تطوير اختبار واطسون - جليسر للتفكير الناقد (WGCT-SF) وفق نموذج راش

Development of the Watson-Glaser Critical Thinking Test (WGCT-SF) Using the Rasch Model

إعداد

عبير عبدالرحمن ابراهيم الشعيفان

Abeer Abdulrahman Ibrahim Alshuayfan

قسم التربية وعلم النفس - كلية التربية والآداب - جامعة تبوك

د. صبري محمد اسماعيل عبدالعال

Dr. Sabry Mohamed Ismail Abdelaal

أستاذ القياس النفسي والتربوي بقسم التربية وعلم النفس في كلية التربية والآداب بجامعة تبوك

Doi: 10.21608/ejev.2025.458396

استلام البحث: ٤ / ٦ / ٢٠٢٥

قبول النشـــــر: ١٣ / ٨ / ٢٠٢٥

الشعيفان، عبير عبدالرحمن ابراهيم وعبدالعال، صبري محمد اسماعيل (٢٠٢٥). تطوير اختبار واطسون - جليسر للتفكير الناقد (WGCT-SF) وفق نموذج راش. ***المجلة العربية للتربية النوعية***، المؤسسة العربية للتربية والعلوم والآداب، مصر، ٩(٤٠)، ٣٢٧ - ٣٧٦.

https://ejev.journals.ekb.eg

تطوير اختبار واطسون - جليسر للتفكير الناقد (WGCT-SF) وفق نموذج راش

المستخلص:

هدفت الدراسة إلى تطوير اختبار واطسون - جليسر للتفكير الناقد (WGCT-SF) وفق نموذج راش لدى طلبة جامعة تبوك، باستخدام المنهج الوصفي، ولتحقيق اهداف الدراسة تم استخدام اختبار واطسون - جليسر للتفكير الناقد (WGCT-SF) تقنين العتيبي (٢٠١٢). طبق الاختبار على عينة مكونة من (٢٠١) طالب وطالبة تم أخذها بالطريقة العشوائية الطبقية من طلبة جامعة تبوك، تم جمع وتحليل البيانات باستخدام برنامج (RStudio). توصلت النتائج الى تحقق افتراضات نظرية الاستجابة للفقرة واستبعاد ١٢ فرد جاءت نتائجهم خارج حدود المطابقة الداخلية والخارجية لنموذج راش ومطابقة جميع فقرات الاختبار (٤٠) للنموذج، وتراوحت قيم صعوبة الفقرات بين (-٢.٣٨١ و٢.٢٥١)، ويعكس هذا التدرج قدرة الاختبار على قياس مستويات متعددة من التفكير الناقد بدقة، وتراوحت قدرات الأفراد بين (-٠.٣٥٣ و٠.٢٩٧)، مما يشير إلى تقارب في مستوى التفكير الناقد لدى طلبة جامعة تبوك، وتم التحقق من صدق البناء للاختبار وثبات الفقرات، كما أظهرت النتائج أن مستوى التفكير الناقد لدى طلبة جامعة تبوك كان أعلى من المتوسط وتوصي الدراسة على قياس قدرات التفكير الناقد بشكل مستمر باستخدام أدوات تحقق شروط الصدق والثبات، خاصة في ضوء أهمية هذه المهارة في دعم الأداء الأكاديمي والمهني واستخدام الصورة القصيرة لاختبار واطسون–جليسر بصيغته المتحققة من افتراضات نظرية الاستجابة للفقرة، خاصة في الدراسات المهتمة بقياس التفكير الناقد بطريقة دقيقة ومبنية على أسس إحصائية سليمة.

الكلمات المفتاحية: نموذج راش – نظرية الاستجابة للفقرة – التفكير الناقد – قياس التفكير الناقد – اختبار واطسون-جليسر

Abstract:

The study aimed to develop the Watson-Glaser Critical Thinking Test (WGCT-SF) based on the Rasch model among students at the University of Tabuk, using a descriptive research approach. To achieve the study's objectives, the WGCT-SF version standardized by Al-Otaibi (2012) was utilized. The test was administered to a stratified random sample of 201 students, and data were analyzed using RStudio. The findings confirmed the assumptions of the Item Response Theory, with 12 participants excluded due to poor internal and external fit to the

ISSN: 2537-0448 *eISSN: 2537-0456*

Rasch model. All 40 test items showed good model fit, with item difficulty values ranging from -2.381 to 2.251, indicating the test's ability to assess multiple levels of critical thinking. The participants' ability estimates ranged from -0.353 to 0.297, reflecting a relatively homogeneous level of critical thinking among the sample. The study confirmed the construct validity and reliability of the test and concluded that students' critical thinking levels were above average. It recommends ongoing assessment using validated and reliable tools, and highlights the short form of the Watson-Glaser test as a statistically sound instrument for accurately measuring critical thinking in future research.

Keywords: Rasch model – Item Response Theory – Critical Thinking - Critical Thinking Measurement – Watson-Glaser Test

المقدمة:

إن التفكير الناقد مهارة أساسية تتضمن عمليات التفكير التأملي والعقلاني، وهي ضرورية للتغلب على تحديات الحياة الواقعية المعقدة، كما انه متعدد الأوجه اذ يقوم على دمج المهارات والتصرفات اللازمة لحل المشكلات واتخاذ القرار بشكل فعال (Sutoyo et al., 2023).

وفي عصرنا الحالي؛ أصبح تدريس التفكير الناقد هدف من اهداف التربية في كثير من دول العالم المتقدمة لإسهامه في تطوير البنية المعرفية للمتعلمين, فهي تسعى لتعليم الفرد كيف يتعلم ويفكر، كما يعد من مهارات القرن ٢١ الذي يحتاج فيه الفرد لمهارات التفكير العميق لمواكبة التغيرات السريعة وتراكم المعلومات (الرباط، ٢٠٢٢).

في عام ٢٠١٨ م تم تدشين برنامج مهارات التفكير الناقد ضمن مقرر المهارات الحياتية، وفي عام ٢٠٢٢ تم إدخاله كمقرر ضمن مقررات وزارة التعليم (للصف الثالث متوسط والأول ثانوي) فهو يركز على تنمية المهارات العقلية للطلاب في طرح الأسئلة ومعالجة المفاهيم والبناء المنطقي للحجج والاسهام في برنامج تعزيز الشخصية السعودية، ومواكبة التطورات الفكرية والثورة المعرفية والتقنية بما يسهم بتحقيق رؤية المملكة ٢٠٣٠ ("عام / وزير التعليم يدشن"، ٢٠١٨؛ الشويرخ، ٢٠٢٢).

ISSN: 2537-0448 eISSN: 2537-0456

يشير ديوي (Dewey, 1910) في كتابه (How we think) أن التفكير الناقد يعني أن تبقى في حالة شك دائماً، فهو يتطلب التغلب على الميل الطبيعي لقبول الأفكار في ظاهرها، والانخراط في التفكير والاستعداد للتشكيك في الافتراضات والاقتراحات الطبيعي قبولها لتعزيز الفهم الحقيقي والنمو الفكري.

كما يعد من المسائل التربوية التي اتجه علماء النفس والتربويون للاهتمام بها لضمان التطور المعرفي كون التفكير الناقد يعد من المفاتيح الهامة له، فالمفكر الناقد يقوم باستغلال أقصى طاقاته العقلية للتفاعل بشكل إيجابي مع ظروف الحياة ومستجداتها (موسى،٢٠٢١). لذا جاءت الحاجة لوضع اختبارات ومقاييس ذات خصائص سيكومترية جيدة لقياس هذا المتغير.

اكتسبت حركة قياس التفكير الناقد اهتمامًا كبيرًا، لا سيما في السياقات التعليمية، كما أن هناك العديد من الاختبارات على مستوى العالم لقياس التفكير الناقد، بما في ذلك مقياس واطسون - جليسر (Watson-Glaser) للتفكير الناقد واختبار مهارات التفكير الناقد في كاليفورنيا، ومع ذلك فقد تم انتقاد العديد من هذه الأدوات بسبب خصائصها السيكومترية، والتي غالبًا ما تكون غير كافية (Duru et al., 2022).

مشكلة الدراسة:

إن حركة قياس التفكير الناقد مدفوعة بالحاجة إلى أدوات تقييم مناسبة ثقافيًا وذات خصائص سيكومترية جيدة نظرًا للانتقادات التي وجهت لهذه المقاييس بسبب خصائصها السيكومترية (Duru et al., 2022)، كما تعد دراسة التفكير الناقد ذو أهمية كونه من مهارات القرن (٢١) حيث يساهم في رفع مستوى الوعي ورفع مستويات التفكير؛ إذ يساعد على تحليل المعلومات واتخاذ قرارات أكثر فاعلية بناء على تقييم الاختيارات بشكل ناقد، إضافة إلى حل المشكلات والتكيف مع التغيرات. ويعد اختبار واطسون – جليسر من أقدم وأشهر الاختبارات التي صُممت لقياس التفكير الناقد، إلا انه بعد البحث والاطلاع وجدت الباحثة ندرة في الدراسات التي تحققت من الخصائص السيكومترية لاختبار الصورة القصيرة لواطسون – جليسر (WGCT-SF) وفق نماذج نظرية الاستجابة للفقرة في مجتمعنا العربي؛ لذا اتجهت الباحثة لدراسة اختبار واطسون – جليسر للتفكير الناقد وفق نظرية الاستجابة للفقرة (IRT) كأول دراسة في المملكة العربية السعودية تتناول هذا المقياس وفق نموذج راش.

ISSN: 2537-0448 eISSN: 2537-0456

أسئلة الدراسة:

١. هل تحقق البيانات المستمدة من الصورة القصيرة لاختبار واطسون – جليسر للتفكير الناقد افتراضات نظرية الاستجابة للفقرة؟

٢. ما درجة مطابقة نموذج راش للبيانات المستمدة من الصورة القصيرة لاختبار واطسون – جليسر للتفكير الناقد ؟

٣. ما قيم معالم صعوبة الفقرة ودقة تقديرها للبيانات المستمدة من الصورة القصيرة لاختبار واطسون – جليسر للتفكير الناقد وفق نموذج راش ؟

٤. ما مستوى قدرة الأفراد ودقة تقديرهم من البيانات المستمدة من الصورة القصيرة لاختبار واطسون – جليسر للتفكير الناقد وفق نموذج راش؟

٥. ما دلالات الصدق والثبات للبيانات المستمدة من الصورة القصيرة لاختبار واطسون – جليسر للتفكير الناقد في ضوء نظرية الاستجابة للفقرة؟

٦. ما مستوى التفكير الناقد لدى طلبة جامعة تبوك؟

الأهداف:

١. التحقق من البيانات المستمدة من اختبار واطسون – جليسر للتفكير الناقد افتراضات نظرية الاستجابة للفقرة.

٢. مطابقة نموذج راش للبيانات المستمدة من الصورة القصيرة لاختبار واطسون – جليسر للتفكير الناقد.

٣. التحقق من قيم معالم صعوبة الفقرة ودقة تقديرها للبيانات المستمدة من الصورة القصيرة لاختبار واطسون – جليسر للتفكير الناقد وفق نموذج راش.

٤. التحقق من مستوى قدرة الأفراد ودقة تقديرهم من البيانات المستمدة من الصورة القصيرة لاختبار واطسون – جليسر للتفكير الناقد وفق نموذج راش.

٥. التحقق من دلالات الصدق والثبات للبيانات المستمدة من الصورة القصيرة لاختبار واطسون – جليسر للتفكير الناقد في ضوء نظرية الاستجابة للفقرة.

٦. التعرف على مستوى التفكير الناقد لدى طلبة جامعة تبوك

الأهمية:

– تبرز أهمية الدراسة في تطوير اختبار واطسون – جليسر للتفكير الناقد على البيئة السعودية، حيث تقدم للباحثين في مجال القياس النفسي والتربوي أداة لقياس التفكير الناقد وفق نموذج راش، كما تساعد المعلمين والخبراء في مجال القياس باستخدام الاختبار في ضوء نظرية الاستجابة للفقرة والتي تعد من الاستخدامات الحديثة في مجال القياس.

– يعزز استخدام نموذج راش تطوير الاختبارات والمقاييس النفسية والتربوية في ظل نظرية الاستجابة للفقرة والتي تسهم في رفع دقة القياس.

ISSN: 2537-0448 eISSN: 2537-0456

– تبرز أهمية الدراسة في إبراز مهارات التفكير الناقد التي تسهم في النمو المعرفي، واستخدام اقصى الطاقات العقلية لدى الفرد؛ لاتخاذ القرارات بشكل دقيق وسليم في ظل التسارع المعرفي في وقتنا الحاضر.

حدود الدراسة:

الحدود الموضوعية: اقتصرت الدراسة على تقدير الخصائص السيكومترية للصورة القصيرة من اختبار واطسون- جليسر للتفكير الناقد (WGCT-SF) النسخة المقننة على البيئة السعودية وفق لنموذج راش.

الحدود البشرية: اقتصرت الدراسة على طلبة جامعة تبوك.

الحدود الزمانية: خلال العام الدراسي ١٤٤٦-٢٠٢٥.

مصطلحات الدراسة:

التفكير الناقد (Critical Thinking):

جاء في قاموس الجمعية الامريكية لعلم النفس (APA): هو تفكير الموجه نحو حل المشكلات، حيث يختبر فيه الشخص الأفكار أو الحلول المحتملة بحثًا عن الأخطاء أو العيوب ويفيد في التحقق من صحة الاحتمالات وتحليل النتائج (APA Dictionary of Psychology, n.d.).

اختبار واطسون-جليسر (Watson-Glaser Test):

تعرفه الباحثة اجرائياً بأنه الدرجة الكلية التي يحصل عليها الفرد في اختبار واطسون – جليسر (WGCT-SF) للصورة القصيرة في الابعاد الخمسة للاختبار:(الاستنتاج – تمييز الافتراضات – الاستنباط- التفسير – تقويم الحجج).

نموذج راش (Rasch Model):

هو طريقة رياضية تقوم على معرفة احتمال إجابة الفرد على سؤال ما إجابة صحيحة. فهو يركز على أمرين رئيسيين، الأول: هو (قدرة الفرد) التي تمكنه من الإجابة بشكل صحيح او خاطئ. الثاني: هو (صعوبة الفقرة) لمعرفة مستوى القدرة لدى الفرد (Bond, 2015).

الإطار المفاهيمي

اولاً: التفكير الناقد

أشار بيكون إلى أن العقل لا يمكن تركه إلى ميوله الطبيعية، ليأتي بعده ديكارت ويقدم منهجًا خاصًا للعقل لتوجيهه في التفكير طوّر فيه التفكير الناقد على مبدأ الشك المنهجي، ويعد هذا السياق واضح في صياغة (الجمعية الفلسفية الأمريكية) للمفكر الناقد المثالي بأنه شخص باحث ومرن ومُطلع ولديه تقبل لوجهات النظر (جميل، ٢٠١٩).

ISSN: 2537-0448 eISSN: 2537-0456

تعريف التفكير الناقد

تعددت التعريفات حول مفهوم التفكير الناقد، إذ جاء كعملية عقلية تقوم على التقويم بشكل دقيق للاستدلال والوصول للأدلة القائمة خلف المعرفة والمحاجة، بذلك فهي تساعد على الدفاع بناء على ما لدينا من معرفة (العقل وآخرون، ٢٠٢٢).

وعرفه روبرت ستيرنبرج (Robert J. Sternberg): بأنه استراتيجيات ذهنية يستعين بها الفرد لتعلم ما هو جديد ومواجهة المشاكل واتخاذ القرارات، **وعرفه روبرت انيس**(Robert Ennis) بأنه تفكير تأملي يعمل على اتخاذ القرارات بناء على ما نفعله وما نؤمن به، **إضافة إلى ذلك تعريف لينذا إيلدر(Linda Elder)** هو تفكير يساعد الطلاب على فهم الروابط المنطقية في اللغة، **وفي تعريف لريتشارد بأول** (Richard Paul) ذكر بأنه تفكير هدفه التغلب على التحيز **كما ذكر آرونس (Arnold Arons)** بانه وعي الفرد بتفكيره ليكون قادر على الانتقال من السياق المألوف للسياق الغير مألوف، **وفي تعريف لروبرت مارتن (Robert Martin)** بأنه التفكير الذي يهدف إلى التكامل **إضافة إلى ذلك تعريف جون ديوي John Dewey)** بأنه التفكير الذي يعتمد توسيع الآفاق ورؤية الصورة الأكبر فهو يتجاوز حدود المصالح الشخصية (جميل، ٢٠١٩). وفي تعريف لواطسون وجليسر بأنه فحص المعتقدات في ضوء الحقائق والشواهد (علي، ٢٠٠٨). كما يمكننا الإشارة لتعريف أكثر عمومية ذكره العقل وآخرون (٢٠٢٢) بأنه عملية عقلية تقوم على تقييم المعلومات التي تهدف لبناء المعرفة، وتصب في قدرة الشخص على معالجة المعلومات وتجميعها والحكم عليها بغرض اتخاذ القرارات المستنيرة وحل المشكلات بفاعلية.

في ضوء ما سبق نلاحظ تعدد وتنوع التعريفات لمفهوم واحد، قد يرجع ذلك لكون المفهوم كامن ويؤخذ من عدة جوانب، وهذا ما يجعل القياس النفسي والتربوي معقداً وبحاجة مستمرة لتطوير المقاييس والاختبارات التي تناسب تلك المفاهيم وتقيسها بدقة.

أهمية التفكير الناقد:

أشار علي (٢٠٠٨) إلى أن الاهتمام بالتفكير الناقد جاء من خلال عدة مجالات، إذ أن هنالك دراسات ركزت على البرامج التدريبية، وطرق واستراتيجيات التفكير الناقد مثل: دراسة (الشنقيطي، ٢٠٢٤؛ إبراهيم، ٢٠٢٣) وهنالك دراسات ركزت على معرفة طبيعة التفكير الناقد وعلاقته ببعض المتغيرات منها الذكاء والتحصيل مثل: دراسة (الزهراني، ٢٠٢٢؛ أيوب، ٢٠١٤). كما جاء الاهتمام بالتفكير الناقد في المراكز الصحية وأهمية دوره في طرح الأسئلة المنطقية للوصول للتشخيص السليم (Reggoug et al., 2024).

ISSN: 2537-0448 eISSN: 2537-0456

مهارات التفكير الناقد:

أشار فاسيوني (Facione, 1998) إلى مهارات التفكير الناقد وهي:

- التفسير: التعبير عن التجارب والمواقف من خلال مهارات فرعية كالتصنيف واستخراج المعاني وتوضيحها.
- التحليل: يشمل الاستقراء والاستنتاج للجمل والاسئلة من خلال مهارات فرعية واكتشاف الحجج وتحليلها.
- التقويم: يشير إلى إدراك الشخص لحكمته من خلال مهارات فرعية منها؛ تقويم الادعاءات.
- الاستدلال: تحديد العناصر التي تساعد في استخلاص النتائج من خلال مهارات فرعية منها؛ فحص الأدلة.
- الشرح: إعلان نتائج التفكير وتبريره في ضوء الأدلة والمفاهيم من خلال مهارات فرعية منها؛ عرض الحجج وتبرير النتائج.
- تنظيم الذات: قدرة الفرد على التساؤل للتأكد من المصداقية من خلال مهارات فرعية منها؛ اختبار الذات وتنظيمها.

وذكر العقل وآخرون (٢٠٢٢) مهارات لتفكير الناقد وهي:

- مهارة تطبيق المنطق: تقوم على بناء الحجة والتأكيد على خلوها من الأخطاء
- مهارة تحديد المحاجة: يركز على المحاجة والتحقق من سلامة بناءها.
- تحديد أسلوب بناء المحاجة: إما يكون برهاني- جدلي- خطابي- سفسطائي.
- مهارة تحليل المبررات: توضيح قوة ارتباطها بالقضية.
- مهارة كشف الغموض: اذ يسعى التفكير الناقد للوضوح التام.
- مهارة تحديد الافتراضات وتقويمها: من خلال المبررات التي تقوم عليها تلك الافتراضات.
- مهارة بناء حجة بديلة: عند نقد حجة من ونفيها او تثبيتها فهو يقوم بطرح حجة بديلة.

ثانياً: قياس التفكير الناقد

ظهرت العديد من الاختبارات والمقاييس التي تقيس القدرة على التفكير الناقد، إذ أن أغلبها ذات طابع أجنبي ومصممة لفئات معينة؛ أي لا تناسب كافة المراحل الدراسية، إضافة إلى تعدد الطرق المستخدمة للقياس، فالبعض يقيسها عن طريق فقرات، أو أسئلة اختيار من متعدد، أو حسب ضمنية المعلومات في نصوص معطاة،

ISSN: 2537-0448 eISSN: 2537-0456

أو عن طريق مواقف معينة يطلب من المفحوص الاستجابة بطرق معينة (الرباط، ٢٠٢٢).

كما جرت محاولات عديدة لقياس التفكير الناقد باختلاف المراحل إذ ارتبط العديد منها في برامج التدريب المعدة للتفكير الناقد، أو الأطر النظرية. ومن أكثر الاختبارات شيوعاً: (العتيبي، ٢٠١٥).

اختبار كورنيل للتفكير الناقد (CCTT)

قام بإعداد الاختبار كل من روبرت انيس و جيسون ميلمان وتوماس تومكو (**Robert Ennis, Jason Millman & Thomas Tomko**) عام (١٩٨٥)، ويقيس الاختبار مهارات التفكير العليا، وهي القدرة على التقييم واتخاذ القرار، ويتكون من مستويين هما مستوى(x) يناسب المرحلة الابتدائية من الصف الرابع فما فوق، ويناسب الموهوبين، ويتكون الاختبار من (٧١) سؤال موضوعي مقسم لأربعة أبعاد تقيس الاستقراء، والمصداقية، والاستنباط، والافتراضات، وجميع الأسئلة تدور حول قصة تسمى (حادث استكشاف كوكب نيكوما)، ومستوى(Z) يتكون من (٥٢) سؤال للطلاب المتميزين والموهوبين البالغين في مرحلة الثانوية، ومرحلة الجامعة و الدراسات العليا، ويقيس الاختبار(٧) أبعاد وهي: الاستقراء، والاستنتاج، و المصداقية، وتحديد الافتراضات، والدلالة، و التعريف، و التنبؤ في تخطيط التجارب(بوبو، منذر ٢٠٢٠؛ Cornell Critical Thinking Test Specimen Set, n.d).

كما أن هناك دراسات قننت الاختبار كدراسة الأحمدي (٢٠١١)، وفيها تم تقنين الاختبار للصورة (z)، ودراسة بوبو (٢٠١٨) قامت بها بتقنين ودراسة الخصائص السيكومترية للمقياس للصورة(x)، ودراسة العنزي (٢٠٢٣) قام فيها بتطوير الاختبار واستخدام نظرية الاستجابة للمفردة.

اختبار كاليفورنيا لمهارات التفكير النقدي (CCTST)

هو اختبار موضوعي تم تطويره بواسطة فاسيوني (Facione) وفريقه في جامعة سانتا كلارا ونشرته "California Academic Press" عام (١٩٩٠). وصُمم الاختبار لقياس المهارات الأساسية للتفكير الناقد، ويتضمن (٣٤) سؤالًا من نوع الاختيار من متعدد، يعتمد الاختبار على تعريف جماعي للتفكير الناقد، تم تطويره من خلال مشروع دلفي بدعم من الرابطة الفلسفية الأمريكية؛ حيث يقيس الاختبار المهارات الأساسية مثل التحليل، التفسير، الاستنتاج، التقييم، الاستنباط، والاستقراء. كما أنه يُستخدم لتقييم الأداء التعليمي، والبحث الأكاديمي، واختيار الكفاءات المهنية، ويتميز بتركيزه على المهارات الإدراكية العليا، ويعتبر أداة معتمدة

ISSN: 2537-0448 eISSN: 2537-0456

لقياس التحصيل في التفكير الناقد، يناسب طلاب الجامعات والدراسات العليا (Facione, 1991).

قام بتقنين هذا الاختبار شطناوي (٢٠٠٣)، كما قام عزيز (٢٠١٢) بتطوير الاختبار لدى طلبة الجامعة وفقاً لنظرية الاستجابة للفقرة.

اختبار نيوجيرسي للمهارات الاستدلالية(NJTRS)

تم تطويره في (١٩٨٣) بواسطة فرجينيا شيبمان(Shipman) لتقييم مهارات التفكير المنطقي والبحثي للأطفال كجزء من برنامج "الفلسفة للأطفال" الذي طوره الدكتور ماثيو ليبمان، ويتضمن الاختبار (٥٠) سؤالًا متعدد الخيارات تقيس (٢٢) مهارة تفكير منطقي مثل: الاستنتاج الاستقرائي، اكتشاف الافتراضات الضمنية، وتجنب التسرع في الاستنتاج. يستهدف الطلاب من (١٠-١٣ سنة) ويمكن استخدامه للصف الرابع بقراءة الأسئلة بصوت عالي، ويستغرق الاختبار (٣٠-٤٥) دقيقة ويهدف إلى قياس التحسن في التفكير الناقد الناتج عن البرامج التعليمية، لكنه لا يُستخدم لتشخيص نقاط الضعف والقوة الفردية بدقة (Shipman, 1983).

اختبار واطسون ـ جليسر للتفكير الناقد

يعد من أكثر الاختبارات شيوعًا، في عام (١٩٦٤) تم إطلاق أول نسختين رسميتين للاختبارين، **(Ym)** و**(Zm)**، حيث تضمنت كل نسخة (١٠٠) سؤال، وركزت على قياس التفكير الناقد في البيئات التعليمية والمهنية، وجاءت النسختان بنفس عدد الأسئلة، ولكنها تختلف من حيث الصياغة أو السيناريوهات المستخدمة، مع الحفاظ على قياس نفس المهارات، وهي تقيس كل من: الاستنتاج، وتمييز الافتراضات، والاستنباط، والتفسير، وتقويم الحجج. وكان الهدف من وجود نسختين؛ استخدامهما بشكل متبادل لضمان عدم تأثير الاعتياد على الإجابات عند إعادة الاختبار.

وفي عام (١٩٨٠) خضع الاختبار لأول تحديث رئيسي، حيث تم تقليص عدد الأسئلة إلى (٨٠) سؤال، وتحسين وضوح اللغة المستخدمة مع إزالة أي تحيزات، وأُعيد تسميته إلى النموذجين (A) و (B) ليتم استخدامهما بشكل أكثر كفاءة في المجالات المهنية، كان الهدف منها اجراء تحسينات وتعديلات تهدف إلى جعل الاختبار أكثر حداثة ودقة (Watson & Glaser, 2009).

وفي عام (١٩٩٤) تم تطوير نسخة (Form S) التي تحتوي على (٤٠) سؤالًا كنسخة مختصرة من الاختبار الأصلي، وكانت هذه النسخة تهدف إلى تلبية احتياجات خاصة تتعلق بتوفير تقييم أسرع وأكثر مرونة في سياقات معينة (Watson & Glaser, 2009).

ISSN: 2537-0448 eISSN: 2537-0456

في عام ٢٠١٠ تم تطوير النموذج الأحمر بعد تحليل الإصدارات السابقة من اختبار(Watson-Glaser (Forms A & B Short Form ، وجد الباحثون أن **هناك ثلاث مهارات رئيسة مترابطة في التفكير الناقد،** وليس خمس مهارات منفصلة كما كان يُعتقد سابقًا بناءً على هذا، تم إعادة تصميم النموذج ليتناسب مع هذه النتائج ليكون أكثر وضوحًا وسهولة في التفسير، حيث جاء التغيير في عدد العوامل الفرعية. وجد الباحثون أن الاستدلال، والتفسير، والاستنتاج جميعها تندرج تحت فئة واحدة وهي (استخلاص الاستنتاجات)(Draw Conclusions) مما يعني أن التفكير الناقد يعتمد على ثلاث مهارات رئيسة فقط، وليس كما كان سابقًا خمس مهارات (Watson & Glaser, 2018).

النموذج الأحمر (RED Model):

هو إطار يستخدم في اختبار Watson-Glaser لقياس مهارات التفكير الناقد، تم تطبيقه في الإصدار الثاني والثالث (Watson-Glaser II & III) بعد إجراء تحليلات وتجارب واسعة لضمان دقته وفعاليته. ساعد هذا التغيير في تحسين دقة التقييم، مما جعل الاختبار أكثر واقعية، وعمل على تقليل التحيزات الثقافية، وأصبح أداةً أقوى، وأكثر صدقًا لاختبار التفكير الناقد في البيئات التعليمية والمهنية حول العالم (Watson & Glaser, 2018).

كما تم إجراء تحليل عاملي استكشافي (EFA) على نتائج الاختبارات السابقة (Forms A& B Short Form) لتحديد عدد العوامل (المهارات الفرعية) التي تفسر الأداء في الاختبار، ووجد الباحثون أن ثلاثة عوامل فقط كانت مستقرة وقابلة للتفسير، وتم استبعاد العوامل الضعيفة (Watson & Glaser, 2018).

بعد تحديد النموذج بثلاثة عوامل، تم استخدام التحليل العاملي التوكيدي (CFA) لاختبار مدى تطابق البيانات مع هذا النموذج الجديد، وجاءت النتائج مسفرة عن تأكيد أن النموذج بثلاثة عوامل يوفر تطابقًا أفضل مع البيانات مقارنة بالنماذج السابقة، لم تكن هناك فروق كبيرة بين النموذج بثلاثة عوامل والنموذج التقليدي ذي الخمسة عوامل، ولكن النموذج الثلاثي كان أكثر استقرارًا وأسهل في التفسير (Watson & Glaser, 2018). كذلك تم تحليل الارتباطات بين الأبعاد (Correlation Analysis) وجد الباحثون أن الاستدلال (Inference)، والتفسير (Interpretation)، والاستنتاج (Deduction) كانت ذو ارتباط قوي، مما أدى إلى دمجها تحت عامل واحد وهو استخلاص النتائج (Draw Conclusions)، وفي المقابل جاءت كل من مهارة تقييم الحجج، والتعرف على الافتراضات أقل ارتباطاً بالمهارات الأخرى، مما أكد استقلالهما كعاملين منفصلين (Watson & Glaser, 2018).

ISSN: 2537-0448 eISSN: 2537-0456

وفي عام (٢٠١٨)، تم تحسين دقة القياس باستخدام نظرية الاستجابة للمفرد ة(IRT)، حيث تم تحليل مستويات صعوبة الأسئلة لضمان تكافؤها عبر النسخ المختلفة من الاختبار، وتطوير بنك أسئلة بحيث لا يحصل أي شخص على نفس الأسئلة بالضبط، مما زاد من مصداقية الاختبار وتقليل احتمالات الغش والاعتماد على القدرة(θ) بدلاً من الدرجة الخام والتحقق من الصعوبة والتمييز (Watson & Glaser, 2018).

تم تقديم النموذج الأحمر (RED Model) ليعكس هذه التغييرات، وأصبح يتكون من ثلاثة أبعاد رئيسة وهي:

- التعرف على الافتراضات (Recognize Assumptions):

القدرة على تحديد الافتراضات غير المصرح بها في الحجج أو الأفكار.

- تقييم الحجج (Evaluate Arguments):

القدرة على تحليل الحجج وتقييم منطقها دون تحيز.

- استخلاص النتائج (Draw Conclusions)

القدرة على اتخاذ قرارات منطقية بناءً على الأدلة المتاحة (Watson & Glaser, 2018).

وصف الاختبار:

يحتوي كل اختبار فرعي على نصوص أو سيناريوهات تتضمن مشكلة، وبيانات، وحجج، وتفسيرات للبيانات، مشابهة لما يواجه الأفراد يوميًا، تتطلب هذه السيناريوهات تحليلًا نقديًا، حيث لا يمكن قبول المعلومات المطروحة دون تقييم دقيق وتتنوع السيناريوهات بين محايد ومثير للجدل، ويعد هذا الاختبار من اختبارات القوة؛ فهو يقيس جودة وعمق التحليل، والاستنتاج، بدلاً من سرعة الأداء، وقد صُمم ليطبّق ضمن إطار زمني مناسب. ضمن مختلف الإصدارات (١٩٦٤-٢٠١٨) بقيت الفئة العمرية متشابهة، حيث تركز على البالغين ١٨ سنة فأكثر (Watson & Glaser, 2018).

ثالثاً: نموذج راش

هو نموذج طوره عالم الرياضيات الدنماركي جورج راش عام (١٩٦٠ Rasch,)، هو أحد نماذج نظرية الاستجابة للفقرة ويعتمد على معلمة واحدة، ويتميز بمرونته حيث يمكن تطبيقه على الفقرات الثنائية، ويحدد احتمالية استجابة معينة بناءً على قدرة المستجيب وصعوبة الفقرة. يقوم نموذج راش بمقارنة الأشخاص والفقرات على نفس المقياس، مما يجعل تقديرات المعالم الخاصة بالأشخاص والفقرات مستقلة؛ بمعنى تحرر القياس من خصائص الفقرات والافراد (Khine, 2020).

ISSN: 2537-0448 eISSN: 2537-0456

يعد نموذج راش أحد أهم نماذج نظرية الاستجابة للفقرة وابسطها، كما يتميز بكثرة التطبيق لتطوير الاختبارات والمقاييس النفسية والتربوية، ويأتي هذا النموذج كأحد النماذج أحادية المعلم لنظرية الاستجابة للفقرة ويستخدم في تحليل الفقرات ثنائية الاستجابة (بن وزة وقماري، ٢٠٢١).

افتراضات نموذج راش:

يفترض نموذج راش ثلاث افتراضات أساسية تضمن دقته وفعاليته في القياس:

أولًا: أحادية البعد(Unidimensionality): حيث يُفترض أن استجابات الأفراد تعتمد على متغير كامن واحد فقط، مما يعني أن جميع العناصر في الاختبار تقيس نفس السمة أو المهارة الأساسية، مثل الكفاءة في الرياضيات.

ثانيًا: الاستقلالية المحلية (Local Independence): والتي تعني أن استجابات الفرد لأي عنصر لا يجب أن تعتمد على استجاباته لعناصر أخرى، بشرط تثبيت مستواه من القدرة (θ)؛ أي أن إجاباته تكون مستقلة عند التحكم في القدرة الكامنة.

ثالثًا: الشكل الوظيفي(Functional Form): حيث يُفترض أن العلاقة بين القدرة الكامنة واحتمالية الإجابة الصحيحة تتبع منحنى لوجستي على شكلS، مما يعني أن احتمالية الإجابة الصحيحة تتغير تدريجيًا وفقًا لمستوى قدرة الشخص، ويتم تمثيل ذلك عبر دالة استجابة العنصر(IRF) (de Ayala, 2022).

الإطار النظري

محور أول: الإطار النظري للتفكير الناقد:

بعد البحث والتقصي اتضح أن التفكير الناقد ليس مفهومًا حديثاً، بل تعود جذوره إلى الفلاسفة، وإذ تتبعنا هذه الجذور تعود بنا إلى (سقراط) الذي أشار إلى التساؤلات السقراطية (Socratic Questioning)، وهي تعد من أفضل الاستراتيجيات في تعليم التفكير الناقد كما أكد على أهمية طرح أسئلة تبحث بشكل عميق في التفكير، بينما في العصر الحديث يعد جون ديوي(Dewey) الرائد في نشأة التفكير الناقد في صورته المعاصرة، إذ كان في تداخل مع مفهوم التفكير التأملي واعتبرهما مترادفين، إلا أنه قام بتطوير المفهوم لاحقاً بأن يكون أساساً للتفكير الناقد، وفيه حث على التحرر من الجمود الذهني، لذا يمكن القول بإنه لم يتشكل كمفهوم مستقل ومجالًا من مجالات التفكير إلا في العصر الحاضر (العنزي، ٢٠٢٣؛ جميل، ٢٠١٩). ويعد كل من وروبرت اينيس (Robert Ennis)، ولندا الدر (Linda Elder) وريتشارد بول (Richard Paul)، من أشهر العلماء اللذين برزت جهودهم في هذا المجال؛ إذ قدموا أطر نظرية لفهم ماهية التفكير الناقد، والتي تساعد وتسهم في تطور مفهوم التفكير الناقد، وقياسه لإدخاله في المناهج التربوية، واستخدامه كمهارة في الحياة اليومية.

ISSN: 2537-0448 eISSN: 2537-0456

أولاً: الميول للتفكير الناقد: (Critical Thinking Dispositions)

المفكر الناقد المثالي يميل إلى:

١. **الاهتمام بأن تكون معتقداته صحيحة وأن تكون قراراته مبررة**: أي السعي للوصول إلى الحقيقة قدر الإمكان، ويشمل:

- البحث عن فرضيات بديلة، وتفسيرات، واستنتاجات، وخطط، ومصادر.
- أخذ وجهات نظر الآخرين بجدية.
- مطلع جيّد.
- قدرته على تبرير المواقف بالمعلومات المتاحة.

٢. **الاهتمام بفهم وتقديم موقفه بوضوح وصدق، سواءً موقفه الشخصي أو موقف الآخرين**؛ ويشمل:

- الاستماع إلى وجهات نظر الآخرين وأسبابهم.
- توضيح المعنى المقصود لما يُقال أو يُكتب أو يتم التعبير عنه بأي طريقة أخرى، مع السعي لتحقيق الدقة بقدر ما يتطلبه الموقف.
- البحث عن الأسباب وتقديمها.

٣. **الاهتمام بالآخرين** (هذا الميل ثانوي ولا يعد أساسي، لكن التفكير الناقد قد يكون صعب دونه) المفكرون النقديون المهتمون بالآخرين:

- يتجنبون تخويف وترهيب الآخرين بقدراتهم النقدية، مع أخذ مشاعرهم ومستوى فهمهم في الاعتبار.
- يهتمون براحة الآخرين (Ennis, 2011a).

ثانياً: المعايير الفكرية

القواعد أو الإرشادات التي تساعد على حل المشكلات واتخاذ القرارات بطريقة ذكية عند استخدامها، تُمَكن الفرد من الفهم بشكل أفضل وإيصال الفكرة بشكل أكثر وضوحًا، فهي تُستخدم لتقييم جودة التفكير.

- **الوضوح**: يعد أهم المعايير لأنه يمثل المدخل الرئيسي لباقي المعايير.
- **الصحة**: أن تكون المعلومات موثوقة.
- **الدقة**: تكون التفاصيل ذات كفاية لنقل المعنى الدقيق وتجنب الغموض.
- **الربط:** العلاقات بين الأسئلة المطروحة والمداخلات والحجج.
- **العمق:** تجاوز سطحية المعالجات والانتقال لما هو أعمق.
- **الاتساع**: أخذ جميع جوانب المشكلة أو الموضوع.
- **المنطق:** مدى تسلسل وانتظام الأفكار المطروحة بحيث تقود إلى نتائج صحيحة.

ISSN: 2537-0448 eISSN: 2537-0456

- **الدلالة أو الأهمية:** مدى قيمة وأهمية المعلومات المطروحة.
- **الانصاف**: عدم الانحياز والميل لطرف دون الاخر (العتيبي، ٢٠١٥؛ Pual & Elder, 2008)

ثالثاً: السمات الفكرية

تطوير السمات الفكرية يسهم في الانفتاح والاطلاع على وجهات النظر الأخرى، فهو يساعد في اتخاذ قرارات منطقية، وحل المشكلات، وتقييم المعلومات والأفكار بدقة، فهي تحافظ على إبقاء التفكير محدداً وواضحًا.

نموذج المفكر الناقد في تطبيق ما سبق ذكره من معايير وعناصر ودورها في تمنية السمات الفكرية

يقدم إطار بول وإيلدر (Pual & Elder) تحديدًا لكل من عناصر وسمات ومعايير التفكير الناقد التي تقدم فهمًا أعمق للتفكير الناقد، والذي بدورة يُمكِننا من تطوير مقاييس واختبارات للتفكير الناقد، وتحسين جودة هذه المقاييس، للاستفادة منها في مختلف المجالات.

محور ثاني: الإطار النظري للقياس النفسي والتربوي

١. نظرية القياس التقليدية (Classical Test Theory):

في عام (١٩٥٠) قام العالم هارولد جولكيسن (Harold Gulliksen) بتطوير نظرية القياس التقليدية التي تعد من أقدم النظريات في القياس النفسي والتربوي، وقد تم استخدامها في بناء وتطوير الاختبارات النفسية والعقلية، مثل اختبارات الميول، والذكاء، والاختبارات التربوية والتحصيلية، وقد أُطلق عليها عدد من التسميات منها نظرية الدرجة الحقيقية ونظرية الاختبار (سليمان وأبو علام، ٢٠١٠).

ولسنوات عديدة استمرت هذه النظرية تمثل الأساس النظري الوحيد للقياس في العلوم السلوكية إذ يعتمد نموذج النظرية على الدرجة الحقيقة والتي تمثل السمة الكامنة للفرد والتي لا يمكن ملاحظتها بشكل مباشر، لذا يتم قياس الدرجة الملاحظة، وعملية القياس هذه تحتوي على قدر من الخطأ في تقدير الدرجة الحقيقية (Hambleton & Zaal, 1991).

افتراضات النظرية التقليدية

تفترض أن الدرجات الملاحظة في أي اختبار هي محصلة مجموع الدرجات الحقيقية بالإضافة إلى عوامل الخطأ، تمثل الدرجة الحقيقية القيمة الفعلية أو المثالية للسمة التي نقيسها، بينما يشير مصطلح الخطأ إلى تأثير العوامل الخارجية غير المرتبطة مباشرة بالسمة المقاسة، لكنها تؤثر على نتائج القياس، وتفترض النظرية أن الأخطاء غير مترابطة مع الدرجات الحقيقية، مما يعني أن تأثير الخطأ يحدث بشكل مستقل وعشوائي، ويتم تحديد العلاقة بين الدرجات الملاحظة من خلال حساب معامل

ISSN: 2537-0448 eISSN: 2537-0456

الارتباط أو تحليل التباين، مما يساعد في تقدير مدى تأثير الخطأ وتحديد نسبة الدرجة المرصودة التي تمثل الدرجة الحقيقية (Franzen, 2011).

يذكر سليمان وابوعلام (٢٠٠٩) مميزات للنظرية التقليدية وهي:

- سهولة استخدام هذه النظرية ووضوح مفاهيمها وبساطتها، يجعلها في متناول الباحثين والممارسين حتى وإن لم يكونوا متخصصين في مجال القياس، فهي توفر إطارًا مرنًا لتحليل البيانات ومعالجتها، مما يسهل استنتاج النتائج وتفسيرها دون الحاجة إلى خلفية إحصائية معقدة.
- بساطة الافتراضات التي تقوم عليها النظرية التقليدية، مما يجعلها مناسبة لتحليل بيانات مختلف أنواع الاختبارات النفسية والتربوية، كما أنها لا تتطلب استخدام نماذج رياضية معقدة، مما يسهل تطبيقها وفهم نتائجها دون الحاجة إلى خبرة متقدمة في مجال القياس أو الإحصاء.
- يمكن تحليل بيانات الاختبارات بسهولة ودون الحاجة إلى استخدام برامج متخصصة أو الاعتماد على خبراء في تحليل البيانات أو الحاسب الآلي، مما يجعلها أكثر مرونة وسهولة في التطبيق.

أوجه القصور في النظرية التقليدية للاختبار:

– تعتمد جميع الخصائص السيكومترية للاختبارات، مثل معامل الصدق والثبات، بالإضافة إلى الخصائص السيكومترية لمفردات الاختبار، مثل معامل الصعوبة ومعامل التمييز، على عينة الأفراد التي يُطبق عليها الاختبار لذا، في حال استخدم باحث آخر نفس الاختبار على عينة مختلفة، يمكن أن تتغير هذه الخصائص بشكل ملحوظ.

– تفترض هذه النظرية أن درجات الاختبار تعكس السمة أو القدرة المقاسة وفق علاقة خطية طردية (Monotone Function)، أي أنه كلما ارتفعت درجة المشارك، دل ذلك على ارتفاع مستوى السمة التي يتم قياسها، ومع ذلك قد يحصل بعض الأفراد ذوي القدرة العقلية العالية على درجات منخفضة في الاختبار، والعكس فقد يحصل بعض الأفراد ذوي القدرات العقلية المنخفضة على درجات مرتفعة، مما يشير إلى وجود استثناءات لهذا الافتراض.

– تنطلق النظرية من افتراض أن الاختبارات تقيس متغيرًا أحادي البعد، وهو فرض صحيح في العلوم الفيزيائية، حيث يتم قياس خاصية واحدة فقط مثل الطول أو الوزن أو الحجم لكن في الاختبارات النفسية مثل: اختبارات الذكاء، قد تتأثر الدرجات بعوامل أخرى غير الذكاء، مثل: البيئة الاجتماعية، والإمكانات المادية بالتالي، فإن

ISSN: 2537-0448 eISSN: 2537-0456

الدرجة التي يحصل عليها الفرد لا تعكس الذكاء فقط، وهذا ما يجعل تحقق أحادية البعد أمرًا صعبًا في العديد من الاختبارات النفسية.

– تعتمد الدرجة الكلية للاختبار على أسئلة الاختبار، مما يعني أن درجة الطالب تتحدد بناءً على طبيعة الأسئلة المستخدمة، وإذا تم تغيير الاختبار مع أنه يقيس نفس السمة، فقد تختلف درجة الطالب.

– تعاني أساليب القياس التقليدية من تغير في معاني مفردات الاختبار مع مرور الوقت، كما أن البنية العاملية للاختبار ليست ثابتة، حيث تتأثر بالظروف البيئية كذلك حذف بعض الفقرات أو تعديلها قد يؤدي إلى تغييرات في درجات الأفراد، مما يزيد من صعوبة التنبؤ بنتائج الاختبار عبر الزمن.

– يستند تعريف ثبات الاختبارات إلى مفهوم الاختبارات المتكافئة، رغم أن تحقيق تكافؤ الاختبارات عمليًا أمرًا معقدًا، بالإضافة إلى أن الأفراد لا يكونون في نفس الظروف الاختبارية عند تطبيق نسخة أخرى من الاختبار؛ لأنهم قد يكتسبون مهارات جديدة أو تتغير مستويات الدافعية والقلق لديهم، مما يؤدي إلى اختلاف في درجاتهم المرصودة.

– لا تقدم النظرية تمييزًا بين أنواع أخطاء القياس، بل تكتفي بمنح قيمة تقديرية إجمالية لمصادر الخطأ المختلفة (سليمان وأبو علام، ٢٠٠٩؛ محاسنة، ٢٠١٣).

٢. نظرية الاستجابة للفقرة (Item Response Theory)

تمثل نظرية الاستجابة للفقرة (Item Response Theory) (IRT) الاتجاه المعاصر في القياس التربوي والنفسي وهي في طبيعتها تركز على تحليل الاستجابات للعناصر (مثل الأسئلة في اختبار معين)، بدلًا من الاعتماد فقط على الدرجة الكلية للاختبار كما كان سائدًا في النظرية التقليدية في القياس. تعتمد نظرية الاستجابة للفقرة (IRT) على استخدام نماذج رياضية لتمثيل العلاقة بين قدرة الشخص (أو مستوى السمة الكامنة لديه، مثل الذكاء أو التحصيل الدراسي او التفكير الناقد) واحتمالية استجابته بشكل صحيح أو معين على عنصر معين (Baker, 2001).

بدأت نظرية الاستجابة للعناصر (IRT) في عام (١٩٤٣) عندما نشر الدكتور ديفيد لاولي (**David Lawley**) من جامعة إدنبرة ورقة بحثية أظهرت إمكانية التعبير عن مفاهيم نظرية الاختبار التقليدية باستخدام منحنى مميز للعناصر (Item Characteristic Curve)، مما وضع الأساس للنظرية كإطار قياس. وفي الخمسينيات والستينيات، ساهم الدكتور لورد (Lord) من خدمة الاختبار التعليمي في تطوير النظرية بشكل كبير، حيث قام بتوسيع مفاهيمها وتطبيقاتها وتطوير برامج حاسوبية لدعم استخدامها، مما أدى إلى إصدار كتبه الكلاسيكية في هذا المجال و في

ISSN: 2537-0448 eISSN: 2537-0456

نفس الفترة، أبرز د. رايت (Right) من جامعة شيكاغو أهمية أعمال الرياضي الدنماركي جورج راش (Rasch)، الذي قدم نموذجًا بسيطًا يعرف بنموذج راش (Rasch Model) كجزء من (IRT)، شهدت النظرية في السبعينيات والثمانينيات تطبيقًا واسعًا في تطوير الاختبارات المعيارية، ومع تطور البرمجيات والقدرات الحاسوبية في التسعينيات، أصبحت النظرية معيارًا أساسيًا لتحليل الاختبارات بفضل دقتها ومرونتها، مما رسخ مكانتها في القياس النفسي والتعليمي على المستوى العالمي (Baker, 2001).

مفردات نظرية الاستجابة للفقرة:

تطرح النظرية عدة مفردات ومفاهيم حديثة ومترابطة، وهي تُستخدم لتحليل العناصر بدقة من خلال دراسة العلاقة بين قدرة المستجيب، واحتمالية استجابته بشكل صحيح، مع التركيز على خصائص الفقرة والمعلومات التي تقدمها، يمكن تعريفها وشرحها كما يلي:

١. القدرة (θ) (Ability):

تشير إلى السمة الكامنة التي تُقاس لدى الفرد، مثل: الذكاء، أو مستوى المعرفة ويتم تمثيل القدرة بمقياس مستمر، وعادةً ما يستخدم من (-٣ إلى +٣)، حيث تشير القيم العالية إلى مستويات أعلى من السمة.

٢. الفقرة (I) (Item):

تشير إلى السؤال في الاختبار الذي يتم تحليله، كل فقرة يتم تقييمها بناءً على خصائصها (مثل الصعوبة، التمييز).

٣. منحنى المميز للفقرة (ICC) (Item Characteristic Curve):

هو منحنى رياضي يمثل العلاقة بين قدرة الفرد واحتمالية إجابته الصحيحة على الفقرة ويأخذ شكل منحنى (S) ويعتمد على معلمات الفقرة.

٤. معالم الفقرة (Parameters): هي التي تصف خصائص الفقرة وتحدد شكلها على المنحنى، هناك ثلاث معلمات رئيسية:

الصعوبة (B) (Difficulty):

تحدد موقع الفقرة على مقياس القدرة وفيها تكون القيمة الإيجابية تعني صعوبة الفقرة ويحتاج الفرد إلى قدرة أعلى للإجابة الصحيحة، بينما القيم السالبة تعني سهولة الفقرة.

التمييز (a) (Discrimination):

يعكس قدرة الفقرة على التمييز بين الأفراد ذوي القدرات المختلفة وفيها تكون الفقرات ذات التمييز العالي أكثر حساسية للفروقات بين القدرات.

التخمين (c) (Guessing):

ISSN: 2537-0448 eISSN: 2537-0456

يعكس احتمال إجابة الفرد على الفقرة بشكل صحيح بالاعتماد على التخمين فقط ويكون غالبًا في الأسئلة متعددة الخيارات (Baker, 2001).

افتراضات نظرية الاستجابة للفقرة:

أحادية البُعد (Unidimensionality):

يُفترض أن الأداء على الفقرات يتم تحديده بواسطة سمة كامنة واحدة، وهذه السمة هي التي تؤثر على احتمالية الإجابة الصحيحة عن الفقرات، وللتحقق من هذا الافتراض نستخدم التحليل العاملي **(Factor Analysis)** لمعرفة ما إذا كان عامل واحد فقط يفسر التباين في البيانات (عبد الجليل، ٢٠٢٤؛ Hambleton& Swaminathan, 1985؛ Lord,2012).

الاستقلال الموضعي (Local Independence):

يُفترض أن تكون استجابات المفحوصين عن الفقرات مستقلة عن بعضها البعض عند تثبيت مستوى القدرة أو السمة الكامنة؛ بمعنى أن الإجابة على فقرة معينة لا تؤثر على الإجابة عن فقرة أخرى (Hambleton et al., 1991؛ Lord, 2012).

أشار كل من هامبلتون وسوامينثان (Hambleton & Swaminathan,1985) إلى تكافؤ في افتراض أحادية البعد والاستقلال الموضعي من الناحية العملية، ويمكن النظر إلى الاستقلال الموضعي كدليل عملي على تحقق أحادية البُعد؛ هذا يعني أنه إذا تم التحقق من افتراض أحادية البعد، أي ان هناك بعد واحد فقط يؤثر على الأداء، فلن تكون هناك ارتباطات بين إجابات الفقرات بعد التحكم في هذا البُعد.

المنحنى المميز للفقرة (Item Characteristic Curves):

يمثل العلاقة بين احتمال الإجابة الصحيحة على فقرة معينة والقدرة الكامنة ويصف كيفية تغير احتمال الإجابة الصحيحة مع تغير مستوى القدرة الكامنة، يمكن أن تأخذ هذه المنحنيات أشكالًا مختلفة بناءً على النموذج المستخدم (مثل النموذج اللوجستي احادي المعلمة أو الثنائي أو الثلاثي (Hambleton & Swaminathan, 1985).

التحرر من السرعة (Non-Speededness):

يُفترض أن جميع المفحوصين لديهم الوقت الكافي للإجابة على كل الفقرات في الاختبار. بمعنى أن الأداء لا يتأثر بضيق الوقت، وبالتالي فإن الاستجابات تعكس القدرات الحقيقية للمفحوصين وليس تأثير السرعة أو الضغط الزمني (Hambleton & Swaminathan, 1985).

نماذج نظرية الاستجابة للفقرة

تُستخدم النماذج لنمذجة العلاقة بين السمة الكامنة المراد قياسها والإجابات عن فقرات الاختبار وهي في طبيعتها تنقسم الى نماذج ثنائية التدريج

ISSN: 2537-0448 eISSN: 2537-0456

(Dichotomously) ونماذج متعددة التدريج (Polytomous Models) يتم تحديدها بناءً على السمة المراد قياسها(Harvey & Hammer, 1999).
يوجد ثلاثة نماذج رئيسية تُستخدم بشكل شائع في تحليل البيانات الثنائية (Dichotomously Scored Items)، وهي:

١- النموذج اللوجستي أحادي المعلمة (١PL):

يعتمد هذا النموذج على معلمة واحدة وهي الصعوبة (difficulty) لكل فقرة، ويفترض أن جميع الفقرات لها نفس القدرة على التمييز (discrimination)، ويُعرف غالباً بإنه نموذج راش لكن هناك فرق بين نموذج راش والنموذج أحادي المعلمة (Harvey & Hammer, 1999).

يرى البعض أن نموذج **راش** يمثل منظورًا فلسفيًا مختلفًا عن النموذج اللوجستي احادي المعلمة، يُقيّد راش التمييز حيث يفترض أنه (a =١)، بينما يسعى النموذج أحادي المعلمة إلى تحقيق أفضل تطابق ممكن مع البيانات دون فرض هذا القيد، أي أن النموذج أحادي المعلمة يفترض التمييز ثابتاً، لكن لم يحدد قيمة محددة له، ومن الناحية الرياضية يعد النموذجان متكافئين، حيث يمكن تحويل القيم من أحدهما إلى الآخر، و يُنظر لنموذج **راش** كأداة لبناء متغير القياس نفسه وفقًا لهذا المنظور، يُعتبر نموذج **راش** الأفضل عند بناء أداة قياس لمتغير معين وهو نهج مشابه لأساليب مثل **(Guttman Scaling)**(de Ayala, 2022).

٢-النموذج اللوجستي ثنائي المعلمة (٢PL):

يضيف هذا النموذج معلمة وهي التمييز (discrimination)، بجانب الصعوبة ويسمح بتمثيل الاختلافات في قوة العلاقة بين كل فقرة والسمة الكامنة كما أن الفقرات ذات التمييز الأعلى تُوفر معلومات أكثر عن السمة الكامنة (Harvey & Hammer,1999).

٣-النموذج اللوجستي ثلاثي المعلمة (٣PL):

يُضيف هذا النموذج معلمة ثالثة تُعرف بمعلمة التخمين (Guessing)، التي تمثل احتمال الإجابة الصحيحة بسبب التخمين العشوائي ويُعتبر مناسبًا للاختبارات التي تتضمن أسئلة اختيار من متعدد (Harvey & Hammer, 1999).

ويمكن التعبير عن هذا النموذج كما جاء في بيكر **(**Baker, 2001**)** بالصيغة الرياضية التالية:

مزايا نظرية الاستجابة للفقرة:

يذكر هامبلتون وسوامينثان(Hambleton & Swaminathan, 1985) أن هنالك ثلاثة مزايا رئيسة هي:

- **استقلالية تقدير القدرة عن عينة المفحوصين:**

ISSN: 2537-0448 eISSN: 2537-0456

تقديرات القدرة (θ) تكون مستقلة عن عينة المفحوصين، بمعنى أن تقدير قدرة المفحوص لا يتأثر بمستوى القدرة للمجموعة التي تم تقديرها منها.

- **استقلالية معالم المفردات عن عينة المفحوصين:**

معالم المفردات مثل الصعوبة والتمييز تكون مستقلة عن عينة المفحوصين، مما يعني أن معالم المفردات تبقى ثابتة عبر مجموعات مختلفة من المفحوصين.

- **دقة تقدير القياس:**

توفر النظرية إحصائية تشير إلى دقة تقدير القدرة (θ) لكل مفحوص، وهذه الدقة يمكن أن تختلف من مفحوص لآخر بناءً على مستوى القدرة.

إضافة إلى ذلك، إن تحقق هذه المزايا مقرونًا بمدى تطابق النموذج (Model) المستخدم مع بيانات الاختبارات وهو ما يشار إليه بجودة المطابقة (goodness of fit) وأنه إذا كان التطابق ضعيف لن تتحقق هذه المزايا أو تحققها سيكون بدرجة محدودة.

من خلال ما تطرقت إليه الباحثة في الإطار المفاهيمي والإطار النظري يتضح بروز أهمية التفكير الناقد في العلوم التربوية كمفهوم معرفي في علم النفس والحاجة لمقاييس واختبارات أكثر تطوراً، ولتعزيز استخدام مفاهيم نظرية الاستجابة للفقرة عوضًا عن مفاهيم النظرية التقليدية التي تم استخدامها في تقنين اختبار واطسون-جليسر على البيئة السعودية، كذلك غالبية الاختبارات والمقاييس النفسية والتربوية؛ لذا سعت الباحثة لتطوير الاختبار من خلال استخدام نظرية الاستجابة للفقرة والتي بدورها تسعى لتحقيق أقصى دقة وجودة للقياس، لتساعد في استخدام اختبار واطسون-جليسر في المجالات التربوية والنفسية داخل البيئات السعودية والعربية لقياس التفكير الناقد.

الدراسات السابقة

بعد البحث والاطلاع وجدت الباحثة ندرة في الدراسات التي تناولت الاختبار بصورته القصيرة من منظور نظرية الاستجابة للفقرة وفي حدود امكانيات الباحثة اتضح انه لا توجد دراسة في المملكة العربية السعودية تناولت الاختبار من منظور نظرية الاستجابة للفقرة (لكافة النماذج) ويوجد دراسة عربية واحدة وهي **دراسة صباح (٢٠٢١)** بعنوان استخدام نموذج راش في تدريج اختبار التفكير الناقد(اختبار واطسون-جليسر للتفكير الناقد الصورة القصيرة) والتي هدفت إلى التحقق من الخصائص السيكومترية لاختبار التفكير الناقد باستخدام نموذج راش تمثلت عينتها في (٢٣١) طالبًا وطالبة من جامعتي سعيدة والشلف بالجزائر واستخدمت (الصورة القصيرة من اختبار واطسون-جليسر) النسخة الموجودة في دراسة (عشوي وآخرون، ٢٠١٧) لجمع البيانات وفق المنهج الوصفي، وتم تحليل البيانات باستخدام برنامج

ISSN: 2537-0448 eISSN: 2537-0456

(Winsteps)، أظهرت نتائج الدراسة أن افتراضات نموذج راش قد تحققت، حيث تم التحقق من أحادية البعد للاختبار، وتبين أن البيانات تتوافق مع النموذج اذ تراوحت قيم الملائمة التقاربية للفقرات بين (٠.٩٠ - ١.٠٦)، وقيم الملائمة التباعدية بين (٠.٨٩ - ١.١٠)، مما يدل على أن جميع الفقرات كانت ملائمة للنموذج ولم يتم استبعاد أي فقرة من الاختبار كذلك للأفراد، تراوحت قيم الملائمة التقاربية بين (١.٢٠ - ١.٥٢)، وقيم الملائمة التباعدية بين (١.٢٠ - ١.٦٩)، مما يشير إلى ملاءمة الافراد للنموذج، وتراوحت قيم الصعوبة بين (-١.٧٩ إلى +١.٧٧)، مع متوسط صعوبة (٠.٠٠) وانحراف معياري (٠.٧٤) ، كما بلغ معامل ثبات الأفراد (٠.٣٢) ومعامل ثبات الفقرات (٠.٩٦)، و بلغ معامل فصل الأفراد (٤.٩٢) ومعامل فصل الفقرات (٠.٦٨)، مما يشير إلى أن الاختبار يتمتع بثبات جيد.

كما وجدت الباحثة دراسات للهدف نفسه لكن باستخدام نسخ أخرى من الاختبار مثل دراسة صالح (٢٠٢٣) بعنوان تدريج مفردات اختبار واطسون-جليسر للتفكير الناقد باستخدام نموذج راش والتي هدفت إلى تدريج فقرات الاختبار باستخدام نموذج راش تمثلت العينة من ٣٨٥ طالبًا وطالبة من كلية التربية بجامعة عمر المختار واستخدمت الدراسة اختبار واطسون- جليسر ويتكون الاختبار من (٩٩) فقرة، وفق المنهج الوصفي وتم تحليل البيانات باستخدام برنامج (Winsteps) للتحقق من مؤشرات المطابقة الداخلية والخارجية لفقرات الاختبار وفقًا لنموذج راش ، وتوصلت النتائج إلى استبعاد (١٤) فقرة غير مطابقة للنموذج، بسبب تجاوز قيم المطابقة الخارجية لها الحدود المقبولة، مما أدى إلى استبعادها بينما تم الاحتفاظ بـ (٨٥) فقرة تحقق افتراضات نموذج راش، كما أظهرت النتائج أن (٢٣) طالباً من عينة الدراسة غير مطابقين، وتم استبعادهم والاحتفاظ باستجابات (٣٦٢) طالبًا، كذلك تراوحت صعوبة الفقرات بين (-١.٦١ إلى ١.٨٣) لوجت، مما يشير إلى تدرج هرمي في صعوبة المفردات، وبلغ معامل الثبات للأفراد (٠.٨٥) وللفقرات (٠.٩٨)، مما يدل على ثبات عالٍ للاختبار وبلغ معامل الفصل للأفراد (٢.٣٩) وللفقرات (٦.٨٠)، مما يشير الى فاعلية الأداة لقياس التفكير الناقد.

ودراسة الجبوري (٢٠١٧) بعنوان الدرجة الحقيقية لاختبار (واطسون-جليسر) للتفكير الناقد باستخدام النظرية التقليدية والنظرية الحديثة في القياس لدى طلبة الجامعة والتي تهدف إلى إعداد الاختبار وفق نظرية القياس التقليدية ونظرية السمات الكامنة وحساب الدرجة الحقيقية للاختبار باستعمال معادلات الدرجة الحقيقية للنظريتين، وتمثلت عينتها (٤٠٠) من طلبة جامعة بابل واستخدمت اختبار واطسون-وجليسر النسخة الاصلية بعد ترجمتها للغة العربية وتكييفها للبيئة العراقية وفق المنهج الوصفي، تم تحليل الفقرات وحساب الدرجة الحقيقية وفق النظرية التقليدية للقياس،

ISSN: 2537-0448 eISSN: 2537-0456

ونموذج راش. وأبرزت النتائج أن طرق تكييف وتطوير الاختبار وفق نموذج راش أكثر موضوعية من خطوات القياس التقليدية إضافة إلى أن الدرجة الحقيقية المحسوبة وفقاً لمعادلة نموذج راش كانت الأكثر دقة ووضوحًا من الدرجة الحقيقية المحسوبة وفق المعادلة التقليدية.

وفي دراسة سويتا (Suwita et al., 2024) بعنوان تقييم مهارات التفكير الناقد لدى طلاب المرحلة الإعدادية في عملية البناء الضوئي: مقاربة نموذج راش هدفت الدراسة إلى قياس مهارات التفكير الناقد لدى طلاب المرحلة الإعدادية في موضوع البناء الضوئي باستخدام نموذج راش، تمثلت العينة من (٤٦٧) طالبًا باستخدام أداة اختبار مستمدة من مقياس واطسون-جليسر للتفكير الناقد، لجمع البيانات وفق المنهج الوصفي، أظهرت نتائج الدراسة أن افتراضات نموذج راش قد تحققت، حيث تم التحقق من أحادية البعد للاختبار، كما أظهرت النتائج أن قيم المطابقة الداخلية والخارجية للفقرات والأفراد كانت ضمن النطاق المقبول، وهي متوافقة مع نموذج راش إضافة الى ذلك تراوحت صعوبة الفقرات من الاسهل (-٠.٨٦) إلى الأصعب (١.٣٦) وأوضحت النتائج أن الاستدلال كان أصعب مهارة للطلاب، بينما كان التقييم أسهل مهارة، مما يشير إلى تباين مستويات التفكير الناقد لدى الطلاب، كما بين التحليل أن الفقرات لم تكن سهلة أو صعبة بشكل مفرط وكانت متناسبة مع قدرات الطلاب، و بلغ معامل الثبات للأفراد (٠.٨٦)، وللفقرات (٠.٩٥)، مما يشير إلى أن الأداة تتمتع بثبات عالٍ في قياس مهارات التفكير الناقد وبلغ معامل الفصل للأفراد (٢.١٤)، وللفقرات (٤.٣٦)، مما يدل على قدرة الأداة على التمييز بين مستويات القدرة المختلفة لدى الطلاب والفقرات.

دراسات تناولت اختبار واطسون وجليسر:

دراسة اريف (Arif, 2024) هدفت إلى تقييم مهارات التفكير الناقد في تعلم العلوم باستخدام اختبار واطسون- جليسر Watson-Glaser Critical Thinking Appraisal (WGCTA)، وتمثلت عينتها في فقرات الاختبار وفق المنهج النوعي المعتمد على تحليل الأدبيات، شملت الدراسة مراجعة تطبيق اختبار (WGCTA) على مفاهيم العلوم، حيث ركزت على خمس اختبارات فرعية: الاستدلال، والتعرف على الافتراضات، الاستنتاج، التفسير، وتقييم الحجج، أظهرت النتائج أن الاختبار مناسب لقياس مهارات التفكير الناقد في تعلم العلوم، ويمكن تطويره واستخدامه بشكل أكبر لتقييم قدرات الطلاب على التفكير الناقد في المواد العلمية.

دراسة أسلان (Aslan et al., 2019) والتي هدفت إلى تطوير اختبار-Watson Glaser - Short Form والتحقق من صدق وثبات الاختبار شملت العينة (٧٧٣) طالبًا جامعيًا من كلية التربية بجامعة إسطنبول، حيث استخدمت أداة -Watson

ISSN: 2537-0448 eISSN: 2537-0456

Glaser Critical Thinking Appraisal - Short Form، التي تتكون من خمس اختبارات فرعية: الاستدلال، والتعرف على الافتراضات، والاستنتاج، والتفسير، وتقييم الحجج لجمع البيانات أظهرت النتائج أن الاختبار يمتلك مستوى جيد من الصدق والثبات، مما يجعلها أداة فعالة لقياس مهارات التفكير الناقد لدى طلبة الجامعات.

دراسة دين (Din, 2020) هدفت الدراسة إلى تقييم قدرة طلبة الجامعات في باكستان على التفكير الناقد كما ينعكس في مهاراتهم في القراءة النقدية، تمثلت عينة الدراسة من (٥٥٠) من طلبة المرحلة الجامعية في عدد من الكليات الحكومية، واستخدمت الدراسة ثلاثة أدوات رئيسة هي: اختبار واطسون-جلاسر للتفكير الناقد (النسخة المختصرة)، واختبار القراءة النقدية الذي صممه الباحث، بالإضافة إلى استبيان الاتجاه نحو التفكير الناقد ، وتم تحليل البيانات باستخدام برنامج (SPSS)، حيث أُجري تحليل الارتباط والانحدار خطي بين نتائج التفكير الناقد والقراءة النقدية و توصلت النتائج إلى أن (٩٨.٥٪) من الطلبة لديهم اتجاه إيجابي نحو التفكير الناقد، في حين أظهرت نتائج اختبار التفكير الناقد أن (١٤.٣٪) من الطلبة حصلوا على مستوى ممتاز، و(٤١.٦٪) على مستوى جيد جدًا، و(٤٢.١٪) على مستوى متوسط، و(٢٪) فقط حصلوا على مستوى ضعيف، مما يشير إلى أن أغلب الطلبة يتمتعون بمستوى تفكير ناقد فوق المتوسط ، لكن لم يتمكن سوى (٣٦.٤٪) من الطلبة من عكس هذه القدرة في اختبار القراءة النقدية، مما يشير إلى وجود فجوة بين الاتجاه الإيجابي ومستوى التطبيق الفعلي للتفكير الناقد، وأظهرت نتائج التحليل وجود علاقة ارتباط ضعيفة ولكن دالة إحصائيًا ($r = 0.152, p < 0.05$) بين التفكير الناقد والقراءة النقدية، وهو ما يبرز الحاجة إلى دمج مهارات التفكير الناقد بشكل عملي في تعليم اللغة.

دراسات تناولت نموذج راش:

دراسة افينك ودوقن (Avinç & Doğan, 2024) هدفت الدراسة إلى تطوير أداة قياس محو الأمية الرقمية (Digital Literacy Scale) (DLS) لتحديد مستويات الإلمام الرقمي لدى طلاب المرحلة الثانوية، مع التحقق من صدق وثبات الأداة باستخدام نموذج راش، تكونت الأداة بصورتها النهائية من (٢٠) فقرة، تمثلت العينة من (٢٥١) طالبًا تم اختيارهم بطريقة عمدية، أظهرت النتائج تحقق افتراضات نموذج راش، حيث تم التأكد من أحادية البعد والاستقلال الموضعي والاتساق الداخلي للأداة باستخدام التحليل العاملي الاستكشافي (EFA) والتوكيدي (CFA)، كما أظهرت النتائج تطابق الأفراد مع نموذج راش، وتطابق الفقرات مع النموذج وكانت ضمن الحد المقبول، وتراوحت قيم صعوبة الفقرات بين (-٠.٤٩) و(+٠.٦٠)

ISSN: 2537-0448 eISSN: 2537-0456

لوجيت، مما يشير إلى أن تفاوت الصعوبة بحد أدنى، بلغ معامل الثبات للأفراد (٠.٩٤) وللفقرات (٠.٨٢) ومعامل الفصل للأفراد (٣.٩٢) وللفقرات (٢.١٣) مما يدل أن الأداة تتمتع بثبات جيد وذات فاعلية في قياس الإلمام الرقمي لدى طلاب المرحلة الثانوية.

دراسة عبد الجليل وآخرون (٢٠٢٤) هدفت إلى التحقق من افتراض أحادية البعد في مقياس الشغف الأكاديمي باستخدام نموذج راش والتحليل العاملي، شملت العينة ٣٠٠ طالب وطالبة من السنة الثالثة في كلية التربية بجامعة قنا، استخدمت الدراسة مقياس الشغف الأكاديمي لجمع البيانات وفق المنهج الوصفي، وتم تحليل البيانات باستخدام (Winsteps) (V.5.6.0.0). أظهرت النتائج تحقق افتراض أحادية البعد للمقياس، مما يعني أنه يقيس بُعدًا واحدًا فقط وهو الشغف الأكاديمي، حيث دعمت مؤشرات التحليل العاملي الاستكشافي (EFA) والتحليل العاملي التوكيدي (CFA) ونموذج راش هذا الافتراض، كما تبين أن البنود مطابقة للنموذج وكانت متسقة مع البيانات، بينما استُبعد خمسة أفراد غير متطابقين وأعيد تطبيق التحليل بعد حذفهم، أكدت النتائج أن فاعلية المقياس في قياس الشغف الأكاديمي.

دراسة عمران وسعد (٢٠٢٣) تهدف الدراسة الى التحقق من الخصائص السيكومترية لمقياس قلق المستقبل تمثلت عينتها (٢٠٠) طالب وطالبة من جامعة غرداية، واستخدمت مقياس قلق المستقبل لجمع البيانات وفق المنهج الوصفي، أظهرت نتائج الدراسة تحقق افتراضات نموذج راش، حيث تم التحقق من أحادية البعد لمقياس قلق المستقبل باستخدام مؤشرات التحليل العاملي للمكونات الأساسية كما تبين أن معظم الفقرات كانت مطابقة للنموذج، بينما تم استبعاد خمسة أفراد من العينة لعدم تطابقهم مع النموذج، وتراوحت صعوبة الفقرات بين (-٠.٦١) و(١.٠٥) لوجيت، مما يدل على تدرج مستوى الصعوبة وقدرة المقياس على التمييز بين الأفراد ذوي المستويات المختلفة من قلق المستقبل، كذلك بلغ معامل الثبات للأفراد (٠.٨٤)، ومعامل الثبات للفقرات (٠.٩٧)، مما يشير إلى ثبات عالٍ يمكن الاعتماد عليه وبلغ معامل الفصل للأفراد (٢.٢٨) ومعامل الفصل للفقرات (٦.٠٩)، مما يدل على قدرة المقياس على التمييز بين الأفراد بشكل فعال.

التعقيب على الدراسات السابقة:

أوجه الاتفاق:

تتفق الدراسة الحالية مع الدراسات السابقة في المنهج المستخدم وعينة الدراسة اذ تم الاعتماد على المنهج الوصفي وطلبة الجامعة باستثناء دراسة اريف(Arif, 2024) استخدمت الدراسة المنهج النوعي وتمثلت عينة هذه الدراسة في مفردات الاختبار, ودراسة سويتا(Suwita et al., 2024) تمثلت عينتها في طلاب مرحلة

ISSN: 2537-0448 eISSN: 2537-0456

الاعدادي, وتتفق أداة الدراسة الحالية مع كل من دراسة (صباح, ٢٠٢١؛ Arif, 2024؛ Aslan et al., 2019؛ Din, 2020) إذ تم استخدام الصورة القصيرة من اختبار واطسون- جليسر كذلك تتفق هذه الدراسة من حيث تبني نموذج راش ومدى مطابقة للأداة مع كل من دراسة (صباح, ٢٠٢١؛ صالح, ٢٠٢٣؛ الجبوري, ٢٠١٧؛ Swita et al., 2024؛ vinç & Doğan, 2024؛ عبدالجليل وآخرون, ٢٠٢٤؛ عمران وسعد, ٢٠٢٣).

أوجه الاختلاف:

تختلف الدراسة الحالية مع الدراسات السابقة من حيث الأداة المستخدمة مع دراسة (صالح, ٢٠٢٣؛ الجبوري, ٢٠١٧) تم فيها استخدام نسخ أخرى من اختبار واطسون -جليسر كذلك دراسة (Suwita et al., 2024) تم انشاء أداة شبيه لاختبار واطسون وجليسر وفي كل من (Avinç & Doğan, 2024؛ عبد الجليل وآخرون, ٢٠٢٤؛ عمران وسعد, ٢٠٢٣) استخدمت هذه الدراسات مقاييس واختبارات أخرى.

أوجه التفرد

تتفرد الدراسة الحالية باستخدام النسخة القصيرة المقننة على البيئة السعودية من اختبار واطسون -جليسر وتطويرها من خلال تقدير الخصائص السيكومترية وفق نموذج راش؛ وهو أحد نماذج نظرية الاستجابة للفقرة، حيث لا توجد دراسة (في حدود علم الباحثة) تناولت الاختبار من هذا المنظور في المملكة العربية السعودية ، كذلك تتفرد هذه الدراسة عن الدراسات السابقة باستخدام برنامج مختلف للتحليل البيانات إذ اعتمدت الدراسات السابقة على برنامج (Spss)، وبرنامج(Winsteps) ، وفي الدراسة الحالية تم استخدام برنامج (RStudio) للبرمجة الإحصائية.

منهجية الدراسة واجراءاتها:

لتحقيق اهداف الدراسة، تم استخدام المنهج الوصفي نظرًا لتوافقه مع طبيعة متغيرات الدراسة من خلال وصف المتغيرات وجمع البيانات والتحقق من مدى تطابق الاختبار لمتطلبات نموذج راش.

مجتمع الدراسة:

تكون مجتمع الدراسة من طلبة جامعة تبوك لعام ١٤٤٦هـ – ٢٠٢٥م والبالغ عددهم (٢٥.١٥٧) طالب وطالبة التي حصلت عليها الباحثة من البيانات المفتوحة لجامعة تبوك.

عينة الدراسة:

تكونت عينة الدراسة من (٢٠١) طالب وطالبة تم اخذهم بالطريقة العشوائية الطبقية، وفقًا للينكر Linacre (1994)، يرى حجم العينة ١٥٠ كافياً لنموذج راش، وفي دراسة (French &Finc, 2019) لمقارنة تقنيات التقدير في نماذجIRT مع

ISSN: 2537-0448 eISSN: 2537-0456

عينات صغيرة، وجد أن عينات بحجم 100 يمكن أن توفر تقديرات دقيقة وفعالة لمعايير النموذج.

أداة الدراسة:

استخدمت الباحثة النسخة القصيرة من اختبار واطسون – جليسر للتفكير الناقد تقنين العتيبي (٢٠١٢) الذي قام بعدد من الإجراءات منها:

- ترجمة الاختبار مع مراعاة المحافظة على بناء الاختبار الأصلي.
- إعادة صياغة العبارات لتتناسب مع الثقافة السعودية على النحو التالي:
  - استبدال أسماء المدن والأشخاص لأسماء تناسب الثقافة السعودية.
  - تغيير صياغة عدد من العبارات المتحيزة ثقافياً لعدم ملاءمتها مع ما يتسم به المجتمع السعودي.
- تجريب الاختبار وذلك من خلال عرضة على عينة استطلاعية قوامها (٢٦) طالباً للوقوف على مدى وضوح العبارات.
- بعد تعديل صياغة بعض الفقرات تم عرض الاختبار على عدد من أعضاء هيئة التدريس متخصصين في دراسة التفكير لمدى مناسبتها الثقافية المحلية وتم تعديل الصياغة اللغوية لتعليمات الاختبار.

ويقاس التفكير الناقد في هذا الاختبار من خلال خمسة أبعاد تتمثل في:

١. الاستنتاج (Inference): يركز على قياس قدرة الفرد على استخلاص نتائج من حقائق مُفترضة والحكم على العلاقة بين المعطيات والاستنتاجات.

٢. تمييز الافتراضات (Recognition Assumptions): يقيس القدرة على تحديد الافتراضات الضمنية وغير المُعلنة التي تقوم عليها الحجة، والتمييز بين الافتراضات الصحيحة والخاطئة أو التي لا يمكن إثباتها كذلك التمييز بين الحقيقة والرأي.

٣. الاستنباط (Deduction): يركز على قياس قدرة الفرد على تحديد ما إذا كانت الاستنتاجات تتبع منطقياً المعطيات ومدى صحة أو خطأ الحجج بناءً على منطقها الداخلي.

٤. التفسير (Interpretation): يقيس قدرة الفرد على تحليل البيانات والمعلومات لاتخاذ قرارات مستندة إلى معقولية التعميمات أو الاستنتاجات.

٥. تقييم الحجج (Evaluate Arguments): يركز على قدرة الشخص على تقييم قوة الحجة من حيث المنطق والحيادية والارتباط بالموضوع المطروح والتمييز بين الحجج ذات الطابع العاطفي والحجج المنطقية وقبولها ورفضها (العتيبي، ٢٠١٥).

ISSN: 2537-0448 eISSN: 2537-0456

إجراءات التطبيق:

تم تصميم الاداة الكترونيًا باستخدام نماذج قوقل لتسهيل نشرها على الطلبة، تكونت الأداة من خمس أبعاد وهي: الاستنتاج وتمييز الافتراضات والاستنباط والتفسير وتقويم الحجج وتم ارسالها للطلبة عبر الرابط https://forms.gle/XRbQxoSKu1ADnsKA8 من ثم تم استخراج الاستجابات وتحليلها عبر البرامج الإحصائية.

تصحيح الاختبار:

يتكون الاختبار بصورته النهائية من (٤٠) فقرة لخمسة أبعاد كما في الجدول (١):

جدول (١) طريقة تصحيح الاختبار

البعد الأول(الاستنتاج)	صادق تماماً	محتمل صدقة	بيانات ناقصة	محتمل خطؤه	خاطئ تماماً
البعد الثاني (تمييز الافتراضات)	وارد			غير وارد	
البعد الثالث (الاستنباط)	صحيح			غير صحيح	
البعد الرابع (التفسير)	مترتبة			غير مترتبة	
البعد الخامس (تقويم الحجج)	قوية			ضعيفة	

يتم التصحيح وفقًا لمعايير الاختبار وفيها يتم حصول درجة (١) للإجابة الصحيحة، وصفر (٠) للإجابة الخاطئة، وبذلك يكون المدى ما بين صفر (٠) إلى أربعين (٤٠) درجة.

نتائج الدراسة:

النتائج المتعلقة بالسؤال الأول: هل تحقق البيانات المستمدة من الصورة القصيرة لاختبار واطسون – جليسر للتفكير الناقد افتراضات نظرية الاستجابة للفقرة؟

للإجابة عن السؤال يتم التحقق من افتراضات نظرية الاستجابة كما يلي:

أولًا: التحقق من افترض أحادية البعد:

للتحقق من افتراض أحادية البعد تم استخدام التحليل العاملي الاستكشافي عن طريق برنامج(RStudio) من حزمة (psych) باستخدام طريقة المكونات الرئيسية (Principal Component Analysis) والتدوير المتعامد (Varimax)، يظهر في الجدول (٢) قيم كل من الجذر الكامن (Eigen Value) للعامل الأول والثاني ونسبة التباين المفسر (Explained Variance).

ISSN: 2537-0448 eISSN: 2537-0456

جدول (٢) نتائج التحليل العاملي الاستكشافي

	الجذر الكامن	نسبة التباين المفسر	الجذر الكامن للعامل الأول / الجذر الكامن للعامل الثاني
العامل الأول	٦.٤٤	١٦.١١	٢.١٥
العامل الثاني	٣.٠٠	٧.٦٥	

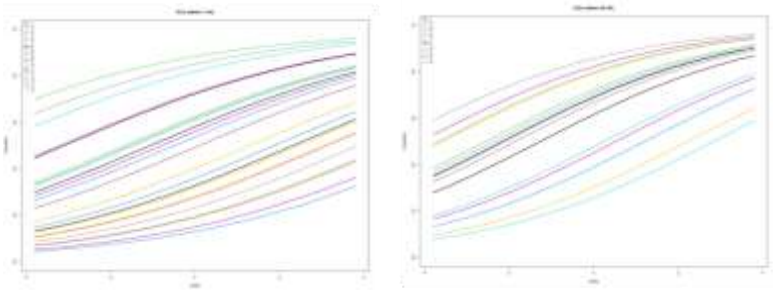
يوضح الجدول (٢) أن قيمة الجذر الكامن للعامل الأول (٦.٤٤) وفسر ما نسبته (١٦.١١%) من التباين الكلي للاختبار، وقيمة الجذر الكامن للعامل الثاني (٣) وفسر ما نسبته (٧.٦٥) من التباين الكلي للاختبار، وبلغت نتيجة قسمة الجذر الكامن للعامل الأول للعامل الثاني (٢.١٥) وهي أكبر من (٢) يعني ذلك وجود سمة سائدة للاختبار، مما يدل على تحقق افتراض أحادية البعد .

ثانيًا: الاستقلال الموضعي:

يشير هذا الافتراض إلى استقلال إجابات المفحوص على نفس الاختبار أي أن إجابته عن الفقرة الأولى لا تتأثر بإجابته عن أي فقرة أخرى، ويشير كل من (Hambleton & Swaminathan, 1985) أن الاستقلال الموضعي يكافئ أحادية البعد وذلك يعني عندما يتحقق افتراض أحادية البعد يتحقق افتراض الاستقلال الموضعي.

الافتراض الثالث المنحنى المميز للفقرة:

يوضح الشكل (١) رسم لمنحى المميز للفقرات (ICCs)، ويظهر أن شكل المنحنيات جميعها تتجه بشكل تصاعدي من اليسار إلى اليمين، مما يدل على تحقق افتراض منحنى الـ (ICC) في نموذج راش، هذا الشكل يوضح أن هناك علاقة إيجابية بين القدرة الكامنة لدى الفرد واحتمالية إجابته الصحيحة على الفقرة، وهو ما يتوافق مع خصائص النموذج، خاصة أن انحدار المنحنيات متقارب، مما يشير إلى ثبات معامل التمييز بين الفقرات، كما يفترضه نموذج راش.



الشكل (١) المنحنى المميز لفقرات اختبار واطسون – جليسر للتفكير الناقد

ISSN: 2537-0448 eISSN: 2537-0456

الخامس: التحرر من السرعة

تم إعطاء الوقت الكافي للاستجابة على الاختبار حيث لم يتم تقييد الاختبار بوقت معين بالتـالي لا يؤثر عامل الوقت على استجابات الطلبة في الاختبار، أي انه تم تحقق افتراضات نظرية الاستجابة للمفردة وتتفق هذه النتيجة مع دراسة (صباح، ٢٠٢١؛ صالح، ٢٠٢٣؛ الجبوري، ٢٠١٧؛ عبد الجليل وآخرون، ٢٠٢٤؛ عمران وسعد، ٢٠٢٣).

النتائج المتعلقة بالسؤال الثاني: ما درجة مطابقة نموذج راش للبيانات المستمدة من الصورة القصيرة لاختبار واطسون – جليسر للتفكير الناقد؟

للتحقق من مطابقة الافراد والفقرات لنموذج راش، تم استخدام برنامج (R) لتحليل البيانات، وفيه استخدمت الباحثة حزمة (mirt) لتحليل البيانات ومطابقتها مع نموذج راش.

أولًا: مطابقة الأفراد لنموذج راش:

للتحقق من مدى ملاءمة الأفراد لنموذج راش، تم حساب قيم المطابقة الداخلية والمطابقة الخارجية ببرمجية (R) باستخدام مكتبة (mirt) لكل فرد من أفراد العينات، حيث تم الاعتماد على قيم المطابقة الداخلية والمطابقة الخارجية لهم خارج الحدود المقبولة وهي (٠.٥ – ١.٥)، فالقيمة الواقعة خارج هذا المدى تدل على عدم مطابقة الفرد للنموذج، وأفرزت نتائج التحليل عدم مطابقة مجموعة من الأفراد لنموذج راش، كما يوضحها الجدول رقم (٣).

جدول (٣) مطابقة الافراد لنموذج راش

	Outfit	Z.Outfit	Infit	Z.Infit		Outfit	Z.Outfit	Infit	Z.Infit
1	1.096	0.555	1.054	0.412	**11**	1.306	1.635	1.197	1.383
2	0.767	-1.358	0.787	-1.580	**12**	1.231	1.265	1.258	1.745
3	0.931	-0.319	0.930	-0.448	**13**	0.961	-0.164	1.045	0.365
4	1.136	0.777	1.168	1.167	**14**	0.765	-1.385	0.815	-1.372
5	0.835	-0.893	0.887	-0.776	**15**	0.973	-0.092	1.033	0.274
6	0.876	-0.657	0.919	-0.543	**16**	0.768	-1.331	0.833	-1.194
7	0.727	-1.640	0.796	-1.523	**17**	0.610	-2.492	0.677	-2.555
8	1.286	1.536	1.310	2.093	**18**	0.999	0.058	1.027	0.236
9	1.806	3.744	1.596	3.669	**19**	0.573	-2.809	0.632	-2.997
10	0.838	-0.886	0.878	-0.844	**20**	0.993	0.026	0.969	-0.169

ISSN: 2537-0448 eISSN: 2537-0456

	Outfit	Z.Outfit	Infit	Z.Infit		Outfit	Z.Outfit	Infit	Z.Infit
21	1.264	1.411	1.201	1.382	**31**	1.020	0.174	0.942	-0.383
22	0.964	-0.140	0.954	-0.275	**32**	0.987	-0.014	0.971	-0.161
23	0.695	-1.871	0.777	-1.685	**33**	0.610	-2.492	0.677	-2.555
24	0.961	-0.162	1.009	0.108	**34**	1.249	1.356	1.200	1.390
25	1.026	0.203	1.044	0.354	**35**	1.166	0.943	1.145	1.035
26	1.080	0.498	1.090	0.676	**36**	0.616	-2.467	0.674	-2.606
27	1.082	0.500	1.103	0.748	**37**	0.841	-0.874	0.883	-0.811
28	1.177	0.990	1.125	0.897	**38**	0.734	-1.592	0.791	-1.559
29	0.852	-0.783	0.921	-0.519	**39**	1.007	0.097	1.000	0.043
30	1.049	0.324	1.110	0.792	**40**	0.916	-0.430	0.930	-0.468
41	0.774	-1.303	0.789	-1.556	**51**	0.903	-0.506	0.957	-0.264
42	0.793	-1.183	0.745	-1.936	**52**	0.610	-2.492	0.677	-2.555
43	0.956	-0.192	0.997	0.026	**53**	1.095	0.575	1.137	0.975
44	1.127	0.746	1.132	0.956	**54**	0.610	-2.492	0.677	-2.555
45	1.466	2.267	1.431	2.693	**55**	0.857	-0.751	0.910	-0.594
46	1.136	0.784	1.186	1.289	**56**	0.610	-2.492	0.677	-2.555
47	0.858	-0.783	0.807	-1.439	**57**	0.883	-0.622	0.925	-0.499
48	0.639	-2.275	0.709	-2.260	**58**	0.988	-0.010	0.976	-0.126
49	0.692	-1.884	0.745	-1.945	**59**	1.071	0.452	1.141	1.010
50	0.610	-2.492	0.677	-2.555	**60**	0.680	-1.959	0.746	-1.928
61	1.122	0.696	1.023	0.207	**71**	0.981	-0.044	0.960	-0.239
62	0.790	-1.153	0.837	-1.147	**72**	1.751	3.474	1.561	3.433
63	0.815	-1.018	0.848	-1.073	**73**	1.108	0.642	1.092	0.682
64	1.213	1.182	1.226	1.551	**74**	1.337	1.778	1.332	2.203
65	0.888	-0.587	0.951	-0.304	**75**	0.884	-0.609	0.891	-0.752
66	0.844	-0.844	0.892	-0.743	**76**	0.872	-0.691	0.921	-0.529
67	1.182	1.030	1.122	0.892	**77**	1.068	0.420	1.068	0.505

ISSN: 2537-0448 eISSN: 2537-0456

	Outfit	Z.Outfit	Infit	Z.Infit		Outfit	Z.Outfit	Infit	Z.Infit
68	0.952	-0.213	0.977	-0.117	**78**	1.080	0.489	1.124	0.877
69	0.969	-0.115	1.016	0.159	**79**	0.707	-1.728	0.769	-1.708
70	1.036	0.251	1.110	0.780	**80**	1.058	0.375	1.080	0.597
81	1.108	0.633	1.131	0.928	**91**	0.755	-1.354	0.820	-1.272
82	1.145	0.838	1.125	0.904	**92**	1.255	1.394	1.158	1.127
83	0.573	-2.809	0.632	-2.997	**93**	1.195	1.086	1.102	0.747
84	0.785	-1.253	0.841	-1.163	**94**	1.622	3.018	1.438	2.811
85	1.204	1.104	1.161	1.116	**95**	1.498	2.485	1.293	1.957
86	0.922	-0.370	0.989	-0.029	**96**	0.610	-2.492	0.677	-2.555
87	1.062	0.400	1.070	0.533	**97**	0.610	-2.492	0.677	-2.555
88	0.881	-0.631	0.955	-0.280	**98**	1.513	2.563	1.414	2.675
89	1.159	0.891	1.126	0.896	**99**	1.261	1.420	1.140	1.011
90	0.792	-1.190	0.859	-1.003	**100**	0.610	-2.492	0.677	-2.555
101	1.319	1.698	1.203	1.421	**111**	1.067	0.427	1.061	0.476
102	1.302	1.618	1.191	1.343	**112**	0.796	-1.139	0.863	-0.956
103	0.822	-0.996	0.871	-0.911	**113**	0.703	-1.754	0.752	-1.854
104	0.716	-1.723	0.750	-1.914	**114**	1.596	2.907	1.344	2.294
105	1.717	3.402	1.613	3.769	**115**	0.573	-2.809	0.632	-2.997
106	1.310	1.656	1.194	1.362	**116**	0.905	-0.485	0.961	-0.231
107	1.602	2.938	1.382	2.494	**117**	0.573	-2.809	0.632	-2.997
108	1.121	0.716	1.138	0.999	**118**	1.140	0.817	1.096	0.714
109	1.346	1.824	1.307	2.064	**119**	0.610	-2.492	0.677	-2.555
110	0.828	-0.973	0.878	-0.859	**120**	1.463	2.348	1.311	2.074
121	1.090	0.551	1.000	0.045	**131**	0.847	-0.827	0.891	-0.779
122	1.110	0.654	1.069	0.529	**132**	1.091	0.559	1.060	0.469
123	1.102	0.616	1.123	0.891	**133**	1.031	0.232	1.011	0.124
124	0.940	-0.284	0.927	-0.495	**134**	1.376	1.844	1.363	2.288

ISSN: 2537-0448 *eISSN: 2537-0456*

	Outfit	Z.Outfit	Infit	Z.Infit		Outfit	Z.Outfit	Infit	Z.Infit
125	1.132	0.771	1.065	0.504	**135**	1.376	1.844	1.363	2.288
126	0.806	-1.103	0.850	-1.075	**136**	1.166	0.940	1.144	1.024
127	0.978	-0.065	0.933	-0.441	**137**	0.971	-0.105	0.970	-0.168
128	1.270	1.458	1.266	1.798	**138**	1.272	1.464	1.243	1.652
129	0.973	-0.092	1.007	0.099	**139**	1.077	0.477	1.047	0.374
130	1.074	0.464	1.031	0.264	**140**	0.893	-0.567	0.908	-0.642
141	0.642	-2.249	0.703	-2.316	**151**	1.197	1.100	1.042	0.341
142	1.134	0.784	1.151	1.087	**152**	1.069	0.443	1.055	0.436
143	0.801	-1.096	0.868	-0.914	**153**	1.073	0.459	1.054	0.427
144	0.573	-2.809	0.632	-2.997	**154**	0.887	-0.604	0.934	-0.438
145	1.059	0.382	1.047	0.375	**155**	0.589	-2.674	0.650	-2.817
146	0.857	-0.766	0.887	-0.779	**156**	0.959	-0.174	0.970	-0.169
147	0.756	-1.395	0.821	-1.285	**157**	0.982	-0.035	0.986	-0.049
148	0.891	-0.582	0.910	-0.627	**158**	0.955	-0.194	1.004	0.075
149	1.122	0.711	1.112	0.812	**159**	0.802	-1.133	0.813	-1.379
150	0.798	-1.134	0.869	-0.914	**160**	1.297	1.593	1.216	1.494
161	1.530	2.533	1.400	2.523	**171**	0.824	-0.900	0.890	-0.725
162	0.956	-0.192	0.991	-0.018	**172**	0.573	-2.809	0.632	-2.997
163	1.104	0.622	1.070	0.533	**173**	1.036	0.250	1.024	0.209
164	0.759	-1.420	0.799	-1.490	**174**	0.918	-0.413	0.929	-0.475
165	1.154	0.886	1.105	0.777	**175**	0.872	-0.687	0.949	-0.321
166	1.095	0.561	1.043	0.341	**176**	1.015	0.145	1.007	0.100
167	1.039	0.278	1.042	0.343	**177**	1.383	1.991	1.342	2.264
168	0.715	-1.719	0.756	-1.849	**178**	0.955	-0.191	0.935	-0.423
169	1.152	0.871	1.136	0.972	**179**	0.610	-2.492	0.677	-2.555
170	0.873	-0.690	0.943	-0.370	**180**	0.884	-0.604	0.942	-0.365
181	1.024	0.191	1.026	0.231	**191**	0.912	-0.456	0.905	-0.662

ISSN: 2537-0448 eISSN: 2537-0456

	Outfit	Z.Outfit	Infit	Z.Infit		Outfit	Z.Outfit	Infit	Z.Infit
182	0.573	-2.809	0.632	-2.997	**192**	0.610	-2.492	0.677	-2.555
183	1.118	0.685	1.098	0.715	**193**	0.692	-1.884	0.763	-1.793
184	0.972	-0.090	1.062	0.469	**194**	1.119	0.705	1.058	0.459
185	0.720	-1.585	0.763	-1.729	**195**	0.843	-0.851	0.852	-1.043
186	0.761	-1.337	0.828	-1.217	**196**	0.944	-0.254	0.975	-0.129
187	0.834	-0.929	0.852	-1.069	**197**	1.287	1.540	1.249	1.714
188	0.693	-1.863	0.754	-1.863	**198**	0.582	-2.734	0.643	-2.899
189	0.610	-2.492	0.677	-2.555	**199**	0.920	-0.391	0.930	-0.454
190	2.352	5.642	2.061	5.950	**200**	1.020	0.170	1.086	0.645
201	1.016	0.149	1.049	0.388					

يوضح الجدول (٣) وجود أفراد غير مطابقين لنموذج راش، وهم: (٩- ٧٢- ٩٤- ٩٨ -١٠٥-١٠٧-١١٤-١١٥-١٦١-١٨٢-١٩٠-١٩٨)، أي وقعت قيم المطابقة الداخلية والمطابقة الخارجية لهم خارج الحدود المقبولة وهي (٠.٥ – ١.٥) وبلغ عددهم ١٢ فردًا، لذا تم استبعادهم واستكمال التحليل على بقية الافراد والبالغ عددهم ١٨٩، وتتفق هذه النتيجة مع دراسة كل من (صالح, ٢٠٢٣؛ عبد الجليل واخرون , ٢٠٢٤؛ عمران وسعد, ٢٠٢٣).

ثانيًا: مطابقة الفقرات لنموذج راش:

للتحقق من ملاءمة الفقرات لنموذج راش تمت عملية معايرة الفقرات مرة أخرى بعد حذف الأفراد غير الملائمين، وتم حساب قيم المطابقة الداخلية والمطابقة الخارجية ببرمجية (R) باستخدام مكتبة (MIRT) لكل فقرة، حيث تم الاعتماد على قيم المطابقة الداخلية والمطابقة الخارجية لهم خارج الحدود المقبولة وهي (٠.٥ – ١.٥)، فالقيمة الواقعة خارج هذا المدى تدل على عدم مطابقة الفقرة للنموذج، كما يوضحها الجدول رقم (٤).

جدول (٤) مطابقة الفقرات لنموذج راش

item	outfit	z.outfit	infit	z.infit	item	outfit	z.outfit	infit	z.infit
1	0.996	-0.056	0.997	-0.040	**21**	0.977	-0.571	0.979	-0.513
2	0.959	-0.294	0.986	-0.077	**22**	0.965	-1.653	0.967	-1.575
3	0.969	-0.214	0.990	-0.038	**23**	0.994	-0.053	0.997	-0.014
4	0.967	-0.846	0.972	-0.738	**24**	0.946	-0.275	0.992	0.003

ISSN: 2537-0448 eISSN: 2537-0456

5	0.982	-1.870	0.982	-1.879	**25**	1.015	0.283	1.012	0.227
6	0.961	-0.214	0.993	0.003	**26**	1.006	0.131	1.005	0.103
7	1.003	0.069	1.003	0.069	**27**	1.016	0.239	1.011	0.171
8	0.989	-0.079	0.995	-0.012	**28**	0.990	-0.141	0.993	-0.099
9	0.974	-0.298	0.985	-0.154	**29**	1.028	0.342	1.018	0.228
10	0.978	-0.269	0.986	-0.161	**30**	0.954	-2.760	0.955	-2.739
11	1.015	0.402	1.013	0.353	**31**	0.941	-0.492	0.975	-0.187
12	1.051	0.314	1.019	0.160	**32**	0.945	-0.268	0.991	0.002
13	1.019	0.186	1.011	0.127	**33**	0.969	-0.901	0.972	-0.820
14	1.001	0.044	1.001	0.045	**34**	0.956	-5.023	0.956	-5.035
15	0.976	-0.887	0.977	-0.854	**35**	0.961	-0.263	0.987	-0.054
16	1.013	0.288	1.011	0.257	**36**	0.977	-0.666	0.980	-0.604
17	0.974	-0.833	0.978	-0.739	**37**	0.961	-3.384	0.961	-3.384
18	0.976	-2.450	0.977	-2.446	**38**	0.970	-0.205	0.991	-0.037
19	0.949	-0.163	1.000	0.073	**39**	0.957	-0.616	0.971	-0.416
20	1.022	1.124	1.021	1.087	**40**	1.009	0.154	1.007	0.131

يبين الجدول (٤) أن جميع الفقرات وعددها (٤٠) فقرة جاءت ضمن الحد المقبول (٠.٥-١.٥) للمطابقة الداخلية والمطابقة الخارجية لنموذج راش، وتتفق هذه النتيجة مع دراسة (صباح, ٢٠٢١؛ عبد الجليل وآخرون, ٢٠٢٤؛ عمران وسعد, ٢٠٢٣؛ Swita et al., 2024؛ vinç & Doğan, 2024) حيث أسفرت نتائج هذه الدراسات عن تطابق جميع الفقرات مع نموذج راش بينما تختلف مع دراسة (صالح, ٢٠٢٣) إذ تم فيها استبعاد (١٤) فقرة.

النتائج المتعلقة بالسؤال الثالث: ما قيم معالم صعوبة الفقرة ودقة تقديرها للبيانات المستمدة من الصورة القصيرة لاختبار واطسون – جليسر للتفكير الناقد وفق نموذج راش؟

للإجابة عن السؤال تم استخدام برنامج (RStudio) من خلال حزمة (Mirt) وتم تحليل البيانات وفق نموذج راش (Rasch Model) وتم حساب قيم معلمة الصعوبة للفقرات، ودقة تقدير معلمة الصعوبة من خلال الخطأ المعياري لتقديرها، وكانت النتائج كما في الجدول (٥).

ISSN: 2537-0448 eISSN: 2537-0456

الجدول (٥) قيم معلمة الصعوبة للفقرات والخطأ المعياري

Item	B	Std.Error	Item	B	Std.Error
A1	0.717	0.155	**C6**	-0.531	0.151
A2	1.517	0.189	**C7**	-0.266	0.147
A3	1.517	0.189	**C8**	1.022	0.165
A4	0.508	0.15	**C9**	-1.928	0.218
A5	-0.074	0.146	**D1**	-0.741	0.156
A6	1.792	0.208	**D2**	-0.741	0.156
A7	0.741	0.156	**D3**	-0.969	0.163
B1	1.284	0.176	**D4**	-0.815	0.158
B2	-1.05	0.166	**D5**	1.134	0.17
B3	0.995	0.164	**D6**	0.202	0.146
B4	-0.508	0.15	**D7**	-1.413	0.183
B5	2.251	0.248	**E1**	-1.976	0.222
B6	-1.629	0.196	**E2**	-0.441	0.149
B7	-1.106	0.168	**E3**	0.053	0.146
B8	0.353	0.148	**E4**	-1.591	0.194
C1	-0.623	0.153	**E5**	0.441	0.149
C2	-0.397	0.148	**E6**	0.117	0.146
C3	-0.074	0.146	**E7**	-1.517	0.189
C4	-2.381	0.261	**E8**	0.891	0.16
C5	-0.245	0.147	**E9**	-0.891	0.16
معلمة الصعوبة			**الخطأ المعياري لتقدير القدرة**		
المتوسط الحسابي		-0.159	**المتوسط الحسابي**		0.170
أقل قيمة		-2.381	**أقل قيمة**		0.146
أكبر قيمة		2.251	**أكبر قيمة**		0.261

يوضح الجدول (٥) قيم معلمة الصعوبة لفقرات الاختبار، والخطأ المعياري لتقديرها، إذ تراوحت قيم معلمة الصعوبة للفقرات من(٢.٢٥١) إلـى (-٢.٣٨١) بمتوسط حسابي بلغ قدره (-٠.١٥٩) وبانحراف معياري (١.١١)، وتشير الفقرة رقم (C4) في بُعد الاستنباط التي نصت على (بعض الشعراء يقضون ساعات طويلة في ترديد القصائد فيشعرون بالمتعة في ذلك)، وهي أسهل فقرات الاختبار وفق نموذج راش، بينما جاءت الفقرة رقم (B5) في بُعد تمييز الافتراضات والتي نصت على (سأقلل فرص إصابتي بالمرض بحصولي على اللقاح ضد حمى التيفوئيد) كانت أصعب الفقرات في الاختبار وفق نموذج راش. وقد توزعت قيم معلمة الصعوبة لفقرات المقياس وفقًا لمعيار تشوي (Choi, 1992) كما هو موضح في جدول (٦).

جدول (٦) توزيع قيم معلمة الصعوبة للمقياس وفقًا لمعيار تشوي (Choi)

مستوى صعوبة الفقرة	سهلة	متوسطة الصعوبة	صعبة
مدى قيم معلمة الصعوبة	أصغر من – ٠.٥٠	من -٠.٥٠ إلى ٠.٥٠	أكبر من ٠.٥٠
عدد الفقرات	١٧	١١	١٢

ISSN: 2537-0448 eISSN: 2537-0456

ويلاحظ من الجدول (٦) أن الفقرات السهلة لاختبار واطسون - جليسر للتفكير الناقد (WGCT-SF) عددها (١٧) فقرة، حيث كانت قيم معلمة الصعوبة لها أقل من (-٠.٥٠)، كما بلغ عدد الفقرات متوسطة الصعوبة (١١) فقرة، حيث تراوحت قيم معلمة الصعوبة لها من (-٠.٥٠) إلى (٠.٥٠)، بينما بلغ عدد الفقرات الصعبة للمقياس (١٢) فقرة، حيث كانت قيم معلمة الصعوبة لها أكبر من (٠.٥٠). وتؤكد النتائج أن فقرات اختبار واطسون - جليسر للتفكير الناقد (WGCT-SF) شملت مدى واسع من الصعوبة، وتميل إلى التدرج في مستوى صعوبة الاختبار.

كما تم حساب دقة تقدير معلمة الصعوبة للفقرات من خلال حساب الأخطاء المعيارية إذ تراوحت الأخطاء المعيارية من (٠.١٤٦) إلى (٠.٢٦١) بمتوسط حسابي (٠.١٧٠) وبانحراف معياري (٠.٠٢٨)، وتشير قيمة الأخطاء المعيارية في تقدير معلمة الصعوبة للفقرات (A6, C3, D6, E3, E6) إلى أنها أكثر الفقرات دقة في تقدير معلمة الصعوبة، بينما الفقرة (C4) أقل الفقرات دقة في تقدير معلمة الصعوبة؛ وقد يكون ذلك بسبب أنها أكثر الفقرات سهولة.

ولابد للإشارة هنا بأن الفقرات مقرونة بعبارات رئيسة تسبقها ونفترض في هذا الاختبار صحة العبارة الرئيسية وبناء على هذه العبارة نحدد مدى صحة أو عدم صحة الفقرات، قد يؤثر صحة الفقرة منطقياً بشكلٍ عام على الافتراض المسبوق والمأخوذ بصحته في الاختبار وهذا ما يؤكد دقة قياس الاختبار للتفكير الناقد وقدرة الفرد في التمييز بين الافتراض بالعبارة الرئيسية ومدى صحة الفقرة بشكل عام.

وتتفق هذه النتيجة مع كل من دراسة (صباح،٢٠٢١) و(صالح، ٢٠٢٣) حيث بلغ متوسط صعوبة الفقرات في الاختبار (٠)، (٠) بانحراف معياري قدرة (٠.٧٤)، (٠.٨٥).

النتائج المتعلقة بالسؤال الرابع: ما مستوى قدرة الأفراد ودقة تقديرهم من البيانات المستمدة من الصورة القصيرة لاختبار واطسون – جليسر للتفكير الناقد وفق نموذج راش؟

للإجابة عن السؤال تم استخدام برنامج (RStudio) من خلال حزمة (Mirt) وتم تحليل البيانات وفق نموذج راش (Rasch Model) وتم حساب قيم معلمة القدرة للأفراد، ودقة تقدير القدرة من خلال حساب الخطأ المعياري، وكانت النتائج كما في الجدول (٧).

جدول (٧) مستوى القدرة للأفراد

	Person	Ability	Std_Error		Person	Ability	Std_Error
1	Person1	0.228	0.361	**21**	Person21	0.126	0.36
2	Person2	0.023	0.358	**22**	Person22	-0.08	0.357
3	Person3	0.16	0.36	**23**	Person23	-0.011	0.358

ISSN: 2537-0448 eISSN: 2537-0456

	Person	Ability	Std_Error		Person	Ability	Std_Error
4	Person4	0.091	0.359	**24**	Person24	-0.011	0.358
5	Person5	0.126	0.36	**25**	Person25	-0.114	0.357
6	Person6	0.057	0.359	**26**	Person26	0.091	0.359
7	Person7	-0.045	0.358	**27**	Person27	0.057	0.359
8	Person8	-0.216	0.357	**28**	Person28	0.16	0.36
9	Person9	0.091	0.359	**29**	Person29	0.091	0.359
10	Person10	-0.148	0.357	**30**	Person30	-0.148	0.357
11	Person11	-0.011	0.358	**31**	Person31	0.023	0.358
12	Person12	-0.011	0.358	**32**	Person32	-0.011	0.358
13	Person13	-0.08	0.357	**33**	Person33	-0.045	0.358
14	Person14	0.091	0.359	**34**	Person34	-0.011	0.358
15	Person15	0.091	0.359	**35**	Person35	-0.114	0.357
16	Person16	-0.011	0.358	**36**	Person36	0.057	0.359
17	Person17	0.126	0.36	**37**	Person37	-0.045	0.358
18	Person18	-0.08	0.357	**38**	Person38	-0.011	0.358
19	Person19	0.16	0.36	**39**	Person39	-0.08	0.357
20	Person20	0.057	0.359	**40**	Person40	0.057	0.359
41	Person41	0.023	0.358	**61**	Person61	0.194	0.361
42	Person42	-0.114	0.357	**62**	Person62	0.126	0.36
43	Person43	-0.114	0.357	**63**	Person63	-0.045	0.358
44	Person44	0.16	0.36	**64**	Person64	0.023	0.358
45	Person45	0.023	0.358	**65**	Person65	0.091	0.359
46	Person46	-0.08	0.357	**66**	Person66	-0.148	0.357
47	Person47	-0.011	0.358	**67**	Person67	0.023	0.358
48	Person48	-0.011	0.358	**68**	Person68	0.023	0.358
49	Person49	-0.011	0.358	**69**	Person69	0.194	0.361
50	Person50	-0.045	0.358	**70**	Person70	0.023	0.358

ISSN: 2537-0448 eISSN: 2537-0456

	Person	Ability	Std_Error		Person	Ability	Std_Error
51	Person51	-0.011	0.358	**71**	Person71	-0.011	0.358
52	Person52	0.023	0.358	**72**	Person72	-0.08	0.357
53	Person53	-0.011	0.358	**73**	Person73	0.057	0.359
54	Person54	0.16	0.36	**74**	Person74	-0.011	0.358
55	Person55	-0.011	0.358	**75**	Person75	0.194	0.361
56	Person56	-0.011	0.358	**76**	Person76	0.126	0.36
57	Person57	0.023	0.358	**77**	Person77	0.126	0.36
58	Person58	-0.08	0.357	**78**	Person78	0.057	0.359
59	Person59	0.023	0.358	**79**	Person79	0.091	0.359
60	Person60	0.16	0.36	**80**	Person80	-0.08	0.357
81	Person81	-0.08	0.357	**101**	Person101	-0.148	0.357
82	Person82	-0.148	0.357	**102**	Person102	-0.148	0.357
83	Person83	0.126	0.36	**103**	Person103	-0.045	0.358
84	Person84	0.16	0.36	**104**	Person104	-0.148	0.357
85	Person85	-0.011	0.358	**105**	Person105	0.126	0.36
86	Person86	0.023	0.358	**106**	Person106	0.126	0.36
87	Person87	0.091	0.359	**107**	Person107	-0.08	0.357
88	Person88	0.023	0.358	**108**	Person108	0.057	0.359
89	Person89	0.228	0.361	**109**	Person109	-0.08	0.357
90	Person90	-0.148	0.357	**110**	Person110	-0.148	0.357
91	Person91	-0.011	0.358	**111**	Person111	-0.011	0.358
92	Person92	-0.011	0.358	**112**	Person112	-0.148	0.357
93	Person93	-0.011	0.358	**113**	Person113	-0.011	0.358
94	Person94	-0.148	0.357	**114**	Person114	-0.08	0.357
95	Person95	-0.011	0.358	**115**	Person115	-0.25	0.357
96	Person96	-0.182	0.357	**116**	Person116	-0.08	0.357
97	Person97	-0.148	0.357	**117**	Person117	0.023	0.358

ISSN: 2537-0448 eISSN: 2537-0456

	Person	Ability	Std_Error		Person	Ability	Std_Error
98	Person98	0.023	0.358	**118**	Person118	-0.045	0.358
99	Person99	-0.045	0.358	**119**	Person119	-0.045	0.358
100	Person100	-0.148	0.357	**120**	Person120	0.057	0.359
121	Person121	-0.045	0.358	**141**	Person141	0.091	0.359
122	Person122	-0.353	0.357	**142**	Person142	-0.045	0.358
123	Person123	-0.216	0.357	**143**	Person143	-0.148	0.357
124	Person124	-0.011	0.358	**144**	Person144	-0.045	0.358
125	Person125	0.228	0.361	**145**	Person145	-0.114	0.357
126	Person126	0.228	0.361	**146**	Person146	-0.045	0.358
127	Person127	0.023	0.358	**147**	Person147	-0.011	0.358
128	Person128	-0.045	0.358	**148**	Person148	0.16	0.36
129	Person129	-0.011	0.358	**149**	Person149	-0.011	0.358
130	Person130	0.023	0.358	**150**	Person150	-0.011	0.358
131	Person131	-0.216	0.357	**151**	Person151	-0.08	0.357
132	Person132	-0.011	0.358	**152**	Person152	-0.011	0.358
133	Person133	-0.216	0.357	**153**	Person153	0.023	0.358
134	Person134	0.16	0.36	**154**	Person154	-0.011	0.358
135	Person135	-0.08	0.357	**155**	Person155	-0.114	0.357
136	Person136	-0.045	0.358	**156**	Person156	0.16	0.36
137	Person137	0.091	0.359	**157**	Person157	-0.182	0.357
138	Person138	0.126	0.36	**158**	Person158	-0.011	0.358
139	Person139	-0.182	0.357	**159**	Person159	-0.011	0.358
140	Person140	0.057	0.359	**160**	Person160	-0.08	0.357
161	Person161	0.297	0.363	**171**	Person171	0.091	0.359
162	Person162	-0.08	0.357	**172**	Person172	-0.08	0.357
163	Person163	0.263	0.362	**173**	Person173	0.091	0.359
164	Person164	-0.045	0.358	**174**	Person174	0.16	0.36

ISSN: 2537-0448 eISSN: 2537-0456

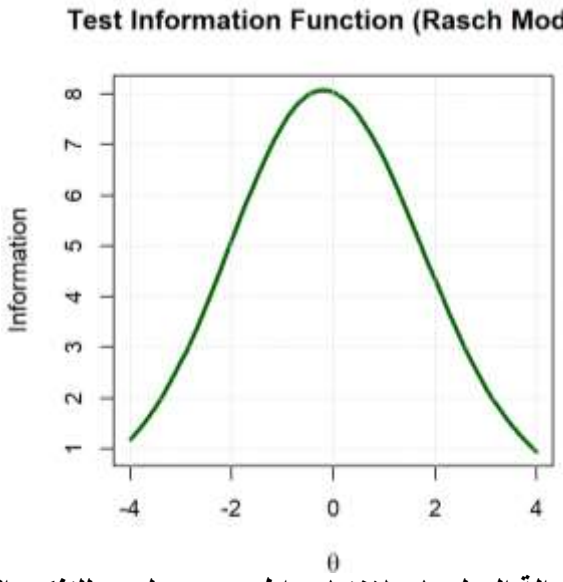
	Person	Ability	Std_Error		Person	Ability	Std_Error
165	Person165	0.023	0.358	**175**	Person175	0.228	0.361
166	Person166	-0.045	0.358	**176**	Person176	0.194	0.361
167	Person167	-0.114	0.357	**177**	Person177	-0.045	0.358
168	Person168	0.057	0.359	**178**	Person178	0.023	0.358
169	Person169	-0.011	0.358	**179**	Person179	-0.011	0.358
170	Person170	0.091	0.359	**180**	Person180	-0.114	0.357
181	Person181	-0.011	0.358	**186**	Person186	-0.216	0.357
182	Person182	-0.011	0.358	**187**	Person187	0.091	0.359
183	Person183	-0.182	0.357	**188**	Person188	-0.045	0.358
184	Person184	0.091	0.359	**189**	Person189	-0.011	0.358
185	Person185	0.091	0.359				

معلمة القدرة		الخطأ المعياري لتقدير معلمة القدرة	
المتوسط الحسابي	0.00٣	المتوسط الحسابي	0.358
أقل قيمة	-0.353	أقل قيمة	0.357
أكبر قيمة	0.297	أكبر قيمة	0.363

يوضح الجدول (٧) مستوى قدرة الافراد وفق نموذج راش إذ تراوحت قدرتهم من (-٠.٣٥٣) إلى (٠.٢٩٧) وبلغ متوسط قدرة الأفراد (٠.٠٠٣) بانحراف معياري قدره (٠.١١٩)؛ أي أن مستوى قدرة طلبة جامعة تبوك في الإجابة عن الاختبار كان متقاربًا جداً، ويشير هذا التجانس إلى أن الطلبة يمتلكون قدرات متشابهة نسبيًا في مستوى التفكير الناقد، يُعزى لوجود مقررات اعداد عام لمهارات التفكير الناقد في جامعة تبوك والتي أدت لتطوير مهارات التفكير الناقد لدى طلبة الجامعة بشكل متقارب، إضافة إلى تقارب وتشابه الخبرات التعليمية لطلبة الجامعة.

كما تراوحت قيم الأخطاء المعيارية لتقدير القدرة من (٠.٣٥٧) إلى (٠.٣٦٣) بمتوسط حسابي بلغ (٠.٣٥٨) وبانحراف معياري قدره (٠.٠٠١) مما يؤكد على أن دقة تقدير معالم القدرة للأفراد كانت متقاربة جدًا. وهذا ما تؤكده دالة معلومات الاختبار كما في الشكل (٢):

ISSN: 2537-0448 eISSN: 2537-0456



شكل (٢) دالة المعلومات لاختبار واطسون – جليسر للتفكير الناقد

حيث يوضح الشكل (٢) أن كمية المعلومات التي قدمها الاختبار تكون أكبر ما يمكن عند مستوى قدرة قريب من الصفر (٠)، وهو ما يتوافق مع قدرة الأفراد، حيث تراوحت قدراتهم من (-٠.٣٥٣) إلى (٠.٢٩٧) وبمتوسط بلغ (٠.٠٠٣)، مما يؤكد تمتع اختبار واطسون – جليسر للتفكير الناقد بدقة عالية في تقدير قدرات الأفراد.

وتتفق هذه النتيجة مع كل من دراسة (صباح،٢٠٢١) ودراسة (صالح، ٢٠٢٣) حيث جاءت بمتوسط للقدرة بلغ (-٠.٤٧)، (-٠.٥٨) وبانحراف معياري (٠.٤٦)، (٠.٦١).

النتائج المتعلقة بالسؤال الخامس: ما دلالات الصدق والثبات للبيانات المستمدة من الصورة القصيرة لاختبار واطسون – جليسر للتفكير الناقد في ضوء نظرية الاستجابة للفقرة؟

يشير إمبريتسون وريز (Embretson & Reise, 2013) إلى أن الصدق البنائي في نظرية الاستجابة للفقرة (IRT) يتحقق عندما تستوفي البيانات افتراضات النموذج، مثل أحادية البعد (Unidimensionality) والاستقلال المحلي (Local Independence)، بالإضافة إلى مطابقة الفقرات والأفراد للنموذج.

وفي نفس السياق تذكر السفياني وآخرون (٢٠٢٢) أنه يشترط للتحقق من صدق الاختبار في نظرية الاستجابة للفقرة يتم التحقق من افتراضات النموذج بالإضافة

ISSN: 2537-0448 eISSN: 2537-0456

لمطابقة الافراد والفقرات بصورته النهائية، تمت الإشارة مسبقاً في السؤال الأول لتحقق كافة افتراضات النظرية وفي السؤال الثاني تمت الإشارة لمطابقة الافراد. مما يؤكد على أن اختبار واطسون – جليسر للتفكير الناقد يتمتع بمؤشرات صدق البناء.
وللتحقق من الثبات تم حساب كل من معامل الثبات للأفراد والفقرات ومعامل الفصل للأفراد والفقرات في الجدول (٨) التالي:

جدول (٨) معاملات الثبات والفصل للأفراد والفقرات

المؤشر	القيمة	التفسير
Item Reliability (ثبات الفقرات)	**0.98**	مرتفع جدًا، يدل أن الفقرات مميزة وتغطي تباين واسع
Item Separation (فصل الفقرات)	**6.16**	ممتاز جدًا، يدل على ان الفقرات تفرق بوضوح بين مستويات القدرة
Person Separation (فصل الأفراد)	**NA**	منخفض جدًا، يدل أن الأفراد متقاربين جدًا في القدرات
Person Reliability (ثبات الأفراد)	**NA**	ضعيف جدًا، يعكس أن العينة ما أظهرت فروق قوية بين الأفراد

يوضح الجدول(٨) نتائج التحليل الاحصائي بأن الفقرات تمتعت بمعامل ثبات للفقرات عالٍ (٠.٩٨) ومعامل فصل للفقرات (٦.٧٦) قوي، مما يشير إلى جودة الاختبار وتمكنه من التمييز بين مستويات الصعوبة المختلفة، بالمقابل اظهر البرنامج قيمة (NA) لكل من معامل فصل الافراد وثبات الافراد، مما يُعزى إلى تجانس لقدرات أفراد العينة، وتتفق هذه النتائج مع دراسة (صالح، ٢٠٢٣) من حيث معامل الثبات للفقرات ومعامل الفصل للفقرات وتختلف من حيث معامل ثبات الافراد ومعامل الفصل للأفراد، بينما أشارت النتائج في دراسة (صباح،٢٠٢١) والتي استخدمت نفس الاختبار بصورته المختصرة حيث بلغ معامل الفصل للأفراد (٠.٦٨) وهو ضعيف جداً كما بلغ معامل الثبات للأفراد (٠.٣٢) وهو أدنى من(٠.٧) أي أنه ضعيف جداً.

النتائج المتعلقة بالسؤال السادس: ما مستوى التفكير الناقد لدى طلبة جامعة تبوك؟

وللإجابة عن السؤال تم حساب درجات الطلبة على اختبار واطسون – جليسر للتفكير الناقد، حيث تم التصحيح وفقًا لمعايير الاختبار وفيها يتم حصول الطالب على درجة (١) للإجابة الصحيحة عن كل فقرة، وصفر (٠) للإجابة الخاطئة، وبذلك يكون المدى ما بين صفر (٠) إلى أربعين (٤٠) درجة. وتراوحت درجات الطلبة من (١١) إلى (٣٠) وبلغ المتوسط الحسابي لدرجات الطلبة (٢١.٣٢) بانحراف معياري (٣.٢٤). كما تم استخدام اختبار (T) لعينة واحدة لمقارنة متوسط درجات طلبة

ISSN: 2537-0448 eISSN: 2537-0456

جامعة تبوك على مقياس التفكير الناقد مع المتوسط الفرضي للاختبار (٢٠) درجة، وجاءت النتائج كما في الجدول (٩):

جدول (٩) اختبار (T)

العينة	المتوسط الفرضي	المتوسط الحسابي	الخطأ المعياري	T	DF	Sig	Cohen's d
189	20	21. 32	0.24	5.62	188	.001	0.41

يوضح الجدول (٩) أن المتوسط الحسابي لأفراد العينة بلغ (٢١.٣٣) بينما بلغ المتوسط الفرضي للاختبار (٢٠) وتم استخدام اختبار (T) لمعرفة دلالة هذا الفرق بين المتوسطين، وبيّنت النتائج أن قيمة اختبارT بلغت (٥.٦٢) وكان الاختبار دال احصائياً عند مستوى دلالة (٠٠١.) بحجم أثر (٠.٤١) متوسط إلى صغير نسبيًا، ويعني ذلك ان مستوى التفكير الناقد لدى طلبة جامعة تبوك أعلى من المتوسط الفرضي، وقد يرجع ذلك لوجود مقررات الإعداد العام لمهارات التفكير الناقد، وتختلف هذه النتيجة مع كل من دراسة (صالح, ٢٠٢٣؛ Din, 2020) إذ جاء مستوى الطلبة في دراسة صالح (٢٠٢٣) دون المتوسط وجاء في دراسة (Din, 2020) أن مستوى الطلبة متوسط.

توصيات الدراسة

من خلال ما توصلت إليه نتائج الدراسة توصي الباحثة بما يلي:

١. الاستمرار في تعزيز مقررات التفكير الناقد ضمن متطلبات الإعداد العام في جامعة تبوك، لما لها من أثر واضح في رفع مستوى تفكير الطلبة وتطوير قدراتهم في التفكير الناقد.
٢. استخدام الصورة القصيرة لاختبار واطسون–جليسر بصيغته المتحققة من افتراضات نظرية الاستجابة للفقرة، خاصة في الدراسات المهتمة بقياس التفكير الناقد بطريقة دقيقة ومبنية على أسس إحصائية سليمة.
٣. استخدام نموذج راش في بناء وتحليل أدوات القياس التربوي والنفسي، نظرًا لفعاليته في تحديد جودة الفقرات وكفاءة القياس في هذه الدراسة بالإضافة لنماذج الأخرى.
٤. التركيز على قياس قدرات التفكير الناقد بشكل مستمر باستخدام أدوات تحقق شروط الصدق والثبات، خاصة في ضوء أهمية هذه المهارة في دعم الأداء الأكاديمي والمهني.

ISSN: 2537-0448 eISSN: 2537-0456

المراجع :

إبراهيم، جيهان. (٢٠٢٣). أثر برنامج تدريبي قائم على الذكاءات المتعددة في تنمية مهارات التفكير الناقد والتحصيل الدراسي لدى تلاميذ ذوي صعوبات التعلم بالمرحلة الابتدائية. *إبداعات تربوية*, ٢٦ (١), ١٧-٥٨.
10.21608/eji.2023.306131.

أيالا، دي (٢٠١٧). **النظرية والتطبيق في نظرية استجابة الفقرة** (ترجمة الكيلاني، عبدالله ؛ والبرصان، إسماعيل). الرياض: جامعة الملك سعود.

أيوب، حسين. (٢٠١٤). علاقة التفكير الناقد بمهارة حل المشكلات والتحصيل الدراسي لطلبة جامعة الإمام محمد بن سعود الإسلامية .*العلوم التربوية، مج٢٢ , ع٣* ، ٢٨١ .320 - مسترجع من
http://search.mandumah.com/Record/649986

البلوي، فرحان علي صالح، و عبدالعال، صبرى محمد إسماعيل. (٢٠١٨). أثر زمن الإختبارعلى تقدير معالم الفقرات وثبات الإختبار: دراسة مقارنة نظرية استجابة الفقرة والنظرية التقليدية .*المجلة الدولية للدراسات التربوية والنفسية، مج٤ , ع٢* ، ٢٥٢ .276 - مسترجع من
http://search.mandumah.com/Record/929823

بن وزة، خديجة و قماري، محمد. (٢٠٢١). تدريج مقياس الذكاء المتعدد باستخدام نموذج راش دراسة ميدانية على عينة من تلاميذ التعليم الثانوي بوادي اهيو غليزان. القياس والدراسات النفسية, ١(٣). ١٤٤-١٦٥.
https://asjp.cerist.dz/en/article/162514

بوبو، منذر، و شبيب، هناء ، وشريبة، بشرى. (٢٠١٨). الخصائص السيكومترية لاختبار كورنيل للتفكير الناقد مستوي "X"ومستوياته لدى عينة من الطلبة الموهوبين: دراسة ميدانية في المركز الوطني للمتميزين في سورية .*مجلة جامعة تشرين للبحوث والدراسات العلمية - سلسلة الآداب والعلوم الإنسانية، مج٤٠ , ع١* ، ٢٠٥ .223 - مسترجع من
http://search.mandumah.com/Record/1186085

بيكر، فرانك ب. (٢٠١٠). *أسس نظرية الاستجابة للمفردة* (عبدالرحمن بن سليمان الطريري، السيد محمد أبو هاشم، سوسن إبراهيم شلبي، مُترجم). جامعة الملك سعود. (العمل الأصلي نشر في ٢٠٠١).

جميل، عصام زكريا. (٢٠١٩). *التفكير الناقد منطق للحياة اليومية*. نيوبوك للنشر والتوزيع.

ISSN: 2537-0448 eISSN: 2537-0456

جندي، محمد. (٢٠٢٣). مقارنة تحليل بنود الأسئلة باستخدام نظرية الاختبار التقليدية ونظرية استجابة البنود لدرس اللغة العربية في المدرسة المتوسطة الإسلامية الحكومية باتو. In http://etheses.uin-malang.ac.id/60055/.

جودة، جيهان. (٢٠٢٢). فعالية برنامج تدريبي قائم على استراتيجيات التفكير الناقد وأثره على التفكير النمطي بالأدوار الاجتماعية الجندرية لدي الأطفال في المرحلة الابتدائية. *المجلة الدولية لدراسات المرأة والطفل*,٢ (٣), ١-٥٦.
10.21608/ijcws.2022.313162

الرباط، بهيره شفيق إبراهيم. (٢٠٢٢). التفكير الناقد رؤى وتطبيقات. دار العالم العربي

الزهراني، أميرة. (٢٠٢٢). العلاقة بين التفكير الناقد والتحصيل الدراسي لدى طالبات الصف الثالث المتوسط بمكة المكرمة .*مجلة الأستاذ للعلوم الإنسانية والاجتماعية، مج٦١ , ع٤* ، ٣٤٧ .363 - مسترجع من
http://search.mandumah.com/Record/1343625

السفياني، رهف بنت سعد، وأبو عيش، بسينة بنت رشاد بن علي. (٢٠٢٤). استخدام نموذج راش في تقدير الخصائص السيكومترية لمقياس قلق المستقبل المهني للطلبة المقبلين على التخرج. المجلة العربية للنشر العلمي، ع٦٧، ٤٧٣ - ٥٢٦. مسترجع من http://search.mandumah.com/Record/1547776

السكري، عماد الدين محمد. (٢٠١٠). بعض الخصائص السيكومترية لاختبار واطسون - جليسر للتفكير الناقد، المصورة "S" لدي عينة من طلاب الجامعة .*المجلة المصرية للدراسات النفسية، مج٢٠ , ع٦١* ، ٢٢٣ .251 - مسترجع من
http://search.mandumah.com/Record/1010142

سليمان، أمين علي وأبو علام، رجاء محمود. (٢٠٠٩). القياس والتقويم في العلوم الإنسانية أسسه وأدواته وتطبيقاته. دار الكتاب الحديث

شطناوي، محمد خالد محمد، وسواقد، ساري سليم. (٢٠٠٣). تقنين اختبار كاليفورنيا لمهارات التفكير الناقد للطلبة الجامعيين في الأردن [رسالة ماجستير غير منشورة]. جامعة مؤتة، مؤتة. مسترجع من
http://search.mandumah.com/Record/572244

الشمري، عفاف بنت عليوي بن سعد، وآل رشيد، هياء معجب مهدي. (٢٠٢١). التفكير الناقد. *المجلة العربية للنشر العلمي، ع٢٩* ، ٦٤٦ - ٦٦٨. مسترجع من http://search.mandumah.com/Record/1435574.

الشنقيطي، سامية. (٢٠٢٤). فاعلية استخدام استراتيجية التفكير الناقد والإبداعي في تحسين مستوى الاستيعاب القرائي لدى طالبات قسم اللغة الانجليزية في المرحلة

ISSN: 2537-0448 eISSN: 2537-0456

الجامعية. *مجلة البحث في التربية وعلم النفس*, ٣٩ (٣)، ١-٤٦ . 10.21608/mathj.2024.308291.1473

الشويرخ، (٢٠٢١). البدء بتدريس مادة "التفكير الناقد" العام القادم. تم الاسترداد من جريدة الرياض: https://www.alriyadh.com/1879288.

الطلى، أحمد. (٢٠٢٢). الخصائص السيكومترية لمقياس بارات للاندفاعية: مقارنة بين النظرية الكلاسيكية ونظرية الاستجابة للمفردة .*مجلة بحوث كلية الآداب. جامعة المنوفية*10.21608/sjam.2022.140725.1568 ,(131.3)33 ,

عام / وزير التعليم يدشن برنامج "مهارات التفكير الناقد والفلسفة" للمرحلة الثانوية. (٢٠١٨). تم الاسترداد من وكالة الانباء السعودية:

https://www.spa.gov.sa/1851747.

العتيبي، خالد بن ناهس الرقاص. (٢٠١٢). الخصائص السيكومترية للصورة القصيرة من اختبار واطسون - جليسر للتفكير الناقد: دراسة على عينة من الطلاب والمعلمين في البيئة السعودية .*مجلة جامعة الملك سعود - العلوم التربوية والدراسات الإسلامية*، مج ٢٤ , ع ٤ ، ١٤٢٧ .1454 - مسترجع من

http://search.mandumah.com/Record/427282

عزيز، صادق عبدالنور، وكرمة، صفاء طارق حبيب. (٢٠١٢). تطوير اختبار كاليفورنيا لمهارات التفكير الناقد لدى طلبة الجامعة وفقاً لنظرية السمات الكامنة. مجلة الأستاذ للعلوم الإنسانية والاجتماعية، ع ٢٠٢، ٦٩٣ - ٧١٨. مسترجع من

http://search.mandumah.com/Record/440877

العقل، عقل بن عبدالعزيز، العبدالكريم، راشد بن حسين، والحسينان، إبراهيم بن عبدالله. (٢٠٢٢). *التفكير الناقد المفهوم والتطبيقات*. شركة وجوه للنشر والتوزيع.

علام، صلاح الدرين. (٢٠٠٠). القياس والتقويم التربوي والنفسي اساسياته وتطبيقاته وتوجهاته المعاصرة. دار الفكر العربي

العنزي، بدرية. (٢٠٢٣). *معوقات تنمية مهارات التفكير الناقد لدى طلبة جامعة تبوك من وجهة نظر خبراء التربية* (رسالة ماجستير غير منشورة). جامعة تبوك

القحطاني, سالم, العامري, أحمد, آل مذهب, معدّي, و العمر, بدران (٢٠٢٠).منهج البحث في العلوم السلوكية (ط.٥). العبيكان.

محاسنة، إبراهيم محمد. (٢٠١٣). *القياس النفسي في ظل النظرية التقليدية والنظرية الحديثة*. دار جرير للنشر والتوزيع.

محمد، سحر جمال، والضوى، محسوب عبد القادر، وسليمان، شيماء سيد. (٢٠٢٤). استخدام نموذج راش والتحليل العاملى لتحرى أحادية البعد فب مقياس الشغف

ISSN: 2537-0448 eISSN: 2537-0456

الأكاديمي: دراسة سيكومترية. *مجلة العلوم التربوية - كلية التربية بقنا*, ٦٠ (٦٠). 10.21608/maeq.2024.349831 .

موسى، مبراك. (٢٠٢١). التفكير الناقد والممارسات التعليمية. *مجلة البحوث التربوية والتعليمية*، مج١٠ , ع٢، ١٥٩ - ١٧٤. مسترجع من http://search.mandumah.com/Record/1213332.

APA Dictionary of Psychology. (n.d.). Https://Dictionary.Apa.Org/Critical-Thinking.

Arif, R. N. H. (2024). Assessment of Critical Thinking Ability in Science Learning Using Watson-Glaser Critical Thinking Appraisal (WGCTA). ARRUS Journal of Social Sciences and Humanities, 4(2), 270–275. https://doi.org/10.35877/soshum2599

Baker, F. B. (2001). The basics of item response theory. ERIC Clearinghouse on Assessment and Evaluation.

Bond, T. (2015). Applying the Rasch Model: Fundamental Measurement in the Human Sciences (3rd ed.). Routledge. https://doi.org/10.4324/9781315814698

Choi, I. C. (1992) An Application of Item Response Theory to Language Testing, Theoretical Studies in Second Language Acquisition Vol. 2, Peter Lang, New York.

de Ayala, R. J. (2022). The Theory and Practice of Item Response Theory. Guilford Publications. https://books.google.com.sa/books?id=Ai1gEAAAQBAJ

Dewey, J. (1910). How We Think. D.C. Heath & Company.

Din, M. (2020). Evaluating university students' critical thinking ability as reflected in their critical reading skill: A study at bachelor level in Pakistan. Thinking Skills and Creativity, 35. https://doi.org/10.1016/j.tsc.2020.100627

Duru, E., Ozgungor, S., Yıldırım, O., Paksu, A. D., & Duru, S. (2022). Pamukkale critical thinking skill scale: a validity and reliability study. International Journal of Assessment Tools in Education, 9(3), 741-771.

ISSN: 2537-0448 eISSN: 2537-0456

Embretson, S. E., & Reise, S. P. (2013). Item response theory for psychologists. In Item Response Theory for Psychologists. Taylor and Francis. https://doi.org/10.4324/9781410605269

Ennis, R. (2011a). Critical Thinking Reflection and Perspective Part I. Inquiry: Critical Thinking Across the Disciplines, 26, 4–18. https://doi.org/10.5840/inquiryctnews20112613

Facione, P. A. (1991). Using the California Critical Thinking Skills Test in Research, Evaluation, & Assessment. California Academic Press. https://eric.ed.gov/?id=ED337498

Facione, P. A. (1998). Critical Thinking: What It Is and Why It Counts. California: California Academic Press.

Franzen, M. D. (2011). Classical Test Theory. In Encyclopedia of Clinical Neuropsychology (pp. 586–587). Springer New York. https://doi.org/10.1007/978-0-387-79948-3_1180

Hambleton, R. K., & Swaminathan, H. (1985). Item Response Theory. Springer Netherlands. https://doi.org/10.1007/978-94-017-1988-9

Hambleton, R. K., & Zaal, J. N. (1991). Advances in Educational and Psychological Testing: Theory and Applications. Springer Netherlands. https://doi.org/10.1007/978-94-009-2195-5

Harvey, R. J., & Hammer, A. L. (1999). Item Response Theory. The Counseling Psychologist, 27, 353–383. https://doi.org/10.1177/0011000099273004

Khine, M. S. (2020). Rasch measurement: Applications in quantitative educational research. Springer Singapore. https://doi.org/10.1007/978-981-15-1800-3

Linacre, J. (1994). *Sample size and item calibration stability*. Https://Www.Semanticscholar.Org/Paper/Sample-Size-and-Item-Calibration-Stability-Linacre/Bd9943b087bfeb266bf8a77290ecb15bc1ccc9be.

ISSN: 2537-0448 eISSN: 2537-0456

Lord, F. M. (2012). Applications of item response theory to practical testing problems. Taylor and Francis. https://doi.org/10.4324/9780203056615

Pual, R., & Elders, L. (2008). Critical & Creativity thinking. The Foundation for Critical Thinking. From: www.criticalthinking.org

Reggoug, N., Boulaz, F.-E., Haouachim, I. El, & Omari, K. El. (2024). Assessing Critical Thinking in a Professional Context: A Case Study of Healthcare Professionals. European Modern Studies Journal, 8(5), 165–175. https://doi.org/10.59573/emsj.8(5).2024.15

Shipman, V. (1983). New Jersey Test of Reasoning Skills - Review. https://digitalcommons.montclair.edu/iapc_nj_test_reasoning skills/3

Sutoyo, S., Agustini, R., & Fikriyati, A. (2023). Online Critical Thinking Cycle Model to Improve Pre-Service Science Teacher's Critical Thinking Dispositions and Critical Thinking Skills. Pegem Journal of Education and Instruction, 13(2), 173–181. https://eric.ed.gov/?id=EJ1386742

Watson, G., & Glaser, E. (2018). Watson-GlaserTM III Critical Thinking Appraisal User's Guide and Technical Manual. www.TalentLens.com

Watson, G., & Glaser, E. M. (2009). Watson-Glaser TM II Critical Thinking Appraisal Technical Manual and User's Guide. www.TalentLens.com

ISSN: 2537-0448 eISSN: 2537-0456