جامعة بنغازي – كلية التّربية
مجلة كلية التربية ... العدد الثالث عشر ... أغسطس 2023

تدريج مفردات اختبار واطسون_جليسر للتفكير الناقد باستخدام نموذج راش

د. سليمان سعد صالح

قسم التربية الخاصة – كلية التربية – جامعة عمر المختار

Suliman.saad@omu.edu.ly

2023

الملخص

هدفت هذه الدراسة إلى تدريج مفردات اختبار واطسون_جليسر للتفكير الناقد باستخدام نموذج راش، ومن أجل تحقيق هذا الهدف تم تطبيق هذا الاختبار على عينة مكونة من (385) طالب وطالبة من طلبة كلية التربية جامعة عمر المختار، وتم تحليل البيانات باستخدام البرنامج الإحصائي Winsteps للتحقق من مطابقة مفردات الاختبار لمؤشرات المطابقة الداخلية والخارجية لنموذج راش، فأظهرت النتائج أن (14) مفردة من مفردات الاختبار المكون من (99) مفردة قد تجاوزت إحصاءات الملاءمة لها حدود الملاءمة المقبولة بالنموذج وهي أقل من (-2.5) أو أكبر من (+2.5)؛ وعليه تم الاحتفاظ بـ (85) مفردة اختبارية توافرت فيها افتراضات نموذج راش للقياس الموضوعي للسلوك، كما تم التحقق من إحصاءات الملاءمة الداخلية والخارجية لأفراد عينة الدراسة فأظهرت النتائج أن (23) طالباً من عينة الدراسة قد تجاوزت إحصاءات الملاءمة لهم الحدود المقبولة للنموذج وهي من أقل من (-2.5) أو أكبر من (+2.5)؛ وعليه تم حذف استجابات هؤلاء الأفراد غير المطابقين، وتم الاحتفاظ باستجابات (362) طالب استخدمت استجاباتهم لتدريج مفردات الاختبار في صيغته النهائية المكونة من (85) مفردة حسب صعوبتها المقاسة بوحدة اللوجيت (Logit)[1]، ثم حولت إلى وحدة المنف[2] بصفته نوعاً من المعايير المستخدمة في تفسير الدرجات حسب نماذج القياس الحديث، كما تم التحقق من ثبات الاختبار لكل من الأفراد والمفردات، فكانت معاملات الثبات (0.85) و(0.98) على التوالي، كما أن ثبات الاختبار يقدر حسب نموذج راش بمعامل الفصل للمفردات، والذي يرمز له بالرمز Gi[3]، وقد بلغ في هذه الدراسة (6.80)، فيما بلغ معامل الفصل للأفراد (2.39)، والذي يرمز له بالرمز Gp[4]، وهذه القيم لمعامل الفصل سواءً للمفردات أو الأفراد إذا زادت عن (2) يُعد دليلاً على أن عينة الأفراد كافية للفصل بين قيم صعوبات المفردات، وبالتالي تدل على أن مفردات الاختبار تتدرج هرمياً من حيث الصعوبة، وتدل أيضاً أن قدرات الأفراد تباينت بما يكفي لتفصل بين صعوبة مفردات اختبار التفكير الناقد لـ واطسون وجليسر، ومن ثم استطاعت تعريف متصل القدرة المراد قياسها.

---

[1] اللوجيت: وحدة قياس تستخدم في نموذج راش، وتُعّرف بأنها اللوغريتم الطبيعي لمرجح نجاح الفرد على المفردات التي تعبر نقطة صفر التدرج على معالم صعوبتها.

[2] المنف: نوع من المعايير تستخدم في نموذج راش، وهي تدريج مئوي متوسطه (50)، وانحرافه (5).

[3] معامل الفصل للمفردات وزيادة هذه القيمة عن (2) يعد دليلاً على قدرة المفردات على الفصل بين قدرات الأفراد.

[4] معامل الفصل للأفراد وزيادة هذه القيمة عن (2) يعد دليلاً على أن عينة الأفراد كافية للفصل بين قدرات الأفراد والمفردات.

الكلمات المفتاحية

– تدريج اختبار. – نموذج راش. – اختبار واطسون جليسر. – التفكير الناقد.

Abstract

This study aimed to Calibrator the Items of the Watson-Glaser test for critical thinking using the Rasch model. Matching the test items with the internal and external matching indicators of the Rasch model, the results showed that (14) of the test items consisting of (99) items had their suitability statistics exceeded the suitability limits accepted by the model, which are less than (-2.5) or greater than (+2.5); Accordingly, (85) test subjects were retained in which the assumptions of the Rasch model for the objective measurement of behavior were met, and the internal and external suitability statistics of the study sample members were verified, so the results showed that (23) individuals from the study sample had exceeded the suitability statistics for them the acceptable suitability limits of the model, which are less than (-2.5) or greater than (+2.5); Accordingly, the responses of those non-conforming individuals were deleted, and the responses of (362) Persons were retained. Their responses were used to Calibrator the items Test in its final form consisting of (85) items according to their difficulty measured by the logit unit, then converted to the Minf unit as a kind of criteria used in the interpretation of Scores, and the reliability of the test was verified for both Persons and items, so the stability coefficients were (0.85) and (0.98), respectively. the validity of the test items according to the Rasch model, it is estimated by the separation coefficient for the items, which is symbolized by Gi, and in this study it reached (6.80), while the separation coefficient for Persons was (2.39), which is symbolized by Gp. If it is more than (2), it is considered evidence that the sample of Persons is sufficient to separate the values of the Items difficulties, and thus indicates that the test items are hierarchically arranged by terms of difficulty, and also indicates that the persons capabilities varied enough to separate the difficulty of the items of the Watson and Glaser critical thinking test.

Key Words

– Calibration of Test. – Reach Model. – Watson-Glaser test. – Critical thinking.

المقدمة:

يُعد التفكير من العمليات العقلية التي حظيت باهتمام المختصين في مجال التربية والتعليم، حيث أصبحت النظم التربوية المعاصرة في كثيرٍ من بلدان العالم تعطي أولوية لتعليم التفكير بأنواعه المختلفة، كالتفكير الاستقرائي، والتفكير الاستدلالي، والتفكير التقاربي والتباعدي، والتفكير الإبداعي، والتفكير الناقد، فهدف التربية الرئيس هو تنمية عقليات مفكرة قادرة على جمع البيانات من مصادرها، وتحليلها للوصول إلى معلومات يمكن الركون إليها في اتخاذ القرارات لحل المشكلات والأزمات بأساليب علمية وموضوعية، وليس أدل على ذلك في رأي الباحث من قول ستيرنبرج Sternberg وهو أحد الرواد في مجال التفكير بقوله " أن المعارف مهمة للطالب بالطبع، ولكنها غالباً ما تصبح قديمة، أما طريقة التفكير حول المعرفة فتبقى متجددة أبداً، فهي التي تمكننا بعد اكتساب المعرفة ومن تمحيصها والاستدلال على صحتها أو بطلانها بغض النظر عن الزمان والمكان" (جروان، 2007: 16).

وعليه فالعملية التعليمية الناجحة يجب أن تركز على إكساب المتعلم مهارات التفكير الناقد التي تمكنه من مواجهة التطورات السريعة والمتلاحقة في الحياة المعاصرة وما ينجم عنها من مشكلات تحتاج إلى هذا النوع من التفكير، وبخاصة في ظل التطور الهائل في وسائل الاتصال التي تجعل الفرد في مواجهة مباشرة مع ثقافات وأفكار واتجاهات عديدة تتطلب منه أن يكون قادراً على فهمها وتمحيصها وإصدار الأحكام المناسبة عليها (السكري، 2010: 225).

ومن هنا أصبح التفكير الناقد من أبرز المهارات التي يحتاجها الطلبة في مواقف التعليم والتعلم، ويرى خبراء التفكير الناقد من أمثال واطسون وجليسر أن القدرة على النقد تتطلب من الفرد عقل متفتح ومرن، ومقدرة على الحكم على الأشياء، وقدرة على التلخيص والتحليل، ويتضمن التفكير الناقد العديد من المهارات أو المكونات أهمها: مهارة معرفة الافتراضات، مهارة التفسير، مهارة التقويم، ومهارة الاستنتاج، ومهارة الاستدلال (خماد، 2021: 239).

وبقدر أهمية التفكير الناقد تزداد الحاجة إلى وجود أدوات لقياسه، والتعرف على مكوناته، فما لا يمكن قياسه لا يمكن تحقيقه، وبمراجعة المحاولات لقياس التفكير الناقد بدأت (حسب علم الباحث) باختبار واطسون ـ جليسر عام 1964للتفكير الناقد، الذي يُعد من أهم أدوات قياس هذه القدرة، وحظي بانتشارٍ واسع في العديد من المجالات البحثية، ومجال الاختيار الوظيفي (عايش، 2021: 102)، كما حظي هذا الاختبار بالعديد من الدراسات للتحقق من خصائصها السيكومترية في ظل نظرية القياس الكلاسيكية، منها دراسة Loo &Thrope 1999، ودراسة Balouglue 2003، ودراسة العتيبي (2012)، وقد أيدت نتائج هذه الدراسات أن هذا الاختبار يتمتع بصدق وثبات

جيد، ألا أنّ ما يؤخذ أو يعاب على هذه الدراسات أن أساليب بنائها وتحليل مفرداتها تمت على أسس ومبادئ النظرية الكلاسيكية للقياس Classical Theory of Measurement، التي تعاني من عيوب كثيرة أهمها عدم خطية القياس، أي عدم وجود وحدة قياس موحدة، وعدم استقلالية القياس عن خصائص كل من مفردات أداة القياس وأفراد العينة (عايش، 2021: 103)، أيضاً تفترض النظرية الكلاسيكية أن درجة الفرد الملاحظة ما هي إلا مجموع درجات جميع مفردات الاختبار الذي يقيس متغيراً ما، فقد يحصل اثنان من المبحوثين على نفس الدرجة رغم اختلاف المفردات التي أجابا عليها، وقد يتساوى الفرق بين درجتي فردين منخفضين في القدرة موضع القياس، مع الفرق بين درجتي فردين مرتفعي القدرة التي يقيسها الاختبار رغم الاختلاف في صعوبة المفردات في الحالتين (Baker, 2001: 7)، ونظراً لهذه الانتقادات التي تعاني منها نظرية القياس التقليدية ظهرت نظرية الاستجابة للمفردة Item Response Theory بمزاياها في مجال القياس النفسي والتربوي، والتي أهمها استقلالية خصائص مفردات أداة القياس كمعاملات الصعوبة والتمييز عن عينة الأفراد المستخدمة استجاباتهم في التحليل، وكذا استقلال قدرات الأفراد عن عينة مفردات أداة القياس، أيضاً تفسير الدرجات في ضوء نظرية الاستجابة للمفردة يتم في ضوء المفردات وليس الجماعة المرجعية كما في نظرية القياس الكلاسيكية، كما أن تقدير الخطأ المعياري يقدر لكل فرد وليس قيمة واحدة لكل الأفراد (علام، 2005: 55).

هذه المزايا لنظرية الاستجابة للمفردة وغيرها أدت إلى استخدامها في مجال بناء وتطوير أدوات القياس النفسي والتربوي، والتأكد من مطابقة مفردات تلك الأدوات لافتراضات أحد نماذج هذه النظرية، حيث انبثق عن هذه النظرية مجموعة من النماذج لتناسب السمات والقدرات المختلفة المراد قياسها، وأنواع وخصائص المفردات المستخدمة في تلك الأدوات، فإذا كان المتغير المراد قياسه يقدر سمة أو قدرة واحدة فقط، استخدمت النماذج الأحادية البعد، أما إذا كان المتغير المراد قياسه مركب من سمات عدة أو قدرات فإنه تستخدم النماذج متعددة الأبعاد، أيضاً هذه النماذج تختلف حسب المفردات الاختبارية إذا كانت ثنائية الاستجابة Dichotomous، أو متعددة الاستجابة Polychotomous، كما تختلف نماذج هذه النظرية حسب عدد بارامترات أو معالم المفردات المراد التدريج في ضوئها كي تلاءم مفرداتها افتراضات النموذج، والتي منها النماذج أحادية البارامتر، وثنائية البارامتر، وثلاثية البارامتر (عبدالوهاب، 2010: 14).

وكون اختبار التفكير الناقد لـ "واطسون وجليسر" ثنائي الاستجابة تم استخدام نموذج راش أحادي المعلم في تدريج مفردات هذا الاختبار.

مشكلة الدراسة:

بناءً على ما سبق يمكن تحديد مشكلة الدراسة في التساؤل الرئيس الآتي:

هل يمكن تدريج مفردات اختبار واطسون جليسر للتفكير الناقد على ميزان واحد مشترك باستخدام نموذج راش بحيث تُعرِّف جميع تلك المفردات نفس المتغير (التفكير الناقد)؟

وتفرع عن هذا التساؤل الرئيس التساؤلات الفرعية الآتية:

1. ما مؤشرات حسن المطابقة (الداخلية والخارجية) الخاصة بمفردات اختبار واطسون _ جليسر للتفكير الناقد؟
2. ما مؤشرات حسن المطابقة (الداخلية والخارجية) الخاصة باستجابات عينة الدراسة على مفردات اختبار واطسون جليسر للتفكير الناقد؟
3. هل يمكن التحقق من شرط أُحادية البعد لاختبار واطسون _ جليسر للتفكير الناقد؟
4. ما قيم صدق وثبات اختبار واطسون _ جليسر للتفكير الناقد بعد تدريجه باستخدام نموذج راش؟
5. ما تقديرات صعوبة مفردات اختبار واطسون _ جليسر للتفكير الناقد؟
6. ما مستوى التفكير الناقد لدى طلبة كلية التربية كما يقاس باختبار واطسون_ جليسر للتفكير الناقد بعد تدريجه بنموذج راش.

أهمية الدراسة

تتحدد أهمية هذه الدراسة فيما يأتي:

1. تضيف هذه الدراسة إلى أدوات القياس اختباراً موضوعياً لقياس التفكير الناقد يتصف بصدق القياس وثباته.
2. إن الاستعانة بأحد نماذج القياس النفسي الحديث (نموذج راش) يعد خطوة أولى لتشجيع الباحثين في ليبيا لتبني هذا الاتجاه البحثي بما يثري الحركة البحثية في مجال القياس النفسي والتربوي.
3. إن ما يوفره نموذج راش من خطية القياس يتيح الفرصة لتقدير معالم المفردات ومعالم الأفراد بوحدة قياس معرفة، ما يحقق دقة القياس وموضوعيته.
4. إلقاء الضوء على متغير معرفي (التفكير الناقد) يُعد من أهم المتغيرات التي تهدف التربية الحديثة إلى تنميتها لدى المتعلمين، لا سيما في المرحلة الجامعية.

أهداف الدراسة

تهدف هذه الدراسة إلى تدريج مفردات اختبار واطسون جليسر للتفكير الناقد على ميزان واحد مشترك باستخدام نموذج راش وذلك من خلال التعرف على:

1. مؤشرات حسن المطابقة (الداخلية والخارجية) الخاصة بمفردات اختبار واطسون _ جليسر للتفكير الناقد.
2. مؤشرات حسن المطابقة (الداخلية والخارجية) الخاصة باستجابات عينة الدراسة على مفردات اختبار واطسون جليسر للتفكير الناقد.
3. أُحادية البعد لفقرات لاختبار واطسون _ جليسر للتفكير الناقد.
4. صدق وثبات اختبار واطسون _ جليسر للتفكير الناقد بعد تدريجه باستخدام نموذج راش.
5. تقديرات صعوبة مفردات اختبار واطسون _ جليسر للتفكير الناقد.
6. مستوى التفكير الناقد لدى طلبة كلية التربية كما يقاس باختبار واطسون_ جليسر للتفكير الناقد بعد تدريجه بنموذج راش.

مصطلحات الدراسة:

تدريج المفردات: Item Calibration

تدريج مفردات الاختبار يعني مدى انتظام تلك المفردات حسب صعوبتها على متصل تعريف المتغير موضع القياس، بحيث يعبر موقع كل مفردة عن مستوى صعوبتها (الخرشة، 2015: 91).

الاختبار: Test

الاختبار مجموعة من المثيرات التي تقدم للفرد لاستثارة استجابات تكون أساساً لإعطاء الفرد درجة رقمية، هذه الدرجة القائمة على عينة ممثلة لسلوك الفرد، تعد مؤشراً للقدر الذي يمتلكه الفرد من الخاصية التي يقيسها الاختبار (سليمان، 2010: 20).

التفكير الناقد: Critical Thinking

التفكير الناقد كما يرى كل من واطسون _ وجليسر "هو فاعلية المفكر في فحص المعتقدات والمقترحات في ضوء الشواهد التي تؤيدها، والحقائق المتصلة بها، بدلاً من القفز إلى النتيجة على نحوٍ غير ناضج، ويتطلب التفكير الناقد قدرة الفرد على فهم اللغة واستخدامها في عملية اتصال دقيقة، وتفكير مميز، مع إدراك العلاقات المنطقية بين القضايا، وكذلك القدرة على تفسير البيانات

واستخلاص النتائج والتعميمات السليمة، وتقويم الشواهد والأدلة والتعرف على المسلمات، وتقويم الأحكام والحجج (عبد السلام، 2020: 45).

تعريف التفكير الناقد إجرائياً: يقصد بالتفكير الناقد في هذه الدراسة هو الدرجة التي يتحصل عليها أحد أفراد العينة من استجاباته على مفردات اختبار واطسون_ جليسر بعد تدريج تلك المفردات باستخدام نموذج راش.

نموذج راش: Rasch Model

هو نموذج رياضي يفترض أن استجابة فرد لمفردة اختبار ما، يُحدث تفاعل بين القدرة أو السمة الكامنة لدى الفرد، وبين صعوبة تلك المفردة، ومن خلال هذا التفاعل يتضح الجانب النفسي في نموذج راش في اعتماده على قضيتين: الأولى احتمال أن يجيب الفرد إجابة مناسبة على مفردة سهلة هو أكبر من احتمال أن يجيب إجابة مناسبة على مفردة صعبة نسبياً. القضية الثانية، يزداد احتمال الإجابة المناسبة على المفردة بزيادة مستوى القدرة على الفرد (وهدي، 2010: 26).

حدود الدراسة:

الحدود الموضوعية: تقتصر الدراسة على تدريج مفردات اختبار واطسون _ جليسر للتفكير الناقد باستخدام نموذج راش أحادي المَعْلم.

الحدود المكانية: أجريت الدراسة بكلية التربية جامعة عمر المختار.

الحدود البشرية: اقتصرت الدراسة على عينة مكونة من (385) طالب وطالبة من الطلبة الدارسين بكلية التربية جامعة عمر المختار.

الحدود **الزمنية** : أجريت الدراسة خلال الفصل الدراسي الأول للعام الجامعي (2022 / 2023).

التفكير الناقد: Critical Thinking

التفكير بشكلٍ عام عملية نمائية شأنها شأن أي مظهر من مظاهر النمو الأخرى عند الإنسان، تعتمد هذه العملية على التدريب والخبرات التي يمر بها الفرد خلال تفاعله مع ما يحيط به من أحياء وأشياء، والتفكير الناقد ليس موجوداً بالفطرة عند الإنسان، فمهاراته متعلمة، وتحتاج إلى تدريب ومران،ويعد التفكير الناقد مطلباً رئيساً لجميع فئات المجتمع، فالفرد الذي يمتلك هذه القدرة يكون مستقلاً في تفكيره، ومراقباً له ومتحرراً من التبعية، وقادراً على اتخاذ القرارات الصحيحة في حياته (شنطاوي، 2003: 15).

وكان التفكير الناقد محور اهتمام الفلاسفة والتربويين منذ أقدم العصور، إذ يمكن تتبع جذوره في أكاديمية أفلاطون في القرن الرابع قبل الميلاد، وفي بداية القرن العشرين جعل جون ديوي تطوير قدرة الفرد على التفكير الناقد هدفاً أساسياً للتربية، أما حديثاً فقد أدى الاهتمام بهذه النوع من التفكير في الولايات المتحدة إلى ظهور العديد العلماء والمنظرين الذين تناولوا هذا الموضوع بالدراسة والبحث للتعرف على ماهيته، ومكوناته باستخدام أساليب قياس موضوعية ومقننة (السيد، 1995: 54)، ومن هؤلاء العلماء على سبيل المثال روبرت إنّس Robert Ennis الذي يعد أحد أشهر من أسهمت في تطوير دراسة التفكير الناقد، وقدم تعريفاً للتفكير الناقد اكتسب شهرة واسعة، وهو أن التفكير الناقد "تفكير منطقي، تأملي، يرتكز على اتخاذ قرارٍ عما يجب اعتقاده أو عمله"، أيضاً من العلماء الذين طوروا التفكير الناقد ريتشارد بول Richard Paul والذي قدم تعريفاً للتفكير الناقد يختلف عن سلفه إنس Ennis، حيث عرّف التفكير الناقد بأنه "تفكير في أي موضوع، أو فحوى، أو مشكلة، حيث يحسن المفكر نوع تفكيره بعنايته الماهرة ببنية التفكير الموروث وتطبيق معايير عقليه عليه (آلك فِشر، 2009: 18).

أما واطسون وجليسر Watson & Glaser فقدما تعريفاً تضمن مكونات التفكير الناقد حسب ما يقاس باختبارهما ذائع الصيت، حيث عرفا التفكير الناقد بأنه "تفكير مركب يتضمن مهارات واتجاهات ومعارف تشتمل على قدرة الفرد على تمييز المشكلات وقبول التعميمات في حالة وجود أدلة وبراهين منطقية تدعمها، ومعرفة مناهج التقصي المنطقي والاستدلالي، والتجريد، والتعميم بغية تحديد قيمة تلك الأدلة والبراهين ومدى صحتها، فضلاً عن الكفاءة في تطبيق تلك الاتجاهات والمعارف (Watson & Glaser, 2008: 16)).

وفي ضوء التعريف السابق تتضح بنية ومكونات التفكير الناقد كما كشف عنها دراسات واطسون وجليسر باستخدام اختبارهما لقياس هذا النمط من التفكير على النحو التالي:

1. الاستنتاج: Inference

الاستنتاج يشير إلى قدرة الفرد على استخلاص نتيجة من حقائق معينة ملاحظة أو مفترضة، ويكون لديه القدرة على إدراك صحة النتيجة أو خطئها في ضوء الحقائق المعطاة.

2. تمييز الافتراضات: Recognition Assumption

ويشير إلى القدرة على التمييز بين درجة صدق معلومات محددة أو عدم صدقها، والتمييز بين الحقيقة والرأي، والغرض من المعلومات المعطاة.

3. الاستنباط: Deduction

يشير إلى قدرة الفرد على تحديد بعض النتائج المترتبة على مقدمات، أو معلومات سابقة لها.

4. التفسير: Interpretation

يعني القدرة على تحديد المشكلة، والتعرف على التفسيرات المنطقية، وتقرير ما إذا كانت التعميمات والنتائج المبنية على معلومات معينة مقبولة أم لا.

5. تقويم الحجج: Evaluation of Arguments

يعني قدرة الفرد على تقويم الفكرة، وقبولها أو رفضها، والتمييز بين المصادر الأساسية والثانوية، والحجج القوية والضعيفة، وإصدار الحكم على مدى كفاية المعلومات (العتيبي، 2012: 1431).

نموذج راش: Rasch Model

قام جورج راش ببناء نموذج رياضي حقق به العلاقة بين قدرة الفرد (bi)، وصعوبة المفردة (δi) والاستجابة الملاحظة (Xvi)، كما حقق به متطلبات القياس الموضوعي للسلوك، ويعد هذا النموذج أبسط نماذج نظرية الاستجابة للمفردة أحادية البعد، ويقوم هذا النموذج على تقدير علاقة احتمالية بين قدرة الفرد وصعوبة المفردة، فعندما يجيب الفرد (v)على المفردة (i) فإن أرجحية حدوث أي من الاستجابتين (نجاح / فشل) يعتمد فقط على المتغيرات الأساسية المؤثرة في الاستجابة، وهما قدرة الفرد وصعوبة المفردة التي يحددها المقابل اللوغارتمي للفرق بين هذين المعْلَمين (bi –δi)، ويُعد المقدار exp (bi – δi) مميزاً أو مرجحاً للنجاح في حالة bi>δi (أي زيادة قدرة الفرد عن صعوبة المفردة) وعندئذٍ يكون احتمال الإجابة الصواب أكبر من 50%، ومن الممكن اعتبار هذا المقدار exp (bi – δi) مميزاً أو مرجحاً للفشل في حال bi<δi عندها يكون احتمال الإجابة الخطأ أكبر من 50%، أما إذا تساوت قدرة الفرد مع صعوبة المفردة أي bi = δi فإن احتمال حدوث الاستجابة الصحيحة يكون 50% (أمينة كاظم، 1996: 446).

الصيغة الرياضية لنموذج راش.

الصيغة الرياضية العامة لنموذج راش، وهي الصورة الأكثر ألفة من نماذج القياس التي ترجع لجورج راش وهي كما في المعادلة التالية:

$$\boldsymbol{Pi(\theta)} = \frac{\mathbf{eD.(\theta - bi)]}}{\mathbf{1 + eD.(\theta - bi)}}$$

($\boldsymbol{Pi}$)=احتمال الإجابة الصحيحة لفرد قدرته ($\boldsymbol{\theta}$) عن مفردة (i)حيث إن (i) تعني رقم المفردة (1 ، 2 ، 3 ، ....).

(e)= تعني الأساس اللوغاريتمي وقيمته (2.718).

(D) = عامل التدريج ويقدر بـ (1.7).

(θ) = معلم قدرة الفرد.

(bi) = معلم صعوبة المفردة (علاونة، 2016: 198)

الدراسات السابقة:

مراجعة الباحث للأدبيات الخاصة باستخدام نموذج راش في تدريج فقرات اختبار واطسون _ جليسر للتفكير الناقد بينت ندرة الدراسات في هذا المجال، لذلك تم عرض الدراسات التي اهتمت بدراسة هذا الاختبار سواءً باستخدام النظرية التقليدية في القياس أم النظرية الحديثة ممثلة في نموذج راش.

دراسة لو وثروب Loo & Thrope 1999

هدفت الدراسة إلى التعرف على الخصائص السيكومترية لاختبار واطسون _ جليسر، وتكونت عينة الدراسة من 271 طالب وطالبة من طلبة الجامعة بكندا، وأشارت النتائج أن المتوسطات الحسابية لدى عينة الدراسة أقل من المتوسطات الحسابية في دليل الاختبار (Manual)، مما يدل على أن مستوى التفكير الناقد لدى الطلبة الذين شملتهم الدراسة أقل من المتوسط الفرضي للاختبار، كما بينت النتائج أن قيم معاملات الارتباط الداخلية بين الاختبارات الفرعية الخمسة قد تراوحت بين (0.53) و(0.64)، وهي معاملات ارتباط دالة إحصائياً ما يعد مؤشراً على الاتساق الداخلي للاختبار، ولم تتوصل الدراسة إلى فروق في التفكير الناقد تعزى لمتغيري النوع والتخصص الأكاديمي (الزغبي، 2017: 61).

دراسة واجنر وهارفي 2003 Wagner & Harvey

هدفت الدراسة إلأى تطوير أداة جديدة لتقييم التفكير الناقد باستخدام نظرية الاستجابة للفقرة، وذلك باستخدام أوجه التفكير الناقد لاختبار واطسون_جليسر، وتكونت عينة الدراسة من 407 طالب وطالبة من الطلبة الجامعة من بينهم 279 طالبة، و128 طالب، تراوحت أعمارهم بين 18 – 27 سنة، تم تطبيق اختبار واطسون_جليسر، بالإضافة للاختبار المطور، وقد أشارت النتائج إلى وجود علاقة بين أداة المفحوصين على الاختبارين ما يعد مؤشراً على تمتع الاختبار المطور بالصدق التلازمي، كذلك تم مقارنة المجموعات الطرفية بوصفه نوعاً من الصدق التمييزي للاختبار المطور، فأظهرت النتائج قدرة الاختبار على التمييز طرفي القدرة موضع القياس (التفكير الناقد).

دراسة الجبوري 2017:

هدفت هذه الدراسة إلى إعداد اختبار واطسون – جليسر للتفكير الناقد على وفق نظرية القياس التقليدية ونظرية السمات الكامنة وحساب الدرجة الحقيقية للاختبار باستعمال معادلات الدرجات

الحقيقية للنظريتين، ولتحقيق هذا الهدف تم ترجمة صورة اختبار واطسون – جليسر للتفكير الناقد الأصلية التي أعدها واطسون _ وجليسر Watson & Glaser، طبق الباحث الاختبار بصيغته النهائية على عينة عشوائية مكونة من (400) من طلبة جامعة بابل، واعتمد الباحث في تحليل فقرات الاختبار وحساب الدرجة الحقيقية وفقاً للنظرية التقليدية للقياس، وأيضاً وفقاً لنموذج راش احد نماذج نظرية القياس الحديثة، الذي اتخذ نموذجاً لتحليل الفقرات وتم التوصل إلى أن طرق تكييف وتطوير الاختبار وفقا لنظرية القياس الحديثة هي الأكثر موضوعية من خطوات نظرية القياس التقليدية لكونها تتعامل بياناتها مع فقرات الاختبار باحتساب قدرة الأفراد وصعوبة الفقرات وهذا يدل على استقلالية القياس بما يحقق موضوعية القياس كما تمثل في نموذج راش، أيضاً الدرجة الحقيقية المحسوبة وفقا لمعادلة نموذج راش كانت الأكثر دقة ووضوح من الدرجة الحقيقية المحسوبة وفق لمعادلة النظرية التقليدية، كما تم التحقق من أحادية البعد؛ والتي تعني أن جميع مفردات الاختبار تعرف متغيراً محدداً، هو التفكير الناقد كما يقس بهذا الاختبار، كما أوضحت النتائج أن مستوى التفكير الناقد لدى طلبة جامعة بابل كما يقاس باختبار واطسون_ جليسر دون المتوسط حين قورن بالمتوسط النظري للاختبار.

دراسة الزعبي 2017:

هدفت هذه الدراسة إلى بناء اختبار لمهارات التفكير الناقد باستخدام نظرية الاستجابة للفقرة، ولتحقيق هذا الهدف تم تطبيق اختبار واطسون _ جليسر المكون من (99) فقرة تغطي مهارات التفكير الناقد، ومن خلال التحليل باستخدام البرنامج الإحصائي Winsteps تم الاحتفاظ (89) فقرة، وطبق الاختبار على عينة مكونة من (444) طالب وطالبة في المرحلة الجامعية، وتم التحقق من افتراض أحادية البعد والصدق والثبات، كما تم التحقق من مطابقة الاستجابات عن فقرات الاختبار لنموذج الاستجابة للفقرة المستخدم في الدراسة، حيث اتصفت معالم الفقرات ومعلم القدرة بخاصية اللاتغاير، وقد قدرت معالم الفقرات ودالة المعلومات لها، وبحساب المطابقة الداخلية والمطابقة الخارجية تم حذف (10) فقرات من الاختبار كانت بيانات الملاءمة لها أكبر من (+2) أو أصغر من (−2.5)، أما بخصوص إحصاءات المطابقة الداخلية والخارجية للأفراد فبينت النتائج أن جميع الأفراد مطابقين لحدود الملاءمة الخاصة بالنموذج، وعليه لم يتم حذف أي فرد من أفراد عينة الدراسة.

دراسة شنان 2019:

هدفت الدراسة إلى الكشف عن بعض العوامل المؤثرة كحجم العينة المستخدمة في التدريج، وأفضل طريقة للربط بين الصور الاختبارية المختلفة، مفردات مشتركة/ أفراد مشتركين، وأفضل عدد للمفردات المشتركة أو الأفراد المشتركين المستخدمة في ربط الصور الاختبارية المختلفة،

بالإضافة إلى أثر مستوى العينة المستخدمة في تدريج المفردات، والتي تؤثر على دقة القياس، وذلك للتوصل إلى أفضل تدريج يوفره نموذج راش، لاختبار القدرة على التفكير الناقد، كما هدفت الدراسة إلى عمل معايير مئينية تفسر على أساسها قدرة الأفراد على الاختبار، وتم استخدام برنامج Winsteps في تدريج فقرات اختبار القدرة على التفكير الناقد، ووفقاً لنموذج راش أجريت الدراسة على عينة مكونة من (829) طالب وطالبة من طلبة المرحلة المتوسطة، وأظهرت النتائج أن أفضل حجم عينة لتدريج مفردات اختبار التفكير الناقد باستخدام نموذج راش هو حجم (800)، وذلك تبعاً لمحكات دقة القياس، وأن أفضل عدد للمفردات المشتركة المستخدمة في ربط الصور الاختبارية المختلفة هو العدد (11) مفردة اختبارية، وأفضل عدد للأفراد المشتركين المستخدمين لربط الصور الاختبارية هو العدد (50)، وذلك تبعاً لمحكات دقة القياس، وأن أفضل طريقة للربط بين الصور الاختبارية هي طريقة المفردات المشتركة، وذلك تبعاً لمحكات دقة القياس فقد اختلفت دقة القياس باختلاف مستوى العينة، وذلك لصالح مستوى العينة منخفضة المستوى، كما تحققت في اختبار التفكير الناقد استقلالية القياس (بمعنى دقة التعادل)، وتم تدريج وتقدير صعوبة مفردات اختبار التفكير الناقد وفقاً لنموذج راش، وذلك بعد توافر محكات دقة وكذلك تقدير قدرة الأفراد المقابلة لكل درجة كلية خام محتملة على اختبار التفكير الناقد وفقاً لنموذج راش، وتم التوصل إلى معايير الرتب المئينية والدرجات التائية التي تفسر على أساسها قدرة الأفراد على الاختبار.

دراسة عايش 2021:

هدفت هذه الدراسة إلى استخدام نموذج راش في تدريج اختبار التفكير الناقد (اختبار واطسون. جليسر للتفكير الناقد، وتكونت عينة الدراسة من (231) طالب وطالبة من جامعتي سعيدة والشلف بالجزائر، وتم استخدام برنامج التحليل الإحصائي Winsteps، وبينت النتائج أن افتراضات نموذج راش قد تحققت، كما أظهرت نتائج تدريج الاختبار عدم استبعاد أي فقرة من فقرات الاختبار، وتوصلت النتائج إلى تمتع الاختبار بخصائص سيكومترية مقبولة، حيث تم التوصل إلى أن قيم الملاءمة الداخلية والخارجية لبيانات عينة الدراسة تحقق افتراضات حسن المطابقة لنموذج راش، وتمتع الاختبار بدرجة صدق وثبات مقبولة، وبينت النتائج الاختبار التائي لعينة واحدة أن مستوى التفكير الناقد لدى أفراد عينة الدراسة أقل من المتوسط النظري للمقياس.

خلاصة الدراسات السابقة

يبدو أن الدراسات التي أجريت في الآونة الأخيرة في بلدان عدة، واهتمت بأدوات قياس التفكير الناقد اتخذت من نظرية الاستجابة للمفردة منهجاً لها في التحقق من صلاحية تلك الأدوات، ومن هذا المنطلق أيضاً تبدو الحاجة محلياً إلى تدريج مفردات اختبار واطسون _ جليسر لقياس التفكير

الناقد على عينة من طلبة كلية التربية جامعة عمر المختار، باستخدام أحد نماذج نظرية الاستجابة للمفردة (نموذج راش)، والتحقق من مطابقة بيانات العينة لافتراضات هذا النموذج.

إجراءات الدراسة

منهج الدراسة:

لتحقيق أهداف الدراسة استخدم الباحث المنهج الوصفي للتعرف على مدى مطابقة مفردات اختبار واطسون _ جليسر للتفكير الناقد لافتراضات نموذج راش، حيث يهتم هذا المنهج بدراسة الظاهرة ووصفها وتفسيرها وتصنيف البيانات والتعبير عنها كمياً وتنظيمها للتوصل إلى استنتاجات وتعميمات قد تساعد في تدريج مفردات هذا الاختبار حسب صعوبة كل منها.

مجتمع الدراسة:

تكون مجتمع هذه الدراسة من جميع طلبة كلية التربية جامعة عمر المختار خلال العام الدراسي 2022 / 2023 البالغ عددهم 1890 طالب منهم (740) طالب و(1150) طالبة.

عينة الدراسة:

العينة الاستطلاعية: تكونت العينة الاستطلاعية من (50) طالباً وطالبة اختيروا بطريقة عشوائية وطبق عليهم اختبار واطسون –جليسر للتفكير الناقد بهدف التعرّف على مدى فهم أفراد العينة لمفردات الاختبار (أداة الدراسة)، ومن خلال التطبيق الاستطلاعي تبين وضوح جميع مفردات الاختبار لدى أفراد العينة، وعليه بدأ الباحث بتطبيق الاختبار على عينة الدراسة الأساسية.

العينة الأساسية: وتكونت من (385) طالب وطالبة اختيروا بطريقة عشوائية، وقد تم تحديد حجم العينة المناسبة لهذا المجتمع باستخدام معادلة كرجسي ومورجان Kergcie Morgan (عزت عبدالحميد، 2011: 532).

$$n = \frac{x^2\ p\ (1-p).N}{d(N-1) + x^2\ p\ (1-p)} = \frac{3.38 \times\ 0.25 \times 1890}{0.0025(1889) + \ 3.38 \times\ 0.25}$$

$$n = \frac{0.96025 \times 1890}{2.9975 + \ 0.96025} = \frac{1814.87}{4.72}$$

$$n = 384.5 \approx 385$$

أداة الدراسة:

اختبار واطسون _ جليسر للتفكير الناقد:

يهدف هذا الاختبار لقياس قدرة الفرد على مهارات التفكير الناقد، وهي (الاستنتاج، التعرف على الافتراضات، الاستنباط، التفسير، تقويم الحجج)، وقد روعي في وضع مفردات الاختبار ألا ترتبط بمقررات دراسية معينة، وقد وضع الاختبار خصيصاً لقياس مهارات التفكير الناقد لمرحلتي التعليم الثانوي والجامعي، ويتكون اختبار التفكير الناقد من (99) مفردة، موزعة على خمسة اختبارات فرعية، يحتوي كل اختبار فرعي على سؤال تدريبي لا تُحْسبُ درجته في تقدير درجة المفحوص، والاختبارات هي:

الاختبار الأول – الاستنتاج: يتكون من 20 مفردة، صممت لقياس القدرة على التمييز بين الدرجات المختلفة من الصدق والكذب، واحتمال التوصل إلى استنتاجات معينة على أساس مقدمات وبيانات متضمنة في قضايا معطاة.

الاختبار الثاني – التعرف على الافتراضات: يتكون من 16 مفردة صممت لقياس القدرة فيالتعرف على الافتراضات متضمنة قضايا معطاة.

الاختبار الثالث – الاستنباط: يتكون من 25 مفردة صممت لقياس القدرة على التفكير الاستنباطي من مقدمات معينة، وللتعرف على العلاقة بين قضيتين، ولتحديد ما إذا كان يبدو الاستنتاج الذي تم الوصول إليه مترتباً بالضرورة على قضية ما.

الاختبار الرابع– التفسير: يتكون من 24 مفردة صممت لقياس القدرة على وزن الأدلة، والتمييز بين التعليمات الواردة وغير الواردة.

الاختبار الخامس– تقويم الحجج: يتكون من 14 مفردة صممت لقياس القدرة على التمييز بين الحجج القوية والمهمة بالنسبة للسؤال المطروح، والحجج الضعيفة وغير ذات الصلة.

تصحيح الاختبار:

تعطى درجة واحدة عن كل مفردة تتم إجابتها إجابة صحيحة، والدرجة صفر عن كل مفردة إجابتها خاطئة، فبدائل الإجابة في هذا الاختبار ثنائية (صح / خطأ)، وتجمع درجات الاختبارات الخمسة لنحصل على درجة المفحوص الكلية للاختبار، وبذلك تكون الدرجة النهائية للاختبار 99 درجة، يحصل عليها المفحوص إذا أجاب إجابات صحيحة عن جميع مفردات الاختبار، وأقل درجة يمكن أن يحصل عليها المفحوص هي صفر، وكلما زادت درجة الفرد في الاختبار كان ذلك دليلاً على قدرته على التفكير الناقد (صالح، 2011: 66، 67).

برامج التحليل الإحصائي:

1- برنامج الحزمة الإحصائية للعلوم الاجتماعية SPSS.

2- برنامج التحليل Winsteps

استخدم الباحث في هذه الدراسة برنامج Winsteps لتحليل بيانات عينة تدريج مفردات اختبار واطسون _جليسر للتفكير الناقد لكون هذا البرنامج من أحدث برامج التحليل باستخدام نموذج راش وله مميزات عديدة منها ما يلي:

أ. يستخدم البرنامج في تدريج بيانات أنواع مختلفة من الاختبارات والمقاييس كالاختبارات ثنائية الاستجابة، وهو ما يستخدمه الباحث في هذه الدراسة، كما يمكن استخدام هذا البرنامج في الجمع بين أكثر من نوع من بيانات تلك الاختبارات في نفس التحليل.

ب. يقدم هذا البرنامج وصفاً دقيقاً للمفردات وخصائص الأفراد، ويحتوي على أدوات قوية للتشخيص.

ج. للبرنامج القدرة على تحديد الأبعاد، وذلك من خلال تحليل المكونات الأساسية للبواقي Principal Component Analysis of Residuals، الذي يقوم بكشف وتكميم الأبعاد الكامنة في بيانات التحليل.

د. يقوم البرنامج بتحديد مواضع وعرض أنماط الاستجابات غير المتوقعة بطرق عديدة.

هـ. للبرنامج القدرة على إجراء تحليلات لعدد كبير من الأفراد يصل إلى مليون فرد، و(30) ألف مفردة متعددة الاستجابة يمكن أن يصل مداه إلى (255) فئة استجابة.

و. يعطى البرنامج مخرجات ثرية تتسم بالوضوح والبساطة، وتتضمن العديد من الجداول والمخططات والرسوم البيانية التي يستعان بها في تفسير النتائج، وتشخيص المشكلات، وتشمل أهم المخرجات جداول لإحصاءات فئات الاستجابة(Linacre, J.M. 2018).

نتائج الدراسة ومناقشتها:

تم عرض النتائج حسب تساؤلات الدراسة كما يأتي:

أولاً: عرض نتائج إجابة التساؤل الأول الذي نصه:

ما مؤشرات حسن المطابقة (الداخلية والخارجية) الخاصة بمفردات اختبار واطسون _ جليسر للتفكير الناقد؟

للتعرف على مؤشرات حسن مطابقة مفرداتItem Fit اختبار واطسون _جليسر للتفكير الناقد لنموذج راش تم استخدام برنامج التحليل الإحصائي Winstepsلتحديد صعوبة كل مفردة من

مفردات هذا الاختبار، كما تم حساب الخطأ المعياري الحقيقي Real S.E لكل مفردة، وأيضاً تم حساب إحصاء المطابقة الداخلية (INFIT)، وإحصاء المطابقة الخارجية (OUTFIT)، لمعالم الصعوبة، والجدول التالي يبين المتوسطات الحسابية والانحرافات المعيارية لكل تقدير تقديرات الصعوبة، والخطأ المعياري في تقديرات صعوبات المفردات، ومتوسطات المربعات (MNSQ) لإحصائي المطابقة الداخلية والخارجية، وقيم إحصاءات المطابقة (ZSTD) لإحصائي المطابقة الداخلية والخارجية لمفردات الاختبار، تبين أن (14) مفردة من مفردات الاختبار المكون من (99)، أنهم غير مطابقين للنموذج، وذلك لتجاوز قيم المطابقة الخارجية لهم عن (+2) أو (−2)، وعليه تم استبعادهم قبل عملية التدريج.

جدول (1) يبين المتوسطات الحسابية والانحرافات المعيارية لكل تقديرات الصعوبة، والخطأ المعياري في تقديرات صعوبات المفردات، ومتوسطات المربعات (MNSQ) لإحصائي المطابقة الداخلية والخارجية، وقيم إحصاءات المطابقة (ZSTD) لإحصائي المطابقة الداخلية والخارجي لمفردات الاختبار.

	الصعوبة	الخطأ المعياري	إحصاء المطابقة الداخلية (INFIT)		إحصاء المطابقة الخارجية (OUTFIT)	
			متوسط المربعات (MNSQ)	قيمة الإحصائي (ZSTD)	متوسط المربعات (MNSQ)	قيمة الإحصائي (ZSTD)
المتوسط الحسابي	147.5	0.12	1	0	1.93	0
الانحراف المعياري	61.5	0.01	0.12	3.3	0.34	3.4

يتضح من جدول (1) أن قيم المتوسط الحسابي لمتوسطات المربعات (MNSQ) الداخلية والخارجية كانت على الترتيب (1) و(1.93)، وهذه القيم قريبة من القيمة المثالية للنموذج، وهي الواحد صحيح، يتضح أيضاً من الجدول (1) أن متوسط قيم المطابقة الداخلية والخارجية لإحصائي (ZSTD) قد بلغت (صفر) وهي القيمة المثالية لنموذج راش، وبعد تفحص قيم المطابقة الخارجية للمفردات Outfit تبين أن (14) مفردة من مفردات اختبار واطسون_جليسر للتفكير الناقد المكون من (99)، أنها غير مطابقة للنموذج، وذلك لتجاوز قيم المطابقة الخارجية لها عن (+2.5) أو (−2.5)، وعليه تم استبعادها قبل عملية التدريج.

جدول (2) يوضح بيانات المفردات غير المطابقة لنموذج راش.

رقم المفردة	الصعوبة	الخطأ المعياري	إحصاء المطابقة الداخلية INFIT		إحصاء المطابقة الخارجية OUTFIT		معامل الارتباط	
			MSQ	ZSTD	MSQ	ZSTD	الملاحظ	المتوقع
45	3.11	0.31	1.05	0.30	3.10	4.10	0.06-	0.28
39	2.7	0.20	1.11	0.70	2.79	6.20	0.06-	0.28
31	1.58	0.16	1.15	1.30	1.91	4.90	0.11-	0.29
70	0.29	0.11	1.23	4.60	1.36	5.40	0.11-	0.29
46	0.90-	0.11	1.30	8.60	1.33	8.20	0.09-	0.29
79	0.29	0.11	1.19	3.90	1.32	3.30	0.09-	0.29
49	0.73	0.12	1.15	2.30	1.31	6.00	0.07-	0.29
98	1.16-	0.11	1.25	6.20	1.30	7.50	0.05-	0.29
41	0.41-	0.11	1.26	7.60	1.28	2.60	0.09	0.27
72	0.81	0.13	1.13	1.80	1.25	3.20	0.08	0.28
96	0.84	0.12	1.16	2.90	1.24	5.00	0.13	0.29
65	0.10-	0.11	1.17	4.90	1.23	6.00	0.15	0.26
95	0.55-	0.11	1.20	6.80	1.21	2.60	0.15	0.26
4	3.58-	0.23	1.21	2.86	1.20	3.22	0.14	0.29

يتضح من جدول (2) أن قيم المطابقة الداخلية (INFIT) للمفردات غير المطابقة مع توقعات نموذج راش قد تراوحت قيم إحصائي (ZSTD) ما بين (0.30- 8.60)، أما قيم المطابقة الخارجية (OUTFIT) لإحصائي (ZSTD) وهي الأهم فقد تراوحت ما بين (2.60-7.50)، أيضا يتبين أن قيم معاملات الارتباط الملاحظة أقل من معاملات الارتباط المتوقعة من النموذج؛

وهذا يعد من أهم المؤشرات التي يتم على أساسها حذف المفردات الاختبارية؛ وعليه تم حذف (14) الواردة بالجدول (2) وعليه أصبح الاختبار مكون من (85) مفردة، وهذه النتيجة تختلف عن نتيجة دراسة الزعبي 2017، الذي توصل إلى إحصاءات ملاءمة تم بناء عليها حذف (10) مفردات، وأيضا تختلف مع نتائج دراسة عايش 2021،التي اهتمت بتدريج اختبار واطسون _جليسر للتفكير الناقد لم تسفر إحصاءات المطابقة عن استبعاد أي فقرة من فقرات الاختبار.

ثانياً: عرض نتائج إجابة التساؤل الثاني الذي نصه:

ما مؤشرات حسن المطابقة (الداخلية والخارجية) الخاصة بأفراد عينة الدراسة؟

للتعرف على مؤشرات حسن مطابقة أفراد عينة الدراسةPersons Fit لنموذج راش تم تحديد قدرة كل فرد من أفراد عينة الدراسة في التفكير الناقد، كما تم حساب الخطأ المعياري الحقيقي Real S.E في قياس قدرة أفراد العينة، أيضاً تم حساب إحصاء المطابقة الداخلية (INFIT)، والذي يعد مؤشراً لسلوك أفراد العينة غير المتوقع، والذي يؤثر في الاستجابات عن المفردات الاختبارية التي تكون قريبة من مستوى معين، كما تم حساب مؤشر المطابقة الخارجية (OUTFIT)، وهو مؤشر للسلوكيات غير المتوقعة من الأفراد عن المفردات التي تبتعد عن مستوى قدراتهم في المتغير موضع القياس، والجدول التالي يبين هذه المؤشرات.

جدول (3) يبين المتوسطات الحسابية والانحرافات المعيارية لكل من تقديرات قدرات الأفراد، والخطأ المعياري في قدرات الأفراد، ومتوسطات المربعات (MNSQ) لإحصائي المطابقة الداخلية والخارجية، وقيم إحصاءات المطابقة (ZSTD) لإحصائي المطابقة الداخلية والخارجي.

	القدرة	الخطأ المعياري	إحصاء المطابقة الداخلية (INFIT)		إحصاء المطابقة الخارجية (OUTFIT)	
			متوسط المربعات (MNSQ)	قيمة الإحصائي (ZSTD)	متوسط المربعات (MNSQ)	قيمة الإحصائي (ZSTD)
المتوسط الحسابي	37.9	0.23	1	0	1.03	0
الانحراف المعياري	11.8	0.02	0.11	1.2	0.36	1.3

يتضح من جدول (3) أن قيم المتوسط الحسابي لمتوسطات المربعات (MNSQ) الداخلية والخارجية كانت على الترتيب (1) و(1.03)، وهذه القيم قريبة من القيمة المثالية للنموذج، وهي الواحد

صحيح، يتضح أيضاً من الجدول (1) أن متوسط قيم المطابقة الداخلية والخارجية لإحصائي (ZSTD) قد بلغت (صفر) وهي القيمة المثالية لنموذج راش، وبعد تفحص قيم المطابقة الخارجية للأفراد Outfit تبين أن (23) طالباً من عينة الدراسة المكونة من (385)، أنهم غير مطابقين للنموذج، وذلك لتجاوز قيم المطابقة الخارجية لهم عن (+2.5) أو (−2.5)، وعليه تم استبعادهم قبل عملية التدريج.

جدول (4) يوضح بيانات الأفراد غير المطابقين لنموذج راش.

رقم الفرد	القدرة	الخطأ المعياري	إحصاء المطابقة الداخلية INFIT		إحصاء المطابقة الخارجية OUTFIT	
			MSQ	ZSTD	MSQ	ZSTD
7	1.94−	0.29	1.18	0.90	3.32	5.1
8	1.94−	0.29	1.18	0.90	3.32	5.1
11	1.77−	0.28	1.23	1.30	3.14	5.3
12	1.77−	0.28	1.23	1.30	3.14	5.3
13	1.77−	0.28	1.22	1.20	3.12	5.2
14	1.77−	0.28	1.22	1.20	3.12	5.2
10	1.77−	0.28	1.21	1.20	3.02	5.1
15	1.70−	0.27	1.20	1.20	2.86	5.0
5	2.12−	0.31	1.05	0.30	1.98	2.4
6	2.03−	0.30	0.30	1.07	1.92	2.4
9	1.85−	0.29	1.06	0.40	1.72	2.2
207	0.49−	0.22	1.16	2.00	1.57	4.2
248	0.25−	0.22	1.19	2.6	1.51	4.2
235	0.35 −	0.22	1.23	3.0	1.33	2.7

2.8	1.31	1.3	1.08	0.23	0.16	337
2.7	1.30	2.01	1.14	0.21	0.12-	286
2.4	1.29	2.5	1.19	0.22	0.39-	226
2.6	1.28	2.00	1.14	0.22	0.16-	272
2.3	1.28	2.9	1.23	0.22	0.44	271
2.3	1.25	1.4	1.10	0.22	0.21-	258
2.2	1.25	1.9	1.14	0.22	0.30-	241
2.3	1.25	3.5	1.25	0.22	0.16-	270
2.1	1.22	2.7	1.18	0.21	0.02	317

يتضح من جدول (4) أن قيم المطابقة الداخلية (INFIT) للأفراد غير المطابقين مع توقعات نموذج راش قد تراوحت قيم إحصائي (ZSTD) ما بين (1.18- 3.50)، أما قيم المطابقة الخارجية (OUTFIT) لإحصائي (ZSTD) وهي الأهم فقد تراوحت ما بين (2.1-5.1)، وبناء على هذه النتائج تم حذف (23) فرداً من عينة الدراسة، وعليه تم تدريج مفردات اختبار واطسون _ جليسر على عينة تكونت من (362) تنطبق عليهم افتراضات نموذج راش،تتفق مع نتيجة دراسة أشنان التي تم بناء على إحصاءات الملاءمة للنموذج حذف (29) فرداً، تختلف نتائج هذه الدراسة مع نتيجة دراسة الزعبي التي لم يتم حذف أي فرد من أفراد عينة الدراسة حسب إحصاءات الملاءمة الداخلية والخارجية التي تبين أنها جميعاً في حدود الملاءمة المطلوبة لنموذج راش.

ثالثاً: عرض نتائج إجابة التساؤل الثالث الذي نصه:

هل يمكن التحقق من شرط أُحادية البعد لمفردات اختبار واطسون _جليسر للتفكير الناقد؟

للإجابة عن هذا التساؤل، قام الباحث بإعادة تحليل استجابات عينة الأفراد المطابقين المكونة من (362) طالب، وطالبة، على اختبار واطسون _ جليسر للتفكير الناقد باستخدام برنامج التحليل الإحصائي Winsteps،والتحقق من توافر شرط أُحادية البعد تطلب فحص لما تتضمنه عينة مفردات الاختبار المطابقة لنموذج راش ومدى إمكانية تعريف هذه المفردات للتغير موضع القياس (التفكير الناقد)، وذلك بأسلوب التحليل العاملي للمكونات الأساسية للبواقي (البواقي المعيارية)

Principal Components Analysis OF Residual (Standardized Residual) والجدول التالي يبين نتائج هذا التحليل.

جدول (5) يوضح نتائج تحليل المكونات الأساسية للبواقي لمفردات اختبار واطسون _ جليسر للتفكير الناقد.

التباين		حجم التباين الملاحظ مقدراً بوحدات القيم المميزة (Eigen Values)	نسبة التباين في الاستجابات	
			تباين ملاحظ	تباين نموذجي
التباين الكلي في الاستجابات		119.4	%100	%100
التباين المفسر بتقديرات نموذج راش (العامل العام)		72.4	%72.4	%71.1
إجمالي التباين غير المفسر في الاستجابات		27.6	%.27.6	%28.9
التباين غير المفسر بالعامل العام	مفسر بالعامل الأول	10.3	%8.7	-
	مفسر بالعامل الثاني	2.7	%2.1	-
	مفسر بالعامل الثالث	2.3	%6.1	-
	مفسر بالعامل الرابع	2.2	%5.3	-
	مفسر بالعامل الخامس	2.1	%4.3	-

يتضح من جدول (5) أن النموذج فسر ما نسبته (0.72.4) من تباينات استجابات أفراد عينة الدراسة، وقد بلغت قيمة الجذر الكامن للعامل الأول (10.3)، وتعد هذه القيمة كبيرة إذا ما قورنت بالعامل الثاني الذي بلغت قيمة الجذر الكامن له (2.7)، أما الجذور الكامنة للعوامل الثالث والرابع والخامس فكانت منخفضة ومتقاربة، وهذا يعد مؤشراً على تحقق شرط أحادية البعد لفقرات اختبار واطسون _ جليسر للتفكير الناقد، وهذه النتيجة تتفق من نتائج دراستي كل من: الجبوري، 2017، والزعبي 2017، اللتين توصلتا إلى تمتع اختبار التفكير الناقد لـ واطسون _ جليسر بافتراض أحادية البعد.

رابعاً: عرض نتائج إجابة التساؤل الرابع الذي نصه:

ما قيم صدق وثبات اختبار واطسون _ جليسر للتفكير الناقد بعد تدريجه باستخدام نموذج راش؟

للإجابة عن هذا التساؤل تم حساب تقديرات معلم القدرة متحررة من صعوبة المفردات، وكذلك تم حساب تقديرات معلم الصعوبة متحررة من قدرات الأفراد، والجدولين التاليين يوضحان الإحصاءات والدرجات الخام، والقدرة مقاسه بوحدة اللوجيت، والخطأ المعياري لكل من معلمي قدرة الأفراد متحررة من صعوبة المفردات، وصعوبة المفردات متحررة من قدرة الأفراد.

جدول (6) تقدير معلم قدرات الأفراد متحررة من صعوبة مفردات الاختبار.

الإحصاءات	الدرجة الخام	القدرة باللوجيت	الخطأ المعياري	قيمة التطابق الداخلي	قيمة التطابق الخارجي
المتوسط الحسابي	37.9	0.58-	0.23	0	0
الانحراف المعياري	11.8	0.61	0.02	1.2	1.3
أعلى درجة	62	0.59	0.22	0.6-	0.8-
أدنى درجة	12	2.22-	0.32	0.2	1.8

يتبين من جدول (6) أن تقديرات معالم قدرات الأفراد متحررة من صعوبة المفردات تراوحت بين (12) و(62)، ومتوسط تقديرات القدرة (0.58-) لوجيت، بانحراف معياري (0.61)، أما الخطأ المعياري للمتوسط الحسابي لتقديرات القدرة (0.23) وهي قيمة قريبة جداً مما يفترضه النموذج، وهي مؤشر على الدقة في تقدير معالم قدرات الأفراد.

جدول (7) تقدير معلم صعوبة مفردات اختبار واطسون _جليسر للتفكير الناقد متحررة من قدرات الأفراد.

الإحصاءات	الدرجة الخام	القدرة باللوجيت	الخطأ المعياري	قيمة التطابق الداخلي	قيمة التطابق الخارجي
المتوسط الحسابي	147.5	0	0.12	0	0
الانحراف المعياري	61.5	0.85	0.02	0.3	0.4
أعلى درجة	353	3.13-	0.19	0.5-	1.3-
أدنى درجة	11	3.11	0.31	0.3	0.4

يتبين من جدول (7) أن تقديرات معالم صعوبة المفردات متحررة من قدرات الأفراد تراوحت بين (11) و(353)، ومتوسط تقديرات القدرة (0) لوجيت، بانحراف معياري (0.85)، أما الخطأ

المعياري للمتوسط الحسابي لتقديرات القدرة (0.12) وهي قيمة صغيرة مما يدل على الدقة في تقدير معالم صعوبة المفردات.

جدول(8) يبين معاملات الثبات والفصل لأفراد عينة الدراسة ومفردات اختبار واطسون_جليسر للتفكير الناقد.

معامل الثبات		معامل الفصل		الانحراف المعياري	
للأفراد	للمفردات	للأفراد	للمفردات	للأفراد	للمفردات
0.85	0.98	2.39	6.80	11.8	61.5

يشير معامل الثبات في نموذج راش إلى دقة تقدير مواقع الأفراد والمفردات على متصل السمة أو القدرة المراد قياسها، ويمكن تحديد هذه الدقة في التقدير بحساب معامل الفصل للمفردات، والذي يرمز له بالرمز Gi، وقد بلغ في هذه الدراسة (6.80)، فيما بلغ معامل الفصل للأفراد (2.39)، والذي يرمز له بالرمز Gp، وهذه القيم لمعامل الفصل سواءً للمفردات أو الأفراد إذا زادت عن (2) يُعد دليلاً على أن عينة الأفراد كافية للفصل بين قيم صعوبات المفردات، وبالتالي تدل على أن مفردات الاختبار تتدرج هرمياً من حيث الصعوبة، وقيمة معامل ثبات المفردات (0.98) وهي قيمة مرتفعة تدل أن قدرات الأفراد تباينت بما يكفي لتفصل بين صعوبة مفردات اختبار التفكير الناقد لـ واطسون وجليسر، زمن ثم استطاعت تعريف متصل القدرة المراد قياسها (التفكير الناقد)، وهذه النتيجة تتفق مع نتيجة دراسة الجبوري 2017، التي بينت نتائجها أن الدرجات المحسوبة بنموذج راش كانت أكثر دقةً وصدقاً، من الدرجات المحسوبة بنظرية القياس التقليدية، وأيضاً تتفق مع نتيجة دراسة الزعبي 2017، التي توصلت إلى التحقق من صدق مفردات اختبار التفكير الناقد، وأن معالم المفردات ومعالم الأفراد تتصف بخاصية اللاتغاير، كما تتفق مع نتائج دراسة عايش 2021 التي توصلت إلى تمتع مفردات اختبار واطسون_جليسر بخصائص سيكومترية مقبولة حسب افتراضات نموذج راش.

خامساً: عرض نتائج إجابة التساؤل الخامس الذي نصه:

ما تقديرات صعوبة مفردات اختبار واطسون _ جليسر للتفكير الناقد؟

للإجابة عن هذا التساؤل تم حساب تقدير معلم صعوبة كل مفردة من مفردات الاختبار، بالإضافة لحساب الأخطاء المعيارية لتلك المفردات بوحدة اللوجيت وهي الوحدة المستخدمة في تقدير معالم الصعوبة باستخدام نموذج راش، ثم تحول تلك الدرجات إلى وحدة قياس معيارية تعرف بوحدة المنف؛ التي تستخدم كمعايير لتفسير أداء المبحوثين على هذا الاختبار.

جدول (9) يبين رموز مفردات اختبار واطسون _ جليسر وصعوباتها وأخطاءها المعيارية بوحدتي اللوجيت والمنف.

م	رمز المفردة	الصعوبة		الخطأ المعياري		إحصاء الملاءمة	
		باللوجيت	بالمنف	باللوجيت	بالمنف	التقاربي	التباعدي
1	4i	-1.78-	32.10	0.23	51.15	-0.17	-1.04
2	73i	-1.61-	41.95	0.12	50.60	2.25	1.66
3	13i	-1.60-	42.00	0.12	50.60	-2.25	-2.02
4	2i	-1.58-	42.10	0.12	50.60	-0.2	-1.09
5	9i	-1.42-	42.90	0.12	50.60	1.25	1.66
6	21i	-1.16-	44.20	0.12	50.60	-2.06	-1.99
7	12i	-1.09-	44.55	0.11	50.55	-1.77	-1.69
8	19i	-1.08-	44.60	0.11	50.55	0.41	0.7
9	50i	-1.04-	44.80	0.11	50.55	1.72	2.34
10	32i	-1.02-	44.90	0.11	50.55	1.84	1.65
11	10i	-99.-	45.05	0.11	50.55	0.76	0.58
12	35i	-98.-	45.10	0.11	50.55	-3.49	-2.62
13	8i	-83.-	45.85	0.11	50.55	-0.45	0.26
14	60i	-81.-	45.95	0.11	50.55	-1.55	-1.97
15	66i	-78.-	46.10	0.11	50.55	1.98	1.78
16	80i	-73.-	46.35	0.11	50.55	6.35	6.08
17	59i	-64.-	46.45	0.11	50.55	-1.37	-1.55
18	25 i	-57.-	46.80	0.11	50.55	-2.02	-1.42
19	71i	-55.-	47.15	0.11	50.55	1.46	5
20	1i	-55.-	47.25	0.11	50.55	-1.76	-0.47
21	27i	-47.-	47.25	0.11	50.55	1.86	2.3

2.22	1.69	50.55	0.11	47.65	−44.−	48i	22
2.22	1.68	50.55	0.11	47.80	−38.−	86i	23
−2.38	−2.32	50.55	0.11	48.10	−38.−	15i	24
1.92	2	50.55	0.11	48.10	−37.−	56i	25
2.07	1.64	50.55	0.11	48.15	−31.−	14i	26
1.02	1.09	50.55	0.11	48.45	−28.−	42i	27
−0.16	0.34	50.55	0.11	48.60	−27.−	57i	28
2.47	2.2	50.55	0.11	48.65	−24.−	38i	29
2.50	1.31	50.55	0.11	48.80	−20.−	61i	30
−0.2	−0.4	50.55	0.11	49.00	−16.−	28i	31
−0.84	−0.41	50.55	0.11	49.30	−0.16	i 22	32
−1.81	−1.39	50.55	0.11	49.30	−0.14	i 11	33
2.18	2.1	50.55	0.11	49.45	−0.14	i 74	34
1.22	0.06	50.55	0.11	49.65	−0.11	i 99	35
−1.66	−0.69	50.55	0.11	49.75	−0.07	i 33	36
0.65	1.37	50.55	0.11	49.75	−0.05	i 7	37
−1.53	−1.95	50.55	0.11	49.80	−0.05	i 64	38
1.25	1.53	50.55	0.11	50.00	−0.04	i 30	39
2.37	2.19	50.55	0.11	50.10	0	i 5	40
−1.66	−2.39	50.55	0.11	50.15	0.02	i 62	41
−1.85	−1.42	50.55	0.11	50.20	0.03	i 17	42
−1.54	−1.92	50.55	0.11	50.20	0.04	i 16	43
1.91	2.1	50.55	0.11	50.45	0.04	i 85	44
−2.05	−1.39	50.55	0.11	50.65	0.09	i 23	45

−0.74	−1.63	50.55	0.11	50.90	0.13	i 55	46
−1.87	−2.39	50.55	0.11	50.90	0.18	i 6	47
−1.29	−0.91	50.55	0.11	51.05	0.18	i 26	48
2.44	1.85	50.60	0.12	51.05	0.21	i 75	49
1.13	1.57	50.60	0.12	51.45	0.21	i 84	50
−1.17	−1.78	50.60	0.12	51.50	0.29	i 58	51
1.73	1.37	50.60	0.12	51.60	0.3	i 87	52
−2.16	−2.09	50.60	0.12	51.75	0.32	i 36	53
1.59	1.75	50.60	0.12	52.10	0.35	i 18	54
0.6	1.22	50.60	0.12	52.10	0.42	i 43	55
1.2	1.38	50.60	0.12	52.10	0.42	i 54	56
−1.88	−1.62	50.60	0.12	52.15	0.42	i 97	57
1.56	1.35	50.60	0.12	52.30	0.43	i 69	58
−0.57	−1.2	50.60	0.12	52.30	0.46	i 47	59
−1.61	−2.07	50.60	0.12	52.35	0.46	i 67	60
−2.12	−2.37	50.60	0.12	52.45	0.47	i 40	61
−1.87	−1.62	50.60	0.12	52.45	0.49	i 63	62
1.77	2.02	50.60	0.12	52.60	0.49	i 94	63
1.17	2.50	50.60	0.12	52.60	0.52	i 3	64
1.49	1.99	50.60	0.12	52.85	0.52	i 92	65
−1.6	−1.57	50.60	0.12	52.85	0.57	i 44	66
1.71	0.95	50.60	0.12	53.10	0.57	i 78	67
1.77	0.14	50.60	0.12	53.25	0.62	i 88	68
1.61	1.33	50.60	0.12	53.40	0.65	i 89	69

1.96	1.47	50.60	0.12	53.50	0.68	i 34	70
1.74	1.27	50.60	0.12	53.70	0.7	i 91	71
−2.35	−1.32	50.65	0.13	54.10	0.74	i 76	72
−2.15	−2.23	50.65	0.13	54.20	0.82	i 53	73
−2.18	−0.93	50.65	0.13	54.30	0.84	i 37	74
−1.06	−0.34	50.65	0.13	54.55	0.86	i 24	75
2.32	1.08	50.65	0.13	54.80	0.91	i 51	76
−2.46	−2.36	50.65	0.13	55.00	0.96	i 83	77
−2.3	−1.07	50.65	0.13	55.35	1	i 90	78
−1.59	−1.64	50.70	0.14	55.35	1.07	i 52	79
−1.58	−0.56	50.70	0.14	55.35	1.07	i 81	80
−1.58	−0.56	50.70	0.14	55.45	1.07	i 82	81
−2.33	−1.15	50.70	0.14	55.90	1.09	i 68	82
−1.64	−2.01	50.70	0.14	56.40	1.18	i 77	83
−2.29	−0.89	50.70	0.14	57.20	1.28	i 93	84
−1.94	−1.57	50.75	0.15	59.15	1.44	i 29	85
−1.96	−1.08	50.85	0.17	49.30	1.83	i 20	86

يذكر هامبلتون وسويمنثان(1991) Hamblton & Swaminathan أن قيمة صعوبة المفردة نظرياً تتراوح بين (+∞، −∞)، إلا أنه عملياً، وحسب برنامج التحليل وينستبس Winsteps المستخدم في هذه الدراسة، فإن قبول صعوبة المفردات يتراوح بين (+2 و−2)، وقد تراوحت قيم صعوبة مفردات اختبار واطسون− جليسر للتفكير الناقد في الجدول (4) بين (−1.61 إلى 1.83)، حيث إنّ المفردات التي تكون درجة اللوجيت ذات الإشارة سالبة تكون مفردات سهلة كما في المفردات من (1 – 39)،

بالجدول السابق، أما والمفردات التي تكون درجة اللوجيت تساوي صفر تكون متوسطة السهولة كما في المفردة (40)، أما المفردات ذات الصعوبة الأعلى من المتوسط فتكون درجة اللوجيت موجبة، وكلما اتبتعدت المفردات عن الصفر بالموجب ازدادت الصعوبة كما في المفردات من (41 إلى 86)، كما يتضح من الجدول السابق أن متوسط درجة صعوبة المفردات بوحدة اللوجيت تساوي صفر، والانحراف المعياري (0.80)، وتدل هذه القيم للمتوسط الحسابي والانحراف المعياري لمعامل صعوبة مفردات اختبار واطسون – جليسر للتفكير الناقد، أن هذه المفردات متسقة ومتجانسة في قياس هذا المتغير، وليس أدل على ذلك في رأي الباحث من وقوع متوسط الصعوبة لمفردات الاختبار في المفردة رقم (40)، والتي تبين أن صعوبتها (0)، وهي تقع تقريباً وسط مفردات الاختبار البالغ عددها (86) مفردة، وهذه النتيجة تتفق مع نتائج دراسة شنان 2019 التي توصلت إلى أن العينة المستخدمة في تدريج مفردات اختبار التفكير الناقد وفرت أفضل تدريج لتلك المفردات باستخدام نموذج راش تبعاً لمحكات دقة القياس التي اختلفت باختلاف مستوى العينة، وذلك لصالح العينة منخفضة المستوى، وهذا ينطبق على مستوى عينة التدريج في الدراسة الحالية، حيث يتضح من جدول (6) أن متوسط أداء عينة الدراسة كان (37.9)، كدرجة خام، و(-0.58) بوحدة اللوجيت، وتقابلها القيمة (47.1) بوحدة المنف، التي تعد نوعاً من معايير تفسير الدرجات حسب نموذج راش، من هذه القيم يتبين أن مستوى التفكير الناقد لطلبة كلية التربية جامعة عمر المختار يعد دون المتوسط، إذا ما قورن بالمتوسط النظري لاختبار واطسون _جليسر كما ورد في دليل تطبيقه أن المتوسط النظري يساوي (49.5) كدرجة خام، أو الدرجة (صفر) بوحدة اللوجيت، والدرجة (50) بوحدة المنف.

سادساً: عرض نتائج إجابة التساؤل السادس الذي نصه:

ما مستوى التفكير الناقد لدى طلبة كلية التربية كما يقاس باختبار واطسون_جليسر للتفكير الناقد بعد تدريجه بنموذج راش؟

للإجابة عن هذا التساؤل تم استخدام اختبار (t) لعينة واحدة فكانت النتائج كما في الجدول الآتي:-

جدول (10) يُبين المتوسط الحسابي لعينة الدراسة، والمتوسط الفرضي لاختبار واطسون_جليسر للتفكير الناقد، وقيمة (t) ودلالتها الإحصائية.

حجم العينة	المتوسط الحسابي	المتوسط الفرضي	قيمة (t)	df	الدلالة الإحصائية
362	37.9	49.5	19.22-	361	0.000

يتضح من الجدول (10) أن المتوسط الحسابي لدرجات أفراد عينة الدراسة على اختبار واطسون_جليسر للتفكير الناقد بلغ (37.9)، في حين أن المتوسط الفرضي للاستبانة (49.5)، وباستخدام اختبار (t) لعينة واحدة لحساب دلالة الفرق بين هذين المتوسطين، بينت نتائج التحليل الإحصائي أن قيمة اختبار (t) بلغت (-19.22)، وهذه القيمة دالة إحصائياً عند مستوى دلالة (0.000)؛ وهذا يعني أن مستوى التفكير الناقد لدى طلبة كلية التربية دون المتوسط، وهذه النتيجة تتفق مع ما توصل إليه كل من لو وثروب Loo & Thrope 1999 من أن مهارات التفكير الناقد لدى عينة دراستهما من الطلبة الجامعيين كانت أقل من المتوسط النظري للاختبار، وأيضا تتفق مع ما توصل إليه الجبوري 2017 من أن مهارات التفكير الناقد لطلبة جامعة بابل أقل من المتوسط النظري للاختبار، وتتفق أيضا مع نتائج دراسة عايش 2021 على طلبة جامعتي سعيدة والشلف في الجزائر.

توصيات الدراسة:

بناءً على النتائج التي خلصت إليها الدراسة يوصي الباحث بما يلي:

1. استخدام اختبار واطسون _ جليسر لقياس التفكير الناقد في صورته النهائية المكونة من (86) مفردة، والتي تبين أنها تحقق شروط نموذج راش؛ أي تحقق شروط القياس الموضوعي لهذا المتغير.
2. استخدام نموذج راش في تطور وتدريج مفردات أدوات القياس لا سيما الاختبارات العقلية شائعة الاستخدام في مراكز ذوي الاحتياجات الخاصة، للتغلب على أوجه القصور التي تعاني منها تلك الأدوات التي صممت باستخدام نظرية القياس التقليدية.

مقترحات الدراسة:

1. إجراء دراسات على عينات أخرى غير طلبة الجامعة، تشمل تلاميذ المراحل التعليمية المختلفة، لأهمية قياس وتطور التفكير الناقد بوصفها قدرة تعد اليوم من أهم القدرات التي تهتم بها النظم التربوية الحديثة لقياسها وتنميتها لدى المتعلمين.
2. إجراء دراسات لتطوير أدوات قياس التفكير الناقد باستخدام نماذج أخرى لنظرية الاستجابة للمفردة كالنموذج ثنائي المعلم أو ثلاثي المعلم.

المراجع

1. آلك فِشر (2009). التفكير الناقد، ترجمة، ياسر العيتي، دار السيد للنشر والتوزيع، الرياض، المملكة العربية السعودية.
2. أمينة محمد كاظم (1996). نماذج السمات الكامنة. في أنور الشرقاوي وآخرون. اتجاهات معاصرة في القياس والتقويم النفسي والتربوي. القاهرة: مكتبة الأنجلو المصرية.
3. الجبوري، حسين موسى (2017). إعداد اختبار واطسون – جليسر للتفكير الناقد على وفق نظرية القياس التقليدية ونظرية السمات الكامنة. رسالة ماجستير غير منشورة، كلية التربية للعلوم الإنسانية، جامعة بابل، جمهورية العراق.
4. جروان، فتحي عبد الرحمن (2007). تعليم التفكير (مفاهيم وتطبيقات). الأردن دار الفكر العربي.
5. الخرشة، طه عقلة (2015). تدريج اختبار أوتيس– لينون للقدرة العقلية المستوى المتقدم الصورة (J) باستخدام نموذج راش. مجلة القراءة والمعرفة، الناشر دار المنظومة، القاهرة.
6. خماد، محمد (2021). مهارات التفكير الناقد لدى تلاميذ السنة الرابعة متوسط، دراسة ميدانية على متوسطية الوئام المدني بولاية الوادي، مجلة العلوم النفسية والتربوية، جامعة خميس مليانة الجزائر.
7. الزعبي، آمال أحمد (2017). بناء اختبار لقياس مهارات التفكير الناقد باستخدام نظرية استجابة الفقرة. مجلة اتحاد الجامعات العربية للتربية وعلم النفس، المجلد الخامس عشر، العدد الثالث، ص ص 55–92.
8. السكري، عماد محمد (2010). بعض الخصائص السيكومترية لاختبار واطسون – جليسر للتفكير الناقد الصور "S" لدى عينة من طلاب الجامعة. المجلة المصرية للدراسات النفسية، العدد 76– المجلد العشرون– أبريل 2010، ص ص 224 – 251.

9. سليمان، سناء محمد (2010). أدوات جمع البيانات في البحوث النفسية والتربوية. القاهرة، مكتبة عالم الكتب.
10. السيد، عزيزة (1995). التفكير الناقد دراسة في علم النفس المعرفي. الإسكندرية، دار المعرفة الجامعية.
11. شنان، علي عباس (2019). استخدام نموذج راش في تدريج مقياس التفكير الناقد لدراسة بعض العوامل المؤثرة على دقة القياس وفق نظرية الاستجابة للمفردة. مجلة كلية التربية، العدد السادس والثلاثون، الجلد الأول، أغسطس 2019، ص ص 439 –466.
12. شنطاوي، محمد خالد (2003). تقنين اختبار كاليفورنيا لمهارات التفكير الناقد للطلبة الجامعيين في الأردن، رسالة ماجستير غير منشورة، الأردن، جامعة مؤتة.
13. صالح، رحاب سعد منيسي (2011). مستوى التفكير الناقد لدى طلبة جامعة عمر المختار وفقاً لبعض المتغيرات، رسالة ماجستير غير منشورة، أكاديمية الدراسات العليا، بنغازي، ليبيا.
14. عايش، صباح (2021). استخدام نموذج راش في تدريج اختبار التفكير الناقد (اختبار واطسون . جليسر للتفكير الناقد الصورة المختصرة)، مجلة العلوم النفسية والتربوية، المجلد، 100، العدد 112، ص ص 100 . 112.
15. عبدالسلام، محمد (2020). مهارات التفكير الناقد، دراسة نظرية وتطبيقات عربية وعالمة. القاهرة، مكتبة نور.
16. عبدالوهاب، محمد محمود (2010). استخدام نماذج نظرية الاستجابة للمفردة الاختبارية في تدريج بعض الاختبارات المعرفية. أطروحة دكتوراة غير منشورة، كلية التربية، جامعة المنيا، جمهورية مصر العربية.
17. العتيبي، خالد ناهس (2012). الخصائص السيكومترية للصورة القصيرة لاختبار واطسون _ جليسر للتفكير الناقد (WGCT– SF) دراسة على عينة من الطلاب المعلمين في البيئة السعودية. مجلة جامعة الملك سعود للعلوم التربوية والدراسات الإسلامية، مجلد 24، العدد 4. ص ص 1427 – 1454.
18. عزت عبد الحميد (2011). الإحصاء النفسي والتربوي تطبيقات باستخدام برنامج SPSS القاهرة: دار الفكر العربي.
19. علام، صلاح الدين محمود (2005). نماذج نظرية الاستجابة للمفردة أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي. القاهرة، دار الفكر العربي.
20. علاونة، معزوز جابر (2016) توظيف نموذج راش في بناء اختبار تحصيلي في منهاج التاريخ للصف الثاني الثانوي الأدبي بفلسطين. منشورات جامعة القدس المفتوحة.

21. وهدي، رانيا ماهر (2010). دراسة سيكومترية مقارنة بين نموذج راش والقياس التقليدي حول دقة التنبؤ بحالة القلق من سمة القلق. أطروحة دكتوراة غير منشورة، القاهرة، كلية البنات للآداب والعلوم والتربية، جامعة عين شمس.

22. Baker, F.(2001): The Basics of Item Response Theory (2ed.). The ERIC Clearinghouse onAssessment and Evaluation, USA

23. Hambleton, R. &Swaminathan ,H. & Rogers. H (1991). Fundamentals of Item response theory. Sage Publications.

24. Linacre, J.M.(2018) WinstepsRasch Measurement Computer Program. Chicago: Winsteps. Com.

25. Wagner, T. & Harvery, R. (2003). Developing A new Critical thinking Test Using item Response Theory. Paper Presented at the 2003 Annual Conference of the Society for Industrial Organization Psychology. Orlando. Retrieved May, 4, 2005. From: WWW.Criticalthinking.org.

26. Watson, G & Glaser, E. Watson-Glaser Critical Thinking Appraisal Short Form Manual. Pearson Education, Inc, (2008).