من طلاب جامعة الملك فيصل "دراسة مقارنة مع النظرية الكلاسيكية"

د. خالد حسن بكر الشريف

قسم علم النفس - كلية التربية - جامعة الملك فيصل

كلية التربية جامعة الإسكندرية

تاريخ تقديم البحث: ٢٢/ ٣/ ١٤٣٩هـ **تاريخ قبول البحث:** ١٢/ ٦/ ١٤٣٩هـ

ملخص الدراسة :

هدفت الدراسة التعرف على الخصائص السيكومترية لاختبار مقرر مهارات التعلم والتفكير على عينة مكونة من (١٦٢) من طلاب كلية التربية بجامعة الملك فيصل باستخدام النظرية الكلاسيكية وكذلك تحديد الخصائص السيكومترية لنفس الاختبار باستخدام نموذج راش والتحقق من إمكانية مقارنة درجات الطلاب على الاختبار في النظريتين .

أشارت النتائج إلى اختلاف المفردات المكونة لاختبار مقرر مهارات التعلم والتفكير باختلاف النموذج الإحصائي المستخدم ؛ حيث أفرزت مؤشرات نظرية الكلاسيكية للصدق والثبات اختباراً مكوناً من (٣٠) مفردة، بينما أفرز نموذج راش اختباراً مكوناً 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 في 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 (- 4.03) لوجيت و (4.04) لوجيت .

الكلمات المفتاحية :

نموذج راش – نظرية القياس الكلاسيكية – طلاب جامعة الملك فيصل

المقدمة:

توفر النماذج الحديثة في القياس التربوي (ومنها نموذج راش) فرصاً أفضل لتدريج مفردات الاختبارات التحصيلية والحصول على مؤشرات أكثر دقة للصعوبة والتمييز عند تحليلها إحصائياً، وهذه الدقة تجعلها أكثر تميزاً عن النظرية الكلاسيكية في القياس التربوي والتي تعتمد بشكل أساسي على الدرجة الكلية للمفحوص في الاختبار كمحك للحكم على أدائه، الأمر الذي يجعل مصمم الاختبار غير واعي بسلوك كل مفردة اختبارية على حدى وبالتالي قد يتم حذف مفردات وإضافة مفردات أخرى دون اتخاذ قرار بشأن محتواها أو إعطاءها فرصة ثانية في محاولة لتحسين الأداء عليها، لكن في النماذج الحديثة يتم وصف سلوك المفردة بدقة من خلال منحنى القدرة الخاص بها.

ورغم اجتهاد الباحثين في التحقق من ملاءمة الاختبارات المصممة في ضوء النظرية الكلاسيكية للتطبيق في مستويات عمرية متباينة ؛ من خلال حساب معاملات الثبات والصدق إلا أن هذه الاختبارات تظل متأثرة بالعينات التي طبقت عليها وخصائصها وهذا يجعلها غير مفيدة إذا أعيد استخدامها في بحوث ودراسات على عينات أخرى دون إعادة تقنين. كما أنه جرت العادة في النظرية الكلاسيكية على تركيز التحليل الإحصائي على الدرجة الكلية للاختبار، وكذلك الاعتماد على تكرارات الإجابات الصحيحة لتقدير معاملات صعوبة المفردات، كذلك يكون التركيز على معامل ثبات الاختبار وحساب معاملات ارتباط كل مفردة بالدرجة الكلية للاختبار ومعاملات تمييز المفردات ؛ وهذه الأساليب تستخدم على مدى واسع إلا أنها محدودة لارتباطها بالعينة وخصائصها. (علام، ٢٠٠٧: ٣١)

لذلك كان لزاماً التنبه إلى خطورة ما ينتاب هذه الاختبارات من مشكلات أفقدتها في كثير من الأحيان الدقة والموضوعية، خاصة وأن تفسير درجات

مثل هذه الاختبارات كان ومازال يعتمد على الدرجة الكلية التي لا تعبر بصدق عن المستوى الحقيقي لقدرة المتعلم ؛ فهي مرتبطة بمعامل صعوبة المفردات، فقد لوحظ أن المفردات المختلفة في مستوى الصعوبة –وبالرغم من أنها تقيس ذات المحتوى – تؤدي إلى مستويات مختلفة من القدرة عند ذات المتعلم، ومن جهة أخرى فإن تقدير معامل صعوبة المفردات في هذه الاختبارات يعتمد أيضاً على مستوى أداء المتعلمين في العينة، فمعاملات صعوبة المفردة الواحدة مختلفة باختلاف مستوى أداء أفراد العينة. (الخولي، ٢٠١٣، ٦٥٤)

ونظرية الاستجابة للمفردة تختلف هنا فهي لا تعتمد بالأساس على الدرجة الكلية للاختبار أو التحليل العاملي وإنما تأخذ في الحسبان الخصائص المميزة لمفردات الاختبار ؛ فنظرية الاستجابة للمفردة تختلف عن نظرية القياس الكلاسيكية في إنها تقوم بنمذجة العلاقات بين مستوى سمة المفحوص (القدرة – الاتجاه) ونموذج استجابات المفردة لذلك فإن تقدير سمات الفرد الكامنة من الممكن أن يختلف حتى بين فردين حاصلين على نفس الدرجة الكلية. (Trang Le,2013 ,7)

وتفترض نظرية استجابة الفقرة Item Response Theory(IRT) أنه يمكن التنبؤ بأداء المفحوصين، أو يمكن تفسير أدائهم في اختبار نفسي أو تربوي، في ضوء خاصية مميزة لهذا الأداء تسمى السمة (Trait)، ويصعب ملاحظة هذه السمة مباشرة ؛ لذلك يجب تقديرها أو الاستدلال عليها من أداء المفحوص الذي يمكن ملاحظته على مجموعة من فقرات المقياس أو الاختبار

وتم تطوير نماذج مختلفة للنظرية الحديثة في القياس تعرف باسم نماذج السمات الكامنة (Latent Trait Models, LTM)، تهدف في مجملها إلى تحديد العلاقات بين أداء الفرد في الاختبار والسمة التي تكمن وراء هذا الأداء

وتفسّره. وتقوم النماذج أحادية البعد في نظرية الاستجابة للمفردة على افتراضات أساسية هي: افتراض أحادية البعد (Unidimensionality)، وافتراض الاستقلال الموضعي (Local Independence)، وافتراض المطابقة لمنحنى خصائص الفقرة (Item Characteristic Curve, ICC)، وافتراض التحرر من السرعة (Speededness).

ويعدّ نموذج المعلمة (البارامتر) الواحدة (One Parameter Model) أو ما يسمى بنموذج راش نسبة للعالم (Rasch) من أكثر نماذج نظرية استجابة الفقرة استخداماً؛ ومن المميزات التي أعطت هذا النموذج أهمية خاصة، هو أنه عندما تتطابق البيانات مع هذا النموذج، فإن معالم الفقرات التي تتمثل في درجة صعوبتها يمكن تقديرها بشكل مستقل عن العينة المستخدمة، كما أن قدرات الأفراد يمكن تقديرها بشكل مستقل عن درجة صعوبة المفردات.(حمادنة وخالد، ٢٠١٣، ٨٠)

ويستخدم نموذج راش أحادي البعد في بناء الاختبارات والمقاييس النفسية والتربوية وتطويرها، فقد برهن كفاءته في المقاييس التحصيلية والقدرات لينعكس ذلك على موضوعية النتائج المشتقة من تلك المقاييس وبالتالي دقة تفسيرها.(العلي، ٢٠١٤، ٥٨)

ويعد نموذج راش أيسر نماذج السمات الكامنة وأكثرها استخداماً فى بناء الاختبارات والمقاييس النفسية والتربوية وتطويرها، والفكرة الأساسية التى يقوم عليها هذا النموذج هى أن كل مفردة تحمل شحنة انفعالية تسهم مع المفردات الأخرى فى المقياس فى تكوين شحنة انفعالية إجمالية تعبر عن اتجاه الفرد بما يتفق مع تقديره لتلك المفردة حسب عدد فئات التدريج المستخدم فى المقياس، ويقوم النموذج بتقدير هذه الشحنة لكل فقرة حسب الدالة الرياضية

الاحتمالية التى يعتمدها النموذج، ثم التحقق من شروط ملاءمة الفقرات للنموذج. (أبوهاشم، ٢٠٠٦ : ٢)

ويقوم نموذج راش على نتائج تفاعل قدرة الأفراد مع صعوبة المفردات، ويتمثل هذا على هيئة استجابات ملاحظة، يمكن بها التوصل إلى تدريجات المفردات وتقديرات الأفراد التي تتحقق بها مطالب الموضوعية في القياس؛ لذلك ينبغي بذل الجهد بحيث تكون كل من : قدرة الفرد وصعوبة المفردة هما العاملين الأساسيين اللذين يؤثران ويتحكمان في استجابات الأفراد على مفردات الاختبار. (كاظم، ١٩٨٨، ٢٣)

لذلك فإن الطريقة الأمثل للتحقق من فروض نموذج معين في البيانات هي أن نستعين بالمنحنيات المميزة للمفردات في التنبؤ بأداء الأفراد الذي يمكن ملاحظته ملاحظة مباشرة من استجاباتهم للمفردات. ودقة التنبؤ الذي يمكن تقديره يتخذ دليلاً على تحقق فروض النموذج المستخدم. (علام، ٢٠٠٧، ٢٢١)

مشكلة البحث:

تفتقر غالبية الاختبارات التحصيلية المصممة وفق النظرية الكلاسيكية في القياس إلى الدقة والموضوعية؛ خاصة في ظل تزايد أعداد الطلاب في الجامعات السعودية إذ جرت العادة على تقدير نتائج الطلاب في الاختبارات من خلال مقارنة أدائهم بأداء زملائهم في الصف ذاته، بحيث يقدم الاختبار فكرة عن موقع الطالب بالنسبة لأقرانه، أو بالنسبة للجماعة المعيارية التي ينتمي إليها، لكن الهدف الرئيسي لغالبية البرامج التعليمية في الجامعات هو إكساب الطلاب المعارف والمهارات اللازمة للوصول إلى مستوى مرتفع من الكفاءة وهو الأمر الذي توفره الاختبارات التحصيلية المصممة وفق نماذج الاستجابة للمفردة. حيث :

تشير نتائج دراسات كل من بروجر وسوكان Progar and Socan(2008)، وفان(1998)Fan، إلى أن بارامترات المفردة والفرد المشتقة من كلا النظريتين (نظرية القياس الكلاسيكية ونظرية الاستجابة للمفردة) كانت قابلة للمقارنة ودرجة الثبات في بارامترات المفردة عبر العينات دائماً تتأثر بالتفوق النظري والإمبريقي لنماذج نظرية الاستجابة للمفردة على نظيرتها النظرية الكلاسيكية في القياس ؛ ويميل عدد كبير من الباحثين التربويين في المنطقة العربية عند تصميم اختباراتهم التحصيلية إلى استخدام نظرية القياس الكلاسيكية في تحليل بيانات اختباراتهم نظراً لسهولتها وانتشارها، دون النظر إلى التفوق النظري والإمبريقي الذي يميز نماذج نظرية الاستجابة للمفردة.

وتستخدم النظرية الكلاسيكية مفهوم الخطأ المعياري للدلالة على أخطاء القياس ؛ لكن في الواقع تستخدم الانحراف المعياري للدرجة الملاحظة وثبات الاختبار كمؤشرات لتقدير حجم الخطأ المعياري في القياس، وكلما زادت قيمة الخطأ المعياري في القياس كلما قلت الثقة في الدرجة المقاسة.

حيث يكون التركيز في التحليل منصباً على الدرجة الكلية للفرد في الاختبار، وتكرارات الإجابات الصحيحة عن كل مفردة كمؤشر لمعامل الصعوبة، الإجابات الأكثر تكراراً كمعيار لجودة المشتتات (البدائل) في مفردات الاختيار من متعدد، وثبات الاختبار، ومعامل ارتباط درجة كل مفردة بالدرجة الكلية للتعبير عن معامل تمييز المفردة ورغم استخدام هذه المؤشرات الإحصائية على مدى واسع إلا أنها مرتبطة بالعينة وخصائصها. (Magno,2009,2)

في حين يكون التركيز في نظرية الاستجابة للمفردة منصباً على الخصائص المميزة للمفردات وسلوكها كأساس للتحليل الإحصائي حيث يكون لكل مفردة منحناها الخاص الذي يصف سلوكها وبارامترات الصعوبة والتمييز

الواضحة ؛ حيث يكون التركيز في النماذج أحادية البعد مثل نموذج راش على بارامترات صعوبة المفردات ؛ حيث ينطلق نموذج راش من فرضية أن كلا من (معامل التخمين، والفروق بين المفردات في معامل التمييز) إما ضئيل يمكن إهماله أو ثابت يمكن تحديد قيمته.

وخلال العقود الأخيرة استخدمت نظرية الاستجابة للمفردة في التقويم التربوي في تطوير العديد من الاختبارات التربوية الأجنبية وأشهرها اختبار الاستعداد المدرسي Scholastic Aptitude Test (SAT)، واختبار تسجيل التخرج Graduation Record Examination (GRE) كلها اختبارات مصممة باستخدام نظرية الاستجابة للمفردة لأن منهجيتها تسمح بدقة القياس والثبات وتوفر الوقت والجهد في عملية القياس كما أن أغلبها يعتمد على الحاسب الآلي في التطبيق والتحليل ؛ وفي الفترة الأخيرة أصبحت نماذج الاستجابة للمفردة أكثر استخداماً كذلك في مجالات أخرى مثل الصحة، وجودة الحياة، وحتى البحوث الإكلينيكية.

وقد اتجهت بعض الدراسات إلى المقارنة بين النظريتين مثل دراسات Courville (2004) الحكماني (٢٠١٠)، وفراج والشريف (٢٠١٣) ؛ حيث أشارت نتائج هذه الدراسات إلى أن البيانات المستمدة من النموذجين قابلة للمقارنة لكن البيئة العربية في حاجة إلى المزيد من الاختبارات التي تقيس القدرات التحصيلية بدقة وتهتم ببارامترات المفردة والفرد كأساس لتفسير درجات المفحوصين ؛ لكن الدراسات المقارنة يجب أن تتخطى نقطة إثبات التفوق النظري والامبريقي لنظرية الاستجابة للمفردة بنماذجها على نظرية القياس الكلاسيكية، فهي بغض النظر عن أنها محسومة أو مثبتة في العديد من الدراسات، لم تؤدي إلى تعميم نظرية الاستجابة للمفردة كأساس لتحليل بيانات الاختبارات التحصيلية في البيئة العربية لذلك فنحن في حاجة لمزيد من

البحوث التي تولد المزيد من الاختبارات التحصيلية المحكية المرجع كخطوة نحو بنوك أسئلة متاحة في كل تخصص أكاديمي بالجامعة.

ويؤكد ذلك (أبوهاشم، ٢٠٠٦) أنه يغلب على الاختبارات والمقاييس المستخدمة فى مجال العلوم الإنسانية فى البيئة العربية اعتمادها عند بنائها وتقنينها على نظرية القياس الكلاسيكية، على الرغم من ظهور بعض العيوب المرتبطة بهذه النظرية، لذا جاءت فكرة استخدام أحد النماذج الحديثة فى القياس وهو نموذج راش من أجل معرفة أهم الخصائص السيكومترية التى يحققها اختبار مقرر مهارات التعلم والتفكير لطلاب كلية التربية جامعة الملك فيصل مقارنة بالنظرية الكلاسيكية للقياس لذلك يلخص الباحث مشكلته في ضرورة الإجابة على الأسئلة التالية :

(١)ما الخصائص السيكومترية التى تتوفر لاختبار مقرر مهارات التعلم والتفكيرباستخدام نموذج راش ؟

(٢) هل تنطبق بيانات العينة في اختبار مقرر مهارات التعلم والتفكير على نموذج راش ؟

(٣) ما معاملات صعوبة مفردات اختبار مقرر التعلم والتفكير بعد تدريجه على نموذج راش ؟

(٤) هل يتغير تدريج مفردات اختبار مقرر مهارات التعلم والتفكير المطورة باستخدام نموذج راش بتغير مستوى قدرات أفراد العينة المستخدمة فى الحصول على هذا التدريج ؟

(٥) ما الخصائص السيكومترية التى تتوفر لاختبار مقرر مهارات التعلم والتفكيرباستخدام النظرية الكلاسيكية ؟

أهداف البحث :

يهدف البحث الحالى إلى :

(١) التعرف على الخصائص السيكومترية لاختبار مقرر مهارات التعلم والتفكير على عينة من طلاب جامعة الملك فيصل باستخدام النظرية الكلاسيكية.

(٢) تحديد الخصائص السيكومترية لاختبار مقرر مهارات التعلم والتفكير على عينة من طلاب جامعة الملك فيصل باستخدام نموذج راش.

(٣) مطابقة بيانات العينة في اختبار مقرر مهارات التعلم والتفكير على نموذج راش.

(٤)الكشف عن الفروق بين مرتفعي ومنخفضي االقدرة التحصيلية في الأداء على اختبار مقرر مهارات التعلم والتفكير باستخدام بارامترات (المفردة/ الفرد) في نموذج راش.

أهمية البحث:

-توجيه أنظار الباحثين والمهتمين بالقياس النفسي والتربوي في المنطقة العربية إلى ضرورة استخدام القياس الموضوعي المتمثل بنظرية الاستجابة للمفردة، والاستفادة من خصائصها ومزاياها الموضوعية.

-يساعد الاختبار الذي تم بناءه في ضوء نظرية الاستجابة للمفردة على فهم كيفية التحقق من ملائمة البيانات لنموذج راش، وكذلك كيفية اختبار شروط نظرية الاستجابة للمفردة وخصوصاً شرطي أحادية البعد والاستقلال الموضعي.

وتكمن أهمية البحث كذلك في إلقاء المزيد من الضوء على فعالية نظرية الاستجابة للمفردة في تصميم الاختبارات التحصيلية بالجامعة وتدريج مفرداتها وفق أحد نماذج النظرية وهو نموذج راش.

وبشكل عام فإن ظهور نماذج نظرية الاستجابة للمفردة (IRT) وتطويرها فيما بعد، أدى إلى حل معظم المشكلات التي واجهت الاختبارات المصممة وفق نظرية القياس الكلاسيكية، والتي تؤثر في موضوعية ودقة القياس ؛

فتقدير مقدار السمة أو القدرة عند المتعلم في اختبار ما وفق أحد نماذج هذه النظرية مستقل عن تقدير معالم مفردات هذا الاختبار وتقدير معالم المفردات وفق أحد نماذج هذه النظرية مستقل عن مستوى قدرة أفراد عينة التقدير، بمعنى أن تقدير قدرة المتعلم في اختبار ما وفق أحد نماذج نظرية الاستجابة للمفردة سوف لن يتغير في حال طبق عليه اختبار آخر يقيس مستوى مختلف من ذات المحتوى، وبالتالي وفق نظرية الاستجابة للمفردة، فإن صعوبة الاختبار لم تعد سبباً في تدني تقدير قدرة المتعلم في هذا الاختبار، وأن سهولة الاختبار لم تعد سبباً في حصول المتعلم على تقدير مرتفع فيه، مما يؤدي إلى تحقيق العدل بين جميع الطلاب أفقياً في ذات العام الدراسي، ورأسياً عبر السنين وخاصة في الاختبارات المصممة أصلاً لتصنيف الطلاب، وتوزيعهم على الكليات المختلفة أو الاختصاصات المختلفة في ذات الكلية. (الخولي، ٢٠١٣، ٦٥٥)

ويشير آن ويونج An and Yung (2014) إلى أن نظرية الاستجابة للمفردة خلقت واقعاً جديداً في مجالات القياس النفسي يهدف إلى تقدير حقيقي للقدرات المقاسة ؛ وتستخدم في مجال التربية وو التعليم على مدى واسع بهدف تدريج وتقويم المفردات الاختبارية في الاختبارات والاستبيانات وأدوات جمع البيانات الأخرى ومن أجل إعطاء الطلاب درجات حقيقية مساوية لقدراتهم واتجاهاتهم وسماتهم الكامنة الأخرى حسب موضوع عملية القياس.

مصطلحات البحث:

نموذج راش : أحد نماذج نظرية الاستجابة للمفردة، وهو نموذج أحادي البعد يقوم على فرضية أن احتمالية الإجابة الصحيحة على كل مفردة دالة في متغير واحد فقط هو صعوبة المفردة، ويستخدم في تحليل المفردات ثنائية

الإجابة، ويقوم على فرضية أن معاملات تمييز المفردات كلها ثابتة يمكن تحديدها أو افتراض أنها تساوي الواحد الصحيح (Magno,2009,3)

تعريف قدرة الفرد : إن قدرة الفرد مقدرة باللوجيت، هي اللوغاريتم الطبيعي لمميز نجاح الفرد على المفردات التي تعبر نقطة صفر التدريج عن صعوبتها.

تعريف صعوبة المفردة : إن صعوبة المفردة مقدرة باللوجيت هي اللوغاريتم الطبيعي لمميز الفشل الأفراد التي تعبر نقطة صفر التدريج عن قدرتهم.

اللوجيت : هي وحدة قياس كل من قدرة الفرد وصعوبة الفقرة، وتعرف باللوغاريتم الطبيعي لمميز نجاح الفرد على المفردات التي تعبر نقطة صفر التدريج عن صعوبتها، عندما يساوي هذا المميز ثابتاً هو الأساس الطبيعي (e) أي (٢,٧٢) ويكون احتمال نجاح هذا الفرد حينئذ هو (٠,٧٣).

(كاظم، 1988، ٢٧ -٢٨)

الإطار النظري:

النظرية الكلاسيكية في القياس

تعرف نظرية الاختبارات الكلاسيكية على أنها نظرية الدرجة الحقيقية وتبدأ النظرية من فرضية أن هناك درجة ملاحظة تختلف عن الدرجة الحقيقية للقدرة المقاسة، والنموذج الأساسي الذي تعتمد عليه نظرية القياس الكلاسيكية هو أن

الدرجة الملاحظة = الدرجة الحقيقية + الدرجة الخطأ

حيث أن الدرجة الحقيقية والدرجة الخطأ مستقلتان عن بعضهما البعض ويمكن اختصاراً التعبير عن العلاقة السابقة بالصيغة التالية :

$$TO = T + E$$

وتفترض النظرية الكلاسيكية أن لكل فرد درجة حقيقية يمكن الحصول عليها إذا لم يكن هناك أخطاء في عملية القياس، لكن بسبب أن أدوات القياس ليست دقيقة بنسبة ١٠٠٪ فإن الدرجة الملاحظة لكل فرد تختلف عن قدرة الفرد الحقيقية، والفرق بين الدرجة الحقيقية والدرجة الملاحظة يأتي من أخطاء القياس، وباستخدام عدد كبير من معايير التحقق من دقة الدرجة فإن الخطأ غالباً يفترض أنه متغير عشوائي يخضع للتوزيع الاعتدالي ومضامين نظرية القياس الكلاسيكية للمفحوصين أن أدوات القياس هي أدوات عرضة للخطأ؛ لذلك فإن الدرجة التي يحصل عليها الفرد في اختبار تحصيلي مصمم ومفسر بمنهجية القياس الكلاسيكية نادراً ما تتطابق مع درجته الحقيقية في القدرة التحصيلية المقاسة؛ لذلك فإن الدرجة الملاحظة هي الدرجة الحقيقية لكن متأثرة ببعض الأخطاء، هذا الخطأ يؤثر في الدرجة الملاحظة : قد يرفعها وقد يخفضها. (Magno,2009,1-2)

نظرية الاستجابة للمفردة

تعد نظرية الاستجابة للمفردة الاختبارية IRT من النظريات السيكومترية المعاصرة التي اتضحت فائدتها في التغلب على كثير من مشكلات القياس التي عجزت النظرية الكلاسيكية عن مواجهتها.

فهذه النظرية تحاول نمذجة العلاقة القائمة بين مستوى سمة معينة لدى الفرد Trait Level التي يقيسها اختبار معين، واستجابته لمفردة من مفردات الاختبار. ونظراً لأن مستويات السمة غير منظورة بطبيعتها وتؤثر في المتغيرات المنظورة، فإنه يتم تقديرها من هذه المتغيرات، أي من الاستجابات الملاحظة. وعادة يفترض أن السمة التي تعبر عنها الاستجابات بمثابة قدرة Ability من نوع معين، أي خاصية 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 استناداً للنموذج الاعتدالي التراكمي Normal Ogive Model وذلك في العقدين الثالث والرابع من القرن الماضي، 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 الكلاسيكية، فإن نظرية الاستجابة للمفردة تستند إلى افتراضات قوية فيما يتعلق بما يحدث عند استجابة الفرد لمفردات اختبارية. وتؤدي هذه الافتراضات إلى ميزات متعددة. فمثلاً يكون وصف المفردة الاختبارية بطريقة مستقلة عن الأفراد الذين اختبروا بها، ووصف الفرد بطريقة مستقلة عن أي عينة من المفردات التي اختبر بها، وكذلك يمكن التنبؤ بخصائص الاختبار قبل تطبيقه.(علام، ٥٣، ٢٠٠٥)

وتتضمن نظرية الاستجابة للمفردة نماذج ثنائية البعد تعتمد على معاملات الصعوبة ومعاملات التمييز، ونماذج أخرى ثلاثية البعد لاحتمالية تطبيق الاختبار على منخفضي التحصيل لأخذ احتمالات التخمين للحصول على إجابات صحيحة، هذا بالإضافة إلى النماذج أحادية البعد التي تعتمد فقط على بارامتر صعوبة المفردة مثل نموذج راش كل هذه النماذج تحدد خصائص المفردة في علاقتها باحتمالية أن يجيب عليها مرتفعي ومنخفضي التحصيل إجابات صحيحة بمستويات مختلفة من الأداء. (Magno,2009,3)

نموذج راش :

يعد نموذج القياس الذي وضعه عالم الرياضيات الدانمركي (جورج راش) عام ١٩٦٠ وطوعه للتطبيق العملي العالم الأمريكي (بن رايت) أهم نماذج السمات الكامنة ؛ حيث تتوفر متطلبات الموضوعية عندما تستوفي فروض النموذج وهي :

١ -**أحادية البعد** : حيث تعرف السمة موضوع القياس بواسطة مجموعة من المفردات ذات صعوبة أحادية البعد، أي أن مفردات الاختبار لاتختلف فيما بينها إلا من حيث مستوى الصعوبة فقط. كما يكون الأفراد ذوي قدرة أحادية البعد تحدد وحدها مستوى أدائهم على الاختبار.

٢ -**استقلالية القياس** : ويعني ذلك أن لا يعتمد تقدير صعوبة المفردة، على تقديرات صعوبة المفردات الأخرى المكونة للاختبار، ولا على تقديرات قدرة الأفراد الذين يجيبون عليها. ولا يعتمد تقدير قدرة الفرد، على تقديرات قدرة أي مجموعة من الأفراد الذين يؤدون الاختبار أو على تقديرات صعوبة المفردات التي يؤدونها

وتتحقق استقلالية القياس هذه عندما تتحقق :

أ -ملاءمة المفردات المكونة للاختبار للنموذج، وهذا يعتمد على صدق المفردات في قياس القدرة موضوع القياس، وعلى صدق تدرجها على مقياس هذه القدرة.

ب -ملاءمة الأفراد الذين يجيبون على مفردات الاختبار للنموذج، وهذا يعتمد على صدق استجابات هؤلاء الأفراد.

ج -توافق مفردات الاختبار مع قدرات الأفراد، ويعتمد هذا على تقارب مستوياتها.

٣ -**توازي المنحنيات المميزة للمفردات** : يوضح المنحنى المميز للمفردة، احتمالات الاستجابة الصحيحة على هذه المفردة للأفراد عند المستويات المختلفة من القدرة. ومعنى توازي المنحنيات المميزة للمفردات، هو أنه إلى الحد الذي تميز فيه المفردات بين الأفراد ذوي المستويات المختلفة من قدرة ما، فإن جميع هذه المفردات ينبغي أن يكون لها نفس القدرة على التمييز. (كاظم، ١٩٨٨، ٢٢ -٢٣)

الاختبارات التحصيلية وفق نموذج راش :

وتعد الاختبارات المصممة وفق نماذج نظرية الاستجابة للمفردة من أكثر أنواع الاختبارات ملاءمة لقياس وتقويم تحصيل الطلاب وتقدير قدرات الأفراد العقلية، وذلك لأنها تقوم على تحديد المهارات والكفايات المطلوب إتقانها بدقة فائقة لكي يتمكن الباحث من قياسها وملاحظتها بشكل مباشر، والمشكلة في النظرية الكلاسيكية للقياس أن الدرجات التي يحصل عليها الأفراد تخضع لخصائص عينة الأفراد التي تستخدم في اشتقاق معايير الاختبار، كما تخضع لخصائص عينة الأسئلة التي يشتمل عليها الاختبار واختلاف هذه الخصائص باختلاف عينة الأفراد المختبرين وعينة الأسئلة يؤثر بلاشك في صدق إجراء هذه الموازنات، وبذلك يصعب تعميم نتائج الاختبارات والاستفادة منها. (علام، ٢٠٠٧، ١٨ -١٩)

ويستخدم نموذج راش في تحليل بيانات المفردات الاختبارية ثنائية الدرجة، وهو نموذج أحادي البعد ويعد من أبسط نماذج الاستجابة للمفردة وهو أحد النماذج التي أطلق عليها راش نماذج القياس Measurement Models وقد طور هذا النموذج بطريقة مستقلة عن بقية نماذج الاستجابة للمفردة، ويمكن اعتباره نموذجاً يكون فيه المنحنى المميز للمفردة تمثله دالة ترجيح لوغاريتمي أحادية البارامترات One-Parameter Logistic Function.(علام، ٢٠٠٥، ٦٩)

ويعتمد نموذج راش على فرضيتين أساسيتين في تحليل مفردات الاختبارات التحصيلية :

-يزيد احتمال الإجابة الصحيحة عن مفردة اختبار بزيادة قدرة الفرد.

-احتمال أن يجيب فرد إجابة صحيحة عن مفردة سهلة، أكبر من احتمال أن يجيب إجابة صحيحة عن مفردة صعبة. (علام، ٢٠٠٧، ٢٣٣)

أوجه الاختلاف بين النظرية الكلاسيكية (التقليدية)، والنظرية المعاصرة في القياس :

(١)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، اعتماداً على درجات أفراد جماعة مرجعية معينة.

وتختلف قيم معاملات الصعوبة والتمييز اختلافاً ملحوظاً 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 تقديره في إطار بعض نماذج النظرية المعاصرة، إذا افترضنا وجود مجتمع كبير من الأفراد.

(٢)تفسير درجات الاختبار

تعتمد النظرية الكلاسيكية في تفسيرها لدرجات الاختبار على معيار جماعة مرجعية معينة Norm Group، وذلك بتحويل الدرجات الخام في الاختبار إلى درجات معيارية أو بأي تحويل خطي استناداً إلى التوزيع الاعتدالي. بينما تعتمد النظرية المعاصرة 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 تفسيراً مباشراً.

فإذا كانت درجة قدرة الفرد مساوية لصعوبة المفردة، فإن احتمال إجابة الفرد 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 أيضاً إجراء تحويلات خطية معينة على الدرجات بحيث يتيسر تفسير الدرجة في إطار الجماعة المرجعية.

٣ -تحقق خصائص الميزان الفتري : تعتمد النظرية الكلاسيكية في انتقاء مفردات الاختبارات السيكولوجية على فرضية أن درجاتها تتخذ شكل التوزيع الاعتدالي في المجتمع المستهدف. ويمكن أن يتحقق ذلك بانتقاء المفردات التي تكون قيم معامل 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 الخصائص 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 قصور النظرية الكلاسيكية، ولكن بعض نماذج النظرية المعاصرة 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 اعتدالياً، حيث أن النموذج يستند إلى مبادىء أساسية في القياس Fundamental Measurement ومقارنات تتصف بعدم التغير Invariance مما يعد أحد ميزات هذه النماذج.

٤ -خصائص الخطأ المعياري للقياس : يعد الخطأ المعياري من الخصائص السيكومترية الأساسية للاختبارات، حيث إنه يستند إليه في تفسير كل درجة من درجاتها. وتعتبر النظرية السيكومترية للقياس أن الخطأ المعياري Standard Error ثابت بالنسبة لمجتمع معين من الأفراد المختبرين. ولكن نظرية القياس المعاصرة لا تفترض تساوي تباين أخطاء القياس لجميع الأفراد الذين يطبق عليهم الاختبار. فبعض الأفراد يكون أداؤهم في الاختبار أكثر اتساقاً من غيرهم من الأفراد، ويختلف هذا الاتساق باختلاف مستوى قدرتهم ؛ فبدلاً من تقديم خطأ 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 مؤشراً لهذه الأخطاء في مجتمع هؤلاء الأفراد.

٥ -تأثر الثبات بعدد مفردات الاختبار : فالنظرية الكلاسيكية للقياس تعتبر أنه كلما زاد عدد مفردات الاختبار ارتفع ثبات درجاته. ويمكن التنبؤ بقيمة الثبات عندما يزداد أو يقل عدد مفردات الاختبار باستخدام الصيغة الرياضية التي اقترحها سبيرمان وبروان Spearman- Brown Formula. أما نظرية القياس المعاصرة فإنها تؤكد أن الاختبارات القصيرة التي تشتمل على نوعية جيدة من المفردات يمكن أن تكون أكثر ثباتاً إذا افترضنا أن تميز مفرداتها ثابت، كما تؤكد أن الدرجات المتطرفة تكون أخطاؤها المعيارية أكبر.(علام، ٥٤، ٢٠٠٥ -٥٦)

ويتضح للباحث مما سبق أن نظرية الاستجابة للمفردة قدمت عدة محكات موضوعية للقياس، وليس محك واحد ؛ حتى مفهوم الخطأ المعياري للقياس

–المحك الذي استخدمته النظريةالكلاسيكية - لم تهمله النظرية وقدمت عدة تقديرات له باختلاف مستويات القدرة، وكذلك استطاعت النظرية الحصول على قدر مناسب من الثبات للقدرات المقاسة باختلاف العينات، والأكثر من ذلك هو الاهتمام بخصائص المفردات وسلوكها من خلال بارامترات (المفردة / الفرد) والحصول على منحنى مميز لكل مفردة باختلاف مستويات قدرة الأفراد؛ هذا المنحنى يمثل راسم الدالة الاحتمالية لإمكانية الحصول على إجابة صحيحة لهذه المفردة، كما أدى ذلك إلى الاهتمام بتحسين نوعية المفردات أكثر من الاهتمام بزيادة عددها.

الدراسات السابقة:

دراسة كارلسون ولوكلن Carlson and Locklin(1995)

وهدفت هذه الدراسة إلى بحث دقة وفعالية أساليب تدريج المفردات تبعاً لنظرية الاستجابة للمفردة في برنامجين من برامج الكمبيوتر هما Micro CAT، وBilog يستخدمان لتدريج الاختبار المستخدم في التعليم العالي، تدريج لاستجابات ١٠٠٠ من الطلاب المختارة بياناتهم عشوائياً في إجاباتهم على ٧٢ سؤال في الرياضيات موجهة لكل الطلاب المستجدين الملتحقين بالجامعة البحث الشرقية، وتم تفسير البيانات باستخدام خيارات متعددة في برنامجين من برامج الكمبيوتر تم مقارنة تدريج المفردات باستخدام البارامترات المحددة الملائمة للمفردات وتقدير القدرات الناتجة ووجد قدر كبير من الاتفاق بين البرنامجين في بارامترات المفردات، وكان برنامج Bilog أكثر ملائمة لنموذج البارامترات الثنائي والثلاثي، وتقدير القدرات كان أيضاً متشابه الفروق الملحوظة كانت واضحة في تقدير قدرة الطلاب ذوي الدرجات المنخفضة. تأثير استخدام عينة من كل استجابات الطلاب الكاملة كانت

متعارضة مع العينة المتضمنة للاستجابات المحذوفة التي تبدو صغيرة بغض النظر عن البرنامج المستخدم في التحليل.

دراسة فان (1998) Fan

هدفت إلى بحث قلة البحوث الامبريقية عن الاختلافات النظرية بين كل من النظرية الكلاسيكية، ونظرية الاستجابة للمفردة، وإلى أي مدى تسلك المفردات وإحصائيات الفرد في كلا الإطارين النظريتين، وبحثت الدراسة سلوكيات المفردة وإحصائيات الفرد المشتقة من كلا النظريتين وإطار عمل كل نظرية وركزت الدراسة على سؤالين الأول : ما العلاقات الامبريقية بين نظرية الاستجابة للمفردة والنظرية الكلاسيكية في القياس في بناء المفردات وإحصاءات الفرد، والثاني : إلى أي مدى تختلف إحصاءات الفرد من نظرية الاستجابة للمفردة والنظرية الكلاسيكية باختلاف العينات ؟

بنيت الدراسة على قاعدة بيانات واسعة وأشارت النتائج إلى أن إحصاءات المفردة والفرد المشتقة من كلا النظريتين كانت قابلة للمقارنة ودرجة الثبات في إحصاءات المفردة عبر العينات دائماً تتأثر بالتفوق النظري والإمبريقي لنماذج نظرية الاستجابة للمفردة على نظيرتها النظرية الكلاسيكية في القياس، واتضح كذلك أن هناك ثبات في إحصاءات المفردات في كلا الإطارين القياسين عبر العينات.

دراسة غادة عيد (٢٠٠٤)

هدفت إلى الكشف عن الفروق في الدرجات الخام لليقظة العقلية بين الذكور والإناث وبين التخصصات العلمية والأدبية، والكشف عن الفروق في الدرجات الحقيقية لليقظة العقلية والمقدرة بنموذج راش بين الذكور والإناث، وبين التخصصات الأدبية والعلمية، والكشف عن الفروق في الدرجات الحقيقية لليقظة العقلية المقدرة بالنماذج الكلاسيكية بينالذكور والإناث، وبين

التخصصات العلمية والأدبية، والكشف عن الفروق بين الدرجة الحقيقية المقدرة بنموذج راش وتلك المقدرة بالنظرية الكلاسيكية. وتكونت عينة الدراسة من ٢٥٠ طالب وطالبة من كلية التربية، بقسميها العلمي ١٥٠ طالب، والأدبي ١٠٠ طالب، وطبق عليهم اختبار اليقظة العقلية كمقياس للذكاء العام.

وباستخدام اختبار ت T-Test بينت النتائج : وجود فروق ذات دلالة إحصائية بين متوسطي الدرجات الخام للطلاب والطالبات في مقياس اليقظة العقلية، وكذلك وجود فروق ذات دلالة إحصائية بين متوسطي التخصصات الأدبية والعلمية في مقياس اليقظة العقلية، وعدم وجود فروق ذات دلالة إحصائية بين متوسطي الدرجات الحقيقية (نموذج راش) للطلاب والطالبات على مقياس اليقظة العقلية، وعدم وجود فروق ذات دلالة إحصائية بين متوسطي الدرجات الحقيقية للتخصصات العلمية والأدبية على مقياس اليقظة العقلية، عدم وجود فروق ذات دلالة إحصائية بين متوسطي الدرجات الحقيقية المقدرة بالنظرية الكلاسيكية (التقليدية) للطلاب والطالبات على مقياس اليقظة العقلية. عدم وجود فروق ذات دلالة إحصائية بين متوسطي الدرجات الحقيقية المقدرة بالنظرية الكلاسيكية للتخصصات العلمية والتخصصات الأدبية في مقياس اليقظة العقلية، وعدم وجود فروق ذات دلالة إحصائية بين متوسطي الدرجات الحقيقية المقدرة بنموذج راش وتلك المقدرة بالمدخل الكلاسيكي لمقياس اليقظة العقلية.

دراسة كورفيل Courville(2004)

هدفت الدراسة إلى المقارنة بين النظرية الكلاسيكية في القياس ونظرية الاستجابة للمفردة، من خلال استخدام مدى واسع من البيانات المرجعة إلى معيار، وكيفية تفسير بارامترات (المفردة /الفرد) في كل نظرية، وتضمنت

الدراسة سؤالين رئيسيين ؛ الأول : إلى أي مدى بارامترات (المفردة/الفرد) المشتقة من النظرية الكلاسيكية للقياس والمشتقة من نظرية الاستجابة للمفردة كانت قابلة للمقارنة.؟

الثاني : مامدى ثبات بارامترات المفردة المشتقة من كلا النظريتين باختلاف عينات المفحوصين؟

أشارت النتائج إلى أنه في ظروف متباينة فإن كلا النظريتين ينتج بارامترات(المفردة / الفرد) متشابهة، الأكثر من ذلك أنه بالرغم من أن المؤيدين لنظرية الاستجابة للمفردة ركزوا نقاشهم حول استخدامها لتفوقها على النظرية الكلاسيكية في خاصية الثبات إلا أن بارامترات المفردات المشتقة من النظرية الكلاسيكية على نفس العينة كانت أيضاً ثابتة.

دراسة أبو هاشم (٢٠٠٦)

وهدفت إلى التعرف على الخصائص السيكومترية لمقياس مداخل الدراسة على عينة من طلاب جامعة الزقازيق باستخدام النظرية التقليدية والتعرف على الفروق بين مرتفعى ومنخفضى التحصيل الدراسى من طلاب الجامعة فى مداخل الدراسة التقليدية، وكذلك تحديد الخصائص السيكومترية لمقياس مداخل الدراسة على عينة من طلاب كلية التربية جامعة الزقازيق باستخدام نموذج راش والتحقق من إمكانية التنبؤ من درجات الطلاب على مقياس مداخل الدراسة المشتقة من نموذج راش بالتحصيل الدراسى لديهم

تكونت العينة من ٢٤٤ طالب من كلية التربية جامعة الزقازيق، وتضمنت الدراسة ترجمة لمقياس مدخل الدراسة، وأشارت النتائج إلى أنه تتوفر درجة مرتفعة من الصدق والثبات لمقياس مداخل الدراسة باستخدام النظرية التقليدية وأن هناك تأييداً قوياً لنموذج "انتوستل" لمداخل الدراسة (العميق - السطحى – الاستراتيجى)، حيث أسفرت نتائج النظرية

التقليدية ونموذج راش وضوحها لدى عينة من طلاب كلية التربية جامعة الزقازيق.

اختلاف المفردات المكونة لمقياس مداخل الدراسة باختلاف النموذج الإحصائى المستخدم، حيث أفرزت المؤشرات التقليدية للصدق والثبات مقياس مكون من (٢٨) مفردة، بينما أفرز نموذج راش مقياس مكون من (٢١) مفردة، كما وجدت فروق ذات دلالة إحصائية بين متوسطات درجات الطلاب مرتفعى ومنخفضى التحصيل الدراسى (كمحك) على مقياس مداخل الدراسة التقليدية. وانطبقت البيانات المستمدة من مقياس مداخل الدراسة على نموذج راش، كما أنه لم يتغير تدريج مفردات مقياس مداخل الدراسة المطورة بتغير مستوى قدرات أفراد العينة المستخدمة فى الحصول على هذا التدريج، وكانت النسخة المطورة لمقياس مداخل الدراسة باستخدام نموذج راش منبئ جيد بالتحصيل الدراسى لدى طلاب الجامعة.

دراسة طيفور (٢٠٠٧)

وهدفت إلى دراسة مقارنة لنماذج نظرية الاستجابة للمفردة في معادلة درجات اختبارين في الجبر أحدهما سهل والآخر صعب ؛ واستخدام نماذج الاستجابة للمفردة وهي النموذج الأحادي المَعْلَم و النموذج الثنائي المَعْلَم و النموذج الثلاثي المَعْلَم، في تدريج مفردات اختبار محكي المرجع وتحديد المفردات الملائمة لكل نموذج من هذه النماذج.

أشارت النتائج إلى أنه عند استخدام تصميم المفردات المشتركة، كان النموذج الأحادي المَعْلَم أكثر النماذج الثلاثة دقة في معادلة درجات الاختبارات، يليه في ذلك النموذج الثنائي المَعْلَم، وكان النموذج الثلاثي المَعْلَم أقل النماذج الثلاثة دقة .النماذج الثلاثة: الأحادي، والثنائي، والثلاثي

المَعْلَم، متكافئة في معادلة درجات الاختبارات باستخدام تصميم الأفراد المشتركين.

عند استخدام تصميم المجموعات المتكافئة، كان النموذج الأحادي المَعْلَم، أكثر النماذج الثلاثة دقة في معادلة درجات الاختبارات، أما النموذجان الثنائي والثلاثي المَعْلَم فكانا متكافئين.

إذا كانت الاختبارات مدرجة وفق النموذج الأحادي المَعْلَم، فإن استخدام تصميم المفردات المشتركة لمعادلة درجات الاختبارات، يعطي نتائج أكثر دقة من تصميمي الأفراد المشتركين والمجموعات المتكافئة، واللذين يكونان متطابقين.

إذا كانت الاختبارات مدرجة وفق النموذج الثنائي المَعْلَم، فإن استخدام تصميم المفردات المشتركة لمعادلة درجات الاختبارات، يعطي نتائج أكثر دقة من تصميمي الأفراد المشتركين والمجموعات المتكافئة، يلي ذلك تصميم الأفراد المشتركين، وأقلها دقة تصميم المجموعات المتكافئة.

استخدام تصميم الأفراد المشتركين يعطي نتائج أكثر دقة في معادلة درجات الاختبارات المدرجة وفق النموذج الثلاثي المَعْلَم، من تصميمي المفردات المشتركة والمجموعات المتكافئة، واللذين يكونان متكافئين.

دراسة بروجر وسوكان Progar and Socan(2008)

وهدفت إلى مقارنة إمبريقية بين نظرية القياس الكلاسيكية ونظرية الاستجابة للمفردة؛ حيث أشارت النتائج إلى أن بارامترات المفردة والفرد في كلا النظريتين قابلة للمقارنة، وأظهرت النظريتان نفس الثبات في تقدير البيانات باختلاف العينات، لكن بيانات بارامترات الفرد بناء على نظرية الاستجابة للمفردة كانت أكثر ثباتاً باختلاف المفردات من نظرية القياس الكلاسيكية.

النتائج أفادت كذلك إلى أنه نظراً إلى خاصية الثبات فإن بارامترات المفردة والفرد بناء على نظرية الاستجابة للمفردة عموماً كانت أكثر تفوقاً إمبريقياً على بارامترات نظرية القياس الكلاسيكية ولكن بشرط اختيار نموذج التحليل المناسب للبيانات من نماذج نظرية الاستجابة للمفردة.

دراسة أبوجراد (٢٠٠٨)

هدفت هذه الدراسة إلى إعادة تدريج اختبار الذكاء الثالث كاتل الصورة (أ) باستخدام نموذج راش الأحادي المعلم، وإلى 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) طالب وطالبة، واستخدم برنامج SPSS13، وبرنامج Winsteps(v.2.88) لإجراء المعالجات الإحصائية اللازمة وفق نموذج راش أحادي 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 إلى معايير للاختبار عن طريق إيجاد الرتب المئينية، والدرجات التائية، ونسب الذكاء الانحرافية المناظرة للتقديرات المختلفة لقدرة الأفراد لكل من الذكور والإناث.

دراسة الثوابية (٢٠١٠)

هدفت هذه الدراسة إلى استقصاء أثر حجم العينة في تقدير معلمة صعوبة الفقرة (Item Difficulty)، والخطأ المعياري في تقديرها (Standard Error of Estimation) باستخدام نظرية الاستجابة للفقرة (Item Response Theory) ولتحقيق أهداف هذه الدراسة تم اشتقاق معلمة الصعوبة، والخطأ المعياري في تقديرها باستخدام اختبار تحصيلي في الرياضيات للصف العاشر الأساسي تكون في صورته النهائية من (٨٠) فقرة من نوع الاختيار من المتعدد.

وطبق الاختبار على عينات عشوائية طبقية تراوح حجمها ما بين (٢٠٠) إلى (١١٢٩٢) طالباً وطالبة وقد استخدم الباحث برمجية (BILOG- MG) لتقدير معلمة الصعوبة، والخطأ المعياري في تقديرها، وقد توصلت الدراسة إلى النتائج الآتية :

تزداد قيمة معلمة صعوبة الفقرة بزيادة حجم العينة. حيث بلغ متوسط صعوبة فقرات الاختبار (٠,٣١) لوجيت عندما كان حجم العينة (٢٠٠) طالب وطالبة ازداد بحيث أصبح (١,١) لوجيت عندما كان حجم العينة (١١٢٩٢) طالب وطالبة.

يتناقص الخطأ المعياري في تقدير معلمة صعوبة الفقرة بزيادة عدد أفراد العينة حيث بلغ متوسط الأخطاء المعيارية في التقدير (0.32)لوجيت عندما كان حجم العينة (٢٠٠) تناقص بحيث (٠,٠٧) لوجيت عندما أصبح حجم العينة (١١٢٩٢).

دراسة العنزي (٢٠١٢)

هدفت إلى استخدام أحد نماذج نظرية الاستجابة للمفردة (النموذج ثلاث المعلم) في تطوير اختبار كورنيل للتفكير الناقد المستوى (X)، وكان مجتمع البحث طلاب المرحلة الثانوية في مدينة عرعر، وتألفت عينة التحليل من

(١٣٠٩) طالباً، وتمثلت أداة البحث في اختبار كورنيل للتفكير الناقد المستوى (X)، من أهم نتائج البحث : تحقق افتراضات نظرية استجابة المفردة في بيانات اختبار كورنيل للتفكير الناقد، تم التعرف على قيم تقديرات معالم مفردات الاختبار (الصعوبة، والتمييز، والتخمين) وفق النموذج ثلاثي المعلم لنظرية استجابة المفردة. وكان هناك تباين في قيم معاملات التمييز، والصعوبة، والتخمين لكل بعد من أبعاد اختبار التفكير الناقد، وتم التدريج النهائي لمفردات اختبار كورنيل للتفكير الناقد وفقاً للنموذج ثلاثي المعلم، واتضح أن بعد المصداقية هو أكثر الأبعاد دقة في القياس، يليه بعد الاستنباط، ثم بعد الاستقراء، ويأتي في النهاية بعد الافتراضات. اتضح أن أعلى مقدار للدالة المعلوماتية لجميع أبعاد الاختبار تقع في الفترة ما بين -٢، +٢ على مقياس القدرة. تم تقدير قدرات الطلاب المقابلة لكل درجة كلية محتملة على اختبار كورنيل للتفكير الناقد وفق النموذج الثلاثي المعلم وكذلك تم تقدير الدرجة الحقيقية لكل درجة كلية خام محتملة على نفس الاختبار وفق النموذج ذاته، وتم التوصل إلى قيم معاملات ثبات مقبولة لأبعاد الاختبار : (الاستقراء، والمصداقية، والاستنباط، والافتراضات) وهي (٠,٨٥، ٠,٨٦، ٠,٧٧، ٠,٧٥) على الترتيب، وتم التحقق من صدق تعريف مفردات الاختبار اعتماداً على التحليل العاملي وصدق التدريج، تم حساب الدرجات المئينية المقابلة لكل درجة خام وكل تقدير للقدرة.

دراسة الخياط (٢٠١٢)

هدفت الدراسة إلى التحقق من فاعلية النموذج اللوغاريتمي ذي المعلمة الواحدة (نموذج راش) ومطابقة البيانات للنموذج لاختبار (TIMSS) والذي يقيس المهارات الرياضية لمستوى الصف الثامن الأساسي، ولتحقيق هدف الدراسة تم تطبيق فقرات الاختبار على عينة من طلبة الصف الثامن الأساسي

بلغ عددها (٥٩٩) طالباً وطالبة، بعد إيجاد الخصائص السيكومترية (الصدق والثبات) لاختبار TIMSS وللإجابة على أسئلة الدراسة تم استخدام برنامج إحصائي (Rumm,2010)؛ وأشارت النتائج إلى تمتع الاختبار المقدم بالصدق والثبات؛ حيث بلغت قيمة معامل الثبات للاختبار باستخدام معامل كرونباخ ألفا (٠,٩) أما معامل الصدق بدلالة المحك بين درجات الطلبة على الاختبار ودرجاتهم في مادة الرياضيات فقد بلغت قيمته (٠,٨٧) وأشارت النتائج إلى مناسبة (٢٥) فقرة من فقرات الاختبار والبالغ عدها (٣٥) فقرة للبيئة المحلية بحيث تمتعت بخصائص إحصائية مناسبة، من حيث معامل صعوبة الفقرة.

دراسة الجبوري (٢٠١٢)

هدفت إلى بناء اختبار تحصيلي لمادة الإدارة والإشراف التربوي باستخدام نظرية السمات الكامنة لطلبة معاهد إعداد المعلمين، تم صياغة (١٨٩) هدفاً سلوكياً للمستويات الثلاثة الأولى من تصنيف بلوم : (المعرفة، الفهم، التطبيق) ومن ثم صياغة مفردة اختبارية واحدة لكل هدف سلوكي من نوع الاختيار من متعدد ذات البدائل الأربعة مع مراعاة الدقة والشمولية لأهداف وفقرات الاختبار.

طبق الاختبار التحصيلي على عينة مكونة من (٣١٠) طالب وطالبة اختيروا بالأسلوب الطبقي العشوائي من (٩) معاهد في محافظة بغداد.

اعتمد الباحث على نموذج راش أحد نماذج نظرية السمات الكامنة في تحليل فقرات الاختبار التحصيلي، وللتحقق من افتراضات النموذج اتبع الباحث ما يلي :

١ -أحادية البعد : التحليل العاملي بطريقة المكونات الأساسية، وذلك للتثبت من البنية العاملية للاختبار، إذ تم الحصول على عامل واحد ذي معنى للاختبار، واعتمد العامل نفسه على الحدود الدنيا ل (جتمان) الذي يعد

العامل دالاً إحصائياً حينما يكون الجذر الكامن الذي يمكن تفسيره يساوي أو يزيد عن (١)، واعتمد الباحث على معادلة الخطأ المعياري (لبرت بانكس) محكاً لتشبع المفردات بالعامل العام والذي بلغ مقداره (٠,١٤٨)، وقد حذفت (١٠) مفردات اختبارية، لأن تشبع المفردات بالعامل العام كان أقل من المعيار المعتمد في تشبع المفردات.

٢ -مطابقة المفردات لنموذج راش باستخدام برنامج راسكال المحوسب استناداً إلى قيمة مربع كاي لحسن المطابقة بمستوى دلالة (٠,٠٥) وكما حسبها البرنامج، إذ استبعد البرنامج (١٤) مفردة من الاختبار لأن قيمتها كانت أكبر من قيمة مربع كاي عند مستوى (٠,٠٥).

٣ -اقتراب قيمة معامل التمييز للاختبار من (١).

٤ -استقلالية القياس بما يحقق موضوعية القياس كما تمثل بنموذج راش، إذ استبعدت (١٨) مفردة اختبارية لعدم تحقيقها استقلالية القياس.

٥ -تكون الاختبار بصيغته النهائية من (١١١) مفردة اختبارية التي أخضعت إلى التحليل الإحصائي لتحويل وحدة اللوجيت إلى وحدة الواط من أجل التخلص من الكسور والإشارات السالبة لتقديرات صعوبة المفردات وقدرة الأفراد، إذ بلغ معامل ثبات الاختبار بالصيغة النهائية (٠,٩٥٤).

دراسة تانج لي Trang Le(2013)

وهدفت إلى استخدام نظرية الاستجابة للمفردة الاختبارية في تحليل نتائج ثلاث أبحاث علمية ؛ حيث تستخدم النظرية الكلاسيكية في عملية تحليل بيانات الاختبار واستجابات الطلاب : المتوسطات الحسابية والدرجات الكلية والتحليل العاملي كمتغيرات تبني عليها التنبؤ بالأداء ؛ لكن هذه الطريقة لا تأخذ في الحسبان الخصائص المميزة لمفردات الاختبار أو الاستبيان، وهي بذلك تختلف عن نظرية الاستجابة للمفردة التي تقوم بنمذجة العلاقات بين

مستوى سمة المفحوص (القدرة، الاتجاه، ...) ونموذج استجابات المفردة. لذلك فإن تقدير سمات الفرد الكامنة من الممكن أن يختلف حتى بين شخصين حصلا على نفس الدرجة الكلية ؛ لذلك فإن نظرية الاستجابة للمفردة تقدم هذا التفسير وبها مميزات أخرى تناولها الإطار النظري للبحث، وبحثت كذلك نماذج نظرية الاستجابة للمفردة : الأحادية والمتعددة وتم تطبيق النظرية على تحليل نتائج ثلاث بحوث.

دراسة حمادنة وبني خالد (٢٠١٣)

هدفت هذه الدراسة 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 مفردة موزعة على الأبعاد بشكل غير متساوٍ. وقد تم تطبيق الأداة على عينة مؤلفة من ٥٥٠ طالباً وطالبة من طلبة جامعة آل البيت ممن يستخدمون مواقع التواصل الاجتماعي .

وعند تحليل استجابات أفراد العينة - وفق نظرية استجابة المفردة -(IRT) لم يتبق من مفردات المقياس سوى ٢٩ مفردة، موزعة في ثلاثة مجالات رئيسة. وتمتع المقياس بصورته النهائية ٢٩ مفردة بخصائص سيكومترية مناسبة، حيث بلغت قيمة معامل الثبات للمقياس ٠,٩٤ ، كما تمتع المقياس بدلالات متعددة للصدق.

دراسة المصري (٢٠١٥)

هدفت 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 (٤٨) مفردة

من نوع الاختيار من أربعة بدائل، وبلغت عينة الدراسة (٢٢٠) طالباً وطالبة من طلبة كلية التربية المحويت باليمن، وأظهرت النتائج تحقق افتراضات النموذج أحادي المعلم ومطابقة (٤٢) مفردة من مفردات الاختبار لتوقعات النموذج، وتم حذف (٤) 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 (٧) أفراد غير مطابقين للنموذج وتراوحت معالم القدرة لأفراد عينة الدراسة بين (- ٢,٦٥٩) لوجيت و (٢,٨١٢) لوجيت.

دراسة دبوس (٢٠١٦)

هدفت الدراسة 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) طالباً 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 أوصى الباحث بمجموعة من التوصيات.

دراسة نور الدين (٢٠١٧)

وهدفت إلى التعرف على مدى فعالية الاختبار التواؤمي المحوسب في تقويم طلاب جامعة الجوف على أحد الاختبارات التحصيلية. تم تطبيق اختباراً تحصيلياً ورقياً تألف من ٤٨ سؤالاً من نمط الاختيار من متعدد على (١١٩) طالباً جامعياً، وبناء على ذلك تم إنشاء اختبار تحصيلي تواؤمي محوسب وفق نموذج راش ذي خصائص سيكومترية جيدة مشتق من الاختبار، وتطبيقه على (٢٥) طالباً جامعياً. وبمقارنة نتائج أدائهم على الاختبار التواؤمي المحوسب والاختبار الورقي التقليدي ؛ أسفرت النتائج عن فعالية الاختبار التواؤمي المحوسب على الاختبار الورقي التقليدي من حيث مقدار أعلى من المعلومات التي يوفرها، وأقل خطأ معياري، مع تقدير للقدرة لا يختلف عن الورقي التقليدي، واختصر أكثر من (٥٠٪) من الأسئلة. وقد اختتمت الدراسة بمجموعة من التوصيات ذات الصلة.

تعقيب على الدراسات السابقة :

اتفقت غالبية الدراسات ذات الصلة بموضوع البحث على أهمية تدريج المفردات الاختبارية باستخدام نماذج الاستجابة للمفردة : الأحادية والثنائية والثلاثية : وكذلك اتفقت الدراسات العربية على استخدام هذه النماذج يسهم في اختفاء جزء كبير من أخطاء القياس ويسهم في تقدير حقيقي للقدرات التحصيلية المقاسة لكن يعاب على بعض الدراسات أنها لم تتحرى الدقة في التفريق بين قياس الاتجاهات وقياس القدرات وقياس الاستعدادات وتعاملت معها كلها على أنها منهجية واحدة، كما أن بعض الدراسات لم توضح مبررات قوية لاختيار نموذج التحليل الإحصائي المستخدم، كذلك أغلب الدراسات المقارنة لم تتجاوز نقطة إثبات التفوق النظري والإمبريقي لنماذج الاستجابة للمفردة على النظرية الكلاسيكية.

فنجد أن بعض الدراسات هدفت إلى تحقيق فعالية نظرية السمات الكامنة في تدريج الاختبارات والتعبير عن القدرات الحقيقية للمفحوصين وتحقيق أهداف عملية القياس بدقة مثل دراسات كل من : المصري (٢٠١٥)، ودبوس (٢٠١٦)، و نورالدين (٢٠١٧)

وبعض الدراسات الأخرى استخدمت نماذج نظرية الاستجابة للمفردة في إعادة تدريج مفردات اختبارات تقيس الذكاء والقدرات العقلية والتفكير وسبق تقنينها من قبل على نظرية القياس الكلاسيكية مثل دراسة أبوجراد (٢٠٠٨) على اختبار كاتل للذكاء ودراسة العنزي (٢٠١٢) على اختبار كورنيل للتفكير الناقد، ودراسات كل من Progar and Socan(2008)، والخياط (٢٠١٢) على اختبار TIMSS. ودراسة عيد (٢٠٠٤) على اختبار اليقظة العقلية.

وبعض الدراسات استخدمت نظرية الاستجابة للمفردة في تقدير درجات الأفراد على مقاييس الاتجاهات وتفضيلات التعلم وسمات الشخصية مثل دراسات كل من : أبوهاشم (٢٠٠٦)، وحمادنة وبني خالد (٢٠١٣). كذلك بعض الدراسات كانت على مجتمع طلاب الجامعات السعودية مثل دراسة العنزي (٢٠١٢)، ونورالدين (٢٠١٧)وبعض الدراسات هدفت إلى تصميم اختبار محوسب وليس ورقي وفق نماذج الاستجابة للمفردة نموذج راش منها دراسة نور الدين (٢٠١٧) وتم تطبيقه وتحليل النتائج باستخدام الحاسب الآلي.

وبعض الدراسات لم تشر إلى سبب اختيار برامج الكمبيوتر المستخدمة في تحليل نموذج نظرية الاستجابة للمفردة المستخدم، وأغلب الدراسات لم تشرح المنهجية المستخدمة في إدخال البيانات على البرامج وكيفية تحليلها وتدريج المفردات واستخلاص مؤشرات المفردات ومنحنياتها بالشكل السليم

حيث لم يجد الباحث سوى دراسة واحدة فقط قدمت شرح للبرنامج المستخدم وهي دراسة محمود (٢٠١٠).ودراسة أخرى قارنت بين برنامجين من برامج الكمبيوتر في تدريج المفردات هي دراسة Carlson and Locklin(1995) وبعض الدراسات استخدمت نموذج راش أحادي البعد فقط دون بقية نماذج نظرية الاستجابة للمفردة مثل دراسة الجبوري (٢٠١٢) على طلاب الجامعات العراقية، و دراسة الخياط (٢٠١٢)، ودراسة أبوجراد (٢٠٠٨). ودراسة عيد (٢٠٠٤).

كذلك بعض الدراسات قارنت بين نماذج نظرية الاستجابة للمفردة : الأحادية والمتعددة مثل دراسات : طيفور (٢٠٠٧)، و Progar and Socan(2008)، و(Trang Le(2013 ودبوس (٢٠١٦)، ويستخدم الباحث نموذج راش أحادي البعد بعد التحقق من شروطه وكذلك فإن مفردات الاختبار من النوع ثنائية الإجابة.

الدراسات المقارنة بين النظرية الكلاسيكية في القياس ونظرية الاستجابة للمفردة :

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، وفي هذا الإطار أجرت الحكماني (٢٠١٠) دراسة هدفت إلى مقارنة النظرية الكلاسيكية للاختبار ونظرية الاستجابة للمفردة متمثلة في النموذج اللوغاريتمي ثنائي البارامتر، من حيث

تقدير مستويات قدرة الطلاب. كما هدفت إلى مقارنة النظريتين من حيث مدى استقرار المؤشرات الإحصائية للمفردات الاختبارية والمتمثلة في صعوبة المفردات وتمييزها، وذلك باختلاف الصفوف الدراسية.

وقد تكونت العينة من (٣٠٨٢) طالب وطالبة من طلاب الصفوف : السابع والثامن والتاسع، بمدارس محافظة مسقط بسلطنة عمان وقد تم جمع بيانات الدراسة باستخدام اختبار المصفوفات المتتابعة العادي لرافن.

وقد تم حساب معامل ارتباط بيرسون ومتوسط مربعات الفروق للكشف عن العلاقة بين درجات ذكاء الطلاب، والتي تم تقديرها باستخدام كل من النظرية الكلاسيكية والنموذج اللوغاريتي ثنائي البارامتر. كما تم استخدام التحليلات الإحصائية السابقة نفسها، للتحقق من مدى استقرار المؤشرات الإحصائية باختلاف الصفوف الدراسية، وذلك في كل إطار نظري.

وقد أظهرت النتائج بأن هناك تشابهاً بين درجات الطلاب المقدرة من النظريتين في الترتيب، إلا أن قيم درجات الطلاب وفق النظرية الكلاسيكية تختلف بشكل كبير نسبياً عن تقديرات درجاتهم المستمدة من النموذج اللوغاريتمي ثنائي البارامتر إذ أنها أكثر استقراراً من تلك المقدرة باستخدام النظرية الكلاسيكية. إضافة إلى ذلك فقد أظهرت النتائج بأن مؤشرات الصعوبة أكثر استقراراً من مؤشرات التمييز، وذلك في كلا الإطارين النظريين.

وفي هذا الإطار قام فراج والشريف (٢٠١٣) 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.

ويتضح من العرض السابق وجود دراسات عربية للمقارنة بين النظريتين ولكن يجب أن تهتم دراسات المقارنة بخلق ثقافة تقويم عربية تشجع على استخدام نظرية السمات الكامنة في الاختبارات العربية على مدى واسع ويجب أن يتم ربط هذه الجهود بشكل مؤسسي من خلال تصميم بنوك أسئلة مدرجة حسب نماذج نظرية السمات الكامنة في التخصصات الأكاديمية المختلفة بالجامعات وربط ذلك بشروط الحصول على الاعتماد البرامجي في التخصصات المختلفة حتى يهتم أعضاء هيئة التدريس بالجامعات بتأهيل أنفسهم في تصميم وتطبيق وتحليل نتائج هذه الاختبارات، فللأسف مازالت الدراسات العربية المقارنة تركز على مقارنة بارامترات (المفردة /الفرد) بين النظريتين واثبات تفوق نظرية السمات الكامنة على الكلاسيكية في تقدير القدرات الحقيقية المقاسة للمتعلمين ومع ذلك لا تستخدم على مدى واسع في تصميم الاختبارات في الجامعات العربية وهو ما أكدته دراسات مثل نورالدين (٢٠١٧)، والجبوري (٢٠١٢)، وأبوجراد (٢٠٠٨)

فروض البحث:

بعد مراجعة الإطار النظري والدراسات السابقة التى تمكن الباحث من الحصول عليها حول نموذج راش الاحتمالى، وكذلك طبيعة اختبار مقرر مهارات التعلم والتفكير الذى تم اختياره للتطبيق وتحليل بياناته : وفق نظرية القياس الكلاسيكية ونموذج راش، لذا فقد تم صياغة فروض البحث الحالي على النحو التالى :

(١) تتوفر درجة مرتفعة من الصدق والثبات لاختبار مقرر مهارات التعلم والتفكير باستخدام النظرية الكلاسيكية في القياس.

(٢) تنطبق البيانات المستمدة من اختبار مقرر مهارات التعلم والتفكير على نموذج راش.

(٣) مفردات الاختبار التحصيلي لمقرر مهارات التعلم والتفكير تتمتع بمعاملات صعوبة ملائمة وفق نموذج راش.

(٤) لا يتغير تدريج مفردات اختبار مقرر مهارات التعلم والتفكير المطورة باستخدام نموذج راش بتغير مستوى قدرات أفراد العينة المستخدمة فى الحصول على هذا التدريج.

منهج البحث:

ينطلق البحث الحالي من المنهج الوصفي التحليلي المقارن : - محاولاً تفسير درجات طلاب كلية التربية جامعة الملك فيصل في اختبار مقرر مهارات التعلم والتفكير وفق محكات موضوعية في ضوء الأطر النظرية المتضمنة في المقارنة وهي نظرية القياس الكلاسيكية ونظرية الاستجابة للمفردة.

أداة البحث:

الاختبار التحصيلي في مقرر مهارات التعلم والتفكير.

اتبع الباحث الخطوات العلمية في بناء الاختبار التحصيلي وبدأت بتحديد مفردات مادة مهارات التعلم والتفكير بالاستناد إلى المفردات التي وضعتها لجنة ضمان الجودة بكلية التربية جامعة الملك فيصل في توصيف المقرر، إذ بلغ عدد فصول المادة (٥) فصول، وبعد توزيع استبانة مفردات مادة مهارات التعلم والتفكير لمدرسي المادة حول المفردات التي يدرسونها بنحو فعلي، تم الاتفاق بنسبة ١٠٠٪ على مفردات المادة، وقد تم صياغة (٣٠) هدفاً سلوكياً للمستويات الثلاثة الأولى من تصنيف بلوم: (المعرفة، الفهم، التطبيق) ومن ثم صياغة مفردة اختبارية واحدة لكل هدف سلوكي من نوع الاختيار من متعدد ذات البدائل الأربعة مع مراعاة الدقة والشمولية لأهداف ومفردات الاختبار.

وللتحقق من صلاحية مفردات الاختبار منطقياً عرضت على (٥) خبراء متخصصين في العلوم التربوية والنفسية، لتثبت مدى صلاحية مفردات الاختبار التحصيلي ومدى ملاءمتها للأهداف السلوكية وبدائل الإجابة، وأخذ الباحث بآرائهم في تعديل بعض المفردات، والاختبار يتكون من (٣٠) مفردة اختبارية بعد تعديل الخبراء.

جدول (١) مواصفات اختبار مقرر مهارات التعلم والتفكير

البعد	الهدف ومستواه	أرقام المفردات	عددها	الوزن النسبي
مفهوم التفكير	أن يذكر الطالب تعريف للتفكير (تذكر)	١	٥	٣,٣٣٪
	أن يميز الطالب تعريف التفكير عن تعريف مهارات التفكير (فهم)	٤		٣,٣٣٪
	أن يتعرف الطالب على أهمية التفكير في حياته اليومية (تطبيق)	٥		٣,٣٣٪
	أن يميز الطالب بين التفكير المعتاد والتفكير المهاري (فهم)	٢		٣,٣٣٪
	أن يميز الطالب بين تعليم التفكير وتعليم مهارات التفكير (فهم)	٣		٣,٣٣٪
تصنيفات التفكير ومهاراته	أن يميز الطالب بين مفهوم العنقود والمهارة والمهارة الفرعية في تصنيف التفكير (فهم)	١٠	٥	٣,٣٣٪
	أن يميز الطالب بين التصنيفات الثنائية والثلاثية والرباعية لمهارات التفكير (فهم)	٢١		٣,٣٣٪
	أن يصنف الطالب	٢٢		٣,٣٣٪

	التفكير الإبداعي ضمن التفكير التباعدي (تحليل)			
	أن يصنف الطالب التفكير (التأملي – العلمي –الناقد) ضمن التفكير التقاربي (تحليل)	٢٣		٪٣,٣٣
	أن يصنف الطالب مهارات التفكير على أساس معالجة المعلومات (تطبيق)	٢٤		٪٣,٣٣
تعليم التفكير	أن يتعرف الطالب على أهمية تعليم التفكير (فهم)	٢٥	٥	٪٣,٣٣
	أن يتعرف الطالب الأسلوب المباشر في تعليم التفكير (تطبيق)	٦		٪٣,٣٣
	أن يتعرف الطالب على أسلوب الدمج في تعليم التفكير (تطبيق)	٧		٪٣,٣٣
	أن يميز الطالب بين الأسلوبين في تعليم التفكير (تحليل)	٨		٪٣,٣٣
	أن يتعرف الطالب على	٩		٪٣,٣٣

	مثال على برامج تنمية التفكير العالمية (تطبيق)			
التفكير التأملي	أن يذكر الطالب مفهوم التفكير التأملي (تذكر)	١٩	٥	٪٣,٣٣
	أن يصنف الطالب التفكير التأملي ضمن مهارات التفكير العليا (تحليل)	١٢		٪٣,٣٣
	أن يتعرف الطالب على مستويات التفكير التأملي (فهم)	١٤		٪٣,٣٣
	أن يذكر الطالب مفهوم أدوات التفكير (تذكر)	١٦		٪٣,٣٣
	أن يوضح الطالب علاقة التفكير بحل المشكلات (فهم)	١٣		٪٣,٣٣
التفكير العلمي	أن يذكر الطالب مفهوم التفكير العلمي (تذكر)	١٥	٥	٪٣,٣٣
	أن يحدد الطالب العلاقة بين التفكير التأملي والتفكير العلمي (تحليل)	١٧		٪٣,٣٣
	أن يميز الطالب عناصر الموقف المشكل (فهم)	٢٠		٪٣,٣٣
	أن يميز الطالب بين المشكلات الواضحة	٢٦		٪٣,٣٣

	والغير واضحة البنية (تطبيق)			
	أن يتعرف الطالب خصائص التفكير العلمي (فهم)	٢٧		٪٣,٣٣
مهارات التعلم	أن يذكر الطالب مفهوم مهارة التعلم الخبري (تذكر)	٢٨	٥	٪٣,٣٣
	أن يذكر الطالب مفهوم صحيفة التعلم (تذكر)	٢٩		٪٣,٣٣
	أن يعرف الطالب مفهوم مهارة حل المشكلات (تذكر)	٣٠		٪٣,٣٣
	أن يحدد الطالب العلاقة بين التفكير والتعلم (تحليل)	١١		٪٣,٣٣
	أن يصنف الطالب مهارة التلخيص كتطبيق على مهارات التفكير والتعلم (تطبيق)	١٨		٪٣,٣٣
مجموع			٣٠	٪١٠٠

قام الباحث بتوليد مصفوفة رقمية مكونة من ١٦٢ سطراً تمثل عدد الأفراد و ٣٠ عموداً تمثل عدد مفردات هذا الاختبار، ثم وباستخدام البرنامج الإحصائي SPSS تم تخزين ملف البيانات المكون من المصفوفة السابقة بعد إضافة ٥ أعمدة، منها أربعة خاصة بتعريف كل حالة وعمود

واحد يفصل بين أعمدة المفردات وأعمدة تعريف الحالة في حين أن عناصر هذه المصفوفة الخاصة بالمفردات تأخذ إحدى القيمتين (٠ ، ١) على افتراض أن المفردات ثنائية الاستجابة ؛ حيث يشير علام (٢٠٠٧) أن نموذج راش يستخدم في تحليل مفردات الاختبارات التي تكون الإجابة عليها إما صحيحة أو خطأ، بمعنى أن درجة كل مفردة إما واحد صحيح أو صفر. ويسمى نموذج البارامتر المحرر Sample-Free Model (علام، ٢٠٠٧، ٢٣٢)

عينة البحث:

طبق الاختبار التحصيلي على عينة مكونة من (١٦٢) طالب اختيروا بطريقة عشوائية من ٣ شعب تدرس مقرر مهارات التعلم والتفكير في كلية التربية جامعة الملك فيصل ؛ الفصل الدراسي الأول ١٤٣٨/١٤٣٩ هجري بمتوسط عمر زمني ١٩,٧ سنة وانحراف معياري ٠,٩.

الأساليب الإحصائية المستخدمة في البحث

بناءً على فروض البحث الحالي، استخدم الباحث البرامج الإحصائية التالية :

أ - البرنامج الإحصائى ٢١ ***SPSS***، وذلك لحساب :

- معاملات ارتباط المفردات بالمكونات الفرعية والرئيسية
- معاملات ارتباط المكونات الفرعية بالرئيسية والدرجة الكلية للاختبار
- معامل ألفا & للتأكد من ثبات الاختبار
- التحليل العاملى للمكونات الرئيسية لاختبار مقرر مهارات التعلم والتفكير

ب - برنامج "جانز -راش" ***(Ganz-Rasch)*** في تحليل مفردات اختبار مقرر مهارات التعلم والتفكير باستخدام نموذج راش ويرجع اختيار هذا البرنامج إلى :

- سهولة التدريب عليه فى فترة زمنية قصيرة واستخدامه فى تحليل فقرات الاختبارات وبنوك الأسئلة بدقة

- سهولة التعامل من خلاله مع البرامج الإحصائية الأخرى ونقل البيانات إليه بسهولة من برنامج SPSS

- يعطى رسماً بيانياً يوضح اختيار أفراد عينة الدراسة للاستجابات على المفردة الواحدة.

ج -برنامج وينستبس Winsteps تم استخدامه في أحد مراحل التحليل الإحصائي نظراً لأنه لا يتيح إمكانية تحليل الاختبارات التي تزيد مفرداتها عن ٢٥ مفردة.

عرض وتفسير نتائج البحث ومناقشتها

النتائج الخاصة بالفرض الأول :

ينص الفرض الأول على أنه

تتوفر درجة مرتفعة من الصدق والثبات لاختبار مقرر مهارات التعلم والتفكير باستخدام النظرية الكلاسيكية في القياس.

(١)صدق المحتوى :

عرضت الصورة الأولية لاختبار مقرر مهارات التعلم والتفكير (٣٠) مفردة على (٥)[١] محكمين من الأساتذة مدرسي مقرر مهارات التعلم والتفكير في كلية التربية جامعة الملك فيصل، وطلب منهم تحديد انتماء أو عدم انتماء كل مفردة إلى البعد الفرعي الذى تنتمى إليه ومدى وضوحها، وإضافة ما يرونه

(١) يشكر الباحث كلاً من د. حسن بدري محمد ، ود. آدم بشير ، ود. شهاب بن شواشة ، ود. أنس الطيب ، ود. عاشور دياب أعضاء هيئة التدريس بكلية التربية جامعة الملك فيصل مدرسي مقرر مهارات التعلم والتفكير للمشاركة في تحكيم مفردات الاختبار.

من تعديلات مناسبة، وقد استفاد الباحث من آراء السادة المحكمين فى إعادة صياغة بعض المفردات لتتناسب صياغتها مع الأسلوب المتبع فى تصميم الاختبارات ذات المفردات ثنائية الدرجة، وقد أسفر صدق المحكمين على أن كل مفردة تنتمى إلى البعد الذى تقيسه، وبعد التحقق من صدق المحكمين تم تطبيق الاختبار على عينة البحث، وذلك تمهيداً للتحقق من فروض البحث الحالي.

(٢) ثبات الاختبار :

تم حساب معامل ثبات ألفا الكلي للاختبار (٠,٧٢١) وفي حالة حذف كل مفردة، وذلك لعينة مكونة من ٦٥ طالب من كلية التربية جامعة الملك فيصل وجاءت النتائج على النحو التالى :

جدول (٢) قيم معامل ارتباط كل مفردة بالبعد وثبات ألفا في حالة حذف درجة المفردة لاختبار مقرر مهارات التعلم والتفكير

	البعد	أرقام المفردات	معامل الارتباط مع درجة البعد	ثبات ألفا
	مفهوم التفكير	١	0.28*	0.641
		٤	0.846**	0.72
		٥	0.333**	0.657
		٢	0.52**	0.641
		٣	0.494**	0.66
	تصنيفات التفكير ومهاراته	١٠	0.46**	0.634
		٢١	0.27*	0.662
		٢٢	0.814**	0.624
		٢٣	0.753**	0.632
		٢٤	0.587**	0.643
	تعليم التفكير	٢٥	0.413**	0.649

	البعد	أرقام المفردات	معامل الارتباط مع درجة البعد	ثبات ألفا
		٦	0.788**	0.64
		٧	0.731**	0.644
		٨	0.538**	0.627
		٩	0.696**	0.638
	التفكير التأملي	١٩	0.446**	0.656
		١٢	0.568**	0.653
		١٤	0.626**	0.634
		١٦	0.54**	0.668
		١٣	0.321**	0.65
	التفكير العلمي	١٥	0.482**	0.659
		١٧	0.468**	0.678
		٢٠	0.62**	0.651
		٢٦	0.21*	0.658
		٢٧	0.428**	0.661
	مهارات التعلم	٢٨	0.536**	0.651
		٢٩	0.531**	0.658
		٣٠	0.437**	0.661
		١١	0.611**	0.642
		١٨	0.487**	0.646

❖ دال إحصائياً عند ٠,٠٥ ❖❖ دال إحصائياً عند ٠,٠١

ويتضح أن قيم معامل ثبات ألفا الكلي تنخفض كلما تم حذف كل مفردة من التحليل، وكذلك فإن معامل ارتباط كل مفردة بدرجة البعد أغلبها مرتفعة وكلها دالة إحصائياً مما يدل على أن الاختبار متسق داخلياً.

النتائج الخاصة بالفرض الثاني

ينص الفرض الثاني على :

تنطبق البيانات المستمدة من اختبار مقرر مهارات التعلم والتفكير على نموذج راش.

لا بدّ من التحقق من افتراض أحادية البعد (Unidimensionality). حيث تفترض نماذج نظرية استجابة الفقرة (IRT)، وجود قدرة واحدة تفسر أداء المفحوص في المقياس، ولذلك تسمى بالنماذج أحادية البعد. وللتحقق من هذا الافتراض تم إجراء تحليل عاملي لبيانات المقياس المتعلقة باستجابات (162) طالب ممن يدرسون مقرر مهارات التعلم والتفكير في كلية التربية بجامعة الملك فيصل وتكون الاختبار من (٣٠) مفردة، باستخدام البرمجية الإحصائية SPSS. الإصدار ٢١

نتيجة اختبار KMO لملائمة البيانات للتحليل العاملي كانت ٠,٦٧٨ وهي قيمة مقبولة أكبر من ٠,٦ ، وكذلك نتيجة اختبار بارتلت كانت دالة إحصائياً مما يؤدي إلى رفض الفرضية الصفرية التي تقول أنه لايوجد ارتباطات بين المتغيرات.

وقد تم استخراج نتائج إجراء التحليل العاملي بطريقة تحليل المكونات الرئيسة (Principle Components Analysis, PCA)، والتدوير تبعاً لمحاور متعامدة (Varimax). حيث أفرز التحليل (١١) عاملا، فسر العامل الأول منها ١٥,٤٥ ٪ من التباين، وفسرت جميع العوامل ٦٣,٢٥ ٪ من التباين

الكلي. ويبين الجدول (٣) قيم الجذور الكامنة (Eigenvalues)، ونسبة التباين المفسر لكل عامل من العوامل المستخلصة بجذور كامنة مقبولة (أكبر من الواحد الصحيح)، وكذلك نسبة التباين المفسر التراكمية.

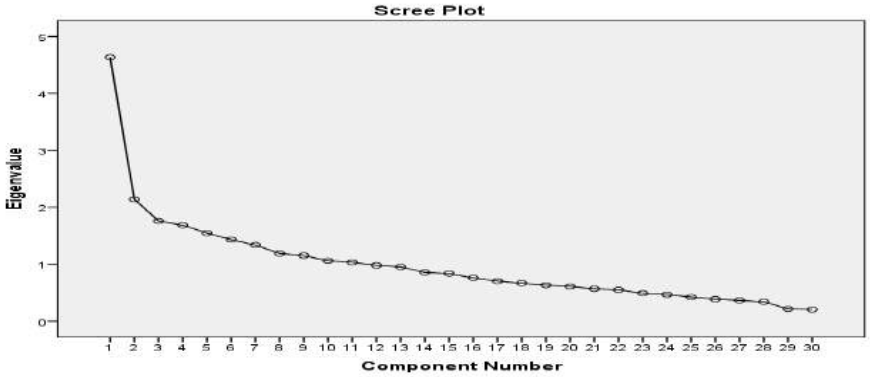
جدول (٣)

خلاصة نتائج التحليل العاملي لاستجابة (١٦٢) فرداً على اختبار مقرر مهارات التعلم والتفكير (٣٠) مفردة

رقم العامل	الجذر الكامن	نسبة التباين المفسر٪	نسبة التباين المفسر التراكمي ٪
١	4.635	15.451	15.451
٢	2.141	7.137	22.588
٣	1.760	5.865	28.454
٤	1.685	5.616	34.070
٥	1.548	5.159	39.228
٦	1.433	4.775	44.004
٧	1.338	4.461	48.465
٨	1.188	3.961	52.426
٩	1.152	3.839	56.265
١٠	1.062	3.540	59.805
١١	1.034	3.446	63.251

يتبين من الجدول (٣)، أن قيمة الجذر الكامن للعامل الأول بلغت ٤,٦٣٥ ويفسر ما نسبته ١٥,٤٥ ٪ من التباين الكلي، وهي قيمة مرتفعة إذا ما قورنت مع قيم الجذور الكامنة لبقية العوامل، أما قيمة الجذر الكامن للعامل الثاني بلغت ٢,١٤١ ويفسر ما نسبته ٧,١٣ ٪ من التباين الكلي، بمعنى أن العامل الأول فسر ما يزيد عن ضعفي ما فسره العامل الثاني. أضف إلى ذلك، يلاحظ بأن نسبة التباين المفسر لكل من العوامل المتبقية متقاربة، بمعنى أنه يوجد شبه استقرار في نسب التباين المفسر لجميع العوامل باستثناء العامل الأول، وهذا مؤشر على تحقق افتراض أحادية البعد للمقياس (Hulin, Drasgow & Parson, 1983; Hattie, 1985)، أي أن الاختبار يقيس سمة واحدة.

ويتعزز افتراض أحادية البعد، للاختبار التحصيلي المعدّ بالبحث الحالي، من خلال تمثيل الجذور الكامنة بيانياً باستخدام ما يعرف باختبار فحص العوامل (Scree plot) الذي يظهر في الشكل (١).



الشكل (١)

التمثيل البياني لقيم الجذور الكامنة للعوامل المكونة للاختبار التحصيلي على البيانات الكلية.

يتبين من الشكل (١)، أن الجذر الكامن للعامل الأول يطغى بشكل واضح على الجذور الكامنة لبقية العوامل، وهذا مؤشر أيضاً على أحادية البعد لبيانات الاختبار.

افتراض الاستقلال المحلي : - ويمكن التحقق منه من خلال فحص معامل الثبات لجميع الدرجات الكلية المختلفة (مستويات القدرة المختلفة)، والذي يعد مؤشراً عن مدى الاتساق الداخلي لمفردات الاختبار، وذلك عند تثبيت الدرجة الكلية المكتسبة من قبل الطلاب على الاختبار، وإن هذا التثبيت للدرجة الكلية سيكون بمنزلة ضبط تأثير السمة الكامنة التي يقيسها الاختبار أوما تسمى بالقدرة، والجدول (٤) يوضح قيم معاملات الثبات بطريقة تحليل التباين للاختبار عند المستويات المختلفة لقدرة الطلاب.

جدول (٤) قيم معاملات الثبات بطريقة تحليل التباين للاختبار عند المستويات المختلفة لقدرة الطلاب

الدرجة الكلية	عدد المفردات	عدد الحالات	قيم الثبات F	الدرجة الكلية	عدد المفردات	عدد الحالات	قيم الثبات F
٢٩	٣٠	٤	0.00	١٦	٣٠	٦	0.00
٢٨	٣٠	١٥	0.00	١٥	٣٠	٥	0.00
٢٧	٣٠	٨	0.00	١٤	٣٠	٦	0.00
٢٦	٣٠	٩	0.00	١٣	٣٠	٢	0.00
٢٥	٣٠	١٠	0.00	١٢	٣٠	٣	0.00
٢٤	٣٠	١١	0.00	١١	٣٠	١	0.00
٢٣	٣٠	١١	0.00	٩	٣٠	١	0.00
٢٢	٣٠	١٣	0.00	٨	٣٠	١	0.00
٢١	٣٠	١٦	0.00				
٢٠	٣٠	١٠	0.00				
١٩	٣٠	١١	0.00				
١٨	٣٠	٨	0.00				
١٧	٣٠	١١	0.00				

يتضح من الجدول السابق أن قيم معامل الثبات بطريقة تحليل التباين لجميع الدرجات تساوي صفر، وعند مقارنة هذه القيم بمعامل ثبات كودر ريتشاردسون لجميع الحالات (ن = ١٦٢)، والذي يساوي (0.742) نستنتج أن المسئول عن ارتفاع معامل ثبات الاختبار هو قدرة الطلاب (الدرجة الكلية لكل طالب)، بحيث إذا أزيل تأثير هذا العامل انعدمت تقريباً الارتباطات

والاتساقات في درجات الطلاب على مفردات الاختبار، وبالتالي يتحقق افتراض الاستقلال المحلي للاختبار.

وللتحقق من افتراض حسن المطابقة (goodness –of –Fit –test) للاستجابات عن مفردات الاختبار وفق نموذج سلم التقدير: أخضعت البيانات للتحليل باستخدام البرمجية الإحصائية (Ganz-Rasch)، وتم حساب إحصاءات الملائمة للمفردات وللأفراد وكانت النتائج كالتالي :

- أن قيم الخطأ المعيارى امتدت من (٠,١٧٧ إلى ٠,٣٨٤) بمتوسط حسابى (٠,٢٢٦)، وهذه القيمة منخفضة جداً مما يدل على دقة الاختبار وثباته.

- أن قيم البواقى امتدت من (-0.3 إلى 0.73) بمتوسط حسابى (0.012)، وهذه القيمة منخفضة جداً حيث كلما قلت هذه الفروق دل ذلك على صدق المفردات و الاختبار ككل.

- أن قيم احتمال الملائمة امتدت من (٠,٠٠٠ إلى ٠,٩٩٨).

- أن قيم كا٢ امتدت من (٨٨,٢ إلى ٢٨٩,١٤)، والمفردات التى جاءت قيم كا٢ لها دالة إحصائياً والذي يدل على عدم تحقق خصائصها السيكومترية طبقاً لنموذج راش هى :

جدول (٥) أرقام المفردات التى تم حذفها طبقاً لنموذج راش

المفردة	الخطأ المعيارى SE	البواقىResidual	كا٢ Chi- Sq	احتمال الملائمة Prob.
٥	٠,٢٩	0.17	١٩٧,٢٦٦	٠,٠٢٧
١٣	٠,٣٨	0.3	١٩٨,٣٣١	٠,٠٢٤
١٧	٠,١٩	0.24	٢١٠,٨٧٨	٠,٠٠٥
١٨	٠,٢٩	0.32	٢٠٠,٩٢٥	٠,٠١٨
١٩	٠,٢	0.11	٢٣١,١١١	٠,٠٠٠٠
٢٦	٠,٣٣	0.73	٢٨٩,١٤٤	٠,٠٠٠٠
٣٠	٠,١٨	0.25	٢٦٦,٧٥٢	٠,٠٠٠٠

هذا يشير إلى أن بقية المفردات وعددها ٢٣ مفردة من أصل ٣٠ مفردة حققت إحصاءاتها لشرط الملائمة طبقاً لمحكات نموذج راش الخاصة بالمفردة

وقد تم استبعاد المفردات الغير ملائمة لنموذج راش من التحليل في المرحلة التالية التي استخدم فيها الباحث برنامج Winsteps حيث عدد المفردات أصبح ٢٣ مفردة فقط بالتالي يمكن استخدام برنامج Winsteps الذي لا يدرج الاختبارات التي تزيد مفرداتها عن ٢٥ مفردة :

ويوضح الجدول التالى ملاءمة البيانات المستمدة من الاختبار لنموذج راش :

جدول (٦) ملاءمة البيانات لنموذج راش

المتغير	المفردات Items		الأفرادPersons	
	الاستقلال المركزى Location	إحصاء الملاءمة Fit Statistic	الاستقلال المركزى Location	إحصاء الملاءمة Fit Statistic
المتوسط Mean	0.00	1.14+	2.36-	0.91
الانحراف المعيارىS D	1.18+	1.86+	0.94+	1.11+
كا٢	**160.25**			
الدلالة	**٠,٠٠٠٠**			
معامل ألفا	**٠,٨٨**			

ويوضح الجدول السابق أن إحصاءات المفردات / الأفراد بعد استبعاد المفردات غير الملائمة لنموذج راش قيمها مناسبة، وقد بلغت قيمة مربع كا باستخدام برمجية Winsteps ١٦٠,٢٥ وهي قيمة دالة إحصائياً، وبلغت قيمة معامل ألفا (٠,٨٨) وهي تدل على ثبات احصاءات المفردات وكذلك ثبات قدرات الأفراد الأمر الذي يوضح أن حذف المفردات غير الملائمة لنموذج راش ادى إلى ارتفاع قيم بارامترات (المفردات / الأفراد) وكذلك ارتفاع مؤشرات النموذج في تقدير القدرة الكامنة وراء الاختبار في ضوء المحكات التي وفرها نموذج راش أحادي البعد.

ويوضح الملحق رقم (٢) أرقام المفردات، والقيمة الاحتمالية لدلالة كا٢ لكل مفردة، والتى تبين مدى ملاءمة المفردة للنموذج، وكذلك الخطأ المعيارى لكل مفردة، حيث أنه كلما قل الخطأ المعيارى للمفردة أو الاختبار ككل دل ذلك على دقة مفردات الاختبار فى قياس القدرة المطلوب قياسها.

النتائج الخاصة بالفرض الثالث

ينص على أن مفردات الاختبار التحصيلي لمقرر مهارات التعلم والتفكير تتمتع بمعاملات صعوبة ملائمة وفق نموذج راش

تم حذف المفردات التي أثبت برنامج جانز راش في بيانات الفرض الثاني أنها غير ملائمة لنموذج راش وهي المفردات أرقام (٥ - ١٣ - ١٧ - ١٨ -١٩ – ٢٦ - ٣٠) والجدول التالي يوضح قيم معاملات صعوبة كل مفردة مقدرة باللوجيت للاختبار التحصيلي في مقرر مهارات التعلم والتفكير :

جدول (٧)

مفردات اختبار مقرر مهارات التعلم والتفكير وفق معاملات صعوبتها بعد تدريجها وفق نموذج راش (ن = ١٦٢)

المفردة	الصعوبة باللوجيت	الخطأ المعياري باللوجيت	إحصاء الملاءمة	
			التقاربى	التباعدي
١	١,٢٩٨-	٠,٢٧٨	٠,٨٤٥	٠,٥٤٤
٢	٠,٨٣١-	٠,٢٤٠	٠,٨٩٢	٠,٧٧٥
٣	١,٢٧٩	٠,١٧٦	٠,٩٣٥	٠,٩٠٨
٤	١,٠١٠	٠,١٧٧	٠,٩٤٩	٠,٩٠٢
٦	٠,٨٣١-	٠,٢٤٠	٠,٨٨٣	١,١٢٧
٧	٠,٧١٩-	٠,٢٣٢	٠,٨٩٥	٠,٧٢٩
٨	٠,٢٤٧-	٠,٢٠٧	٠,٩١١	٠,٧٨٦
٩	٠,٧٠٤	٠,١٨٠	٠,٨٧٢	٠,٧٦٦
١٠	٠,١٤٠	٠,١٩٣	٠,٨٥٦	٠,٦٩٩
١١	٠,٩١٩	٠,١٧٨	١,٠٣١	١,٠٣٧
١٢	١,٨٢٣	٠,١٨١	٠,٩٧٤	١,٠٨٨
١٤	٠,٥٤٦	٠,١٨٣	٠,٩٣٦	٠,٨٦٩
١٥	٠,١٦٥-	٠,٢٠٣	١,٠١٣	١,٠٨٥
١٦	٠,٨٥٩	٠,١٧٨	١,٠٤٩	١,٠١٤
٢٠	٠,٣٨٣	٠,١٨٦	٠,٩٧٢	٠,٨٢٢
٢١	١,٤٦٥-	٠,٢٩٤	٠,٩٩٨	٠,٨٧٢
٢٢	٠,٤٤٩	٠,١٨٥	٠,٨٠١	٠,٦٧٩
٢٣	٠,٨٢٨	٠,١٧٩	٠,٨١٣	٠,٧١١

المفردة	الصعوبة باللوجيت	الخطأ المعياري باللوجيت	إحصاء الملاءمة	
			التقاربى	التباعدي
٢٤	٠,٨٢٨	٠,١٧٩	١,٠١٧	٠,٩٣٣
٢٥	٢,١٥٥-	٠,٣٨٤	٠,٨٩٥	٠,٦٥٢
٢٧	١,٤٦٥-	٠,٢٩٤	١,٠٢٤	١,٠٧٢
٢٨	٠,٢٤٧-	٠,٢٠٧	٠,٩٦٩	٠,٨٤٥
٢٩	٠,٣٨٣	٠,١٨٦	١,٠٤٧	٠,٩١٧

يتضح من الجدول السابق أن قيم تقديرات صعوبة المفردات امتدت من (-٢,١٥٥ إلى ١,٨٢٣) باللوجيت وامتدت قيم الأخطاء المعيارية لتقديرات صعوبة المفردات من (٠,١٨ إلى ٠,٣٨)

وكانت قيم إحصاءات الملاءمة (التقاربية – التباعدية) مناسبة، ويلاحظ اختلاف تدريج مفردات الاختبار حسب اختلاف مستويات صعوبتها.

النتائج الخاصة بالفرض الرابع

ينص الفرض الرابع على أنه

لا يتغير تدريج مفردات اختبار مقرر مهارات التعلم والتفكير المطورة باستخدام نموذج راش بتغير مستوى قدرات أفراد العينة المستخدمة في الحصول على هذا التدريج.

يشير (أبوهاشم، ٢٠٠٦) إلى أن التحقق من صحة فرض كهذا يستلزم وجود مجموعتين من الأفراد الملائمين لأداء نفس مجموعة المفردات المدرجة بواسطة نموذج راش حيث يؤدى هذا إلى التحقق من أن تقديرات القدرة المقابلة لكل درجة محتملة على الاختبار والناتجة من تحليل أداء أفراد إحدى المجموعتين على هذا الاختبار تتكافأ إحصائياً مع تقديرات القدرة المقابلة لكل

درجة كلية محتملة على الاختبار والمشتقة من تحليل أداء أفراد المجموعة الثانية مع الأخذ فى الاعتبار الخطأ المعيارى لهذه التقديرات، فإن هذا يعنى أن تقدير قدرة الفرد الذى حصل على درجة كلية معينة على القائمة لا تتأثر باختلاف مستوى أداء عينة التحليل، وبمعنى آخر، أن قدرة الفرد المقدرة بمستوى تحصيله متحررة أو مستقلة من قدرة أفراد العينة المستخدمة فى الحصول على هذه التقديرات؛ وهو أحد الافتراضات الأساسية لنموذج راش أحادي البارامتر.

ولتحقيق ذلك استرشد الباحث ب أبوهاشم (٢٠٠٦) و قام بحساب الوسيط لدرجات أفراد العينة على اختبار مقرر مهارات التعلم والتفكير ثم تصنيفهم إلى مجموعتين ثم حساب القيم التقديرية لقدرة أفراد كل مجموعة على حدة معتمد فى ذلك على الدرجات الكلية التى حصل عليها كل طالب في اختبار مقرر مهارات التعلم والتفكير والملحق (٣) يوضح القيم لقدرة الأفراد المناظرة للدرجات الكلية فى الاختبار لدى المجموعتين ويمكن وضع تقديرات القدرة المقابلة للمجموعتين فى الجدول التالى :

جدول (٨) تقديرات القدرة المقابلة لكل درجة كلية محتملة على الاختبار والمشتقة من كل من : العينة المرتفعة والعينة المنخفضة، والأخطاء المعيارية

مجموعات الدرجات الكلية	العينة المرتفعة ن=82		العينة المنخفضة ن=80	
	تقدير القدرة	الخطأ المعيارى	تقدير القدرة	الخطأ المعيارى
1	٤,٠٦٦-	١,٠٤٤	٤,٠٦١-	١,٠٤٢
2	٣,٢٨٣-	٠,٧٦٨	٣,٢٨١-	٠,٧٦٧
3	٢,٧٨٧-	٠,٦٥١	٢,٧٨٦-	٠,٦٥١
4	٢,٤٠٩-	٠,٥٨٤	٢,٤٠٨-	٠,٥٨٥

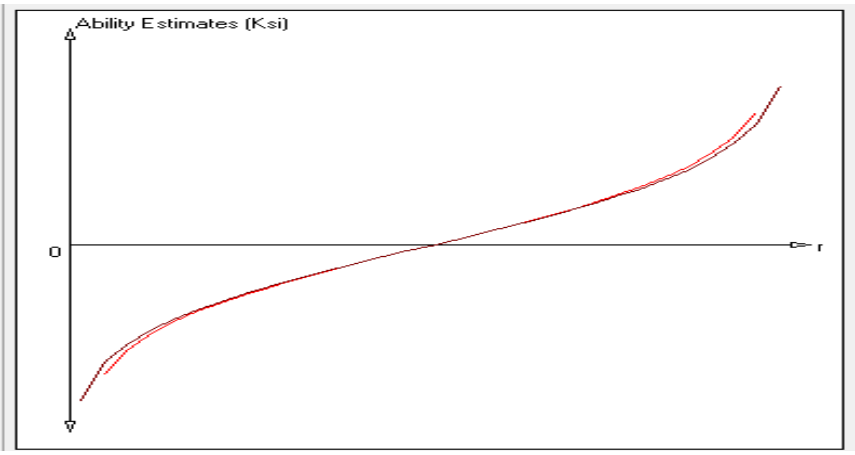
مجموعات الدرجات الكلية	العينة المرتفعة ن=82		العينة المنخفضة ن=80	
	تقدير القدرة	الخطأ المعيارى	تقدير القدرة	الخطأ المعيارى
5	٢,٠٩٥-	٠,٥٣٩	٢,٠٩٢-	٠,٥٤٢
6	١,٨٢٢-	٠,٥٠٨	١,٨١٥-	٠,٥١٢
7	١,٥٧٦-	٠,٤٨٤	١,٥٦٥-	٠,٤٨٩
8	١,٣٥١-	٠,٤٦٧	١,٣٣٤-	٠,٤٧٣
9	١,١٤٠-	٠,٤٥٣	١,١١٧-	٠,٤٦٠
10	٠,٩٣٩-	٠,٤٤٢	٠,٩١١-	٠,٤٤٩
11	٠,٧٤٧-	٠,٤٣٤	٠,٧١٣-	٠,٤٤١
12	٠,٥٦١-	٠,٤٢٩	٠,٥٢١-	٠,٤٣٥
13	٠,٣٧٩-	٠,٤٢٥	٠,٣٣٤-	٠,٤٣٠
14	٠,٢٠٠-	٠,٤٢٣	٠,١٥٠-	٠,٤٢٧
15	٠,٠٢١-	٠,٤٢٢	٠,٠٣١	٠,٤٢٥
16	٠,١٥٨	٠,٤٢٤	٠,٢١١	٠,٤٢٤
17	٠,٣٣٨	٠,٤٢٦	٠,٣٩٢	٠,٤٢٥
18	٠,٥٢٢	٠,٤٣١	٠,٥٧٤	٠,٤٢٨
19	٠,٧١١	٠,٤٣٨	٠,٧٥٩	٠,٤٣٢
20	٠,٩٠٦	٠,٤٤٧	٠,٩٤٨	٠,٤٣٩
21	١,١١١	٠,٤٥٩	١,١٤٤	٠,٤٤٨
22	١,٣٢٨	٠,٤٧٤	١,٣٥٠	٠,٤٦٠
23	١,٥٦١	٠,٤٩٣	١,٥٦٨	٠,٤٧٦

مجموعات الدرجات الكلية	العينة المرتفعة ن=82		العينة المنخفضة ن=80	
	تقدير القدرة	الخطأ المعيارى	تقدير القدرة	الخطأ المعيارى
24	١,٨١٦	٠,٥١٨	١,٨٠٥	٠,٤٩٨
25	٢,١٠٢	٠,٥٥٢	٢,٠٦٧	٠,٥٢٨
26	٢,٤٣١	٠,٥٩٨	٢,٣٦٨	٠,٥٧٢
27	٢,٨٢٨	٠,٦٦٧	٢,٧٣٢	٠,٦٣٩
28	٣,٣٤٧	٠,٧٨٥	٣,٢١١	٠,٧٥٧
29	٤,١٥٨	١,٠٥٨	٣,٩٧٧	١,٠٣٥

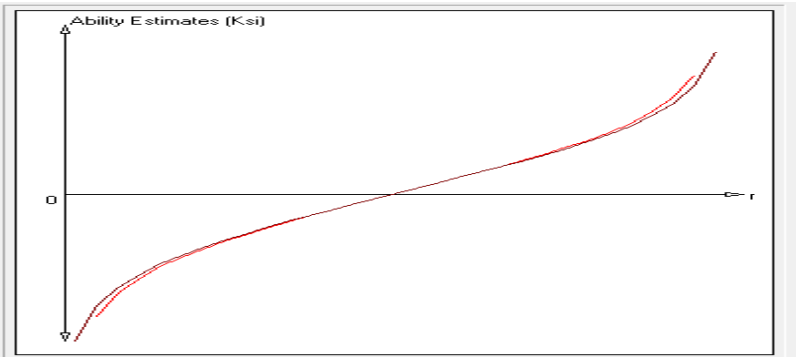
ويتضح من الجدول (٨) تكافؤ التقديرات المتناظرة المشتقة من تحليل أداء أفراد المجموعتين المرتفعة والمنخفضة القدرة – مع الأخذ فى الاعتبار الأخطاء المعيارية لتقديرات القدرة – وذلك بالنسبة لقدرة الأفراد الحاصلين على كل درجة كلية ممكنة أو محتملة على الاختبار وهذا يعنى عدم تأثر تلك التقديرات باختلاف مستوى عينة التحليل (المرتفعة أو المنخفضة في القدرة المقاسة)

ويتضح من الجدول كذلك أنه كلما زادت الدرجة الكلية زادت القدرة المقابلة لها، وأن أعلى درجة مقابلة لأعلى قدرة.

وللتحقق من مدى انطباق كل من توزيعى القدرة" المرتفعة والمنخفضة" الموضحين بالجدول (٨) يمكن تمثيل كل منهما بيانياً على شكل منحنى مميز للقدرة (شكل رقم٢ ، ٣) يعبر عن العلاقة القياسية **Calibration Relationship** بين الدرجات الكلية المحتملة على الاختبار والقيم التقديرية للقدرة المناظرة لها، كما اشتقت من كل من عينتى المجموعة المرتفعة القدرة والمجموعة المنخفضة القدرة



شكل (٢) منحنى مميز لتقدير القدرة للعينة المرتفعة



شكل (٣) منحنى مميز لتقدير القدرة للعينة المنخفضة

ويلاحظ تطابق الشكلين تقريباً فيما يخص تقدير القدرات رغم اختلاف العينات (مرتفعة –منخفضة)

تفسير النتائج والتعليق عليها

نجح نموذج راش في تدريج مفردات أداة البحث : الاختبار التحصيلي في مقرر مهارات التعلم والتفكير لدى طلاب كلية التربية جامعة الملك فيصل، وتكونت الصورة النهائية للاختبار من ٢٣ مفردة؛ كانت إحصاءات (المفردات / الأفراد) لها ملاءمة لنموذج راش، وتم استبعاد المفردات الغير ملاءمة للنموذج؛ واتفقت النتائج في ذلك مع نتائج دراسة أبو هاشم (٢٠٠٦) مع اختلاف القدرات المقاسة، وكان هناك قدر من التشابه بين كلا النظريتين : الكلاسيكية والاستجابة للمفردة في نتائج تحليل المفردات لكن نموذج راش (كأحد نماذج الاستجابة للمفردة –أحادية البعد) كان أكثر دقة في تقدير القدرات وتحليل المفردات وتدريجها وفق مستويات القدرة المناظرة لها لدى الأفراد وظهر قدر واضح من الثبات في ذلك التدريج رغم اختلاف العينات (منخفضة – مرتفعة القدرة)، وهو ما اتفق فيه كورفيل (٢٠٠٤) Courville مع البحث الحالي حيث أظهر المسح الذي أجراه أن نماذج الاستجابة للمفردة كانت أكثر ثباتاً في التعبير عن بارامترات (المفردة/الفرد) من النظرية الكلاسيكية. وأشارت نتائج أبوهاشم (٢٠٠٦) إلى نفس النقطة حيث أفرزت النظرية الكلاسيكية مقياساً يتكون من (٢٨) مفردة في حين افرز نموذج راش مقياساً يتكون من (٢١) مفردة. وكذلك نتائج دراسات فان Fan(١٩٩٨)

وأشارت نتائج دراسة أبو جراد (٢٠٠٨) أن نموذج راش حذف مفردتين عند إعادة تدريج اختبار الذكاء الثالث كاتل الصورة (أ) باستخدامه عن النظرية الكلاسيكية.

واتفقت نتائج دراسة Progar and Socan(2008) مع هذه النتيجة مع التركيز على أن إحصاءات الفرد كانت أكثر ثباتاً باستخدام الاستجابة للمفردة من النظرية الكلاسيكية.

وكانت بيانات العينة المستمدة من كلا الإطارين القياسيين قابلة للمقارنة، واتفق البحث في ذلك مع دراسة عيد (٢٠٠٤)، وملحق ٣ يوضح قيم القدرات المقاسة باللوجيت للأفراد على الاختبار والدرجات الكلية المقابلة لها، واتفق البحث الحالي في ذلك مع نتائج دراسات كل من : الحكماني (٢٠١٠)، وفراج والشريف (٢٠١٣).

واختلف البحث الحالي مع دراسة طيفور (٢٠٠٧) التي قارنت بين نماذج (احادية – ثنائية – ثلاثية) الأبعاد في حين استخدم البحث الحالي نموذج راش أحادي البعد فقط لأن تفسير سلوك المفردة وفق أكثر من نموذج من نماذج الاستجابة للمفردة ربما يشتتنا عن اتخاذ قرار صائب خاصة إذا كانت النتائج لنفس المفردة متعارضة بين أكثر من نموذج. نفس الشيء على دراسة العنزي (٢٠١٢).

واتفق البحث الحالي مع نتائج دراسة الثوابية (٢٠١٠) أن الخطأ المعياري في تقدير بارامترات صعوبة المفردات يتناقص بزيادة عدد أفراد العينة ؛ حيث بلغ متوسط صعوبة مفردات الاختبار (0.01-) لوجيت عندما كان حجم العينة (٨٢)، في حين ازداد بحيث أصبح (٠,٠٠) لوجيت عندما أصبح حجم العينة (١٦٢).

واختلف البحث الحالي مع دراسة محمود (٢٠١٠) الذي استخدم برنامج راسكال في تدريج اختبار وفق نموذج راش وقدم دليلاً مختصراً لاستخدام البرنامج في حين كانت نسخ البرنامج الإلكترونية غير متاحة على الانترنت ؛ في حين استخدم البحث الحالي برنامج جانز راش، وبرنامج وينستبس وهي

برامج أسهل ومتاحة وفعالة في تدريج المفردات.وقد استخدمت دراسة Carlson and Locklin(1995) برامج قديمة نسبياً مثل Bilog وظهر في السنوات الأخيرة مع التطور التكنولوجي برامج أحدث وأسهل في استخدامها وأكثر توافقاً مع البرمجيات الأخرى مثل برنامج SPSS.

الدراسات العربية التي استخدمت نموذج راش : مثل أبوجراد (٢٠٠٨)، وعيد (٢٠٠٤)، والخياط (٢٠١٢)، والجبوري (٢٠١٢)، والمصري (٢٠١٥) كلها اتفقت مع البحث الحالي في المنهجية من حيث التقنين وفق النظرية الكلاسيكية ؛ ثم تدريج الاختبارات وإجراء اختبارات المطابقة مع نموذج راش وفق محكاته على اختلاف برامج التحليل الإحصائي المستخدمة

البحوث والدراسات المقترحة والتوصيات

-ضرورة تضمين بحوث نظرية الاستجابة للمفردة المنشورة نبذة عن برنامج التحليل الإحصائي المستخدم في تدريج مفردات أدوات كل بحث ؛ حيث تتعدد البرامج ولا يوجد – في حدود علم الباحث - مصادر واضحة باللغة العربية لشرح كيفية استخدام هذه البرامج من قبل الباحثين والتربويين والإحصائيين العرب، فمن بين البحوث التي راجعها الباحث الحالي لم يجد سوى بحث واحد فقط قدم شرح للبرنامج المستخدم وهو دراسة محمود (٢٠١٠).

-استخدام الاختبار التحصيلي بصورته النهائية (٢٣) مفردة للكشف عن القدرات التحصيلية التي يتمتع بها الطلاب في مقرر مهارات التعلم والتفكير؛ نظرا لتمتع الاختبار المبني والمطابق لتوقعات نموذج سلم التقدير التابع لنظرية استجابة الفقرة (IRT) بدلالات صدق وثبات مقبولة.

-إجراء المزيد من الدراسات على الصورة الأولية للاختبار التحصيلي المبني في البحث الحالي، باستخدام نماذج السمات الكامنة المختلفة (ثنائية

وثلاثية الأبعاد)، وإجراء مقارنة بينها، وتطبيقه ليشمل جامعات أخرى، من أجل تأكيد الثقة حول صدق وثبات الاختبار.

- الاستناد على البرمجيات الإحصائية المستخدمة بالبحث الحالي (SPSS & Winsteps & Ganz-Rasch) في بناء الاختبارات التحصيلية الجامعية محكية المرجع وفق نظرية استجابة الفقرة، ومحاولة الاستفادة من المؤشرات الأخرى التي تقدمها هذه البرمجيات في بناء الاختبارات بصفة عامة.

-ضرورة أن تتبنى الجامعات السعودية ضمن منظومة التأهيل للجودة والاعتماد الأكاديمي عمل خريطة معرفية للقدرات التحصيلية المقاسة تنبع من المجالات المعرفية والمهارية والوجدانية المقاسة لدى الطلاب، وتتوافق وتنبع من توصيفات المقررات بحيث يكون الأمر نواة لبنوك الأسئلة التي يجب تطويرها مع المقررات وتوصيفاتها في ضوء نتائج تقييم المقررات الدراسية في نهاية كل عام جامعي ؛ الأمر الذي سيفرض على الأساتذة في الجامعات التوسع في بناء وتصميم اختباراتهم وفق منهجية نظرية الاستجابة للمفردة ونماذجها المتنوعة بالتالي نشر ثقافتها في الجامعات يتيح زيادة في دقة تقدير مستويات قدرات الطلاب المقاسة ويحررها من أثر العينات ويتيح قدر من التوافق بين أهداف التدريس الجامعي، وأهداف القياس التربوي، ويحل مشكلات اختلاف وتعدد نماذج وتصميمات الاختبارات المليئة بالأخطاء ويتم تصميمها وتحليلها وفق المنهجية التقليدية من بعض غير ذوي التخصص فتحدث أخطاء القياس الشائعة.

المراجع:

- أبو جراد، حمدي يونس (٢٠٠٨). استخدام نموذج راش في تطوير اختبار كاتل الثالث للذكاء الصورة أ، **مجلة الجامعة الإسلامية** (سلسلة الدراسات الإنسانية)، ١٦(١)، ٥٥٥ -٥٨٣.

-أبوهاشم، السيد محمد (٢٠٠٦). دراسة مقارنة بين النظرية التقليدية ونموذج راش في اختيار قائمة مداخل الدراسة لدى طلاب الجامعة، **مجلة كلية التربية بالزقازيق**، ٥٢(١)، ١ -٣٧، يناير ٢٠٠٦.

-الثوابية، أحمد محمود (٢٠١٠). أثر حجم العينة على تقدير صعوبة الفقرة والخطأ المعياري في تقديرها باستخدام نظرية الاستجابة للمفردة، **مجلة جامعة دمشق**، ٢٦(٢)، ٥٢٥ -٥٥٩.

-الجبوري، رشيد صالح مهدي (٢٠١٢). بناء اختبار تحصيلي لمادة الإدارة والإشراف التربوي على وفق نظرية السمات الكامنة لطلبة معاهد إعداد المعلمين، **مجلة الأستاذ**، العدد (٢٠٣)، ١٣٩٢ – ١٤٢٧.

-الحكماني، رحاب سعيد (٢٠١٠). مقارنة بين النظرية الكلاسيكية للاختبار ونظرية الاستجابة للمفردة في تقدير قدرات الأفراد ومدى استقرار مؤشرات المفردات الاختبارية، **رسالة ماجستير**، كلية التربية، جامعة السلطان قابوس. تم استرجاعه بتاريخ (٢٠١٧/١٠/٥)

thesis.al-kawakab.com/thesis/518.

-الخولي، زياد عبد الحسيب (٢٠١٣). اتجاهات معاصرة في القياس الموضوعي (نموذج التقدير الجزئي متعدد التدريج)، **مجلة الآداب**، جامعة البعث، ١٠٦ (١)، ٦٥٣ -٦٨٠.

-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.

-العلي، ديالا.(٢٠١٤).استخدام نموذج راش في تدريج مقياس تقبل إدمان وإدمان المخدرات من اختبار مينيسوتا"MMPI-2" لدراسة بعض العوامل المؤثرة على دقة القياس.**مجلة جامعة البعث**، ٣٦(١٧)، ٥٧ - ٩٢.

-العنزي، سعود بن شايش (٢٠١٢). **استخدام نظرية استجابة المفردة في تطوير اختبار كورنيل للتفكير الناقد**، رسالة دكتوراه، قسم علم النفس التربوي، كلية التربية، جامعة طيبة.

-المصري، محمد عبد المجيد (٢٠١٥). الخصائص السيكومترية لاختبار محكي المرجع في البحث التربوي وفق نظرية الاستجابة للفقرة، **مجلة كلية التربية الأساسية**، ٢١(٨٩)، ٧٠١ - ٧٢٧.

-حمادنة، إياد و خالد، محمد بني (٢٠١٣). بناء مقياس اتجاهات نحو العنف الإلكتروني لدى عينة من مستخدمي مواقع التواصل الاجتماعي بجامعة آل البيت، **مجلة المنارة**، ١٩(٣)، ٧٧ -١٠٠.

-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)، ١ -٣٧

-طيفور، مصطفى أحمد (٢٠٠٧). **دراسة مقارنة لنماذج نظرية الاستجابة للمفردة في معادلة درجات الاختبارات**، رسالة دكتوراه، معهد الدراسات التربوية، جامعة القاهرة (تم استرجاعه بتاريخ ٢٠١٧/١٠/٢١)

http://www.alnodom.com.sa/index.php/الملخصات -الجامعية/٣٥٢١.html

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

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

-عيد، غادة خالد (٢٠٠٤). الدرجة الحقيقية المقدرة باستخدام نظرية السمات الكامنة والنظرية الكلاسيكية. (دراسة سيكومترية)، ***مجلة جامعة أم القرى للعلوم التربوية والاجتماعية، ١٦ (٢)***، ٢٢٩ -٢٦٤.

-فراج، محمد أنور والشريف، خالد حسن (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**، جامعة الإمام محمد بن سعود الإسلامية، ٢٨(١)، ٩١ -١٤٧.

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

-محمود، محمد أنور (٢٠١٠). بناء دليل استخدام برنامج راسكال في تحليل بيانات الاختبارات المحكية، ***مجلة القادسية للعلوم الإنسانية، ١٣ (١)***، ٢٩٣ - ٣١٣.

-نور الدين، أمين صبري (٢٠١٧). **فعالية اختبار تحصيلي تواؤمي محوسب في تقويم طلاب جامعة الجوف**، بحث مقدم إلى ندوة التقويم في التعليم الجامعي ؛ مرتكزات وتطلعات، كلية التربية جامعة الجوف، ٤٦ -٧٩. (تم استرجاعه بتاريخ ٢٠١٧/٩/١٣)

http://www.ju.edu.sa/index.php?id=1883

-An , X. and Yung ,Y.(2014). Item Response Theory : What It Is and How You Can Use the IRT Procedure to Apply It , **SAS Institute** ,Paper SAS364-2014. 1-14.

- Carlson, Randal D.; Locklin, Ralph H. (1995). Item Response Theory : Comparing Bilog and MicroCAT calibration for a Mathematics Ability Test.**,** Educational Resources Information Center **, ERIC no : ED 393 881(**Retrieved 29/9/2017)

eric.ed.gov/?q=Item+Response+Theory+%3a+Comparing+Bilog+and+MicroCAT+calibration+for+a+Mathematics+Ability+Test&id=ED393881

-Courville , T.(2004). An Empirical Comparison of Item Respons Theory and Classical Test Theory Item/Person Statistics. Submitted for the degree of DOCTOR OF PHILOSOPHY , Texas A&M University.

-Fan ,Xitao (1998). Item Response Theory and Classical Test Theory : an empirical comparison of their item/person statistics , ***Educational and Psychological Measurement*** , 58(3) ,357-383.

- Gruijter, D. and Kamp, L.(2005). Statistical Test Theory for Education and Psychology, Retrieved December 30, 2005,Retreived(21/9/2017)

From: www.leidenuniv.nl/-gruijterdnmde.

- Hambleton, R., Swaminathan, H. and Rogers, J.(1991). ***Fundamental of Item Response Theory***, Second Edition, Newbury Perk California, 1991.

-Hattie, J. A.(1985).Methodology review : Assessing unidimensionality of tests and items, ***Applied psychological Measurement*** ,9, 139-164.

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

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

- Progar,S. and Socan,G.(2008). An empirical Comparison of Item Response Theory and Classical Test Theory **, Horizons of Psychology ,**17(3) ,5-24.

-Trang Le,D.(2013). Applying item Response Theory Modeling in Educational Research , A dissertation for the degree of DOCTOR OF PHILOSOPHY , Iowa State University , Ames ,Iowa. Retrieved(20/9/2017):

Lib.dr.iastate.edu/etd/13410.

ملاحق البحث:

ملحق ١

الاختبار التحصيلي في مقرر مهارات التعلم والتفكير

ملحوظة (المفردات المظللة بالأصفر هي المفردات التي تم استبعادها من الصورة النهائية للاختبار بسبب عدم ملاءمتها لنموذج راش)

أمامك ٣٠ سؤال من نوع الاختيار من متعدد والمطلوب منك **اختيار الإجابة الصحيحة لكل سؤال والتظليل على رقمها في ورقة الإجابة الإلكترونية**

١ -يعرف على أنه المهارة الفعالة التي تدفع بالذكاء إلى العمل : .....

(أ) الذكاء. (ب) التعلم.

(ج)التفكير. (د) الإدراك.

٢ -تفكير مكتسب من خلال التدريب والبرامج المتخصصة ويتميز عن تفكير الحياة اليومية : .....

(أ) التفكير التقليدي (ب) التفكير المعتاد.

(ج) التفكير المهاري. (د) التفكير الخرافي.

٣ -إعطاء الطالب الفرصة والمجال كي يفكر دون وجود برنامج متخصص للتفكير : .....

(أ) تعليم التفكير. (ب) تعليم مهارات التفكير.

(ج) تعليم أساليب التفكير. (د) تعليم القراءة.

٤ - تعرف على أنها كل ما ينقل عملية التفكير من كونها عملية عفوية تتم بشكل تلقائي إلى عملية فعالة متميزة : .....

(أ) أساليب التفكير. (ب) مهارات التفكير.

(ج) عمليات التفكير. (د) أدوات التفكير.

٥ - كل ما يأتي يوضح أهمية تعليم التفكير في حياتنا اليومية ماعدا : .....

(أ)يسهم في حل مشكلات الحياة اليومية

(ب) ضروري للنجاح في الدراسة

(ج) يشجع في الطلاب روح الاستقلالية (د) يـزرع في الطـلاب روح السلبية

٦ - أبرز رواد الأسلوب المباشر في تعليم مهارات التفكير هو: .....

(أ) ديبونو. (ب) جون ديوي.

(ج) باير. (د) مارزانو.

٧ -أبرز رواد أسلوب الدمج في تعليم مهارات التفكير هو : .....

(أ) ديبونو. (ب) جون ديوي.

(ج) باير. (د) مارزانو.

٨ - كل مما يأتي من خصائص الأسلوب المباشر في تعليم التفكير ما عدا: -

(أ)دمج مهارات التفكير في المقرر الدراسي.

(ب) هناك برامج متخصصة للتفكير.

(ج) هناك مدرس مستقل لمادة التفكير.

(د) هناك دروس مخصصة للتفكير.

٩ -برنامج عالمي لتنمية مهارات التفكير صممه ديبونو..... :

(أ) الماستر ثنكر. (ب) تريز.

(ج) فيورشتين. (د) كورت.

١٠ - يعرف على أنه تجمع لعدد كبير من المهارات التي تجمعها خاصية معينة : .....

(أ) مهارة فرعية. (ب) مهارة رئيسية.

(ج) عنقود مهارات. (د) تحت مهارات.

١١ - العلاقة بين التفكير والتعلم هي علاقة : .....

(أ) علاقة تنافر. (ب) علاقة الجزء بالجزء.

(ج) علاقة السبب بالنتيجة. (د) علاقة سالبة.

١٢ - تصنف مهارات التفكير التأملي في تصنيفات مهارات التفكير ضمن : .....

(أ) مهارات التفكير الأساسية. (ب) مهارات التفكير العليا.

(ج) مهارات التفكير الميتامعرفية. (د) مهارات التفكير الدنيا.

١٣ - التفكير عملية هادفة، ويكون الهدف منها هو : .....

(أ) حل المشكلات. (ب) خلق المشكلات.

(ج) تصنيف المشكلات. (د) تقسيم المشكلات.

١٤ - ميز في تصنيف مستويات التفكير التأملي بين التفكير التأملي أثناء العمل، والتفكير التأملي بعد العمل : .....

(أ) جون ديوي. (ب) ديفيد كولب.

(ج) دونالد شون. (د) جيني موون.

١٥ - نمط التفكير الذي يعمل عندما نطلب من الفرد التخلي عن الآراء والتقيد بالحقائق فقط : .....

(أ) التفكير التأملي. (ب) التفكير التخيلي.

(ج) التفكير الإبداعي. (د) التفكير العلمي.

١٦ - كل ما يلي من أدوات التفكير <u>ماعدا</u> : .....

(أ) صحيفة التعلم. (ب) ملفات الأداء.

(ج) الألعاب التعليمية. (د) المدرب.

١٧ - العلاقة بين التفكير التأملي والتفكير العلمي هي علاقة : .....

(أ) الجزء بالكل. (ب) الكل بالجزء.

(ج) الجزء بالجزء. (د) لكل بالكل.

١٨ - مهارة من مهارات التعلم تتمثل في القدرة على استخلاص الفكرة الرئيسية في النص والتعبير عنها بألفاظ أقل : .....

(أ) التخطيط. (ب) التنبؤ.

(ج) التلخيص. (د) التصنيف.

١٩ - يعرف على أنه صيغة من تجهيز العقل نستخدمه لتحقيق هدف أو الوصول إلى نتائج متوقعة ويطبق على المشكلات التي ليس لها حل واضح : .....

(أ) التفكير الإبداعي. (ب) التفكير الناقد.

(ج) التفكير العلمي (د) التفكير التأملي.

٢٠ - جميع ما يلي من عناصر الموقف المشكل ماعدا : .....

(أ) المعلومات. (ب) الحلول.

(ج) الأهداف. (د) العقبات.

٢١ - تصنيف مهارات التفكير إلى مهارات (أساسية – عليا – ميتا معرفية) من تصنيفات التفكير: .....

(أ) الثنائية. (ب) الثلاثية.

(ج) الرباعية. (د) السداسية.

٢٢ - يصنف التفكير الإبداعي ضمن قدرات التفكير..... :

(أ) التقاربي. (ب) الخطي.

(ج) الهرمي. (د) التباعدي.

٢٣ - كل ما يلي من قدرات التفكير التقاربي ماعدا : .....

(أ) التفكير التأملي. (ب) التفكير العلمي.

(ج) التفكير الإبداعي. (د) التفكير الناقد.

٢٤ - في تصنيف معالجة المعلومات تصنف مهارة التنبؤ ضمن مهارات : .....

(أ) جمع المعلومات. (ب) توليد المعلومات.

(ج)اختزال المعلومات. (د) تنظيم المعلومات.

٢٥ - وجود (تعليم التفكير) كهدف رئيسي في غالبية برامج التعليم المدرسي : .....

(أ) دليل أهمية التفكير. (ب) يمثل مضيعة للوقت.

(ج)تقليد لبرامج أجنبية. (د) ترجع للصدفة.

٢٦ - تعرف المشكلات التي تتسم بالغموض ولا تتضح فيها المعلومات عن العقبات بأنها مشكلات : .....

(أ) واضحة البنية. (ب) سهلة التكوين.

(ج) غير واضحة البنية. (د) متوسطة البنية.

٢٧ - جميع ما يلي من خصائص التفكير العلمي ماعدا : .....

(أ) الموضوعية. (ب) السلبية.

(ج) التراكمية. (د) المنطقية.

٢٨ -أطلق ديفيد كولب على الاستفادة من الخبرات التي تمر بالمتعلم أثناء عملية التعلم مصطلح : ....

(أ) التعلم التعاوني. (ب)التعلم الذاتي.

(ج) التعلم الخبري. (د)التعلم الاجتماعي.

٢٩ - سجل ذاتي خاص بالمتعلم يكتبه عن نفسه وبنفسه هي : .....

(أ) صحيفة التعلم. (ب) ملف الأنشطة.

(ج) السجلات المدرسية. (د) لعب الدور.

٣٠ - مهارة من مهارات التعلم يتمثل فيها تطبيقات التفكير العلمي : .....

(أ) مهارة الاستماع. (ب) مهارة الميتا معرفة.

(ج)مهارة حل المشكلات. (د) مهارة العد.

انتهت الأسئلة

كل سؤال درجة واحدة

ملحق ٢

إحصاءات الملاءمة لمفردات اختبار مقرر مهارات التعلم والتفكير طبقاً لنموذج راش

هى :

المفردة	الخطأ المعيارى SE	البواقى Residual	كا٢ Chi- Sq	احتمال الملائمة Prob.
١	٠.28	0.30-	٨٨,٢٠٨	١,٠٠٠
٢	٠.24	0.12-	١٢٥,٥٠٥	٠,٩٨٢
٣	٠.18	0.03-	١٤٧,١١١	٠,٧٧٦
٤	٠.18	0.05-	١٤٦,١٥١	٠,٧٩٣
٦	٠.24	٠.24	١٨٢,٦١٦	٠,١١٧
٧	٠.23	0.17-	١١٨,١٤٢	٠,٩٩٥
٨	٠.21	0.13-	١٢٧,٣٧٥	٠,٩٧٦
٩	٠.18	0.11-	١٢٤,٠٧٤	٠,٩٨٦
١٠	٠.19	0.16-	١١٣,٢٦١	٠,٩٩٨
١١	٠.18	٠.01	١٦٧,٩٨٦	٠,٣٣٧
١٢	٠.18	٠.11	١٧٦,٢٦٧	٠,١٩٤
١٤	٠.18	0.07-	١٤٠,٧٠٨	٠,٨٧٤
١٥	٠.20	٠.07	١٧٥,٨٠٧	٠,٢٠١
١٦	٠.18	0.03-	١٦٤,٢٥٣	٠,٤١٤
٢٠	٠.19	0.15-	١٣٣,٢٣٥	٠,٩٤٦
٢١	٠.29	0.13-	١٤١,٢٦٩	٠,٨٦٧

المفردة	الخطأ المعيارى SE	البواقى Residual	كا٢ Chi- Sq	احتمال الملائمة Prob.
٢٢	٠.19	0.12-	١١٠,٠٥٨	٠,٩٩٩
٢٣	٠.18	0.10-	١١٥,١٤٣	٠,٩٩٨
٢٤	٠.18	0.08-	١٥١,٢٢٤	٠,٦٩٨
٢٥	٠.38	0.24-	١٠٥,٦٨٥	١,٠٠٠
٢٧	٠.29	0.05	١٧٣,٦١٠	٠,٢٣٥
٢٨	٠.21	0.12-	١٣٦,٨٩٤	٠,٩١٦
٢٩	٠.19	0.13-	١٤٨,٥٧٤	٠,٧٥٠

ملحق ٣

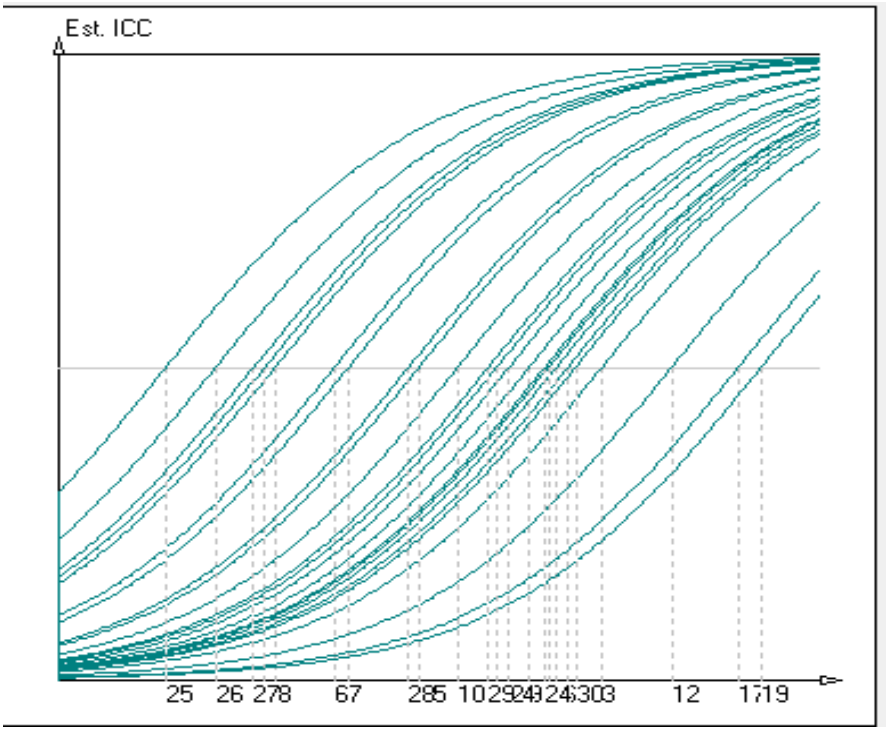
تقديرات القدرة المقابلة لكل درجة كلية محتملة على الاختبار والأخطاء المعيارية

مجموعات الدرجات الكلية	العينة كلها ن=١٦٢	
	تقدير القدرة باللوجيت	الخطأ المعيارى باللوجيت
١	٤,٠٢٧-	١,٠٣٩
٣	٣,٢٥٢-	٠,٧٦٣
٤	٢,٧٦٤-	٠,٦٤٦
٤	٢,٣٩٢-	٠,٥٧٩
٥	٢,٠٨٢-	٠,٥٣٦
٦	١,٨١١-	٠,٥٠٦
٧	١,٥٦٦-	٠,٤٨٤
٨	١,٣٤٠-	٠,٤٦٨
٩	١,١٢٨-	٠,٤٥٥
١٠	٠,٩٢٥-	٠,٤٤٥
١١	٠,٧٣١-	٠,٤٣٧
١٢	٠,٥٤٢-	٠,٤٣٢
١٣	٠,٣٥٧-	٠,٤٢٨
١٤	٠,١٧٦-	٠,٤٢٥
١٥	٠,٠٠٤	٠,٤٢٤
١٥	٠,١٨٣	٠,٤٢٤
١٧	٠,٣٦٤	٠,٤٢٥

مجموعات الدرجات الكلية	العينة كلها ن=١٦٢	
	تقدير القدرة باللوجيت	الخطأ المعيارى باللوجيت
١٨	٠,٥٤٦	٠,٤٢٩
١٩	٠,٧٣٢	٠,٤٣٤
٢٠	٠,٩٢٣	٠,٤٤١
٢١	١,١٢٢	٠,٤٥١
٢٢	١,٣٣٢	٠,٤٦٥
٢٣	١,٥٥٦	٠,٤٨٢
٢٤	١,٧٩٩	٠,٥٠٦
٢٥	٢,٠٧١	٠,٥٣٨
٢٦	٢,٣٨٣	٠,٥٨٣
٢٧	٢,٧٦١	٠,٦٥١
٢٨	٣,٢٥٨	٠,٧٦٩
٢٩	٤,٠٤٤	١,٠٤٦

ملحق (٤)

منحنى الخصائص المميزة لمفردات الاختبار على العينة الكلية (ICC)



* * *

- Hamadneh, Iyad and Khalid, Mohamed Bani (2013). Building a measure of trends towards cyber violence in a sample of users of social networking sites at Al-Albayet University, Al-Manara Magazine, 19 (3), 77-100.

-Debos, Mohamed (2016). The use of the theory of response to the paragraph in the construction of the paragraphs of the test of reference in mathematics in terms of dual-staging and multi-staging according to the binary model of logistics, Al-Najah University Journal of Research (Humanities), 30 (7), 1-37

-Tifour, Mustafa Ahmed (2007). A comparative study of the theory of individual response in the equation of test scores, PhD thesis, Institute of Educational Studies, Cairo University (retrieved on 21/10/2017)

http://www.alnodom.com.sa/index.php/ Abstracts - University / 3521 .html

-Alam, Salaheddine Mahmoud (2005). Models of response to individual test; one-dimensional and multidimensional and its applications in the measurement of psychological and educational, Cairo, Arab Thought House.

-Alam, Salaheddine Mahmoud (2007). Diagnostic tests Reference: in the fields of education, psychology and training, Cairo: Arab Thought House.

- Eid, Ghada Khaled (2004). The real degree is estimated using the theory of latent features and classical theory. (Socometric Study), Umm Al-Qura University Journal of Educational and Social Sciences, 16 (2), 229-264.

-Farag, Anwar and Elsherif, Khaled(2013)."The Effectiveness of Using the Theory Classical Test and the Latent Traits Theory in the Measurement of Educational Achievement in Algebra in the Preparatory Stage - Comparative Study, Journal of Human and Social Sciences, Imam Muhammad bin Saud Islamic University, 28 (1)), 91-147.

- Kazem, Amina Mohamed (1988). Using Rush Model in Building Achievement Test in Psychology and Achieving Objective Interpretation of Results, Kuwait University Press, Kuwait.

- Mahmoud, Mohamed Anwar (2010). Building RASCAL's Manual on Analysis of Audited Test Data, Qadisiyah Journal of Human Sciences, 13 (1), 293- 313.

- Nouruddin, Amin Sabri (2017). The Effectiveness of the Computerized Adaptive Achievement Test in Jouf University Students' Calendar, Research Presented at the Symposium on Evaluation in University Education, Perspectives and Aspirations, College of Education, Al-Jouf University, 46-79. (Retrieved on 13/9/2017)

http://www.ju.edu.sa/index.php?id=1883

❊ ❊ ❊

List of References:

-Abu Jarad, Hamdi Younis (2008). The use of the Rush model in the development of the CATEL III test of intelligence A, Journal of the Islamic University (Series of Humanities), 16 (1), 555-583.

-Bohashim, Elsayed Mohamed (2006). A comparative study between the traditional theory and the Rush model in selecting the list of study entries among university students, Zagazig College of Education Journal, 52 (1), 1-37, January 2006.

- Al-Thawabiya, Ahmed Mahmoud (2010). The effect of the sample size on the assessment of the difficulty of the paragraph and the standard error in its estimation using the individual response theory, University of Damascus Journal, 26 (2), 525-559.

-Jalbouri, Rashid Saleh Mahdi (2012). Building a test for the achievement of the subject of management and educational supervision according to the theory of the characteristics of the students of the teacher training institutes, Journal of Professor, No. (203), 1392 - 1427.

Al-Hakmani, Rehab Saeed (2010). Comparison between the classical theory of testing and the theory of individual response in estimating the abilities of individuals and the stability of the indicators of experimental vocabulary, Master Thesis, Faculty of Education, Sultan Qaboos University. Retrieved on (5/10/2017)

thesis.al-kawakab.com/thesis/518.

- AlKhouli, Ziad Abdel Hassib (2013). Contemporary Trends in Objective Measurement (Multimodal Stochastic Estimation Model), Journal of Arts, Al-Albayet University, 106 (1), 653-680.

- Al-Khayat, Majid Mohammed (2012). Degree of Achievement Achievement according to a Single Parameter Model in the Detection of the Level of Scientific Knowledge in the Mathematical Skills of Eighth Grade Students, Al-Aqsa University Journal (Humanities Series), 16 (1), 87-111.

Al-Ali, Diala (2014). Using the Rush model in the scale of the acceptance of drug addiction and addiction test from the Minnesota MMPI-2 to study some factors affecting the accuracy of measurement. Al-Albayet University Journal, 36 (17), 57-92.

- Al-Anzi, Saud bin Shaish (2012). The use of individual response theory in the development of the Cornell test of critical thinking, PhD thesis, Department of Educational Psychology, Faculty of Education, University of Taibah.

- Elmasry, Mohamed Abdelmajid (2015). The Psychometric Characteristics of the Estimator of the Reference in Educational Research according to the Response Theory of the Paragraph, Journal of the College of Basic Education, 21 (89), 701-727.

The effectiveness of using Rasch Model in calibration of Test Descriptors for the Learning and Thinking skills Course Applied to a sample of undergraduate students at King Faisal University Comparative Study with Classical Test Theory

Dr. Khalid Hassan Bakr AlShareef
Department of Psychology
College of Education
King Faisal University
Alexandria University

Abstract:

The aim of the study is to identify the psychometric characteristics of the Learning and Thinking Skills Test on a sample of 162 students of the college of Education, King Faisal University, using Classical Theory, as well as determining the psychometric characteristics of the same test using the Rasch model to verify the possibility of comparing students' grades in the test within both theories.

The results indicate that the number of items used to test the skills of learning and thinking differ according to the theory used in analysis. The classical theory showed a thirty-item test, while the Rasch model produced a thirty-two -item test. The calibration of the test items does not change according to changes in sample achieving ability (high – low). Seven items were omitted according to Rasch analysis. The ability parameters for persons ranged between (-4.03) logit and (4.04) logit.

Key Words: Rasch Model – Classical Test Theory – King Faisal University Students