

استخدام نموذج راش في تدريج مقياس التفكير الناقد لدراسة بعض العوامل المؤثرة على دقة القياس وفق نظرية الاستجابة للمفردة

م. علي عباس شنان الزامل

المديرية العامة لتربية واسط

ملخص البحث

تهدف هذه الدراسة إلى الكشف عن بعض العوامل المؤثرة (كحجم العينة المستخدمة في التدريج، وأفضل طريقة للربط بين الصور الاختبارية المختلفة وتدرج ميزان تدرج واحد مفردات مشتركة/أفراد مشتركين، وأفضل عدد للمفردات المشتركة أو للأفراد المشتركين) (المستخدمين في ربط الصور الاختبارية المختلفة) بالإضافة إلى أثر مستوى العينة المستخدمة في تدريج المفردات). والتي قد تؤثر على دقة القياس، وذلك للتوصل إلى أفضل تدريج يوفره نموذج راش لمقياس القدرة على التفكير الناقد، كما تهدف الدراسة إلى عمل معايير (الرتب المئينية، والدرجات التائية) والتي تفسر على أساسها قدرة الأفراد على المقياس و استخدام برنامج الحاسب الآلي ونستبس Winsteps في تدريج مقياس القدرة على التفكير الناقد وفقاً لنموذج راش وأجريت الدراسة على عينة من الأفراد بلغ عددهم ٨٢٩ طالب وطالبة من طلبة المرحلة الابتدائية (رابع، خامس) والمتوسطة (أول وثاني) وأظهرت النتائج البحث إن أفضل حجم من بين أحجام العينة المستخدمة في تدريج مقياس القدرة على التفكير الناقد باستخدام نموذج راش هو حجم (٨٠٠) وذلك تبعاً لمحاكات دقة القياس. وإن أفضل عدد للمفردات المشتركة والمستخدم في ربط الصور الاختبارية المختلفة هو العدد (١١) وذلك تبعاً لمحاكات دقة القياس وإن أفضل عدد للأفراد المشتركين والمستخدمين في ربط الصور الاختبارية هو العدد (٥٠) وذلك تبعاً لمحاكات دقة القياس. وإن أفضل طريقة للربط بين الصور الاختبارية هي طريقة

المفردات المشتركة وذلك تبعا لمحكات دقة القياس ووتبعاً لمحكات دقة القياس فقد اختلفت دقة القياس باختلاف مستوى العينة وذلك لصالح العينة منخفضة المستوى. وتحققت في مقياس القدرة على التفكير الناقد استقلالية القياس (بمعنى دقة التعادل) وتم تدرّج وتقدير صعوبة المفردات مقياس القدرة على التفكير الناقد وفقاً لنموذج راش وذلك بعد توفر محكات الدقة وكذلك تقدير قدرة الأفراد المقابلة لكل درجة كلية خام محتملة على مقياس القدرة على التفكير الناقد وفقاً لنموذج راش بعد توفر محكات الدقة وتم التوصل إلى معايير الرتب المئينية والدرجات التائية التي تفسر على أساسها مستويات قدرة الأفراد على المقياس.

Abstract

This study aims to reveal some of effective factors such as: the size of used sample in grading, best way to join between different test images, grading one calibration gradation (common vocabularies/common individuals), best number of contributed vocabularies or contributed individuals (who used to join the different test images). In addition to the impact of used sample in grading vocabularies that effect on the accuracy of measures criterion. In order to reach to best gradation of Rasch model to measure criticizing intellectuality ability. It also aims to study the calibration performance (Percentile Ranks and T-scores) to interpret individuals capacity on calibration and using Winsteps program in grading the criticizing measurement of intellectual capacity according to Rasch model. The study was made on 829 female and male students (4th and 5th primary class students) and (1st and 2nd intermedi

The results showed that the best size between used sample sizes in grading measurement of criticizing intellectual capacity by using Rasch model is 800 according to measurement quality. The best number of contributed and used vocabularies in joining different test images is 11 according to the measurement accuracy, the number of contributed and used individuals in joining test images is 50 according to the measurement accuracy criterion, and the best way to join between test

images is joint vocabularies according to measurement accuracy criteria. In accordance to measurement accuracy criterion, the measurement accuracy was different according to sample difference for the sample benefit with low level. Measurement independence in measurement of capability was achieved (equivalence accuracy). The vocabularies difficulty and estimation were graded of capability capacity on intellectual criticizing according to Rasch model after the availability of accuracy criterion. Also, estimating all interviewed individual capability for each total raw score on capability measurement for criticizing intellectuality according to Rasch model after criterion accuracy existent. It was succeeded to reach Percentile Ranks and T-scores that explain individuals capability levels on measurement

الفصل الأول التعريف بالبحث

مشكلة البحث :

لقد قصرت أساليب القياس التقليدية السابقة عن تحقيق الدقة والموضوعية المرجوة في

القياس السلوكي، حيث وجهت لها العديد من الانتقادات أهمها:-

أ- اعتماد القياس على خصائص الاختبار المستخدم (الصياغة، مستوى الصعوبة

، الصدق، الثبات، غير ذلك) بمعنى اختلاف نتيجة القياس باختلاف الاختبار

المستخدم .

ب- اعتماد القياس في تقنيته على عينة معينة من الأفراد، بمعنى اختلاف نتيجة القياس

باختلاف مستوى جماعة التقنين وخصائصها.

ت- عدم خطية القياس، أي عدم وجود وحدة قياس متساوية على المتصل الذي

يتدرج عليه المتغير موضوع القياس. (كاظم، ١٩٩٥: ٢٥٢-٢٥٣)

وفي محاولة للتغلب على تلك المشكلات ومع بدء الدعوة لتحقيق الموضوعية في القياس السلوكي

كما هو الحال في القياس الفيزيائي، فقد ظهر اتجاه حديث في القياس النفسي والتربوي يتمثل في

نظرية السمات الكامنة latent traits theory والتي أطلق عليها فيما بعد نظرية الاستجابة للمفردة Item Response Theory ويعد نموذج راش من أشهر نماذج هذه النظرية حيث تندرج مفردات الاختبار وفقاً لهذا النموذج من حيث صعوبتها فقط بحيث تعرف المتغير موضوع القياس (أحادية البعد)، كما إن تقدير صعوبة المفردة أو قدرة الفرد لا تتأثر باختلاف عينة التدرج أو مجموعة المفردات المستخدمة في القياس (استقلالية القياس). ويكون التعبير عن كل من صعوبة المفردة وقدرة الفرد بوحدة قياس معرفة (خطية القياس)، مما يتيح الفرصة لدراسة نمو القدرة موضوع القياس، كما يتيح عقد المقارنات بين الشرائح المختلفة من المجتمع على القدرة وغير مما تتجه أدوات القياس السلوكي الشائعة (كاظم، ١٩٨٨: ١٩-٢٠).

وتركز مشكلة البحث بالأسئلة الآتية :

- ١- هل تختلف دقة القياس باختلاف حجم العينة المستخدمة في تدرج مقياس القدرة على التفكير الناقد باستخدام نموذج (راش)؟
- ٢- هل تختلف دقة القياس باختلاف عدد المفردات المشتركة في ربط الصور الاختبارية المختلفة على ميزان تدرج واحد مشترك باستخدام نموذج (راش)؟
- ٣- هل يختلف شكل توزيعات القدرة للأفراد باختلاف نموذج التحليل المستخدم ؟
- ٤- هل تختلف دقة القياس باختلاف عدد الأفراد المشتركين في ربط الصور الاختبارية المختلفة وتدرجها على ميزان تدرج واحد مشترك باستخدام نموذج (راش)؟
- ٥- هل تختلف دقة القياس باختلاف طريقة الربط وهل يؤثر الاختلاف في مستوى العينة ؟
- ٦- هل تحققت في مقياس القدرة على التفكير الناقد استقلالية القياس أي دقة التعادل؟
- ٧- ما تدرج وتقدير صعوبات مفردات مقياس القدرة على التفكير الناقد وفقاً لنموذج راش بعد توفر محكات الدقة ؟ وما هي معايير الرتب المئينية والدرجات التائية التي تفسر أساسها مستويات القدرة؟

أهمية البحث :

يعد تحقيق الدقة الموضوعية من الأهداف الهامة في القياس النفسي والتربوي، وقد استحوذت دقة القياس وموضوعيته في مجال الظواهر السلوكية على أذهان العلماء والمفكرين على غرار ما يتمتع به القياس في مجال العلوم الطبيعية (الفيزيائية)، وكان الهدف دائماً هو مواكبة التقدم الذي حصل في مجال هذه الظواهر الأخيرة وما يستتبع ذلك من دقة في الفهم والتنبؤ وبالتالي إمكانية الضبط والتحكم .

وتعتبر نظرية الاستجابة المفردة احد الاتجاهات الحديثة للقياس ولقد حققت النماذج الرياضية اللوغاريتمية المرتبطة بها طفرة هائلة ومتطورة في مجال القياس النفسي والتربوي والسلوكي، ولقد أكدت العديد من الدراسات الأجنبية والعربية تفوق هذه النظرية في تحقيق الموضوعية المأمولة منها في هذا المجال إذا ما قورنت بالنظرية الكلاسيكية كما إن هذه النظرية تعتمد على افتراضات وشروط اقوي من الافتراضات والشروط ذات الصلة بالنظرية الكلاسيكية التي هيمنت على مجريات وإجراءات القياس السلوكي والتحليلات الخاصة بأدوات القياس لفترة زمنية .

(Hulin,1987:131) ونتيجة لما حققته هذه النظرية وكذلك النماذج الرياضية المرتبطة بها في تحقيق الموضوعية في مجال قياس السلوك الإنساني، وفي بناء وإعداد اختبارات حققت تطوراً وتقدماً هائلاً في شتى ميادين القياس فقد اتجهت العديد من الدراسات الأجنبية والعربية إلى بناء وتطوير اختبارات جديدة أو إعادة تطوير بعض الاختبارات العالمية المشهورة وشائعة الاستخدام في ضوء النظرية الحديثة للقياس والتحليلات الإحصائية المتعلقة بأحد النماذج الرياضية المرتبطة بها بهدف تحقيق المزيد من الموضوعية Feingold,1994:34

وتعد مقاييس التفكير الناقد من أهم الأدوات لقياس القدرات العقلية بما تتضمنه من عمليات متعددة كالاستدلال والتفسير وغيرها من العمليات التي تعتبر مكونات أساسية للتفكير الناقد، لذا فقد كان من المهم استخدام احد النماذج الهامة لنظرية الاستجابة للمفردة مثل نموذج راش في تدريج مثل هذه المقاييس وتطويرها، حيث تتيح خاصية استقلالية القياس- التي يتميز بها نموذج

راش من إعداد صور اختباريه مختلفة متعادلة القياس حتى وان اختلفت من حيث المستوى في تعريفها للمتغير (القدرة) موضوع القياس (عبدالله، ٢٠٠٣ : ٥).

الأهمية النظرية

١. قد تساهم هذه الدراسة في مزيد من الفهم والإيضاح لما قد يحققه نموذج راش من موضوعية القياس .
٢. قد تشجع هذه الدراسة المهتمين بالقياس النفسي على استخدام الأساليب الحديثة للقياس في بناء وتطوير اختبارات نفسية أخرى .
٣. تضيف هذه الدراسة إلى مركز المقاييس النفسية مقياسا مدرجا بوحدة الصعوبة والقدرة ومزوداً بمعايير الرتب المئينية والدرجات التائية .

الأهمية التطبيقية:

١. يتيح نموذج (راش) المستخدم في هذه الدراسة لتدريج مقياس على التفكير الناقد، الفرصة لتقسيم المقياس إلى صورة اختبارية فرعية متعادلة القياس ،بمعنى انه يمكن التوصل إلى تقديرات لقدرة الفرد على مقياس التفكير الناقد لا تختلف باستخدام أي من الاختبارات الفرعية أو المقياس الكلي الذي يضمها،ويمكن الاكتفاء باستخدام الاختبار الفرعي بدلا من المقياس الكلي وبهذا توفير للوقت والجهد والمال.
٢. إن التدريج واسع المدى الذي يوفره نموذج راش لقياس القدرة على التفكير الناقد يتيح الفرصة لقياس النمو أو التغيير الحادث في هذه القدرة لدى الفرد أو الأفراد على مدى زمني معين وفي مواقف تجريبية مختلفة.
٣. إن استخدام نموذج راش في تدريج مقياس التفكير الناقد يوفر متطلبات الموضوعية في قياس القدرة موضوع القياس ويعني هذه توفر شرطي الصدق والثبات لتقديرات كل من صعوبة المفردات وقدرات الأفراد إي تحقق صدق وثبات القياس.

أهداف الدراسة

- ١- اختيار انسب حجم للعينة المستخدمة في التدرج وذلك بناءً على محكات الدقة المتمثلة في:-
 - الخطأ المعياري لتقدير الصعوبة .
 - الخطأ المعياري لتقدير القدرة.
 - معامل الثبات لتقدير صعوبات المفردات .
 - معامل الثبات لتقدير قدرة الأفراد
 - دالة المعلومات الاختبار .
- ٢- اختيار أفضل عدد للمفردات الرابطة بناءً على محكات الدقة السابقة .
- ٣- اختبار أفضل عدد للأفراد المشتركين بناءً على محكات الدقة السابقة .
- ٤- اختيار أفضل طريقة للربط (مفردات مشتركة/ أفراد مشتركين).
- ٥- تحديد اثر الاختلاف في مستوى العينة (مرتفعة/متوسطة/منخفضة) على دقة القياس.
- ٦- تحديد مدى تحقق استقلالية القياس في مقياس القدرة على التفكير الناقد.
- ٧- تدرج وتقدير صعوبة مفردات مقياس القدرة على التفكير الناقد وفقاً لنموذج راش بعد توفر محكات الدقة .
- ٨- تقدير قدرة الأفراد المقابلة لكل درجة كلية خام محتملة على مقياس القدرة على التفكير الناقد وفقاً لنموذج راش بعد توافر محكات الدقة.
- ٩- عمل معايير الرتب المئينية والدرجات التائية التي تفسر على أساسها مستويات قدرة الأفراد.

حدود البحث :

بالرغم من بناء المقاييس النفسية وتدرجها باستخدام نموذج راش يكون مستقلاً عن توزيع العينة وخصائصها (Sample Free) إلا إن المعايير المستخرجة التي تفسر بها المعنى السيكولوجي للدرجات تحدد بعينة الدراسة وكما يأتي:-

- اقتصر البحث على بيانات مقياس التفكير الناقد.
- اقتصر البحث تطبيقه على تلاميذ المرحلة الابتدائية (رابع -خامس) وطلبة المرحلتين (الأول -الثاني) المتوسط من طلبة المدارس الابتدائية و المتوسطة في محافظة واسط .
- اقتصر البحث على تطبيق الاداة الخاصة به في العام الدراسي ٢٠١٨ - ٢٠١٩ م.

تحديد المصطلحات :

أولاً: نظرية الاستجابة للمفردة

نظرية حديثة في القياس النفسي والتربوي يتم فيها تحديد العلاقة بين أداء المفحوص والسمة الكامنة موضع القياس وفق دالة رياضية محددة وتعتمد هذه النظرية على عدد من النماذج تسمى نماذج السمات الكامنة ،التي يتم من خلالها الربط بين الأداء على الفقرة وقدرة المفحوص. (Hambelton&Swaminathan,1985:32)

ثانياً: التفكير الناقد :

بأنه نمط من التفكير قوامه فحص الحلول المعطاة لمشكلة ما لبيان ما إذا كانت هذه الحلول مقبولة ومنطقية ومتسقة مع المعطيات المعروفة في ميدان المشكلة (الوقفى، ٢٠٠٣: ٥٠٧).

الفصل الثاني

الإطار النظري والدراسات السابقة

أولاً: مفهوم القدرة على التفكير الناقد:

مفهوم مركب مرتبط بعدد غير محدد من السلوكيات في عدد غير محدود من المواقف أو الموضوعات وله ارتباط بمفاهيم أخرى كالمنطق وحسب المشكلة والتعلم ونظرية المعرفة ويرى جون ديوي (Dewey) بأنه التريث في إعطاء الأحكام لحين التحقق من أمرها ويتكون من عناصر تشمل على صياغة تعميمات بحذر والنظر الحاذق في الاحتمالات والبدائل وتعليق الحكم لحين توافر معلومات وأدلة كافية ولقد استخدم سقراط هذا الفكر عن طريق توليد أفكار عند محاوريه ومن ثم توجيه النقد لها عن طريق الأسئلة السابرة لتغيير وجهات نظرهم ومع انه كان يعتمد على المجادلة والحكم مع أهميتها إلا أنها ليست كافية لافتقارها لعناصر من مثل التفكير المنتج ، ويرى جود (Good) بأنه يقوم على أساس التقويم الدقيق للمعلومات بهدف الوصول إلى أفضل النتائج، ويرى واطسن بأنه محاولة مستمرة لاختيار الحقائق والآراء في ضوء الأدلة التي تقدمها وتدعمها بدلاً من القفز المتسرع للنتائج، إما فيشر يرى بأنه القدرة على قياس المعلومات والأفكار وتقويم المناقشات بهدف الوصول إلى أحكام متوازنة. (عبد العزيز ٢٠٠٩: ١٠٧-١٠٩) وقد ورد مفهوم التفكير الناقد في المصادر التربوية بتفسيرات متعددة منها حيث يبين انه إصدار الأحكام أو المقارنة بين موضوعين أو أكثر للمفاضلة بينهما وأساس التفكير الناقد تطبيق المعرفة في موقف معين وزمن محدد. (نوفل وأبو عواد، ٢٠١١: ٢٩).

أ- أنماط التفكير الناقد

هي نشاطات معرفية تشير إلى وجود عمليات عقلية داخلية عند الطالب ولها عدة مستويات :-

- ١- المستوى الحسي :وهو إدراك الأشياء المحسوسة .
- ٢- المستوى التطوري :وهو إدراك الأشياء بالصور الذهنية .
- ٣- المستوى المجرد:وهو معرفة معاني الأشياء وترميزها.

٤- المستوى القاعدي :وهو الذي يستند إلى قواعد وأسس ومبادئ وفيه يقوم الطالب العلاقات بين الأشياء .(عبد الهادي وعياد، ٢٠٠٩ : ٩٩-١٠٠).

ب-مهارات التفكير الناقد :

عند اطلاع الباحث على أدبيات وبحوث التفكير الناقد وجد هناك العديد من التصنيفات لمهارات التفكير الناقد تبعا لتعدد تعريفاته والأطر النظرية المفسرة له ومن هذه التصنيفات هو تصنيف أوول ودانيالز (Udall&Daniels,1991) لمهارات التفكير الناقد فهي :-

١- مهارات تفكير الاستقرائي.

٢- مهارات التفكير الاستنباطي.

٣- مهارات التفكير التقييمي. (الخرجي، ٢٠١١: ٨٩)

ثانيا :نظريات القياس

أ- قياس القدرة على التفكير الناقد في إطار نظريات القياس

استخدم الباحثون في مجال قدرات التفكير الناقد مقاييس كثيرة لقياس القدرة على التفكير الناقد وتنوعت تلك المقاييس بتنوع المهارات المتضمنة فيه، ويتركز الاهتمام في الدراسة الحالية على قياس قدرة التفكير في إطار نظريات القياس ،ولذا يستعرض الباحث فيما يلي أساليب القياس التقليدية والموضوعية الحديثة المستخدمة في نتائج الاختبارات النفسية والعقلية .

ب- أساليب القياس السلوكي المستخدمة في تحليل نتائج المقاييس والاختبارات

ويمكن تصنيف أساليب القياس المستخدمة في تحليل بيانات الاختبارات والمقاييس النفسية والتربوية إلى:

١- أساليب القياس التقليدية التي تعتمد على معيار واحد معين في تفسير درجة

الفرد على الاختبار والتعبير عنها(معيار الجماعة /معيار المحك/ أو مستوى

(الإتقان)

٢- أساليب القياس الموضوعية التي تعتمد على ما يسمى بنظرية الاستجابة للمفردة

Item Response theory (IRT)(كاظم ١٩٩٥ : ٢٤٩).

أولاً: أساليب القياس التقليدية التي تعتمد على معيار واحد في تفسير الدرجة :-

وتشمل هذه الأساليب على نوعين من أنواع القياس هما

أ- القياس جماعي المرجع Norm Referenced measurement

نشأ هذا الأسلوب في القياس في إطار القياس النفسي، وكان ذلك بهدف التمييز بين المستويات العقلية للأفراد. وعلى هذا يقوم القياس جماعي المرجع على تقدير الفروق بين الأفراد، والمقارنة بينهم. لذا تعتبر أفضل مفردات الاختبار هي تلك التي تميز بين الأفراد بشكل أفضل وذلك عندما يتساوى تقريباً -معامل سهولة وصعوبة المفردة. ويمثل صدق وثبات الاختبار في صدقه وثباته في قياس الفروق بين الأفراد (كاظم، ١٩٩٦: ٢٨٥-٢٨٧).

وتعتبر الدرجة الخام الكلية أساس للقياس. حيث يعبر عن قدرة الفرد بالدرجة الخام الكلية التي يحصل عليها من أدائه على الاختبار (Wright & Masters, 1982: 33). بينما تقدر صعوبة المفردة من عدد الاستجابات الصواب على تلك المفردة، ويعتمد تفسير درجة الفرد على الاختبار على مقارنتها بمعايير مشتقة من أداء جماعة التقنين التي ينتمي إليها الفرد على هذا الاختبار مثل الدرجات المعيارية والرتب المئينية (أبو حطب وآخرون، ١٩٩٣: ٤١٩).

القياس محكي المرجع

نشأ هذا الأسلوب في القياس في إطار القياس التربوي، وذلك عندما تزايد الاهتمام بالتعليم من أجل الإتقان، وبالتالي برزت الحاجة إلى مقارنة أداء الفرد بما ينبغي تحقيقه، خاصة وان الاعتماد على الجماعة كمرجع وحيد لتفسير الدرجة قد أدى إلى انخفاض مستوى التعليم (فام، ١٩٩٧: ٢٢). ويتمثل الهدف من القياس محكي المرجع في قياس مدى الاكتساب، ووصف أداء الفرد بغض النظر عن مكانته بين إقرانه. لذا يعتبر أفضل مفردات الاختبار هي التي تكون بالغة الصعوبة قبل البدء في البرنامج التعليمي، ثم تصبح بالغة السهولة بعد انتهائه (عبد السلام، ١٩٩٦: ٦٨). ولقد أشار علماء القياس إلى هذا النوع من القياس بسميات عديدة، مثل الاختبار هدي المرجع أو الاختبار المرجع إلى مجال، أو اختبارات الإتقان، وتختلف هذه الاختبارات فيما بينها بدرجة تحديد النطاق أو المجال السلوكي الذي يستند إليه بناء كل منها (علام، ١٩٩٥: ٢٤).

ولقد استندت معظم المقاييس في مجال علم النفس على المنحنى جماعي المرجع الذي يقوم على التمييز بين الأفراد في الأداء وقد واجهت تلك المقاييس كثيرا من أوجه النقد والتي من أهمها:-

- ١- عدم استقامة هذا الأسلوب من أساليب القياس مع متطلبات الموضوعية ويتمثل ذلك في :-
 - أ- اعتماد القياس في ظل هذا الأسلوب على خصائص الاختبار المستخدم (الصياغة ومستوى الصعوبة، والصدق والثبات، وغير ذلك).
 - ب- اعتماد القياس على معيار الجماعة في تفسير الدرجة، بمعنى ذلك اختلاف نتيجة القياس باختلاف مستوى الجماعة وخصائصها .
- ٢- يقتصر هذا الأسلوب على تحقيق هدف واحد فقط من أهداف القياس، هو التمييز بين الأفراد مع إغفال الأهداف الأخرى مثل مقارنة أداء الفرد الحالي بأدائه السابق، أو تقدير مستوى أداء الفرد بالنسبة لمستوى جماعة أخرى لا ينتمي إليها، أو تقدير مستوى أداء الفرد بالنسبة لأداء أي فرد من الجماعة التي ينتمي إليها أو بالنسبة لأداء أي فرد ينتمي لأي جماعة أخرى... الخ.
- ٣- عدم توفر خطية القياس، بمعنى عدم وجود وحدة قياس متساوية على مدى المتصل الذي يتدرج عليه المتغير موضوع القياس، مما يترتب عليه صعوبة المقارنة بين الأفراد وكذا صعوبة قياس التغير في السمة المقاسة .
- ٤- يعمل هذا الأسلوب على مستوى الاختبار وليس على مستوى المفردة، فهو لا يهتم بكيفية استجابة المفحوصين على مفردة ما بعينها؛ ومن ثم يتعذر التنبؤ بكيفية أداء الفرد على مفردة معينة وبذلك لا تقدم المقاييس التقليدية تفسيراً لنمط في أداء الفرد على مفردات الاختبار .
- ٥- أفضل المفردات من وجهة نظر المقاييس جماعية المرجع هي المفردات الأقدر على التمييز (المفردات متوسطة الصعوبة)، وتستبعد المفردات مرتفعة أو منخفضة الصعوبة عند إجراء عند أعداد هذا النوع من المقاييس حتى وإن كانت تقيس أهم الأهداف.

٦-يفترض هذا الأسلوب ثبات خطأ القياس عبر المفحوصين الذي يطبق عليهم الاختبار، وهذا الافتراض لا يمكن إن يتحقق لان درجات المفحوصين ذوي القدرة المختلفة على نفس الاختبار تحتوي على مقادير مختلفة من خطأ القياس.

٧-عدم إمكانية إضافة بنود جديدة إلى الاختبارات المعدة لأساليب القياس التقليدية إذا دعت الحاجة إلى ذلك.

٨-إضافة إلى أوجه القصور السابقة فان أساليب القياس التقليدية بوجه عام لم تقدم حلولاً شافية للكثير من المشاكل الاختبارية مثل تحديد البنود المتحيزة،الاختبارات الموائمة،معادلة درجات الأفراد على الاختبارات المختلفة (عبد الحجازين،٢٠١٠: ٨٧).

وانطلاقاً مما سبق تتضح الحاجة الملحة إلى قياس التفكير الناقد بشكل أكثر موضوعية ويقتضي تحقيق الموضوعية في القياس ما يلي :

١- عدم تأثير القياس باختلاف الأداة المستخدمة في القياس (خصائص الاختبار) طالما أنها أداة مناسبة .

٢- عدم تأثير القياس باختلاف العناصر (الأفراد)التي استخدمت أداة القياس في تقديرها.

٣- تدرج أداء القياس بوحدة مطلقة ثابتة تتوافق مع تدرج مستويات المتغير موضوع القياس وجميع هذه المتطلبات توفرها نظرية الاستجابة للمفردة Item Response Theory .

ولقد صنفت نماذج الاستجابة للمفردة إلى نوعين من النماذج إما النوع الأول فهي النماذج التجميعية والتي اعتبرت الأساس النظري لنماذج النوع الثاني من النماذج اللوغاريتمية ،وتعتبر النماذج اللوغاريتمية هي النماذج التطبيقية لنماذج الاستجابة للمفردة (Suen,1990 :92-90pp).

وتصنف النماذج اللوغاريتمية إلى عدة نماذج وذلك تبعاً لعدد المعالم (البارامترات)التي تصف البند،وهي على الشكل التالي:-

١- النماذج اللوغاريتمية أحادي المعلم (نموذج راش):-حيث تتدرج المفردات تبعاً لصعوبتها، ويفترض هذا النموذج انعدام التخمين وتساوي معاملات تمييز جميع الفقرات

- ٢- النموذج اللوغاريتمي ثنائي المعلم:- حيث تدرج المفردات تبعا لصعوبة المفردات وتمييزها ويفترض هذا النموذج انعدام التخمين.
- ٣- النموذج ثلاثي المعلم (نموذج برنبوم) حيث يأخذ هذا النموذج في الاعتبار كل معالم المفردة الثلاث (الصعوبة، التخمين، التمييز) ويعتبر نموذج راش حالة خاصة من النموذجين الثنائي والثلاثي المعلم. (علام، ٢٠٠٥: ٦٩-٧٢)
- وهناك أيضا نماذج أخرى لوغاريتمية مثل النموذج اللوغاريتمي رباعي البارامتر ونموذج الاستجابة المتدرجة ونموذج الاستجابة المتصلة ونموذج التقدير الجزئي (عبد الحجازين، ٢٠١٠: ١٠٥).

ثالثا: نموذج راش

وهو النموذج الأبسط من بين نماذج منحنى خصائص الفقرة ويعرف بنموذج راش (Georg Rasch) جورج راش عالم الرياضيات الدنماركي حيث كان أول من قام بنشره وذلك اقترن هذا النموذج باسمه منذ الستينات في القرن العشرين، إذا استخدم نظرية الاحتمالات في تحليل البيانات، وكان هذا الانجاز مغايراً لما كان مؤلفاً آنذاك، حيث إن النموذج الناتج لمنحنى خصائص الفقرة هو نموذج لوغاريتمي، وحسب هذا النموذج سيتم تثبيت قيمة معلمة التمييز (ت) على افتراض أنها متساوية لجميع الفقرات وقيمتها العددية تساوي (١) إما المعلمة التي سوف تأخذ قيمة مختلفة فهي معلمة الصعوبة (ص) ولذلك سمي هذا النموذج بالنموذج أحادي المعلمة . ويمكن التعبير عنه رياضياً كالتالي :-

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

حيث :-

$p_i(\theta)$: هو احتمال الإجابة الصحيحة للفرد الذي قدرته θ على المفردة (i)

b_i : مستوى صعوبة المفردة i

θ : مستوى قدرة الفرد

e: الأساس اللوغاريتمي الطبيعي ويساوي ٢,٧١٨

D: عامل التدرج ويساوي ١,٧ وهو القيمة التي تجعل شكل المنحني اللوغاريتمي يقترب من شكل اقتران المنحني الطبيعي . (محاسنة، ٢٠١٣: ١٨٢).

شروط نموذج راش

يعد نموذج راش أهم وأبسط نماذج السمات الكامنة حيث يمكن إن تتوفر متطلبات الموضوعية عندما تستوفي فروض النماذج وهي :-

١- **أحادية البعد:** حيث تعرف السمة موضوع القياس بواسطة مجموعة من البنود ذات صعوبة أحادية البعد، أي أن بنود الاختبار لا تختلف فيما بينها إلا من حيث مستوى صعوبة فقط، كما يكون الأفراد ذوو قدرة أحادية البعد تحدد وحدها مستوى أدائهم على الاختبار.

٢- **استقلالية القياس:** ويعني ذلك إن لا يعتمد تقدير صعوبة المفردة على تقديرات صعوبة البنود الأخرى المكونة للاختبار، ولا على تقديرات قدرة الأفراد الذين يجيبون عليها، ولا يعتمد تقدير قدرة الفرد على تقديرات قدرة الفرد على تقديرات قدرة إي مجموعة من الأفراد الذين يؤدون الاختبار، أو على تقديرات صعوبة البنود التي يؤدنها وتحقق استقلالية القياس هذه عندما تحقق:-

أ- ملائمة البنود المكونة للاختبار للنموذج، وهذه يعتمد على صدق البنود في قياس موضوع القياس وعلى صدق تدرجها على مقياس هذه القدرة .

ب- ملائمة الأفراد الذين يجيبون على بنود الاختبار للنموذج وهذا يعتمد على صدق استجابات هؤلاء الأفراد .

ت- توافق بنود الاختبار مع قدرات الأفراد ويعتمد هذا على تقارب مستوياتها

٣- **توازي المنحنيات المميزة للبنود:** يوضح المنحني المميز احتمالات الاستجابة الصحيحة على هذه المفردة للأفراد عند المستويات المختلفة من القدرة، وبمعنى توازي المنحنيات المميزة للبنود هو أنه إلى الحد الذي يميز فيه البنود بين الأفراد ذوي المستويات المختلفة من قدرة ما فإن جميع هذه البنود ينبغي أن يكون لها نفس القدرة على التمييز .

٤- خطية القياس: أي إن يكون هناك معدل ثابت لتدريج القياس باستخدام وحدة قياس محددة حيث يكون عند أي مستوى من مستويات المتغير يكون تقدير الفرق بين أي قياسين متتالين في هذا التدريج ثابتاً. (منير، ٢٠٠٠: ٦٤-٦٥).

رابعاً: تعادل القياس

إن تعادل تقدير قدرة الفرد عبر اختبارات مختلفة بقياس نفس القدرة يمثل قضية ذات أهمية كبيرة للقائمين على تطوير الاختبارات والمتخصصين في القياس وكذلك المتقدمين للاختبار، فعند مقارنة درجات فرد على اختبارين مختلفين، فإنه لا بد من القيام بعملية التعادل لتلك الدرجات على الاختبارين والتي من خلالها يتم تطابق درجات الاختبار الأول مع درجات الاختبار الثاني وذلك بتوحيد صفر التدريج لكل من الاختبارين .

(Hambleton and Others, 1991: 123)

أنواع التعادل

هناك نوعان من التعادل هما، التعادل الأفقي والتعادل الرأسي ويشير التعادل الأفقي إلى تعادل اختبارين أو أكثر يقيسان نفس السمة في نفس المستوى من الصعوبة. أما التعادل الرأسي فيشير إلى تعادل اختبارين أو أكثر يقيسان نفس السمة في مستويات مختلفة من الصعوبة (Suen, 1990: 199).

التعادل باستخدام نظرية الاستجابة للمفردة

وتتقسم طرق التعادل إلى قسمين الطرق التقليدية والمتضمنة بتساوي المئينات (تساوي الرتب المئينية وطرق التعادل الخطية أما التعادل باستخدام نظرية الاستجابة للمفردة حيث بينت الدراسات أنه يكون أكثر دقة من الطريقة التقليدية، ويستطيع التعادل القائم على نظرية الاستجابة للمفردة من التغلب على كثير من مشكلات التعادل التقليدية وذلك إذا كان نموذج الاستجابة للمفردة الذي تم اختياره ملائماً للبيانات . (Baker F ..2001. P154)

بعض المفاهيم والموضوعات التي ترتبط ارتباطا وثيقا بنموذج راش

وفيما يلي بعض المفاهيم والموضوعات التي ترتبط ارتباطا وثيقا بالنموذج والتي يرى الباحث ضرورة ذكرها حتى تتضح الصورة :

• أساليب التأكد من ملائمة البيانات للنموذج المستخدم **Finess of IRT Models**

وتعني ملائمة المفردات لتوقعات النموذج، أي التحقق من تطابق خصائص العينات مع فروض النموذج، أي تحقيق قيم بارامترات النموذج التي تم تقديرها من خلال النتائج المستخلصة من الاختبار لهذا الفروض . (مسعود، ٢٠٠٤: ٨٥) وحقيقة الأمر ان هناك عدة طرق لقياس مدى ملائمة الأفراد أو المفردات للنموذج سوف يذكر الباحث طريقتين منها هما الأكثر شيوعا واستخداما:

(أ) **طريقة كا ٢ Chi-Square**: وتستخدم هذه الطريقة لحساب ملائمة الفرد للنموذج

في كثير من البرامج الإحصائية مثل Bilog-MG ٢٠١٠ Bilog-3 RUMM إلا انه يجب الحذر عند استخدامها حيث ان أسلوب كا ٢ الإحصائي له بعض المحاذير، حيث انه شديد الحساسية لحجم العينة، وهذا ينصح خبراء القياس بعدم الاعتماد عليه كمحك وحيد لحذف المفردات، بل يجب ان يكون للباحث نظرة متحفظة بالإضافة إلى نتائج كا ٢ .
(Embretson&Reise ,2000 p230-238)

(ب) **حساب الملائمة عن طريق مقياسي الملائمة التقاربي Infit والتبايدي Outfit**: - ويستخدم هذان الأسلوبان في برامج Microscale , winsteps وتحدد جودة المفردات بدرجة ملائمتها للنموذج، وذلك من خلال نوعين من احصاءات الملائمة أولها هو إحصاء الملائمة التقاربي وثانيها إحصاء الملائمة التبايدي . وتعتمد الملائمة التقاربية على أداء الأفراد الذين تقترب قدراتهم من صعوبة مفردات القياس بينما تعتمد الملائمة التبايدي على أداء الفرد الذين تبتعد قدراتهم (أي تزيد كثيرا أو تقل كثيرا) عن صعوبة المفردة (مسعود، ٢٠٠٤: ٨٧).

دالة المعلومات Information Function: - يمثل دالة المعلومات في نظرية الاستجابة للمفردة احد المفاهيم الأساسية للنظرية الحديثة في القياس، حيث أنها تساعد المهتمين في بالقياس التربوي

في تحديد قدر المعلومات التي تقدمها مفردة ما عن فرد ما وذلك من خلال تحديد أقصى ارتفاع للمنحني الممثل لدالة المعلومات لتلك المفردة على متصل القدرة، وبهذا يمكننا ان نحدد أي المفردات الاختبارية تقيس المتغير المراد قياسه (السمة) بدرجة أفضل عند مستويات محددة للقدرة، ويرجع الفضل في المعنى الإحصائي لدالة المعلومات الى فيشر الذي عرفها على أنها مقلوب قياس دقة تقدير البارامترات وقياس الدقة يعني التباين بين عناصر التقدير ويقصد به مربع الانحراف المعياري (Suen, 1990: 199).

• **الخطأ المعياري للقياس:** معامل الثبات خاصيتان تكملان بعضهما البعض في التعبير عن ثبات المقياس والدقة التي يمكن بها تحديد الدرجة الحقيقية للفرد والتي ترجع إلى حجم الخطأ المعياري الذي يتأثر بمعامل ثبات المقياس، وعندما نبحث عن اختبار ما لاستخدامه في جمع البيانات يجب ان يكون معامل الثبات لهذا الاختبار مرتفعاً مع انخفاض الخطأ المعياري له، ويستدل على الخطأ المعياري لتقدير القدرة عندما يحصل على تقديرات الاحتمال الأقصى، وكذلك يحصل على الخطأ المعياري لتقدير صعوبة كل مفردة وكذلك لمفردات الاختبار ككل، ويعتبر الثبات والخطأ المعياري من المفاهيم الشائعة والمعروفة لدى الأغلبية وانه كلما ارتفع الثبات وقل الخطأ المعياري كلما عظمت دقة القياس (Wright B.D, 1990. P293).

الدراسات السابقة

(أ)دراسة واجنز وهارفي Wagner T.and Harvey R.J (٢٠٠٣)

وهي بعنوان: تطوير اختبار جديد للتفكير الناقد باستخدام نظرية الاستجابة للمفردة

هدفت الدراسة إلى تطوير اختبار واجنز (The Wagner assessment test (WAT) الشبيه باختبار واطسن للتفكير الناقد وذلك باستخدام نظرية الاستجابة للمفردة التي تسمح برؤية أكثر تفصيلاً لمفردات الاختبار، وقد افترضت هذه الدراسة ان زيادة عدد البدائل المفردات ذات الاختبار من متعدد سوف يخفض التخمين وكذلك سوف يحسن دالة المعلومات والخطأ المعياري ومؤشرات الثبات، اما بالنسبة لعينة الدراسة فقد بلغت ٤٠٧ طالبا وطالبة من طلبة الجامعة، حيث طبق الاختباران

(WAT) و(WGCTA) باستخدام البرنامج الحاسب الآلي Bilog وبعد ذلك تم تشكيل اختبار فرعي مشتق من الاختبار WAT تكون من ٨٠ مفردة متدرجة الصعوبة ومماثل لطول الاختبار WGCTA الكلي. وقد أظهرت الدراسة اختبار للتفكير الناقد WAT الأصلي (١١١ مفردة) والمعدل (٨٠ مفردة) دالة معلومات اكبر ومؤشرات ثبات اعلي وأخطاء معيارية اقل مما أنتجها الاختبار WGCTA (عبدالله، ٢٠٠٣: ٩٦-٩٧)

(ب)دراسة سيكومترية لتطوير اختبار رسم الرجل باستخدام نموذج راش

هدفت الدراسة إلى استخدام نموذج راش في إعادة تدرج مفردات اختبار رسم الرجل لجودانف -هريس تبعاً لمستوى صعوبتها مع ما قد يستدعيه ذلك من حذف بعض المفردات أو تعديلها، استخدمت الدراسة برنامج الحاسب الآلي Winsteps الإصدار ٢,٨٨ في إجراء المعالجات الإحصائية لبيانات أداء أفراد العينة على اختبار رسم الرجل وفقاً لنموذج راش وأهم ما توصلت إليه الدراسة هو تدرج اختبار رسم رجل بعد تعديل بعض المفردات وحذف بعض المفردات و تحقيق شرط استقلالية القياس في المقياس في صورته الجديدة، وتحقق صدق وثبات القياس في صورته الجديدة بعد التدرج باستخدام نموذج راش وحساب معايير الاختبار بعد التدرج باستخدام نموذج راش عن طريق إيجاد رتب مئينية ودرجات تائية ونسب ذكاء انحرافية، وحساب معدل نمو قدرة الأفراد تبعاً للعمر على اختبار رسم الرجل، وتمثيل منحنيات النمو التي توضح العلاقة بين مستويات العمر المختلفة وتقديرات القدرة، وأخيراً تكوين صورة مختصرة من الاختبار متحررة من اثر اختلاف الزي بين الثقافتين الأوروبية والعربية (مسعود، ٢٠٠٤: ٢٧١-٢٨٢)

الفصل الثالث

منهج البحث وإجراءاته

مجتمع البحث :

اعتمد البحث الحالي في إجراءاته على المنهج الوصفي (التحليل -المقارن) وتألف مجتمع البحث الحالي من طلبة المرحلة الابتدائية و المتوسطة لمحافظة واسط للعام الدراسي (٢٠١٨/٢٠١٩) حيث بلغ المجموع الكلي للطلبة (١٥٥٩٣).

عينة البحث :

تكونت عينة الدراسة من (٨٢٩) فرداً من طلبة وطالبات المدارس الابتدائية (الرابع،الخامس) والمتوسطة (الأول المتوسط) في محافظة واسط.

أداة البحث :

فهي عبارة عن مقياس في التفكير الناقد قام بصياغته مفرداته فريق من الباحثين بالمركز القومي للامتحانات والتقويم عام ١٩٩٧ وتضمن المقياس الأبعاد التالية :

- إدراك التناقض الداخلي .
- الاستدلال والاستنتاج من المعطيات.
- تقويم الأفكار في مقابل المحكات الخارجية والخبرة الشخصية .
- تقويم الحجج .

وتم اختيار البعد الثاني،حيث تضمن ٥٥ مفردة متدرجة الصعوبة،وقد تم تقسيم المقياس إلى ثلاثة صور مكونة من (٢٦،٢٦،٢٥) مفردة على الترتيب،وهناك ١١ مفردة رابطة بين كل صورتين وفي هذه الدراسة أعيد تدريج المقياس لخدمة الأغراض الأساسية لهذه الدراسة .

الإجراءات

إجراءات جمع البيانات

- التجريب الاستطلاعي.
- تطبيق الاختبار تبعاً للتعليمات.
- تصحيح البيانات وإدخالها على الحاسب .

إجراءات تدريج المقياس

في تدريج الاختبار فإن البيانات التي نحصل عليها نتيجة لعمليات التدريج ستختلف باختلاف عدد المفحوصين الذين يجيبون أو لا يجيبون على أي من الفقرات الاختبار، وفي كل عملية تدريج فإن المفحوصين الذين يحصلون على نفس العلامة الخام سوف يحصلون على نفس تقدير القدرة في ظل نموذج راش لكن ذلك لا يحدث في ظل النموذجين ثنائي وثلاثي الملمات. (محاسنة، ٢٠١٣: ٢٨٣).

١- تدريج الاختبار باستخدام الحاسب الآلي: لتدريج مقياس التفكير الناقد تم اختيار برنامج الحاسب الآلي Winsteps Version 3.67 وهو برنامج لتحليل البنود بواسطة نموذج راش الأحادي المعلم الذي يتدرج تبعاً لمعلم واحد هو الصعوبة، وتم إجراء التحليلات باستخدام البرنامج التحليل الإحصائي Winsteps Version 3.67 عن طريق حذف الدرجة التامة والصفرية قبل إجراء عملية التحليل بمعنى حذف كل فرد حصل على الدرجة النهائية والدرجة صفر حيث يعتبر تقدير قدرة الفرد خارج نطاق القياس، وحذف البنود التي اجمع أفراد العينة على إجاباتها صواباً أو خطأ حيث تعتبر هذه البنود غير ملائمة أي غير مناسبة لاستجابة الفرد فهي لا تستطيع إن تميز بين مستوياتهم على متصل القدرة، وحذف الأفراد غير الملائمين، وأعيد التحليل للمرة الثانية بهدف حذف الأفراد غير الملائمين تبعاً لمحكات الملائمة:

- استبعدت الحالات غير الملائمة من الأفراد والتي تكون قيم إحصاءات الملائمة لهم أقل من القيمة (٢-) حيث يعني ذلك تشابه إجابات الأفراد، مما قد يشير إلى عدم صدق الاستجابات.

- استبعدت الحالات غير الملائمة من الأفراد والتي تكون قيم إحصاءات الملائمة لهم اكبر من القيمة (+٢) حيث يعني هذا انه هؤلاء الأفراد يجيبون صواباً على البنود التي تزيد في مستوى صعوبتها عن مستوى قدرتهم إن يفشلوا في الإجابة الصواب على بنود أسهل من قدرتهم، ويتفق ذلك مع ما اقترحه رايت ولينسر في دليل برنامج... (Wright&Linacre, 2001p4-13).

- حذف البنود غير الملائمة: أعيد التحليل للمرة الثالثة بهدف البنود غير الملائمة لعملية التدرج وفقاً لإحصاءات الملائمة عن طريق حذف البنود التي تقل قيم إحصاءات الملائمة لها عن القيمة (-٢,٥) حيث يعني هذا عدم استقلالية البند عن باقي البنود وحذف البنود التي تزيد قيم إحصاءات الملائمة لها عن القيمة (+٢,٥) حيث يعني هذا عيباً في بنائها أو عدم صدقها في قياس ما تقيسه باقي البنود وقد قام الباحث بتجربة حذف البنود ذات الملائمة السالبة التي تجاوزت القيمة (-٢,٥) حيث قام بإعادة التحليل، ثم قام مرة أخرى بتجربة حذف البنود ذات الملائمة الموجبة التي تجاوزت القيمة (+٢,٥) ثم قام بإعادة التحليل مرة أخرى.. وهكذا أجريت سلسلة من التحليلات حتى أمكن الوصول إلى أحسن وضع ممكن من حيث ملائمة بنود الاختبار. حيث ينتهي عمل البرنامج بالحصول على جدولين يمثلان التدرج النهائي لكل من (صعوبة البنود حيث تدرج بنود الاختبار في صورته النهائية تبعاً لتقدير الصعوبة باللوغيت) و (قدرة الأفراد المقابلة لكل درجة كلية محتملة يحصل عليها الأفراد على الاختبار) و (تحويل تقديرات الصعوبة والقدرة من وحدة اللوجيت إلى وحدة القياس منف باستخدام معادلة التحويل الخطي).

١- **عمل المعايير:** يقف استخدام نموذج راش عند تكوين ميزان تدرج لكل من صعوبة البنود وقدرة الأفراد، يأتي بعد ذلك عمل المعايير التي تفسر على أساسها المستويات المختلفة للأفراد وذلك في إطار أسلوب القياس جماعي المرجع، عن طريق تحويل الدرجات الخام الكلية لجميع أفراد العينة إلى تقديرات القدرة المقابلة لها، وحساب مقاييس النزعة المركزية ومقاييس التشتت.

٢- تمثيل منحنيات النمو: حيث تم رسم منحنيات نمو القدرة العقلية كما هي مقياسة بمقياس القدرة العقلية وهذه المنحنيات توضح العلاقة بين مستويات العمر المختلفة وتقديرات القدرة مقدرة بالمنف.

٣- ترجيح الاختبار الفرعي المناسب لشرائح العمر المختلفة : وتم ذلك عن طريق حساب متوسط صعوبة الاختبارات الفرعية وكذا مدى صعوبة البنود وكذلك حساب متوسط الانحراف المعياري لتقديرات قدرة الأفراد، وتحديد مدى القدرة الذي يمكن للاختبار الفرعي تقديره والذي على أساسه يمكن ترجيح الاختبار الفرعي المناسب لتقدير مدى معين من القدرة.

الفصل الرابع

نتائج البحث

- ١- إن أفضل حجم من بين إجماع العينة المستخدمة في تدرج مقياس القدرة على التفكير الناقد باستخدام نموذج راش هو حجم (٨٠٠) وذلك تبعاً لمحكات دقة القياس.
- ٢- إن أفضل عدد للمفردات المشتركة والمستخدم في ربط الصور الاختبارية المختلفة هو العدد (١١) وذلك تبعاً لمحكات دقة القياس.
- ٣- إن أفضل عدد للأفراد المشتركين والمستخدمين في ربط الصور الاختبارية هو العدد (٥٠) وذلك تبعاً لمحكات دقة القياس.
- ٤- إن أفضل طريقة للربط بين الصور الاختبارية هي طريقة المفردات المشتركة وذلك تبعاً لمحكات دقة القياس.
- ٥- وتبعاً لمحكات دقة القياس فقد اختلفت دقة القياس باختلاف مستوى العينة وذلك لصالح العينة منخفضة المستوى.
- ٦- تحققت في مقياس القدرة على التفكير الناقد استقلالية القياس (بمعنى دقة التعادل).
- ٧- تم تدرج وتقدير صعوبة المفردات مقياس القدرة على التفكير الناقد وفقاً لنموذج راش وذلك بعد توفر محكات الدقة .

- ٨- تم تقدير قدرة الأفراد المقابلة لكل درجة كلية خام محتملة على مقياس القدرة على التفكير الناقد وفقاً لنموذج راش بعد توفر محكات الدقة .
- ٩- تم التوصل إلى معايير الرتب المئينية والدرجات التائية التي تفسر على أساسها مستويات قدرة الأفراد على المقياس.

التوصيات

في ضوء نتائج البحث يوصي الباحث بما يأتي :

١. على أساس ما يحققه نموذج راش من خطية التدرج وأحادية القياس واستقلالية، وما تتميز به أداة القياس التي طورت باستخدام النموذج من تحقيق لموضوعية القياس، توصي الدراسة باستخدام نماذج نظرية الاستجابة للمفردة في تطوير وبناء الاختبارات العقلية المعروفة والشائعة الاستخدام للتغلب على أوجه النقد التي وجهت لتلك المقاييس.
٢. كانت من المشكلات التي واجهت الباحث عند بداية الدراسة العينات المناسبة للتحليل وعلاقتها بعدد المفردات وعلاقة ذلك بعدد الأفراد المشتركين اللازم. لذلك توصي الدراسة الحالية في التوسع في الدراسات الخاصة بشروط التحليل باستخدام الأفراد المشتركين.
٣. توصي الدراسة باستخدام النموذج الثنائي المعلم في تدرج مقياس القدرة للتفكير الناقد.
٤. توصي الدراسة الحالية بإجراء دراسات تكون بمثابة امتداد لهذه الدراسة بهدف التوسع في المدى العمري للعينات المستخدمة لوضع معايير للاختبار في مراحل عمرية أكبر وزيادة حجم العينات الصغيرة في الدراسة الحالية للتأكد من المعايير الخاصة بها.
٥. توصي الدراسة الحالية بإجراء دراسة بمثابة امتداد لهذه الدراسة تقوم بتعديل المفردات المحذوفة وإضافة مفردات جديدة للاختبار وإعادة التدرج للاختبار.

(أنموذج) من التدرج النهائي لمفردات مقياس التفكير الناقد تبعاً لصعوبتها

| م | رمز المفردة | الصعوبة | | الخطأ المعياري | | إحصاء الملائمة | |
|----|-------------|----------|--------|----------------|--------|----------------|----------|
| | | باللوجيت | بالمنف | باللوجيت | بالمنف | التقاربي | التبايدي |
| ١ | X1 | ٤,٤- | ٢٨ | ٠,٥٢ | ٣ | ٠ | ٠,٣- |
| ٢ | X8 | ٣,٦٣- | ٣٢ | ٠,٣٨ | ٢ | ٠ | ٠,٦- |
| ٣ | X10 | ٣,٣٨- | ٣٣ | ٠,٣٤ | ٢ | ٠ | ٠,٣- |
| ٤ | X14 | ٣,٦- | ٣٥ | ٠,٣١ | ٢ | ٠,٣- | ١- |
| ٥ | C2 | ٣,٠١- | ٣٥ | ٠,٢٨ | ١ | ٠,٢- | ٠,٣- |
| ٦ | X13 | ٢,٩٧- | ٣٥ | ٠,٣ | ٢ | ٠,١ | ٠,٣ |
| ٧ | X6 | ٢,٨- | ٣٦ | ٠,٢٨ | ١ | ٠,٨- | ١,٦- |
| ٨ | C5 | ٢,٧٨- | ٣٦ | ٠,٢٦ | ١ | ٠,٨- | ٠,٦- |
| ٩ | X9 | ٢,٢٢- | ٣٦ | ٠,٢٨ | ١ | ٠,٥- | ٠,٧- |
| ١٠ | X2 | ٢,٧٢- | ٣٦ | ٠,٢٨ | ١ | ٠,٦ | ٠,٢ |
| ١١ | X5 | ٢,٥٨- | ٣٧ | ٠,٢٧ | ١ | ٠,٨ | ١,٥ |
| ١٢ | X12 | ٢,٥١- | ٣٨ | ٠,٢٦ | ١ | ٠,٨- | ١,١- |
| ١٣ | X3 | ٢,٣٨- | ٣٨ | ٠,٢٥ | ١ | ٠,٥ | ١,١ |
| ١٤ | C6 | ٢,٣٥- | ٣٨ | ٠,٢٣ | ١ | ٠,٦ | ٠,٦- |
| ١٤ | Y13 | ٢,٣٥- | ٣٩ | ٠,٢٣ | ١ | ٠,٦- | ٠,١- |
| ١٥ | C1 | ٢,٢١- | ٤٠ | ٠,٦١ | ٣ | ٠,١- | ٢,٤ |
| ١٦ | X11 | ٢,١- | ٤٠ | ٠,٢٢ | ١ | ٠,٤- | ٠,٩ |
| ١٧ | C12 | ٢,٠٨- | ٤٠ | ٠,٢٣ | ١ | ٠,٨ | ٠ |
| ١٨ | C10 | ٢,٠٥- | ٤٠ | ٠,٢٢ | ١ | ٠,٧ | ٠,٨ |
| ١٩ | X7 | ٢,٠٣- | ٤١ | ٠,٢٣ | ١ | ٠,٤ | ٠,٦ |
| ٢٠ | C4 | ١,٧٥- | ٤٣ | ٠,٢ | ١ | ٠,٢ | ٠,٥ |
| ٢١ | C8 | ١,٤٥- | ٤٣ | ٠,١٩ | ١ | ١- | ٠,٥ |
| ٢٢ | Y15 | ١,٤٢- | ٤٤ | ٠,٤٥ | ٢ | ٠ | ١,٧ |

| | | | | | | | |
|------|------|---|------|-----|-------|-----|----|
| ٢ | ٠,٦- | ١ | ٠,١٨ | ٤٤ | ١,١٨- | C7 | ٢٣ |
| ٠,٥ | ٠,٩- | ١ | ٠,١٨ | ٤٥ | ١,١٦- | C4 | ٢٤ |
| ٢,٣ | ٠,١ | ١ | ٠,١٧ | ٤٤٠ | ٠,٩٦- | C11 | ٢٥ |
| ٠,٥- | ١,٩ | ١ | ٠,١٦ | ٤٧ | ٠,٥٩- | C9 | ٢٦ |
| ٢,٩ | ٠ | ٢ | ٠,٣ | ٥٠ | ٠,٠٧- | C16 | ٢٧ |
| ٠,٩- | ١,٦ | ٢ | ٠,٣ | ٥١ | ٠,١ | Y12 | ٢٨ |
| ٠,٩- | ٠,١- | ٥ | ١,٠٩ | ٥١ | ٠,٢ | Z21 | ٢٩ |
| ٠,٩- | ٠,١- | ٥ | ١,٠٢ | ٥١ | ٠,٢ | Z18 | ٣٠ |
| ٠,١ | ٠,١ | ٣ | ٠,٦١ | ٥٧ | ١,٣٦ | Z26 | ٣١ |
| ٢,٥- | ٣,٢- | ١ | ٠,٢٤ | ٥٧ | ١,٤٢ | C21 | ٣٢ |
| ٠ | ٢- | ١ | ٠,٢٤ | ٥٨ | ١,٦ | C18 | ٣٣ |
| ٠٠,٥ | ٠,١ | ٣ | ٠,٥٣ | ٥٨ | ١,٦٨ | Z12 | ٣٤ |
| ٠,٥- | ٠,٢- | ١ | ٠,٢٣ | ٥٩ | ١,٨٧ | C22 | ٣٥ |
| ٠,٧- | ٠ | ٢ | ٠,٤٨ | ٦٠ | ١,٩٤ | Z24 | ٣٦ |
| ٠,٢- | ٠,١ | ١ | ٠,٢٣ | ٦٠ | ٢,٠٨ | C19 | ٣٧ |
| ٠,٢- | ١,٦- | ١ | ٠,٢٢ | ٦١ | ٢,٢٨ | C12 | ٣٨ |
| ٢,٢- | ٢,٣- | ١ | ٠,٢٢ | ٦٢ | ٢,٣٣ | C14 | ٣٩ |
| ٠,٢ | ٠,١- | ٢ | ٠,٤ | ٦٣ | ٢,٥ | Z13 | ٤٠ |
| ٠,٤ | ٠,٤ | ١ | ٠,٢١ | ٦٤ | ٢,٧ | C17 | ٤١ |
| ٠,٤ | ٠,٣ | ١ | ٠,٢١ | ٦٤ | ٢,٨٤ | C20 | ٤٢ |
| ١,٥- | ١,٢- | ٢ | ٠,٣٣ | ٦٦ | ٣,٢٥ | Z20 | ٤٣ |
| ٠,١- | ٠,٥ | ٢ | ٠,٣٣ | ٦٦ | ٣,٢٥ | Z19 | ٤٤ |
| ٠,٥ | ٠,٨ | ٢ | ٠,٣ | ٦٩ | ٣,٨٢ | Z17 | ٤٥ |
| ١,٣- | ١,٢- | ٢ | ٠,٢٨ | ٧٢ | ٤,٤ | Z16 | ٤٦ |
| ٠,٣ | ٠,٤ | ١ | ٠,٢٨ | ٧٢ | ٤,٤٨ | Z22 | ٤٧ |
| ٠,٤ | ٠,٣ | ٢ | ٠,٢٩ | ٧٥ | ٥,٠٤ | Z25 | ٤٨ |
| ١,٧ | ١,٣ | ٢ | ٠,٣ | ٧٧ | ٥,٣٧ | Z15 | ٤٩ |
| ٠,٤- | ٠,٢- | ٢ | ٠,٣ | ٧٨ | ٥,٥٦ | Z23 | ٥٠ |

* المصادر العربية

١. الوقفي، راضي، ٢٠٠٣: مقدمة في علم النفس، دار الشروق، عمان، ط ١.
٢. الخرجي، عزيز حسن جاسم، ٢٠١١: بناء برنامج تعليمي على وفق استراتيجيات معالجة المعلومات وأثره في التحصيل والتفضيل المعرفي وتنمية التفكير الناقد لدى طالبات قسم علوم الحياة أطروحة دكتوراه غير منشورة، جامعة بغداد، كلية التربية ابن الهيثم .
٣. الزيد، جواهر محمد (٢٠٠٧) : فاعلية نظرية الاستجابة للمفردة في بناء الاختبارات (إعداد بنك أسئلة في مقرر علم النفس التربوي) أطروحة دكتوراه غير منشورة ، جامعة الإمام محمد بن سعود الإسلامية ،كلية العلوم الاجتماعية.
٤. أبو حطب، فؤاد وعثمان، سيد احمد وآخرون (١٩٩٣):التقويم النفسي، مكتبة الانجلو المصرية ، القاهرة .
٥. عبد العزيز، سعيد (٢٠٠٩) : تعليم التفكير ومهاراته تدريبات وتطبيقات عملية ، ط١، عمان، دار الثقافة للنشر والتوزيع
٦. عبدالله، اعتدال غازي (٢٠٠٣):استخدام نموذج راش في تدرج مقياس للقدرة العقلية لدراسة بعض العوامل المؤثرة على دقة القياس ،رسالة ماجستير ،كلية البنات ،جامعة عين الشمس .
٧. عبد العزيز، سعيد(٢٠٠٩):تعليم التفكير ومهارته تدريبات وتطبيقات عملية ، ط١، الإصدار الثاني، دار الثقافة للنشر والتوزيع، عمان، الأردن.
٨. عبد الهادي، نبيل، وعياد ،وليد(٢٠٠٩):استراتيجيات تعلم مهارات التفكير بين النظرية والتطبيق ، ط١، دار وائل، عمان ،الأردن.
٩. عبد السلام ،نادية (١٩٩٦):حول معايير التقويم نظرة جديدة في :أنور الشرقاوي،وسليمان الخضري وآخرون، اتجاهات معاصرة في القياس النفسي والتربوي. القاهرة: مكتبة الانجلو المصرية : ٤١-٦٢
١٠. عبد الحجازين ،نايل (٢٠١٠) نماذج نظرية الاستجابة للمفردة متعددة التدرج ومتعددة الإبعاد وبرامجها الحاسوبية ، ط١، دار جليس الزمان للنشر والتوزيع ،بيروت.
١١. علام ،صلاح(١٩٩٥):الاختبارات التشخيصية مرجعية المحك في المجالات التربوية والنفسية والتدريبية .دار الفكر العربي، القاهرة .
١٢. علام،صلاح الدين (٢٠٠٥):نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الإبعاد وتطبيقها في القياس النفسي والتربوي، ط١، دار الفكر العربي، القاهرة
١٣. كاظم، أمينة (١٩٨٨):دراسة نظرية نقدية حول القياس(نموذج راش)،مؤسسة الكويت للتقدم العلمي،سلسلة الكتب المتخصصة.
١٤. كاظم ،أمينة (١٩٩٥):اتجاهات معاصرة في بنوك الأسئلة في :الأسس التربوية لإعداد المعلم الجامعي،جامعة عين الشمس، القاهرة، الطبعة الثانية.
١٥. كاظم، أمينة (١٩٩٦):نماذج السمات الكامنة في :أنور الشرقاوي وآخرون:اتجاهات معاصرة في القياس والتقويم النفسي والتربوي، الانجلو المصرية، القاهرة.
١٦. مسعود ،وليد احمد (٢٠٠٤):دراسة سيكومترية لتطوير اختبار رسم الرجل باستخدام نموذج (راش)،رسالة ماجستير ،كلية البنات للآداب والعلوم التربوية ،جامعة عين الشمس.

١٧. منير، منى ربيع طنطاوي (٢٠٠٠) : دراسة سيكومترية حول اختبار المصفوفات المتتابعة لرافن باستخدام نموذج راش، رسالة ماجستير ،كلية البنات للآداب والعلوم والتربية ،جامعة عين الشمس .
١٨. محاسنة،إبراهيم محمد (٢٠١٣):القياس النفسي في ظل النظرية التقليدية والنظرية الحديثة ،ط١،دار جرير للنشر والطباعة ،عمان
١٩. نوفل ،محمد بكر وأبو عواد،فريال محمد ،٢٠١١:التفكير والبحث العلمي ،ط١،دار المسيرة للنشر والتوزيع،عمان.

المصادر الأجنبية

- 20-Hulin,C.L.(1987).A Psychometric theory of evaluations of item and scal translations:fidelity across languages.Journal of cross-cultural psychology,18.(2),115-142
- 21-Embretson Susane and Reise.stven (2000):Item Response Theory for Psychologists ,Lawrence Erlbawm associates publshers New Jersy
21. Feingold,(1994).Gender defferences in personality:A meta analysis.psychological Bulletin 116,429-456
22. Backer,frank(2001):The Basies of Item Response Theory ,eric clearinghouse on assessment and evaluation, university of Maryland,college park:MD
23. Suen.Hoik(1990):Principles of Test Theories Lawrence Erlbaum associates publishers ,Newjersy.
24. Hambleton&Swaminathan, H. (1985):Item Response Theory Principle and Applicationns,The Netherland,Kluwer Nijihoff Publishing
25. Hambleton& Swaminathan, H. (1987):Item Response Theory Principle and Applicationns,The Netherland,Kluwer Nijihoff Publishing,United States of America, 3rd
26. Hambleton.R.K,Swaminthan.H. and Rogers .J (1991):Fundamentals of Item Response Theory. Sage publications the international professional publishers Newbury park
27. Wright. B. and masters G.(1982):Rating Scale Analysis Mesa press Chicago
28. Sen Hoik (1990):Principles of Test Theories Lawrence Erlbaum associates publishers, Newjersy.
29. Wright ,B.(1992):The Internationnal Objective Measurement Workshops,past and future in markwilson Objective measurement "theory into practiele" Ablex publishing corporation Norwood Newjersey.
30. Wright,B and Linacre ,J (2001):A user Guide to Bigsteps (winsteps).Rasch model computer program mesa press