مجلة جامعة الانبار للعلوم الانسانية
University of Anbar Journal for Humanities
P. ISSN: 1995-8463 E.ISSN: 2706-6673
المجلد ١٩- العدد ٢ – حزيران ٢٠٢٢ | Volume 19- Issue 2- June 2022

العلاقة بين معاملي الصعوبة والتمييز وفقاً لنظريتي القياس التقليدية والسمات الكامنة لفقرات اختبار الاستدلال المنطقي لدى طلبة الصف السادس الاعدادي

د. بلال طارق حسين
جامعة الانبار- كلية التربية للعلوم الانسانية
bthussain@uoanbar.edu.iq

DOI
10.37653/juah.2022.174790

تم الاستلام: ٢٠٢٢/١/١٢
قبل للنشر: ٢٠٢٢/٣/٩
تم النشر: ٢٠٢٢/٦/١

الكلمات المفتاحية
معاملي الصعوبة والتمييز
نظريتي القياس التقليدية
السمات الكامنة

الملخص:

هدفت هذه الدراسة إلى ايجاد العلاقة الارتباطية بين معاملات الصعوبة والتمييز بالنظريتين التقليدية والسمات الكامنة (الانموذج ثنائي المعلم) ، واعتمد الباحث على اختبار الاستدلال المنطقي المعد من قبل مؤسسة (assmentday) المختصة في اصدار مختلف الاختبارات النفسية في بريطانيا عام ٢٠١٨ . ترجم الباحث تعليمات الاختبار من اللغة الانكليزية إلى اللغة العربية لان الفقرات عبارة عن اشكال لذلك لا يحتاج الى التحقق من صدق الترجمة لانها عابرة للثقافة . وقد طبق الاختبار على عينة مكونة من (٦٥٠) طالب وطالبة اختيروا بطريقة عشوائية من طلبة الصف السادس الاعدادي. وقد تم اجراء التحليل الإحصائي للفقرات وفقاً لنظريتي القياس التقليدية والسمات الكامنة بالاعتماد على الانموذج ثنائي المعلم في تحليل فقرات الاختبار وكما حسبه برنامج بايلوج (Bilog-mg3) ولم تستبعد اي فقرة من الاختبار وفقاً لنظريتي القياس. وعلى وفق نظريتي القياس كشف الباحث عن العلاقة الارتباطية لمعلم الصعوبة وقد بلغ (٠,٩٠٧) ، في حين بلغ معامل الارتباط لمعلم التمييز (٠,٨١٣) وجميع معاملات الارتباط موجبة ودالة احصائياً عند مستوى دلالة (٠,٠٥) . وفي ضوء نتائج البحث الحالي توصل الباحث إلى بعض الاستنتاجات والتوصيات والمقترحات.

The relationship between the difficulty and discrimination coefficients according to the two traditional measurement theories and the latent features of the items of the logical reasoning test among the sixth grade students

Dr. Bilal T. Hussein
University of Anbar –College of Education for Humanities

Abstract:
This study aimed to find the correlation between the factors of difficulty and Discrimination in the Classical Test theories and the Latent Trait Theory (Two-Parameter Logistic Model), and the researcher relied on the logical reasoning test prepared by the Assmentday Foundation, which is specialized in issuing various psychological tests in Britain in 2018. The researcher carried out the translation of the test instructions from English into Arabic because the paragraphs are forms, so he does not need to verify the validity of the translation because it crosses culture. The test was applied to a sample of (650) male and female students who were randomly selected from the sixth grade students. The statistical analysis of the items was carried out according to the two traditional measurement theories and the latent features, depending on the two-parameter model in analyzing the test items, as calculated by the Bilog-mg3 program, and no item was excluded from the test according to the two measurement theories. And discrimination according to the two measurement theories and it was found that the value of the correlation coefficient for the parameter of difficulty according to the two measurement theories (0.907) and the value of the correlation for the teacher of discrimination according to the two measurement theories (0.813), and all the correlation coefficients are positive and statistically significant at the level of significance (0.05). Some conclusions, recommendations and suggestions.

Submitted: 12/01/2022
Accepted: 09/03/2022
Published: 01/06/2022

Keywords:
Difficulty and discrimination coefficients
traditional measurement theories
latent features.

©Authors, 2022, College of Education for Humanities University of Anbar. This is an open-access article under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

مشكلة البحث :

رغم اتساع نطاق استخدام النظرية التقليدية ونظرية الاستجابة للمفردة في تطوير أدوات القياس إلا إنهما يمثلان إطاري عمل مختلفين في مجال القياس النفسي والتربوي (fan 357: 1998,). فبالرغم من الاتفاق بينهما في كثير من الخصائص، إلا أن هناك اختلافا في معالجة هذه الخصائص مثلاً صعوبة الفقرة في القياس التقليدية، تدل على متوسط احتمالات الإجابة الصحيحة على الفقرة ، لجميع الأفراد الذين اشتركوا في الإجابة على فقرات الاختبار، وتتغير قيمة صعوبة الفقرة بتغير العينة أو الأفراد (Dick & Hagert,1971:13).

اذ اشار هاشوي (Hashway , 1977) في دراسته عن المقارنة بين الاختبارات الناتجة من أنموذج راش ومن الطرق التقليدية عند تطبيقها لانتقاء فقرات من نطاق شامل للفقرات الا ان كلا الاسلوبين لا يختلف في انتقائه للفقرات من حيث المحتوى

كما اشارت دراسة عودة (١٩٩٢) التي بحثت في مدى التوافق بين أنموذج راش والمؤشرات التقليدية في اختيار فقرات مقياس اتجاه ، إلى درجة عالية من التوافق في عدد الفقرات ومعامل الثبات للمقياس على عينة أفراد الذين انسجمت اجاباتهم مع الأنموذج الا ان اختيار الفقرة على اساس معامل الارتباط المعدل لا يعني بالضرورة انسجامها مع الأنموذج في مقياس يفترض بأنه احادي البعد (عودة ، ١٩٩٣ : ١٥٣).

اذ أُعتمد في البيئة العربية عند بناء وتقنين الاختبارات والمقاييس على نظرية القياس التقليدية على الرغم من ظهور العيوب المرتبطة بهذه النظرية وأن ظهور نظرية السمات الكامنة قد أحرز تقدماً وأصبح إطار القياس المفضل للوصول إلى الموضوعية في القياس مما ساعد الباحث على التفكير بأيجاد العلاقة في الصعوبة والتمييز لاختبار الاستدلال المنطقي وفقاً لنظريتي القياس التقليدية والسمات الكامنة .

لذا يسعى البحث الحالي إلى استخراج الخصائص القياسية لاختبار الاستدلال المنطقي للمرحلة الإعدادية وايجاد العلاقة في الصعوبة والتمييز وفقاً لنظريتي القياس التقليدية والسمات الكامنة .

اهمية البحث :

تؤدي نظريات القياس دورا بارزاً في منهجية البحث وذلك بطرحها أساليب عامة في إجراءات القياس أو قياس المتغيرات التي نهتم بها ، وذلك من أجل اختبار دقة نماذج القياس التي تم تطويرها (كروكر، الجينا ،٢٠٠٩ : ٣٠).

حظي مجال القياس والتقويم التربوي والنفسي باهتمام بحثي متزايد من علماء النفس والتربية، وذلك للكشف عن الفروق الفردية بأنواعها سواء كانت فروقاً بين الافراد أو بين الجماعات أو فروقاً في ذات الفرد، وقد تنوعت أساليب وأدوات القياس لتحقيق هذا الغرض، حيث كانت جودة التقويم تعتمد على جودة الأداة وجودة عملية القياس في نظرية القياس التقليدية (Classical Theory) وقد أثمرت هذه الجهود إلى الانتقال من النظرية التقليدية في تصميم الاختبارات، والتي استخدمت لفترة طويلة في التقويم التربوي والنفسي، إلى النظرية أو نظرية السمات Item Response Theory الحديثة المتمثلة بنظرية الاستجابة للمفردة (العبد الله ، ٢٠٠٩: ٥)

كما تعد نظرية الاستجابة للفقرة من التطورات والاتجاهات الحديثة المهمة في مجالات القياس التربوي والنفسي لما قدمته هذه النظرية من طرائق جيدة ومهمة في تفسير الدرجات على المقاييس التربوية والنفسية عند مقارنتها مع النظرية التقليدية في القياس، وقد قدمت أيضا هذه النظرية الحديثة في القياس العديد من الطرائق السيكومترية الجيدة والمفيدة والفعالة في بناء المقاييس التربوية والنفسية (Mislevy.& ,Bock, 1990)

لذا تبرز أهمية البحث الحالي من خلال ما يأتي:

١. الإسهام في توفير اختبار يتمتع بخصائص قياس جيدة من خلال تطوير اختبار الاستدلال المنطقي.

٢. أهمية الموضوع من خلال إجراء العلاقة في الصعوبة والتمييز وفقاً لنظريتي القياس التقليدية والسمات الكامنة .

٣. هذه الدراسة أول دراسة في العراق (حسب علم الباحث) تضاف إلى الدراسات التي اهتمت بمعالجة مشكلات القياس في نظريتي القياس التقليدية والسمات الكامنة .

هدف البحث :

يهدف البحث الحالي الى التعرف على العلاقة بين معامل الصعوبة والتمييز وفقاً لنظريتي القياس التقليدية والاستجابة للفقرة لاختبار الاستدلال المنطقي لدى طلبة المرحلة الاعدادية وينبثق من هذا الهدف الفرضيات الاتية:-

١. لا توجد علاقة ذو دلالة احصائية في معامل الصعوبة وفقاً لنظريتي القياس التقليدية والسمات الكامنة لفقرات اختبار الاستدلال المنطقي .

٢. لا توجد علاقة ذو دلالة احصائية في معامل التمييز وفقاً لنظريتي القياس التقليدية والسمات الكامنة لفقرات اختبار الاستدلال المنطقي .

حدود البحث

يتحدد البحث الحالي بالاتي :

- طلبة الصف السادس الاعدادي للعام الدراسي (٢٠٢٠- ٢٠٢١) في محافظة بغداد.
- اختبار الاستدلال المنطقي الصادر من مؤسسة AssessmentDay, (2012) على طلبة المرحلة الإعدادية.

تحديد المصطلحات : فيما يأتي تعريف لأهم المصطلحات التي وردت في البحث

اولاً: معامل الصعوبة وفقاً لنظرية القياس التقليدية عرفه :

- سليمان و رجاء (2010) : بأنه نسبة الطلبة الذين لم يجيبوا عن الفقرة الى العدد الكلي للطلاب المشاركين في اداء الاختبار" (سليمان و رجاء، 313:2010).

ثانياً: معامل التمييز وفقاً لنظرية القياس التقليدية عرفه:

- الدليمي وعدنان (٢٠٠٥) : وهو قدرة المفردة على كشف الفروق الفردية بين المجموعة العليا والدنيا اي بين الطلبة الذين يملكون السمة او متمكنين من الاجابة وبين الذين لا يملكون السمة او غير متمكنين من الاجابة (الدليمي وعدنان ،٢٠٠٥: ٦٦).

ثالثاً: معامل الصعوبة وفقاً لنظرية السمات الكامنة عرفه :

- (Reeve, 2003) :ان معلمة صعوبة الفقرة (β)، أو العتبة، هي نقطة على مقياس السمة أو القدرة الكامنة (θ) حيث تكون لدى الفرد المفحوص فرصة (50%) الاستجابة بالإيجاب لمقياس أو ميزان الفقرة (السؤال). والفقرات ذات العتبات العالية هي الأقل تأييداً أو إيجاباً في الغالب (Reeve, 2003: 7).

رابعاً: معامل التمييز وفقاً لنظرية السمات الكامنة عرفه :

- (Baker, 2001) : هي مقدار قدرة فقرة ما على التمييز الأفراد المفحوصين ذوي القدرة دون موقع الفقرة وأولئك الأفراد ذوي القدرة فوق موقع الفقرة. وهذه الخاصية الأساسية تعكس شدة انحدار Steepness الفقرة (Baker, 2001: 7).

خامساً : نظرية القياس التقليدية عرفها:

- **كروكر وألجينا Crocker & Algina (1986):** "النظرية التي تقوم على افتراض أساسي مفاده أن الدرجة الملاحظة على الاختبار تتكون من حاصل جمع كل من الدرجة الحقيقية ودرجة الخطأ، وتهتم هذه النظرية بتقدير المصادر المتعدد للخطأ التي تؤثر في درجة الفرد في الاختبار من أجل التحقق من دقة القياس" (Crocker & Algina, 1986: 11).

سادساً : نظرية السمات الكامنة عرفها :

- **علام (٢٠٠٧):** وهي من التطورات الحديثة والمعاصرة ببناء الاختبارات حظيت بأهتمام العديد من العلماء القياس ، وهي نظرية تحاول ان تربط بين سمة معينة عن الافراد التي يقيسها اختبار ما وبين استجابتهم لفقرة من فقرات الاختبار . اي وجود علاقة بين مستويات القدرة عن الافراد واحتمال اجابتهم على الفقرات الاختبارية اجابة صحيحة (علام، ٢٠٠٧: ٢١٠).

سابعاً: الاستدلال المنطقي يعرفه:

- كابلن (Chaplain, 1974): عمليات عقلية تبدأ من عدة قضايا وتنتهي بنتائج جديدة أي اكتشاف معلومات جديدة (Chaplain, 1974: 85).
- عطيفة والسرور (٢٠١١): نشاط عقلي يتطلب مجموعة من القدرات مثل (الأستدلال، الأستنتاج ، المقارنة ، التصنيف ، التعميم ، التحليل ، التركيب ، التقويم ، التخيل) لمواجهة موقف او مشكلة تحتاج الى حل.(عطيفة والسرور،٢٠١١: ١٣٨).

اطار نظري ودراسات سابقة

اولاً: نظرية القياس التقليدية

تسمى هذه النظرية احيانا بنظرية الدرجة الحقيقية ودرجة الخطأ Theory of True (and Error Scores) لانها فسرت التباين الذي نجده بين مرات القياس للفرد الواحد بوجود درجة الخطا الذي منشؤه عوامل غير منتظمة يتداخل مع الدرجة الحقيقية التي تعكس ما يمتلكه الفرد من المتغير المقيس والعوامل المنتظمة الاخرى (Ghiselli et al , 1981 : 195)، لقد قامت نظرية الاختبار التقليدية على فلسفة قياس الفروق بين الافراد ، ولكن بالاستناد الى معيار جماعة الاقران ، وهذا يعني ان نتائج قياس سمة باختبار ما ترجع الى الى معيار المجموعة التي ينتمي اليها الفرد لذلك يطلق على هذا النوع من القياس في كثير من الاحيان بالقياس مرجعي المعيار (Norm – Referenced) لذلك فأن الاختبار الجيد هو الذي يميز بين الافراد في السمة المقاسة اذ يتوزع الافراد على المستويات المختلفة من الاداء على تلك السمة ويتميز هذا النوع من القياس بتباين المعايير تبعا لارتفاع مستوى الجماعة المرجعية التي ينسب لها اداء الافراد وانخفاضها ، كما تقوم هذه النظرية على فرضية ان توزيع درجات الافراد في السمة المقاسة يتخذ شكل المنحنى الطبيعي (Normal Curve). (Caller, 1966:272).

من هنا استندت النظرية التقليدية في القياس الى ثلاثة مفاهيم اساسية هي الدرجة الحقيقية والدرجة الملاحظة ودرجة الخطأ ، يمكن توضيحها فيما يأتي :

أ- الدرجة الحقيقية True Score:

يمتلك الفرد خصائص او سمات مستقرة تبقى مستمرة عبر الزمن يعبر عنها من ناحية القياس "بالدرجة الحقيقية" ،وتعرف الدرجة الحقيقية بأنها الدرجة التي من الممكن ان يحصل عليها المفحوص اذ اعطي له كل الفقرات التي تقيس السمة ويتم التعبير عنها احصائيا بانها "متوسط درجة الفرد في عدد كبير من التطبيقات المستقلة للاختبار نفسه " ، ويشير جيلفورد (Guilford) الى ان هذه الدرجة هي درجة الشخص التي يحصل عليها تحت ظروف نموذجية اذا كانت اداة القياس (الاختبار) تامة أي تقيس بدون اخطاء Ghiselli) (et al , 1981 : 197.

ب- درجة الخطأ Error Score :

ان الخطأ الذي عالجته نظرية الاختبار التقليدية يختلف عن غيره من الاخطاء ، فهو ليس خطأ (Mistake) في تسجيل الدرجات او اعداد مفتاح التصحيح على سبيل المثال

ومن ثم يمكن تفاديه او تصحيحه ، وانما هو خطا (Error) ملازم لعملية القياس لا يمكن تجنبه وانما يمكن تقديره (Lyman , 1963 : 62-63). يقول جيلكسن ,Gulliksen (1967) ان نسبة الخطأ في القياس النفسي اكبر بكثير منه في القياس المادي، لذلك يعد تقدير الخطأ في درجات الاختبار هي من المشكلات المركزية في القياس العقلي (Gulliksen,1967 : 4).

ج- الدرجة الملاحظة :

وهي الدرجة التي يتم الحصول عليها بعد تطبيق الاختبار ، وهذه الدرجة لا توضح بدقة تامة الكمية الحقيقية او الفعلية " الدرجة الحقيقية " للسمة التي يمتلكها الفرد بسبب تأثيرات العوامل غير المنتظمة او عوامل المصادفة ، لذلك تسمى احيانا الدرجة التي تحتوي على خطأ القياس ، وبسبب خطأ القياس فأن الدرجة الملاحظة التي يحصل عليها المفحوص في مجموعة من الاختبارات قد تتوزع وتنتشر بطريقة عشوائية حول درجته الحقيقية (Ghiselli et al , 1981 : 169) .

تحليل الفقرات وفق نظرية القياس التقليدية Item Analysis :

ان القضية النظرية الأساسية لإجراءات تحليل الفقرات هي إيجاد العلاقة الوظيفية بين معالم الاختبار ككل ومعالم الفقرات المختارة بشكل مناسب . مع ضرورة التحقق من مختلف العوامل التي تؤثر في معالم الفقرات مثل الخطأ العشوائي والتغيرات المنتظمة الناتجة من التغيرات في عوامل اخرى مثل طول الاختبار وتجانس عينة تحليل الفقرات وحجمها (Gulliksen ,1967 : 14).

وفيما يأتي اهم الخصائص السيكومترية التي تعتمد عليها النظرية التقليدية في تحليل الفقرات :

١- صعوبة الفقرة Item difficulty (P) : بشكل عام تركز النظرية التقليدية على اختيار معامل الصعوبة للفقرات بما يحقق اقصى تمييز بين الافراد ويحقق من ثم افضل توزيع لدرجات الاختبار مشابه للتوزيع الطبيعي. ويتحقق ذلك عندما يكون متوسط قيم معاملات الصعوبة هو (٠.٥٠) والذي يعطي اقصى معلومات حول أي مستوى لإداء المفحوص ، كما ان قيم معاملات الصعوبة مرتبطة بثبات الاختبار ، أي ان معامل الفا وكيودر – ريتشاردسون ٢٠ يميلان ليكونان اقل في الاختبارات المكونة من فقرات لها معدل

متطرف لقيم الصعوبة اكثر منه في الاختبارات التي لها قيم صعوبة قريبة من منتصف المدى(Nunnaly,1981: 270) (Anastasi,1988:203-210).

ونظرا لصعوبة ايجاد المتوسط المطلوب لمعامل الصعوبة فأن صعوبة الفقرات على وفق النظرية التقليدية ينبغي ان تمتد من (٠.٣٠ الى ٠.٧٠) لتعطي الحد الاقصى من معلومات الاختبار حول الفروق بين المفحوصين (Allen & Yen , 1979 : 121) .

٢- تمييز الفقرة Item Discrimination (D):

هنالك ارتباط كبير بين تمييز الفقرات وصعوبة الفقرات ، فإذا كان هدف الاختبار هو التمييز بين المتمكنين والاقل تمكناً ، فأن الفقرة المميزة هي ما تحقق الى هذا الهدف ، اذ ان مهمة التمييز تتمثل في تحديد مدى فاعلية فقرة ما في التمييز بين طلاب المجموعة العليا والمجموعة الدنيا بالقدر نفسه الذي يفرق الاختبار بينهما في الدرجة النهائية بصورة عامة (علام ، ٢٠٠٠)

ثانيا: نظرية السمات الكامنة:

ان بداية ظهور النظرية الحديثة في القياس ترجع الى سنة (١٩٢٥) على يد (Thurstone) الذي يعد من أوائل من وضعوا أساسا تصورياً لهذا الاتجاه في بحث بعنوان (A Method of Scaling Psychological and Educational Tests) حين قدم فيه تقنية لتسكين مفردات اختبار سيمون – بينيه للتطور العقلي عند الأطفال في أعمار متدرجة، بعد ذلك ظهر النموذج الاعتدالي التراكمي والذي يستخدم في وصف العلاقة بين مستوى قدرة الأفراد واحتمالية إجابتهم عن فقرات اختبارية مختلفة إجابة صحيحة. وان الفضل في هذا المدخل للقياس يعود الى كل من (1968) Lord & Novic الذي بني عليه كثير من التطورات المتسارعة في كتابهما "النظريات الإحصائية لنتائج الاختبارات العقلية Statistical Theories of Mental Test Scores" حول المفاهيم والمبادئ (عبد الوهاب ، ٢٠١٠: ٢٢)

افتراضات نظرية السمات الكامنة

بنيت هذه النظرية وتم اعتمادها على عدد من الافتراضات القوية التي يجب تحقيقها في البيانات المستمدة من الأختبار، لكي تؤدي إلى العديد من النتائج الموثوق بها وفيمايلي توضيح لهذه الافتراضات التي أشار إليها معظم الباحثين ومنهم:

(Crocker and Algina, 1986) (Hambleton & Swaminathan,1985)
(Wiberg, 2004) (Warm, 1978)

١. **احادية البعد Unidimensionality:** يدل هذا الافتراض على وجود قدرة أو سمة واحدة كامنة لدى الفرد تقوم على تفسير أداء الفرد على الاختبار.

٢. **الاستقلال الموضعي Local Independence:** ويدل هذا الافتراض على أن إجابة الفرد على زوج من فقرات الاختبار مستقل استقلال إحصائي عند مستوى قدرة معين، أي استجابة الفرد على فقرة معينة ليس له أي تأثير على استجابته على فقرة أخرى، وافتراض الاستقلال الموضعي مكافئ لافتراض أحادي البعد.

٣. **السرعة في الاداء Speediness:** يشير هذا الافتراض إلى أنّ السبب الرئيس وراء إخفاق الطلاب في الإجابة عن فقرات الاختبار هو أن قدرتهم محدودة، فقدرة الفرد هي العامل الرئيس لاستجابته على الاختبار وليس بسبب عامل السرعة الكافية من أجل إنجاز هذا الاختبار.

٤. **منحنى خصائص الفقرة (ICC: Item characteristics curve):** وهو دالة رياضية تربط احتمال الإجابة عن الفقرة إجابة صحيحة وبين مستوى القدرة ويشير إلى أن احتمال إجابة الفقرة إجابة صحيحة يزداد بزيادة قدرة المفحوص، وهو منحنى صاعد تراكمي ويوصف بدلالة معلمة واحدة هي صعوبة الفقرة، أو معلمتين صعوبة الفقرة وتمييزها، أو ثلاث معالم هي الصعوبة والتمييز والتخمين، وذلك تبعًا لنموذج نظرية الاستجابة للفقرة المستخدم، وهناك نماذج متعددة الاستجابة للفقرة، ونماذج ثنائية الاستجابة للفقرة

نماذج نظرية الاستجابة للفقرة المعلمية ثنائية التدريج

اختلفت نماذج نظرية الاستجابة للفقرة وتشتمل على ثلاث نماذج أساسية وتهتم هذه النماذج بالقياس في وقت معين، وهذه النماذج تندرج تحت نظرية الاستجابة للفقرة وتمكن من تحديد احتمالية إجابة الفرد ذو القدرة (θ) على الفقرة وهى كما يرى كل من:

Lord, 1980) (Baker, 2001) (Hambleton and Swaminathan,)
(1985

- **الأنموذج اللوغارتمي أحادي المعلمة** One- Parameter Logistic Model: وسمي بنموذج راش Rasch model ، ويقدر هذا النموذج ان الفقرات تختلف في

معامل الصعوبة فقط (**b**) ، ويفترض انعدام معامل التخمين لجميع الفقرات واعتباره صفر (**c_i=0**) ويقوم على افتراض تساوي معاملات التمييز لجميع الفقرات

$$P(x_{ij}=1|\theta_i,\beta_j)=\frac{\exp(\theta_i-\beta_j)}{1+\exp(\theta_i-\beta_j)}$$

- **الأنموذج اللوغارتمي الثنائي المعلم** (Two- Parameter Logistic Model): وتمت تسميته بنموذج بيرنبوم Birnbaum model ويفترض ان التخمين منعدم وان الفقرات تختلف في معامل الصعوبة (b) ومعامل التمييز (a) .

ويتم التعبير عن الانموذج رياضياً بحسب المعادلة (3) الاتية:

$$P_1(\theta)=\frac{\exp[a_i(\theta-\beta_j)]}{1+\exp[a_i(\theta_i-\beta_j)]}$$

الأنموذج اللوغارتمي الثلاثي المعلم Three- Parameter Logistic Model:

وتمت تسميته بنموذج لورد (Lord model) ، ويفترض ان الفقرات مختلفة في معامل الصعوبة (b) ، ومعامل التمييز (a) ، ومعامل التخمين (c) ، حيث اضيف معلم التخمين الى هذا النموذج وحسب المعادلة الاتية:

$$Pi(\theta) = ci + (1- ci)\frac{\exp[a_i(\theta-\beta_i)]}{1+\exp[a_i(\theta-\beta_i)]}$$

ثالثاً : الاستدلال المنطقي

الاستدلال : ـ

أولى علماء النفس والمنطق والفلسفة ، الاستدلال قدراً كبيراً من الاهتمام منذ زمن بعيد إلى الدرجة التي وصف بأنه الفن الذي يكفل لعمليات العقل قيادة منظمة ومسيرة خالية من الأخطار ، وعلى الرغم من هذه العناية لا يقصد العلماء التقليل من شأن الأنماط الأخرى للتفكير ، وإنما كان اهتمامهم لهذا النوع من التفكير ، لأنه يؤدي إلى كشف عن الحقائق وتنمية المعرفة . (صبري ، ٢٠٠٢ : ٤٨) .

والاستدلال نوع من التفكير يستهدف حل مشكلة ، حلاً ذهنياً وذلك عن طريق استخدام الرموز والخبرات السابقة للوصول إلى نتيجة مجهولـة مـن مقدمات معلومة .

(خيرالله ، ١٩٧٨ : ٣٩٦) .

الاستدلال والمنطق :

ينصب الاهتمام في المنطق على عمليات الاستدلال وفق القواعد المنطقية ، لان علم المنطق هو في الأصل علم الاستدلال الصائب ، ويميز علماء المنطق بين ثلاثة أنواع من الاستدلال هي الاستنباط والاستقراء والتمثيل .

أ – الاستدلال الاستنباطي ويعني الاستدلال من العام إلى الخاص .

ب– الاستدلال الاستقرائي ويعني الاستدلال من الخاص إلى العام .

ج– الاستدلال التمثيلي وهو استدلال من الخاص إلى الخاص ، ويتم عن طريق إجراء مماثله بين شيئين أو حالتين بينهما أوجه شبه ، ويترتب على عمليه المماثلة الوصول إلى نتيجة مفادها نقل حكم أو وصف من احد المتماثلين إلى الأخر . وليس هناك شك في أن فهم قواعد المنطق الصوري ومبادئه والألفة ببعض مصطلحات المنطق ، ومعرفه أنواع الاستدلال المنطقي تمثل عوامل معينه ومفيدة لممارسه الاستدلال بفاعليه ، ولكن ينبغي ملاحظة أن المنطق ليس إلا مظهرا واحدا من مظاهر الاستدلال ، لان الاستدلال بمعناه الواسع لا يقتصر على استخدام قواعد المنطق وتحري المغالطات المنطقية ، بل يتعدى ذلك إلى أشكال عديدة تستخدم فيها قواعد علم النفس المعرفي والتطوري وعلم نفس التعلم ومعالجة المعلومات وغيرها ، كما أن المنطق لا يقدم الكثير في موضوع الكشف عن صحة الأشياء والادعاءات أو زيفها ، بينما الاستدلال معني بالتوصيل إلى الحقيقة عن طريق توليد الفرضيات وفحصها وموازنة البدائل ، وهي عمليات للابداع دوراً جوهرياً فيها ، لأنها تتجاوز ما هو متاح من معلومات (جروان ، ١٩٩٩) (الحجازين ، ٢٠١٢ : ٣٢ – ٣٣).

علاقة الاستدلال ببعض المتغيرات

هنالك علاقة لمفهوم الاستدلال مع العديد من المتغيرات منها التخصص (علمي ، أدبي) والجنس (ذكر ، أنثى) .

أولا : التخصص ، تعتمد المواد التعليمية المختلفة على اطر تنظيمية متنوعة ، والنجاح في كثير من المواد مثل العلوم ، والتاريخ ، والأدب يتوقف على قدرة الطالب على التفكير الاستدلالي ، من خلال جمع البيانات واكتشاف العلاقات بين الحقائق والتوصل إلى النتائج التي تدعمها ، أما النجاح في مواد أخرى كالرياضيات ، والعلوم الطبيعية فيتطلب التفكير الاستقرائي (الأعسر ، 1998) . وأما بالنسبة لعلاقة التخصص (علمي ، أدبي)

بالاستدلال فقد أجمعت كل الدراسات على أن طلبة التخصص العلمي أكثر قدرة على الاستدلال من طلبة التخصص الأدبي.

ثانياً : الجنس ، هنالك دراسات رجحت بأن التفكير الاستدلالي عند الإناث أكثر منه عند الذكور ، وأخرى خلصت إلى عدم وجود فروق بين الجنسين .

فالدراسة التي قام بها غرايبة (2001) بينت أن أداء الإناث على اختبار الاستدلال المنطقي بأبعاده (ضبط المتغيرات ، الاستدلال ألارتباطي ، الاستدلال الاحتمالي) أفضل منه عند الذكور ، في حين دراسة قام بها أبو الجديان (1999) بينت عدم وجود فروق ذات دلالة إحصائية في قدرات التفكير الاستدلالي (الاستقرائي ، الاستنباطي ، ألاستنتاجي) تعزى للجنس . (الحجازين ، ٢٠١٢ : ٣٤ – ٣٥) .

الدراسات السابقة : لم يجد الباحث دراسة سابقة تناولت العلاقة بين معامل الصعوبة والتمييز وفقاً لنظريتي القياس التقليدية ونظرية السمات الكامنة ، فيما وجد مجموعة من الدراسات تناولت صعوبة الفقرات وتمييزها وفق نظرية القياس التقليدية ونظرية السمات الكامنة كل على حدة ، لذا سيتم عرض مجموعة منها وكالاتي:

- **دراسة جمحاوي (٢٠٠٠):** هدفت الدراسة إلى المقارنة بين النظرية التقليدية والنظرية الحديثة لخصائص فقرات مقياس القدرة الرياضية والتي كان عد دها (٣٩) فقرة، من حيث المحتوى، والعدد، ومدى الاتفاق في تقدير صعوبة الفقرات وتمييزها، وفي توزيع القدرة للمفحوصين، وفي الثبات. ولتحقيق هذا الهدف طبقت الباحثة المقياس على عينة بلغ حجمها (١٠٦١) مفحوصا من طلبة الصف التاسع في اربد. ودلت النتائج إلى وجود اتفاق عالي بين الاسلوبين في تقدير صعوبة وتمييز الفقرات، وفي توزيع القدرة للمفحوصين. كما بينت النتائج أن عدد الفقرات التي تم اختيارها وفق المؤشرات التقليدية كان (٣٣) فقرة، مقابل (٢٠) فقرة انسجمت مع نموذج راش، و (٣٥) فقرة انسجمت مع نموذج المعلمتين، و(٣٨) فقرة انسجمت مع نموذج المعالم الثلاثة. كما أشارت النتائج إلى عدم وجود فروق بين معاملي ثبات المقياسين الناتجين من المؤشرات التقليدية ونموذج المعلمتين، في حين وجود فروق بين هذا المعامل الناتج من المؤشرات التقليدية ونموذج المعالم الثلاثة ونموذج راش (جمحاوي، ٢٠٠٠).

- **دراسة (Magno,2009)** : تستهدف الدراسة إيجاد الفرق بين نظرية القياس التقليدية (CTT) ونظرية الاستجابة للفقرة (IRT) عبر عينتين مختلفتين وصورتين متكافئين لمادة الكيمياء، المشاركين في هذا الدراسة هم (٢١٩) طلبة في المدارس الثانوية من عاصمة فلبين من (٨) مقاطعات . وقد تم اختيار هؤلاء الطلبة بشكل عشوائيا ، تم جمع البيانات باستخدام صورتين متكافئين لاختبار لمادة الكيمياء وتم بناء صورتين وفق جدول المواصفات ، وتالف كل اختبار على (٧٠) فقرة ويتكون من بدائل متعددة الاختيار ، **توصلت نتائج الدراسة الى :-**

- أن تقديرات (IRT) من صعوبة الفقرة لا تتغير عبر عينات مختلفة بالمقارنة مع (CTT) التي تتغير تقدير معلم الصعوبة.
- ارتفع مؤشرات الصعوبة تكون أكثر استقرارا عبر أشكال الاختبارات في النظرية الاستجابة للفقرة من نظرية القياس التقليدية.
- أظهرت نتائج (IRT) أن الاتساق الداخلي يكون مستقرة جدا عبر العينات، بينما نتائج (CTT) الاتساق الداخلي غير مستقرة عبر العينات.
- أظهرت نتائج إن أخطاء القياس في (IRT) أقل بكثير من اخطاء القياس في (CTT) (Magno,2009: 1-11) .

منهجية البحث واجراءاته :

اولاً : منهجية البحث: اعتمد الباحث في البحث الحالي على المنهج الوصفي الارتباطي، الذي يفترض وجود علاقة ارتباطية بين نظريتي القياس التقليدية والحديثة في معلمي الصعوبة والتمييز.

ثانياً : مجتمع البحث: تألف مجتمع البحث من طلبة الصف السادس الاعدادي بفرعيه العلمي (احيائي وتطبيقي) والادبي في المدارس الاعدادية والثانوية موزعين في مديريات التربية في محافظة بغداد ، والبالغ عددهم (٨٤٩٦٩) طالب وطالبة للعام الدراسي (٢٠٢٠- ٢٠٢١) ، والجدول (١) يوضح ذلك.

الجدول (١) مجتمع البحث بحسب الجنس والمديرية

مديريات التربية حسب الموقع الجغرافي	الجنس				المجموع	%
	الذكور	%	الإناث	%		

تربية الكرخ الاولى	٧٠٥٧	%٩	٦٤٧٨	%٧	١٣٥٣٥	%١٦
تربية الكرخ الثانية	٩٨٩٧	%١٢	١٠٣٢١	%١٢	٢٠٢١٨	%٢٤
تربية الكرخ الثالثة	٥٤٦٣	%٦	٥٥٧٨	%٧	١١٠٤١	%١٣
تربية الرصافة الاولى	٧٧٤٦	%٩	٨٢٥٤	%١٠	١٦٠٠٠	%١٩
تربية الرصافة الثانية	٩٤٧٣	%١١	٦٢٨٨	%٧	١٥٧٦١	%١٨
تربية الرصافة الثالثة	٤١٨٥	%٥	٤٢٢٩	%٥	٨٤١٤	%١٠
المجموع	٤٣٨٢١	%٥٢	٤١١٤٨	%٤٨	٨٤٩٦٩	%١٠٠

ثالثاً : عينة البحث : اختار الباحث بطريقة عشوائية من طلبة الصف السادس الاعدادي عينة مكونة من (٦٥٠) طالب وطالبة، تشكل نسبة (١%) تقريباً من المجتمع (الطلبة) موزعين بحسب متغير الجنس والمديرية، والجدول (٢) يوضح ذلك.

الجدول (٢) عينة البحث بحسب الجنس والمديرية البحث

مديريات التربية حسب الموقع الجغرافي	الجنس		المجموع
	الذكور	اناث	
تربية الكرخ الاولى	٥٩	٤٥	١٠٤
تربية الكرخ الثانية	٧٨	٧٨	١٥٦
تربية الكرخ الثالثة	٣٩	٤٦	٨٥
تربية الرصافة الاولى	٥٨	٦٥	١٢٣
تربية الرصافة الثانية	٧٢	٤٥	١١٧
تربية الرصافة الثالثة	٣٢	٣٣	٦٥
المجموع	٣٣٨	٣١٢	٦٥٠

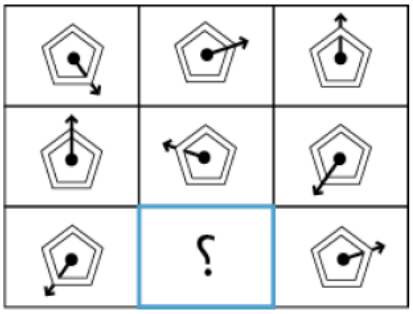
رابعاً : اداة البحث :

استعمل الباحث اختبار الاستدلال المنطقي الصادر من مؤسسة (AssessmentDay , 2012)) على طلبة المرحلة الاعدادية. المعد وفقاً لنظرية ثرستون للعوامل المتعددة.

وصف الاختبار:

يتكون الاختبار من (٣٠) فقرة تحتوي كل فقرة على مجموعة من الرموز وفي كل فقرة هنالك واحد من الرموز مفقود ، مهمة الطالب هي اختيار الرمز الانسب من بين (١٢) اجابة محتملة ، لا يوجد حد زمني اجمالي للاختبار ، ولكن هنالك حد (٧٠) ثانية لكل فقرة ، حاول ان تجد وقتاً ومكاناً لن تتم مقاطعتك فيه اثناء الاختبار . ومن امثلة الفقرات ما يأتي

- **ضع علامة صح فوق الرمز الذي يحل محل علامة الاستفهام**



إجراءات اعداد الاختبار :

تمت ترجمة الاختبار وتعليماته اللغة الإنكليزية إلى اللغة العربية عن طريق الاستعانة بمترجم متخصص[*] علماً بأن محتوى فقرات الاختبار جميعها غير لفظية اي عبارة عن اشكال باستثناء منطوق السؤال.

التحليل الإحصائي للاختبار:

اذ يتضمن عدة إجراءات عملية مثل: (التحليل الإحصائي وفق النظرية التقليدية واستخراج الخصائص القياسية، والتحليل الإحصائي وفق نظرية الاستجابة للفقرة واستخراج الخصائص القياسية وايجاد العلاقة الارتباطية في الصعوبة والتمييز وفقاً لنظريتي القياس وكالاتي :

[*] – أ.م.د. علي صباح جميل – تخصص طرائق تدريس اللغة الانكليزية

أولا:- التحليل الاحصائي للاختبار وفقاً لنظرية القياس التقليدية:- طبق الباحث الاختبار على عينة حجمها (٦٥٠) طالب وطالبة من طلاب الصف السادس الاعدادي والغرض من هذا التطبيق هو لمعرفة خصائصه القياسية والتأكد من صحة الفقرات وفقاً لنظريتي القياس التقليدية ونظرية السمات الكامنة ، وقد اشرف الباحث بنفسه على تطبيق الاختبار وبعد تطبيق الاختبار صححت الاجابات ثم رتبت ترتيب تنازلي وقسمت الى مجموعتين عليا ودنيا بنسبة (27%) بعد ذلك تم تحليل الفقرات لإيجاد معامل الصعوبة وقوة تمييز الفقرات وكالاتي.

اولا : معامل صعوبة الفقرة Item Difficulty : لحساب صعوبة فقرات الاختبار اعتمد الباحث على المجموعتين العليا والدنيا، كما حُسبت معاملات الصعوبة لكل فقرة من فقرات الاختبار إذ تراوحت قيم الصعوبة للفقرات بين (٠,٣٣- ٠,٥٢) وهذه القيم مقبولة كما يراها (الفقي،٢٠١٤) الذي يرى ان صعوبة الفقرات تعد مقبولة اذا تراوحت ما بين (٠,٨٠ – ٠,٢٠) (الفقي،٢٠١٤: ١٩٨) والجدول (٣) يوضح ذلك

ثانيا :- القوة التمييزية للفقرات Discrimination power of Items : قام الباحث بترتيب درجات أفراد العينة البالغ عددهم (٦٥٠) طالب وطالبة من أعلى درجة إلى أقل درجة وحددت المجموعتان المتطرفتان في الدرجة الكلية بنسبة (٢٧%) في كل مجموعة وبعد ذلك حُسب معامل التمييز لكل فقرة من فقرات الاختبار باستخدام معادلة التمييز ، كما اعتمد الباحث على محك أيبل لاختيارات الفقرات بحسب قوتها التمييزية، إذ يشير أيبل إلى أن الحد الأدنى لقبول الفقرة على أنها فقرة مميزة في الاختبارات يجب أن لا يقل عن (٠,٢٠) وقام الباحث بحساب القوة التمييزية لكل فقرة من فقرات الاختبار فوجد ان قيمتها تتراوح بين (٠,٣٦ – ٠,٥٥) ، جدول (٣)، وبذلك تعد فقرات الاختبار جيدة ومعامل تمييزها مقبول

جدول (٣)معاملات الصعوبة والقوى التمييزية لفقرات الاختبار وفقاً لنظريتي القياس التقليدية

رقم الفقرة	الاجابات الصحيحة للمجموعة العليا	الاجابات الصحيحة للمجموعة الدنيا	معامل الصعوبة	القوة التمييزية	رقم الفقرة	الاجابات الصحيحة للمجموعة العليا	الاجابات الصحيحة للمجموعة الدنيا	معامل الصعوبة	القوة التمييزية
١	١٥٣	٨٢	٠,٣٣	٠,٤٠	١٦	١٤٧	٧٥	٠,٣٧	٠,٤١

رقم الفقرة	الاجابات الصحيحة للمجموعة العليا	الاجابات الصحيحة للمجموعة الدنيا	معامل الصعوبة	القوة التمييزية	رقم الفقرة	الاجابات الصحيحة للمجموعة العليا	الاجابات الصحيحة للمجموعة الدنيا	معامل الصعوبة	القوة التمييزية
٢	١٣٦	٤٨	٠,٤٨	٠,٥٠	١٧	١٥٢	٧٨	٠,٣٥	٠,٤٢
٣	١٣٢	٣٧	٠,٥٢	٠,٥٤	١٨	١٤٩	٧٩	٠,٣٥	٠,٤٠
٤	١٤٥	٤٨	٠,٤٥	٠,٥٥	١٩	١٤٥	٥٥	٠,٤٣	٠,٥١
٥	١٥٧	٧٩	٠,٣٣	٠,٤٤	٢٠	١٣٧	٥١	٠,٤٧	٠,٤٩
٦	١٤٧	٥٢	٠,٤٣	٠,٥٤	٢١	١٢٨	٦٢	٠,٤٦	٠,٣٨
٧	١٤٦	٦٣	٠,٤١	٠,٤٧	٢٢	١٢٩	٤٢	٠,٥١	٠,٤٩
٨	١٣٤	٥٩	٠,٤٥	٠,٤٣	٢٣	١٢٥	٦٢	٠,٤٧	٠,٣٦
٩	١٣٧	٦٣	٠,٤٣	٠,٤٢	٢٤	١٣٧	٧١	٠,٤١	٠,٣٨
١٠	١٣٣	٤٨	٠,٤٩	٠,٤٨	٢٥	١٣٤	٥٦	٠,٤٦	٠,٤٤
١١	١٤٧	٥٣	٠,٤٣	٠,٥٣	٢٦	١٤٥	٦٨	٠,٣٩	٠,٤٤
١٢	١٣٩	٤٨	٠,٤٧	٠,٥٢	٢٧	١٤٦	٦٧	٠,٣٩	٠,٤٥
١٣	١٣٨	٦٦	٠,٤٢	٠,٤١	٢٨	١٤٧	٦٣	٠,٤٠	٠,٤٨
١٤	١٤١	٦٨	٠,٤١	٠,٤١	٢٩	١٤٣	٥٠	٠,٤٥	٠,٥٣
١٥	١٤٦	٦٣	٠,٤١	٠,٤٧	٣٠	١٥٧	٧١	٠,٣٥	٠,٤٩

ثانياً: ثبات الاختبار وفقاً لنظريتي القياس التقليدية :

إن الثبات يعني اتساق درجات فقرات الاختبار التي يفترض أن تقيس ما يجب أن تقيسه (Marshally, 1972: 104) وقد قام الباحث بحساب الثبات للاختبار بطريقة كيودر ريتشاردسون ٢٠ وكما يأتي :-

استخدم الباحث معادلة كيودر – ريتشاردسون ٢٠ لحساب ثبات الاختبار المطبق على عينة تبلغ (٦٥٠) طالب وطالبة اذ بلغ معامل الثبات (٠,٨٦) وهذا يدّل على ان الاختبار له معامل ثبات جيد ، يعد الاختبار جيداً اذا كان معامل ثباته (٠,٦٧) فأكثر. (النبهان ، ٢٠٠٤ ، ٢٤٠).

ثانياً:- التحليل الاحصائي للاختبار وفقاً لنظرية السمات الكامنة :

١. التحقق من افتراضات الانموذج

الافتراض الاول:- التحقق من افتراض أحادية البعد : قام الباحث باستخدام التحليل العاملي الاستكشافي (Principal Component)، لاستجابات الطلاب عن فقرات الاختبار ، لغرض التحقق ما اذا كان الاختبار يقيس سمة واحدة ، وقد تم حساب قيمة الجذر الكامن ونسبة التباين المفسر للجذر الكامن واظهرت النتائج وجود عامل واحد له جذر كامن تبلغ قيمته (١٠,٦٨٣) ويفسر (٣٥,٦١%) من التباين وكما موضح في جدول (٤)

جدول (٤)الجذر الكامن للعامل السائد لاختبار الاستدلال المنطقي

رقم العامل	الجذر الكامن	نسبة التباين المفسر
١	١٠,٦٨٣	٣٥,٦١

ويتضح من الجدول السابق هنالك عامل واحد يقف وراء استجابات الافراد لفقرات اختبار الاستدلال المنطقي اذ تجاوزت قيمة التباين المفسر للعامل الاول عن (٢٠%) كما يرى ريكاس (Reckas, 1979 as cited in : onder, 2007: 213)، كما قام الباحث بالتحقق من معيار تشبع الفقرات بالعامل العام بالاعتماد على قيمه التشبع التي وفق معيار جيلفورد (٠,٣٠) (فرج، ١٩٩١: ١٥١)، ويتضح من جدول (٥) مدى تشبع فقرات الاختبار بالعامل المسيطر .

جدول (٥) تشبع فقرات الاختبار بالعامل المسيطر

رقم الفقرة	Component ١	رقم الفقرة	Component ١	رقم الفقرة	Component ١	رقم الفقرة	Component ١	رقم الفقرة	Component ١
١	٠,٧٦٥	٧	٠,٥٣٧	١٣	٠,٥٨٦	١٩	٠,٨٤٣	٢٥	٠,٥٧٦
٢	٠,٤٨٦	٨	٠,٧٢٣	١٤	٠,٧٣٨	٢٠	٠,٧٥٤	٢٦	٠,٤٩٥
٣	٠,٤٩٣	٩	٠,٥٣٣	١٥	٠,٥٢٧	٢١	٠,٥٢٣	٢٧	٠,٦٥٦
٤	٠,٥٢٣	١٠	٠,٥	١٦	٠,٤٥٦	٢٢	٠,٣٩٧	٢٨	٠,٤٧٥
٥	٠,٥٠٩	١١	٠,٤٥٥	١٧	٠,٧١١	٢٣	٠,٤٥٣	٢٩	٠,٧٦٥
٦	٠,٦١١	١٢	٠,٧٣٨	١٨	٠,٤٦٣	٢٤	٠,٦٨٧	٣٠	٠,٥٦٣

ومن الجدول اعلاه نجد ان الفقرات جميعها مقبولة اذ ان تشبعها اعلى (٠,٣٠) وفقاً لمحك (جيلفورد)، وعليه لم تستبعد أي فقرة من فقرات الاختبار.

ب. افتراض الاستقلال الموضعي :

للتحقق من هذا الفرض استخدم الباحث برنامج (eirt) من خلال مؤشر (G^2) والنتائج اظهرت ان قيم ازواج الفقرات البالغة (٤٣٥) زوج تراوحت بين (٠,٠٠١ الى ١٨,٣٧٦) وعند مقارنتها بالقيمة الجدولية التي تبلغ (٣,٨٤) بدرجة حرية (١) وبمستوى دلالة (٠,٠٥) اتضح عدد الازواج المعتمدة والمستقلة موضعياً من خلال مؤشر (G^2) وكما موضحة بالجدول (٦) ادناه

جدول (٦) مؤشر إحصائي مربع (G^2 statistics)

حالات الاستقلال الموضعي	عدد الازواج	النسبة المئوية
معتمد	٣٧	%٩
مستقل	٣٩٨	%٩١
الكلي	٤٣٥	%١٠٠

ج. استخراج الصعوبة والتمييز لفقرات الاختبار والتحقق من ملائمة البيانات لانموذج السمات الكامنة :

استخدم برنامج بايلوك (Bilog – MG3) وباستخدام طريقة الارجحية القصوى الهامشية (MML) في دراسة مطابقة فقرات الاختبار للانموذج المستخدم في البحث الحالي ، وهو الانموذج الثنائي البارامتر، كما تم استخراج الصعوبة والتمييز لفقرات الاختبار وفيما يلي جدول (٧) يظهر قيم مربع كاي (كا^2) لفقرات الاختبار للحكم على مدى ملائمتها للانموذج ثنائي المعلم وقيم الصعوبة والقوة التمييزية لفقرات الاختبار.

جدول (٧) قيم مربع كاي لفقرات الاختبار للحكم على مدى ملائمتها للأنموذج ثنائي المعلم ومعامل الصعوبة والقوة التمييزية

تسلسل الفقرة	الصعوبة	التمييز	مربع كاي (كا)^2	درجة الحرية	الحكم	ت	الصعوبة	التمييز	مربع كاي (كا)^2	درجة الحرية	الحكم
١	-١,١٢٧	٠,٦٣٣	٥,٤	٨	غير دالة	١٦	-٠,٩٩٣	٠,٧١٠	١٢,٠	٩	غير دالة

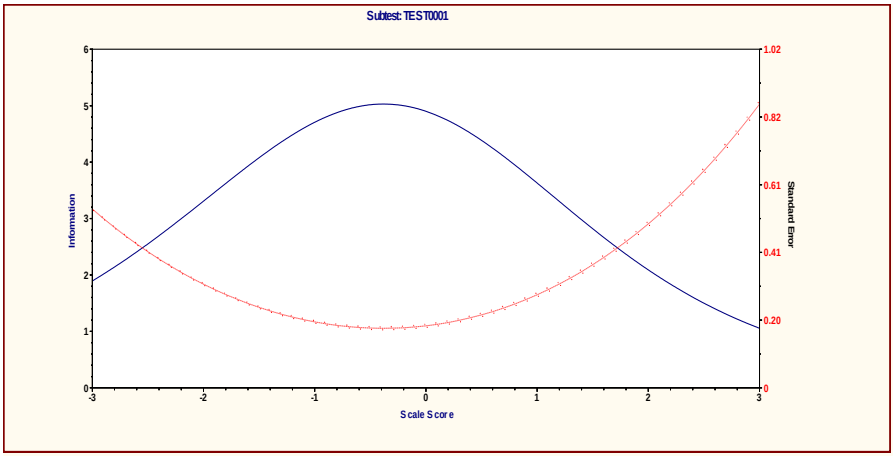
٢	٠,٠٥٠	٠,٧٤ ١	٩,٠	٩	غير دالة	١ ٧	– ٠,٨٢٤	٠,٥٧ ٣	٧,٤	٩	غير دالة
٣	٠,٢٧٢	٠,٧٦ ٢	١١, ٨	٩	غير دالة	١ ٨	– ٠,٩٤٤	٠,٦٤ ٣	٦,٦	٩	غير دالة
٤	– ٠,١١٧	٠,٨٧ ٨	٥,٨	٨	غير دالة	١ ٩	– ٠,٤٢٤	٠,٨١ ٢	٧,٠	٨	غير دالة
٥	– ٠,٩٠٠	٠,٧٣ ٢	٨,٣	٨	غير دالة	٢ ٠	– ٠,٠٩٨	٠,٦٩ ٣	١١, ٦	٩	غير دالة
٦	– ٠,٠١٤	٠,٧٦ ٥	١٠, ٩	٩	غير دالة	٢ ١	– ٠,٣٨١	٠,٥٠ ٠	١٤, ٤	٩	غير دالة
٧	– ٠,٦٩٠	٠,٦٦ ٥	١١, ٥	٩	غير دالة	٢ ٢	٠,١٣٩	٠,٦٩ ٤	١٠, ٩	٩	غير دالة
٨	– ٠,٤٠٠	٠,٦٤ ٧	٨,٥	٩	غير دالة	٢ ٣	– ٠,٣١٠	٠,٥٠ ٠	١٣, ٩	٩	غير دالة
٩	– ٠,٤٦٨	٠,٧٠ ٣	٦,٠	٨	غير دالة	٢ ٤	– ٠,٤٠٢	٠,٥٩ ٢	٦,٨	٨	غير دالة
١٠	– ٠,٠٩٨	٠,٦٩ ٨	٤,٧	٩	غير دالة	٢ ٥	– ٠,٣٠٨	٠,٥٧ ٧	٢,٠	٩	غير دالة
١١	– ٠,٤٣٤	٠,٧١ ٦	٧,٠	٩	غير دالة	٢ ٦	– ٠,٧١٤	٠,٦٣ ٧	٣,٦	٩	غير دالة
١٢	– ٠,٣٤٤	٠,٦٩ ٧	١١, ٦	٩	غير دالة	٢ ٧	– ٠,٦١٩	٠,٧٤ ٨	١٣, ٨	٩	غير دالة
١٣	– ٠,٧٤٧	٠,٥٧ ٢	١٤, ٨	٩	غير دالة	٢ ٨	– ٠,٣٩٧	٠,٧٣ ٨	١٤, ٢	٩	غير دالة
١٤	– ٠,٦٧٤	٠,٥٦ ٢	٤,١	٩	غير دالة	٢ ٩	– ٠,١٧٠	٠,٧٥ ٢	٨,٧	٩	غير دالة
١٥	– ٠,٧١٣	٠,٦٧ ٧	٣,٩	٩	غير دالة	٣ ٠	– ٠,٨٣٢	٠,٧٧ ٥	٩,٨	٩	غير دالة

يتضح من الجدول (٧) الى ان جميع الفقرات كانت ملائمة للانموذج الثنائي لان قيمها المحسوبة اصغر من قيمة مربع كاي الحرجة (١٥,٩٢، ١٥,٥١) عند درجة حرية مقدارها (٩، ٨) .

د. دالة معلومات الاختبار وخطأ القياس

تستخدم دالة معلومات الاختبار (Test Information Function) مؤشراً لثبات تقدير القدرة على الاختبار ، وقد اعتمد الباحث على دالة معلومات الاختبار في تحديد ثبات القدرة والشكل (٢) يوضح دالة معلومات الاختبار.

شكل (١) دالة المعلومات والخطأ المعياري للاختبار



يلاحظ من الشكل (١) إن القيمة العليا لدالة معلومات الاختبار بلغت (٦) وهي تناظر القيمة الدنيا للخطأ المعياري التي بلغت (١,٠٠)، وهاتان القيمتان تناظران القيمة (صفر) على متصل القدرة، حيث بلغ الخطأ المعياري للقياس (٠,٤٠٨)، وبلغت قيمة الثبات لتقدير القدرة (٠,٨٣٤) وفق دالة المعلومات.

الوسائل الاحصائية :

اولا : برنامج SPSS لاستخراج :

- التحليل العاملي الاستكشافي للتحقق من احادية البعد .
- معامل ارتباط بيرسون لحساب العلاقة بين معاملات الصعوبة والتمييز .

ثانيا : برنامج (BILOG – MG3) لاستخراج :

١- معامل صعوبة الفقرات .

٢- معامل تمييز الفقرات .

٣- اختبار مربع كاي لمطابقة الفقرات.

٤- درجة الحرية لكل مفردة .

ثالثا:- وسائل إحصائية، استعمها الباحث يدوياً، بمساعدة برنامج اكسل وكما يأتي:

- معامل الصعوبة = ١- معامل السهولة
- معامل التمييز
- معامل الثبات المستمد من الخطأ المعياري للقياس، ودالة المعلومات، بحسب المعادلة:

$$R = 1 - (SEE)^2 \ldots\ldots \text{or} \ldots\ldots R = 1 - (1/\ I(\theta))$$

عرض النتائج وتفسيرها:-

الفرضية الاولى : لا توجد علاقة ذو دلالة احصائية في معامل الصعوبة وفقاً لنظريتي القياس التقليدية والسمات الكامنة لفقرات اختبار الاستدلال المنطقي .

من اجل التحقق من هذه الفرضية تم تحليل اختبار الاستدلال المنطقي وفقاً لنظريتي القياس التقليدية ونظرية السمات الكامنة واستخراج معامل الصعوبة وفقاً لنظريتي القياس ومن ثم استخدام معامل ارتباط بيرسون لاستخراج العلاقة الارتباطية في الصعوبة بالنظريتين التقليدية والحديثة فكانت النتائج كما مبينة في الجدول(٨).

الجدول(٨)العلاقة بين معاملات الصعوبة بالنظريتين التقليدية والسمات الكامنة

عدد الفقرات	قيمة معامل الارتباط بين معاملات الصعوبة بالنظريتين	القيمة التائية		مستوى الدلالة (٠,٠٥)
		المحسوبة	الجدولية	
٣٠	٠,٩٠٧	١١,٧٧٩	٢,٠٤٨	دالة

من الجدول اعلاه يتضح ان العلاقة موجبة ودالة احصائياً عند مستوى دلالة (٠,٠٥) بين معاملات الصعوبة بالنظريتين التقليدية والسمات الكامنة وتتفق هذه النتيجة مع العديد من الدراسات منها دراسة (Magno,2009) ودراسة جمحاوي (٢٠٠٠) التي توصلتا الى وجود انسجام وعلاقة بين معاملات الصعوبة وفقاً لنظريتي القياس وايضاً اتفقت مع دراسة عودة

(١٩٩٢) التي اشارت نتائجها التوافق بين أنموذج راش والمؤشرات التقليدية في اختيار الفقرات.

الفرضية الثانية : لا توجد علاقة ذو دلالة احصائية في معامل التمييز وفقاً لنظريتي القياس التقليدية والسمات الكامنة لفقرات اختبار الاستدلال المنطقي .

من اجل التحقق من هذه الفرضية تم تحليل اختبار الاستدلال المنطقي وفقاً لنظريتي القياس التقليدية ونظرية السمات الكامنة واستخراج معامل التمييز وفقاً لنظريتي القياس ومن ثم استخدام معامل ارتباط بيرسون لاستخراج العلاقة الارتباطية في التمييز بالنظريتين التقليدية والحديثة فكانت النتائج كما مبينة في الجدول(٩).

الجدول(٩)العلاقة بين معاملات التمييز بالنظريتين التقليدية والسمات الكامنة

عدد الفقرات	قيمة معامل الارتباط بين معاملات التمييز بالنظريتين	القيمة التائية		مستوى الدلالة (٠,٠٥)
		المحسوبة	الجدولية	
٣٠	٠,٨١٣	٧,٣٩٠	٢,٠٤٨	دالة

من الجدول اعلاه يتضح ان العلاقة موجبة ودالة احصائياً عند مستوى دلالة (٠,٠٥) بين معاملات التمييز بالنظريتين التقليدية والسمات الكامنة وتتفق هذه النتيجة مع العديد من الدراسات منها دراسة (Magno,2009) ودراسة جمحاوي (٢٠٠٠) التي توصلتا الى وجود انسجام وعلاقة بين معاملات التمييز وفقاً لنظريتي القياس وايضاً اتفقت مع دراسة عودة (١٩٩٢) التي اشارت نتائجها التوافق بين أنموذج راش والمؤشرات التقليدية في اختيار الفقرات.

الاستنتاجات :

١. يتميز اختبار الاستدلال المنطقي بخصائص قياسية جيدة ومطابقة ممتازة وفقاً لنظريتي القياس التقليدية والسمات الكامنة.

٢. وجود علاقة موجبة ودالة احصائياً بين معاملات الصعوبة والتمييز وفقاً لنظريتي القياس تؤكد على استخدام اي من النظريتين في ايجاد الصعوبة والتمييز للاختبارات والمقاييس.

التوصيات:

١. استخدام اي من النظريتين في حساب الخصائص القياسية للاختبارات والمقاييس يعطي نتائج ممتازة ومتسقة.

٢. استخدام اختبار الاستدلال المنطقي كأداة بحث في دراسات مستقبلية لما يتماز به من خصائص قياسية ممتازة.

المقترحات :

١. اجراء دراسة لايجاد العلاقة في معاملات الثبات وفقاً لنظريتي القياس التقليدية والحديثة.

٢. اجراء دراسة لايجاد العلاقة في معاملات الصعوبة والتمييز للمقاييس النفسية.

قائمة المصادر :

- جروان ، فتحي عبد الرحمن ،(١٩٩٩) ، **تعليم التفكير مفاهيم_ وتطبيقات** ، الامارات ، دار الكتاب الجامعي .
- جمحاوي، ايناس (٢٠٠٠). **مقارنة خصائص الفقرات وفق النظرية التقليدية ونظرية الاستجابة للفقرة في مقياس للقدرة الرياضية** ، رسالة ماجستير غير منشورة ، جامعة اليرموك، الاردن.
- الحجازين ، نايل عيد (٢٠١٢) : **التفكير الاستدلالي** ، ط١ ، دار جليس الزمان ، عمان .
- الحجازين ، نايل عيد (٢٠١٢) : **التفكير الاستدلالي** ، ط١ ، دار جليس الزمان ، عمان.
- الدليمي، إحسان عليوي وعدنان محمود المهداوي (٢٠٠٥): **القياس والتقويم في العملية التعليمية**، كلية التربية/ ابن الهيثم، جامعة بغداد.
- سليمان، امين علي محمد و رجاء محمود ابو علام (٢٠١٠) : **القياس والتقويم في العلوم الانسانية اسسه وادوات وتطبيقاته**، الطبعة الاولى، دار الكتاب الحديث، القاهرة، مصر.

- صبري، وعد محمد نجاة (٢٠٠٢) : **أثر استخدام أنموذجي سكمانورايجلوث في التفكير الاستدلالي والتحصيل العلمي لدى طلاب الصف الخامس العلمي في مادة الفيزياء،** أطروحة دكتوراه غير منشورة، كلية التربية/ أبن الهيثم، جامعة بغداد.
- العبد الله، زياد أحمد (٢٠٠٩) : أثر حجم العينة وعدد المفردات المشتركة على تدريج مفردات بنك الأسئلة باستخدام نموذج التقدير الجزئي، رسالة ماجستير غير منشورة، كلية التربية، جامعة القاهرة.
- عبد الوهاب، محمد محمود محمد. (٢٠١٠) . **استخدام نماذج الاستجابة للمفردة الاختبارية في تدريج مفردات بعض الاختبارات المعرفية.** المنيا: مذكرة دكتوراه قياس نفسي ، غير منشور.
- عطيفة، حمدي ابو الفتوح والسرور ، عايدة عبد الحميد .(٢٠١١). **تعليم العلوم في ضوء ثقافة الجودة** . دار النشر للجامعات. القاهرة.
- علام ، صلاح الدين (٢٠٠٧): **القياس والتقويم التربوي في العملية التدريسية،** ط١، دار الميسرة للطباعة والنشر والتوزيع، عمان.
- علام ، صلاح الدين (٢٠٠٠). **القياس والتقويم التربوي والنفسي** ، دار الفكر العربي ، القاهرة – مصر .
- عودة، احمد سليمان : (١٩٩٣)، مدى التوافق بين أنموذج راش والمؤشرات التقليدية في اختبار مقياس اتجاه سباعي التدريج، **مجلة كلية التربية**، ع٨، الامارات.
- فرج، صفوت (١٩٩١). **التحليل العاملي في العلوم السلوكي،** ط٢، القاهرة، مكتبة الأنجلو المصرية.
- الفقي، عبد اللاه إبراهيم (٢٠١٤) : **الاحصاء التطبيقي باستخدام Spss،** الطبعة الاولى، دار الثقافة، عمان، الاردن.
- كروكر، ليندا، وألجينا، جيمز .(٢٠٠٩). **مدخل إلى نظرية القياس التقليدية والمعاصرة،** ترجمة: دعنا، زينات يوسف. ط١. دار الفكر ناشرون وموزعون.
- النبهان ، موسى (٢٠٠٤) : **أساسيات القياس في العلوم السلوكية** ، دار الشروق ، عمان .
- وزارة التربية ، (٢٠١١)، مراحل التعليم وأنواعه، المادة ٨، ب، والمادة ١١، ثانياً، ج.

- Allen. M, J & Yen, W.M. (1979) **Introduction to measurement theory**, California, Books/Cole Publishing company.
- Baker, F.B. (2001). **The Basics of Item Response Theory**. (2nd ED). College Park, MD: ERIC Clearing house on Assessment and Evaluation. (ERIC Document Reproduction Service No. ED 458 219).
- Baker, Frank, B. (2001): **The Basics of item responses Theory**, 2ed, ERIC Clearinghouse on Assessment and Evaluation.
- Caller, E.A. (1966) Studies in Psychomuietinc theory, **Journal of Experent of psychology**, Vol 9, No2, P271-298.
- Chaplain, J. (1974). **System & Theories of Psychology & Development**. New York. John Willy & Sons, Inc
- Crocker, L. & Algina, J. (1986). ***Introduction to Classical and Modern Test Theory***. NY. Holt, Rinhart & Winston.
- Dick , C. Hagert . Y. (1971) :**Topics in Measurement**. New York , Mc Graw . Hill .
- Fan, X. (1998): Item Response theory and class test theory An empirical comparison of their item/ person statistics. **Educational and psychological measurement**, Vol. 158, No. 3, pp. 357 – 381.
- Ghiselli, E.E. Campbell. J.P & Zedeck, S (1981) **Measurement Theory for the Behavioral Science**, San Francisco. W.H- Frehman & Company.
- Gulliksen, H. (1967): **Theory Mental Test**, New York. John Wiley & Sons. Inc.
- Hambleton, R.K. & Swaminathan, H. (1985). **Item Response Theory: principles and applications**. Boston, MA: Kluwer -Nijhoff.

- Hashway , R.M. (1977) : **A Comparison of test derived Using rasch and traditional Psychometric paradigms** (Dissertation Abstract International 38, 2, 744A.
- Lord, F.M. (1980). **Application of Item Response Theory to Practical Testing Problems**. New Jersey: Lawrence Erlbaum Associates, Publishers.
- Lyman. H.B. (1963): **Test Scores and what they Mean**. New Jersey, Prentic- Hall.Inc.
- Magno, Carlo (2009): Demonstrating the Difference between Classical Test Theory and Item Response Theory Using Derived Test Data, **The International Journal of Educational and Psychological Assessment**, pp. 1-11.
- Mislevy, R.J. & Bock, R.D. (1990). **Item analysis and test scoring with binary logistic models**. (2^{nd} edition). Scientific software, Inc.
- Nunnally, C. (1981) **Psychometric theory**, 2^{nd} ed, New Delhi, Tata McGrown-Hill Publishing Company LTD.
- Onder, 1. (2007). An investigation of goodness of model data fit model veriuyumunun arastirilmasi. ***Hacettepe Universitesi Egitim Fakültesi Dergisi, (32),*** 210-220.
- Reeve, Bryce B. (2003). **An Introduction to Modern Measurement Theory**. Outcomes Research Branch, Applied Research Program, National Cancer Institute.
- Warm, T. A. (1978). **A primer of Item Response Theory**. Oklahoma City : U.S. Coast Guard Institute.

- Wiberg, M. (2004). **Classical Test Theory vs. Item Response Theory, An evaluation of the theory test in the Swedish driving-license test**. UMEA University, EMNO 50, ISSN 1103-2685.

English Reference

- Jarwan, Fathi Abdul Rahman, (1999), teaching thinking concepts and applications , UAE, University Book House .
- Jemhawi, Ines (2000). Comparing the characteristics of paragraphs according to the traditional theory and the theory of responding to the paragraph in a measure of mathematical ability, unpublished master's thesis, Yarmouk University, Jordan.
- Ah al-hijazin, Nayel Eid (2012): heuristic thinking , Vol .1 , Dar Jalis Al-Zaman , Amman.
- Ah al-hijazin, Nayel Eid (2012): heuristic thinking , Vol.1 , Dar Jalis Al-Zaman , Amman.
- Al-Dulaimi, Ihsan Aliawi and Adnan Mahmoud al-Mahdawi (2005): measurement and evaluation in the educational process, Faculty of Education/ Ibn al-Haytham, University of Baghdad.
- Suleiman, Amin Ali Mohammed and Raja Mahmoud Abu Allam (2010): measurement and evaluation in the humanities, foundations, tools and applications, first edition, Dar Al-Kitab al-Hadith, Cairo, Egypt.
- Sabri, Waad Mohammed Najat (2002): the impact of using the scmanorigluth model on deductive thinking and educational achievement among fifth grade students in physics, unpublished doctoral dissertation, Faculty of Education/ Ibn al-Haytham, University of Baghdad.
- Al-Abdullah, Ziad Ahmed (2009): the impact of the sample size and the number of shared vocabulary on the graduation of the vocabulary of the question bank using the partial estimation model, unpublished master's thesis, Faculty of Education, Cairo University.
- This is Abd al-Wahhab, Mohamed Mohamed Mohamed. (2010) . Using the response model for the test vocabulary to include the vocabulary of some cognitive tests. Mona: psychological measurement PhD note, unpublished.
- Along with ataifa, Hamdi Abu Al-Fatouh and Al-Surour, Aida Abdel Hamid .(2011). Science education in the light of quality culture . University Publishing House. Cairo.
- Alam, Salah al-Din (2007): educational measurement and evaluation in the teaching process, i1, Al-Maysara house for printing, publishing and distribution, Amman.
- De Alam , Salahuddin (2000). Educational and psychological measurement and evaluation, Dar Al – Fikr Al-Arabi, Cairo-Egypt .

- Aouda, Ahmed Suleiman: (1993), the extent of compatibility between the rush model and traditional indicators in the test of a seven-step trend scale, Journal of the College of Education, P.8, UAE.
- Faraj, Safwat (1991). Factor analysis in behavioral sciences, i2, Cairo, Anglo-Egyptian library.
- elfeki, Abdallah Ibrahim (2014): applied statistics using Spss, first edition, House of culture, Amman, Jordan.
- Crocker, Linda, and Algina, James .(2009). Introduction to traditional and contemporary measurement theory, translation: let's, Zeinat Yusuf. I1. Dar Al-Fikr publishers and distributors.
- Al-Nabhan, Musa (2004): fundamentals of measurement in behavioral sciences , Dar Al-Shorouk, Amman .
- Ministry of Education, (2011), stages and types of education, Article 8, B, and Article 11, II, C.
- Allen. M, J & Yen, W.M. (1979) **Introduction to measurement theory**, California, Books/Cole Publishing company.
- Baker, F.B. (2001). **The Basics of Item Response Theory**. (2nd ED). College Park, MD: ERIC Clearing house on Assessment and Evaluation. (ERIC Document Reproduction Service No. ED 458 219).
- Baker, Frank, B. (2001): **The Basics of item responses Theory**, 2ed, ERIC Clearinghouse on Assessment and Evaluation.
- Caller, E.A. (1966) Studies in Psychomuietinc theory, **Journal of Experent of psychology**, Vol 9, No2, P271-298.
- Chaplain, J. (1974). **System & Theories of Psychology & Development**. New York. John Willy & Sons, Inc
- Crocker, L. & Algina, J. (1986). ***Introduction to Classical and Modern Test Theory***. NY. Holt, Rinhart & Winston.
- Dick , C. Hagert . Y. (1971) :**Topics in Measurement.** New York, Mc Graw . Hill .
- Fan, X. (1998): Item Response theory and class test theory An empirical comparison of their item/ person statistics. **Educational and psychological measurement**, Vol. 158, No. 3, pp. 357 – 381.
- Ghiselli, E.E. Campbell. J.P & Zedeck, S (1981) **Measurement Theory for the Behavioral Science**, San Francisco. W.H- Frehman & Company.
- Gulliksen, H. (1967): **Theory Mental Test**, New York. John Wiley & Sons. Inc.
- Hambleton, R.K. & Swaminathan, H. (1985). **Item Response Theory: principles and applications**. Boston, MA: Kluwer -Nijhoff.
- Hashway , R.M. (1977) : **A Comparison of test derived Using rasch and traditional Psychometric paradigms** (Dissertation Abstract International 38, 2, 744A.
- Lord, F.M. (1980). **Application of Item Response Theory to Practical Testing Problems**. New Jersey: Lawrence Erlbaum Associates, Publishers.

- Lyman. H.B. (1963): **Test Scores and what they Mean**. New Jersey, Prentic-Hall.Inc.
- Magno, Carlo (2009): Demonstrating the Difference between Classical Test Theory and Item Response Theory Using Derived Test Data, **The International Journal of Educational and Psychological Assessment**, pp. 1-11.
- Mislevy, R.J. & Bock, R.D. (1990). **Item analysis and test scoring with binary logistic models**. (2nd edition). Scientific software, Inc.
- Nunnally, C. (1981) **Psychometric theory**, 2nd ed, New Delhi, Tata McGrown-Hill Publishing Company LTD.
- Onder, 1. (2007). An investigation of goodness of model data fit model veriuyumunun arastirilmasi. ***Hacettepe Universitesi Egitim Fakültesi Dergisi, (32),*** 210-220.
- Reeve, Bryce B. (2003). **An Introduction to Modern Measurement Theory**. Outcomes Research Branch, Applied Research Program, National Cancer Institute.
- Warm, T. A. (1978). **A primer of Item Response Theory**. Oklahoma City : U.S. Coast Guard Institute.
- Wiberg, M. (2004). **Classical Test Theory vs. Item Response Theory, An evaluation of the theory test in the Swedish driving-license test**. UMEA University, EMNO 50, ISSN 1103-2685