جامعة سوهاج

كلية التربية

مجلة شباب الباحثين

الخصائص السيكومترية لاختبار التفكير الاستدلالي باستخدام نظرية الاستجابة للمفردة

إعداد

أ/ ملاك علي محمد الحربي
باحثة دكتوراه في القياس والتقويم بقسم علم النفس– كلية التربية
جامعة أم القرى

د. محمد رزق الله رزق الله الزهراني
أستاذ القياس النفسي والتقويم التربوي
المشارك بجامعة أم القرى

تاريخ استلام البحث : ١٦ فبراير ٢٠٢٣م – تاريخ قبول النشر: ٨ مارس ٢٠٢٣م

DOI: 10.21608/JYSE. 2023

مستخلص البحث:

هدفت الدراسة الحالية إلى التحقق من الخصائص السيكومترية لاختبار التفكير الاستدلالي باستخدام نظرية الاستجابة للمفردة. ولتحقيق هذا الهدف طبقت الدراسة النسخة العربية من اختبار التفكير الاستدلالي لحجازين (2003)، على عينة تكونت من (2023) من طلبة المرحلة الثانوية في العام الدراسي 1443هـ وقد تم تحليل البيانات في الدراسة الحالية باستخدام نظرية الاستجابة للمفردة. يتكون الاختبار من 70 مفردة اختبارية من نوع الاختيار من متعدد، وقد تم التحقق من افتراضات نماذج نظرية الاستجابة للمفردة أحادية البعد. وتم تحليل بيانات الدراسة باستخدام برنامجي SPSS وBilog Mg، وقد أسفرت النتائج عن تطابق 53 مفردة من مفردات الاختبار مع النموذج الثلاثي البارامتر. كذلك أسفرت النتائج عن قيم مقبولة لبارامترات الصعوبة تراوحت بين (0.1- و 1.399) ولبارامترات التمييز تراوحت بين (0.792 و 1.994) وبارامترات التخمين تراوحت بين (0.103 و0.386) وتعتبر جميعها ضمن المدى المقبول. كما أشارت النتائج إلى ثبات مقبول للاختبار بلغ (0.929).

الكلمات المفتاحية: الخصائص السيكومترية، اختبار التفكير الاستدلالي، نظرية الاستجابة للمفردة.

Psychometric Properties of The Deductive Thinking Test Using Item Response Theory

Abstract:

The current study aimed to investigate the psychometric properties of the deductive thinking test using Item Response Theory. To achieve this goal, the study applied the Arabic version of deductive thinking test for Hijazin (2003), on a sample of (2023) secondary school students in the academic year 1443 AH. And the data of the current study had been analyzed by using Item Response Theory (IRT). The test consists of 70 items of the multiple-choice type, and univariate IRT Models assumption had been verified. The data of the study were analyzed using the SPSS and Bilog Mg programs, and the results revealed that 53 of the test items matched the three-parametric model. The results also showed acceptable values for the difficulty parameters ranged between (-0.1 and 1.399) and for the discrimination parameters ranged between (0.792 and 1.994). The guessing parameters were ranged between (0.103 and 0.386), and all of which considered within the acceptable range. The results also indicated an acceptable reliability coefficient reached to (0.929)

Keywords: Psychometric Properties, Deductive Thinking Test, Item Response Theory.

مقدمة البحث:

تعد مهارات التفكير الاستدلالي من المهارات الأساسية التي تسعى المؤسسات التربوية لتطويرها وتدعيمها بهدف جعلها عادة ومهارة يمتلكها الطالب في تحليليه وتعاطيه مع المواقف اليومية في حياته وذلك لأن هذا الأسلوب من التفكير يساعده على الفهم السريع والمتعمق للمعطيات من حوله والتفاعل معها بشكل إيجابي وبصورة فورية.

وتبدو أهمية هذا النمط من التفكير في أنه يساعد الفرد في اتخاذ قرارات مستقبلية في الكثير من المواقف الحياتية ليحصل على نتائج صحيحة تجعل حياته أكثر تنظيماً، فهو لا يقتصر على اكتساب المعرفة، بل يصل إلى الحياة الشخصية للفرد وذلك لحل المشكلات التي تواجهه وتحليل أبعادها لمعرفة السبب والنتيجة للوصول إلى الحلول المناسبة أو التنبؤ بما سيحدث مستقبلاً. (رزوقي وعبد الكريم، 2015)

ويعد الاستدلال نشاطاً فكرياً يهدف إلى اتخاذ قرار أو حل مشكلة حلاً ذهنياً عن طريق الرموز والخبرات السابقة، فالتفكير الاستدلالي هدفه الوصول إلى نتيجة لم تكن معلومة من مقدمات معلومة وهذا ما يميز الاستدلال عن غيره من أنواع التفكير فما يميزه هو الانتقال من معلوم إلى مجهول (الحجازين، 2014). وتبرز أهمية دراسة هذا النوع من التفكير كونه يعد ضرورة من الضرورات التي تقوم عليها أنواع التفكير الأخرى، ويكاد يتفق الكثير من الباحثين في مجال أنماط التفكير على أن التفكير الاستدلالي هو العنصر المشترك بين كل أنواع التفكير الأخرى كالتفكير العلمي والتفكير التأملي والتفكير الناقد وأسلوب حل المشكلات (رزوقي وعبد الكريم، 2015). كما أنه يرتبط ارتباطاً وثيقاً بالذكاء، وقياسه يساعد في قياس وتحديد القدرات العقلية للأفراد فالناس جميعاً لديهم القدرة على التفكير الاستدلالي، ولكن بدرجات متفاوتة. (راجح، 1977)

وتتأكد أهمية أسلوب التفكير الاستدلالي في التنبؤ بمدى نجاح الموظفين في أداء مهامهم الوظيفية حيث تعتمد وزارة الخدمة المدنية في المملكة العربية السعودية في إجراء المفاضلات الوظيفية على وظائف السلم العام في نظام الخدمة المدنية على نتائج اختبار القدرة المعرفية والذي يتكون من أربع قدرات رئيسة أحدها القدرة الاستدلالية وهي القدرة على التفكير المنطقي من خلال تطبيق قواعد الاستدلال الاستنباطي والاستقرائي على البيانات أو

المعلومات الرمزية المجردة أو المعنوية للتوصل إلى استنتاجات بشأنها. (هيئة تقويم التعليم والتدريب، 2021)

ويستخدم الكثير من الباحثين مقاييس واختبارات لقياس مهارات التفكير الاستدلالي ويطبقونها على فئات مختلفة من المفحوصين ويدرسون الفروق بينهم، إلا أن تحققهم من جودة الخصائص السيكومترية لها من صدق وثبات، وتحليل للمفردات يتم بصورة سطحية وليست متعمقة ووفقاً لطرق النظرية التقليدية في القياس.

وفي هذا الصدد أشار علام (2015) إلى أن النظرية التقليدية في القياس النفسي والتربوي تعد من النظريات محدودة الدقة في معالجتها لمفهومي الصدق والثبات نظراً لأنها لا تميز بين أخطاء القياس وإنما تعطي قيمة تقديرية كلية لمصادر أخطاء القياس المتعددة، كما أن التعبير عن قدرة الفرد يتم من خلال الدرجة الحقيقية التي تتضح من خلال أداء الفرد على الاختبار ككل وليس على مستوى المفردة وبالتالي سيتغير وضع قدرة الفرد حسب تغير مستوى الاختبار، وبالإضافة إلى ذلك فإن خصائص الاختبار والمفردات تتغير بتغير خصائص الأفراد كما أن خصائص الأفراد تتغير بتغير خصائص الاختبار من حيث الصعوبة والسهولة وهذا يتعارض مع موضوعية القياس واستقلاليته.

ونظراً لما تعاني منه النظرية التقليدية من جوانب قصور واضحة وعدم تحقيقها لموضوعية القياس، فقد تزايد اهتمام العلماء في الدول المتقدمة بضبط إجراءات بناء المقاييس والتحقق من جودة الخصائص السيكومترية لها باستخدام النظريات الحديثة في القياس ومنها نظرية الاستجابة للمفردة، كما قاموا بجهود كبيرة لتفعيل استخدام هذه النظرية وتطبيقاتها المختلفة وما ينبثق عنها من نماذج في محاولة للتغلب على عيوب النظرية التقليدية التي سيطرت على منهجيات القياس طوال القرن الماضي، وما تزال أغلب الدراسات والقياسات في العالم العربي تعتمد على مبادئها وقوانينها في بناء وتقنين الاختبارات والمقاييس المختلفة. (عبدالوهاب، 2010)

وتقوم نظرية الاستجابة للمفردة على فرضية أساسية وهي أن القيمة الاحتمالية لاستجابة فرد ما عن مفردة اختبارية تكون عبارة عن دالة في متغيرين رئيسين هما: المتغير المراد قياسه، وخصائص المفردة التي يحاول الفرد الإجابة عنها (علام، 2001). وتنقسم نماذج هذه النظرية إلى نماذج أحادية البعد ونماذج متعددة الأبعاد، إلا أن النماذج الأحادية

البعد هي الأكثر شيوعاً واستخداماً وتتطلب هذه النماذج مجموعة من الافتراضات الواجب توفرها في البيانات المستمدة من الاختبار وهي: افتراض أحادية البعد ويعني وجود قدرة واحدة كامنة تفسر أداء الفرد في الاختبار، وافتراض الاستقلال الموضعي ويعني أن استجابات الفرد على مفردات الاختبار المختلفة مستقلة عن بعضها البعض عند مستوى قدرة معين بمعنى أن أداؤه على مفردة ما لا يتأثر سلباً ولا ايجاباً بأدائه على مفردة أخرى، بالإضافة إلى افتراضي التحرر من السرعة وتحقق منحنى خصائص المفردة (Hambleton & Swaminathan, 1985). ويوجد ثلاثة نماذج لتحليل مفردات الاختبارات أحادية البعد ثنائية الاستجابة وهي النموذج الأحادي البارامتر والنموذج الثنائي البارامتر والنموذج الثلاثي البارامتر والفارق الأساسي بين هذه النماذج هو عدد البارامترات التي يتم تقديرها في كل مفردة ففي النموذج الأحادي يتم تقدير بارامتر الصعوبة فقط، وفي النموذج الثنائي يتم تقدير بارامتري الصعوبة والتمييز، أما في النموذج الثلاثي فيتم تقدير بارامتر التخمين بالإضافة إلى بارامتري الصعوبة والتمييز (أبو فودة، 2016).

وقد أوضح عبد الوهاب (2010) حاجة الباحثين والمتخصصين في التربية وعلم النفس في البيئة العربية إلى قياس وتشخيص الجوانب المعرفية والوجدانية على أساس علمي دقيق من خلال توفير مقاييس واختبارات ذات خصائص سيكومترية جيدة للاستفادة منها في أبحاثهم العلمية وإرشاد وتوجيه طلابهم. ويقوم الكثير من الباحثين ببناء أدوات بحثية والتحقق من جودتها السيكومترية لإجراء دراساتهم، فقد قام المنشدي (2019) ببناء اختبار لقياس القدرة الاستدلالية لدى طلبة المرحلة الإعدادية، وكانت الخصائص السيكومترية له جيدة. كما قام باسيرير BASERER (2020) ببناء اختبار يهدف لتحديد مستوى التفكير المنطقي لدى الطلاب المتدربين على مهنة التعليم من التخصصات المختلفة. وسعت الحجيلي (2022) إلى الكشف عن البناء العاملي لاختبار الاستدلال المنطقي الذي قامت ببنائه والتحقق جودة خصائصه السيكومترية، إلا أنه في جميع تلك الدراسات تم تحليل البيانات والتحقق من جودة الخصائص السيكومترية للأدوات المستخدمة بها في ضوء نظرية القياس التقليدية.

ونظراً لما يشوب نظرية القياس التقليدية من عيوب وقصور وما تمتاز به نظرية الاستجابة للمفردة من مزايا ودقة في القياس فقد أوصى عبد الحافظ (2016) بضرورة

الاستفادة من نظرية الاستجابة للمفردة ونماذجها الرياضية وتطبيقاتها العملية في تطوير الاختبارات والمقاييس النفسية والتربوية لما لها من فائدة في تحقيق دقة وموضوعية القياس. وقد سعى عيال وجاسم (2019) إلى تطوير اختبار لويولا لمهارات الاستدلال الاستقرائي لـ Charles Doyle لدى طلبة الجامعة في العراق، وقد قام بتحليل مفرداته في ضوء نظرية الاستجابة للمفردة وباستخدام النموذج الثلاثي البارامتر. كما قام فيرولا وآخرون (Vierula, et. al, 2021) بتطوير اختبار يقيس مهارات التفكير الاستدلالي لدى المتقدمين لكلية التمريض في مرحلة البكالوريوس وذلك لاختيار الأفضل منهم، وقد تم تحليله من الناحية السيكومترية باستخدام نظرية الاستجابة للمفردة من خلال تحديد بارامترات الصعوبة والتمييز.

ويتضح من خلال مراجعة أدوات القياس في الأدبيات التربوية السابقة، وجود ندرة في تحليل تلك الأدوات والمقاييس في ضوء نظرية الاستجابة للمفردة بصفة عامة وفي الاختبارات والمقاييس الخاصة بالتفكير الاستدلالي بصفة خاصة، مما يؤكد أهمية إجراء دراسة للتحقق من الخصائص السيكومترية لاختبار التفكير الاستدلالي بالاعتماد على نظرية الاستجابة للمفردة في البيئة السعودية على وجه التحديد وهو ما تسعى إليه الدراسة الحالية.

مشكلة الدراسة

يعد التفكير الاستدلالي من مهارات التفكير المهمة التي يسعى الكثير من الباحثين إلى تسليط الضوء عليها ودراستها من عدة جوانب، فقد سعى العديد من الباحثين إلى دراسة التفكير الاستدلالي وعلاقته بعدد من المتغيرات المختلفة، كما سعى البعض منهم إلى بناء اختبارات وأدوات تقيس التفكير الاستدلالي وسعوا إلى التحقق من جودتها من خلال تحليلها ودراسة الخصائص السيكومترية لها، إلا أن أغلب تلك الدراسات كانت تتم في ضوء نظرية القياس التقليدية.

ونظراً لما تعاني منه نظرية القياس التقليدية من أوجه قصور تقلل من دقة النتائج المستمدة منها من خلال اعتمادها على الدرجة الكلية الملاحظة التي يحصل عليها الفرد بإجاباته على جميع مفردات الاختبار والحكم على مستواه بغض النظر عن صعوبة كل مفردة (Baker, 2001). وكذلك من خلال تقديرها لخطأ معياري واحد لجميع الأفراد المختبرين، بالإضافة إلى اعتمادها على عينة الأفراد الذين تم تطبيق الاختبار عليهم في حساب الثبات بالإضافة إلى بارامترات الصعوبة والتمييز مما يعني عدم توفر القياس الموضوعي حيث يتأثر

الثبات بدرجة تجانس العينة، وتتأثر بارامترات الصعوبة والتمييز بمستوى الأفراد الذين تم تطبيق الاختبار عليهم وبالتالي تختلف خصائص الاختبار باختلاف العينة التي يتم تطبيق الاختبار عليها (Reeve, 2004).

في حين تتميز نظرية الاستجابة للمفردة بتحقيق استقلال خصائص المفردات كالصعوبة والتمييز عن عينة الأفراد المستخدمة في التحليل، وكذلك استقلال تقدير قدرات الأفراد عن عينة المفردات المكونة للمقياس. كما أن تفسير درجات الأفراد يتم في ضوء المفردات وليس في ضوء الجماعة المرجعية كما في نظرية القياس التقليدية، ويتم تقدير خطأ معياري لكل فرد على حده وليس خطأ معياري واحد لكل الأفراد مع إمكانية المقارنة بين أداء الأفراد الذين اختبروا باختبارات مختلفة تقيس نفس السمة أو القدرة (علام،2005).

ومع التقدم الحاصل في علم القياس وكثرة الباحثين في هذا المجال، فإن تطوير الاختبارات النفسية والتربوية في ضوء نظرية الاستجابة للمفردة أصبح من الضروريات الملحة اليوم وبخاصة في العالم العربي، فقد أثبتت هذه النظرية تفوقها ومعالجتها لكثير من أوجه القصور في نظرية القياس التقليدية عند التحقق من جودة الخصائص السيكومترية لهذه الاختبارات والمقاييس وتحقيقها لشروط القياس الموضوعي للسمة المراد قياسها والمتمثلة في استقلال تقدير خصائص المفردات عن عينة الأفراد المطبق عليهم المقياس، واستقلال تقديرات الأفراد عن مجموعة المفردات المكونة للمقياس (الشرقاوي وآخرون،1996). وبالتالي يمكن الثقة بدرجة أكبر في النتائج المستمدة منها. ومن خلال مراجعة الأدبيات التربوية السابقة والاطلاع على العديد من الدراسات الخاصة بالتفكير الاستدلالي، اتضح ندرة الدراسات التي اهتمت بتحليل الاختبارات وأدوات القياس باستخدام نظرية الاستجابة للمفردة ولذلك تأتي هذه الدراسة للكشف عن الخصائص السيكومترية لاختبار التفكير الاستدلالي في البيئة السعودية والتحقق من جودتها في ضوء نظرية الاستجابة للمفردة وذلك من خلال تدريج مفردات الاختبار باستخدام النموذج الثلاثي البارامتر. وتحاول الدراسة الإجابة عن السؤال الرئيس التالي:

ما الخصائص السيكومترية لاختبار التفكير الاستدلالي باستخدام نظرية الاستجابة للمفردة؟

ويتفرع من هذا السؤال الأسئلة التالية:

- **هل تتحقق افتراضات نماذج نظرية الاستجابة للمفردة أحادية البعد في بيانات الدراسة الحالية؟**
- **هل تتطابق مفردات اختبار التفكير الاستدلالي مع النموذج الثلاثي البارامتر؟**
- **ما تقدير بارامترات الصعوبة والتمييز والتخمين لمفردات اختبار التفكير الاستدلالي؟**
- **ما تقدير دالة المعلومات لاختبار التفكير الاستدلالي وللمفردات؟**
- **ما تقدير بارامتر القدرة للأفراد على اختبار التفكير الاستدلالي؟**
- **ما تقدير قيمة معامل الثبات التجريبي لاختبار التفكير الاستدلالي؟**

أهداف الدراسة

تهدف الدراسة إلى التحقق من جودة الخصائص السيكومترية لاختبار التفكير الاستدلالي في البيئة السعودية في ضوء نظرية الاستجابة للمفردة، مما يساعد الباحثين والمؤسسات التي تهتم بقياس هذه القدرة لدى الطلبة باستخدام هذا الاختبار في أبحاثهم ودراساتهم وكذلك في معرفة مستوى مهارة التفكير الاستدلالي لدى الطلبة بدرجة عالية من الثقة بعد التأكد من جودة الاختبار، ويمكن تحديد أهداف الدراسة في النقاط التالية:

- **التحقق من افتراضات نماذج نظرية الاستجابة للمفردة أحادية البعد في بيانات الدراسة الحالية.**
- **التحقق من مطابقة مفردات اختبار التفكير الاستدلالي للنموذج الثلاثي البارامتر.**
- **التعرف على بارامترات الصعوبة والتمييز والتخمين لمفردات اختبار التفكير الاستدلالي.**
- **التعرف على دالة المعلومات لاختبار التفكير الاستدلالي وللمفردات.**
- **تقدير بارامتر القدرة للأفراد على اختبار التفكير الاستدلالي.**
- **تقدير قيمة معامل الثبات التجريبي لاختبار التفكير الاستدلالي.**

أهمية الدراسة

يمكن توضيح أهمية الدراسية الحالية من جانبين وهما الأهمية النظرية والأهمية التطبيقية كالتالي:

الأهمية النظرية:

- **يتسلط الضوء في الدراسة الحالية على نمط مهم من أنماط التفكير وهو التفكير الاستدلالي الذي يعتبر من الأمور المهمة لنجاح الطلاب في دراستهم وكذلك في تحليلهم وتعاملهم مع العديد من المواقف الحياتية التي تتطلب الربط بين المعطيات والاستنتاج والتوصل لأفضل الحلول.**
- **تأتي الدراسة الحالية استجابة لتوصيات العديد من الدراسات السابقة التي نادت بضرورة التوجه إلى تطبيق نظرية الاستجابة للمفردة في التحقق من جودة الخصائص السيكومترية لأدوات القياس المختلفة وذلك لاتسامها بقدر كبير من الدقة وتحقيقها لموضوعية القياس.**
- **تعتبر الدراسة من الدراسات القليلة – في حدود علم الباحثة – التي اهتمت بدراسة اختبار التفكير الاستدلالي والتحقق من خصائصه السيكومترية بالاعتماد على نظرية الاستجابة للمفردة في البيئة السعودية على وجه التحديد.**

الأهمية التطبيقية:

- **يعد التفكير الاستدلالي من أساليب التفكير المهمة التي ينبغي الاهتمام بدراسته من عدة جوانب، حيث يمكن أن يساعد الطلبة في اتخاذ الكثير من القرارات المتعلقة بحياتهم المستقبلية وبخاصة طلاب المرحلة الثانوية.**
- **يمكن أن تثري الدراسة الحالية المكتبة العربية باختبار يقيس القدرة على التفكير الاستدلالي لطلبة المرحلة الثانوية، يتسم بدرجة عالية من الدقة في خصائصه السيكومترية.**
- **دراسة الخصائص السيكومترية لاختبار التفكير الاستدلالي من خلال نظرية الاستجابة للمفردة يمكن أن يعطي نتائج أكثر دقة وجودة لهذا الاختبار، مما يمنح الباحثين ثقة أكبر في استخدامه مستقبلاً وفي الاعتماد عليه في إجراء المزيد من الدراسات المستقبلية المتعلقة بمتغير التفكير الاستدلالي.**

- **يمكن أن تساعد الدراسة الحالية في تسليط الضوء على نظرية الاستجابة للمفردة وأهميتها، وقد تشجع الكثير من الباحثين على استخدامها وتطبيقها عند التحقق من جودة أدواتهم البحثية ودراسة الخصائص السيكومترية لها.**

حدود الدراسة:

اقتصرت الدراسة على عينة من طلبة المرحلة الثانوية وقد شملت التخصصين العلمي والأدبي والمسار المشترك، وقد تم التطبيق خلال الفصل الدراسي الثالث من العام الدراسي 1443هـ.

مصطلحات الدراسة:

الخصائص السيكومترية: يعرفها علام (2001) بأنها بارامترات الصعوبة والتمييز والتخمين للمفردة ومعاملات الصدق والثبات للاختبار.

ويمكن تعريفها اجرائياً بأنها تقدير لبارامترات الصعوبة والتمييز والتخمين لمفردات الاختبار، وأيضاً تقدير لصدق الاختبار من خلال تحقق افتراضات النموذج المستخدم وملاءمة المفردات له، وكذلك تقدير لمعامل الثبات وذلك في ضوء نظرية الاستجابة للمفردة.

التفكير الاستدلالي: يعرفه حجازين (2014) بأنه عملية عقلية تتطلب مقدراً من المقدمات المعلومة والتي يتم من خلالها التوصل إلى حقيقة جديدة أو استخلاص نتيجة معينة أو حل مشكلة ما أو اتخاذ قرار معين.

ويعرف اجرائياً بأنه القدرة على التفكير بطريقة منطقية استدلالية: لفظياً، عددياً ولفظياً عددياً من خلال مفردات الاختبار المستخدم في الدراسة الحالية.

نظرية الاستجابة للمفردة: يعرفها علام (2001) بأنها نظرية سيكومترية معاصرة تحاول نمذجة العلاقة بين مستوى سمة معينة لدى الفرد، والتي يتم قياسها من خلال اختبار معين، وبين درجات استجابة الفرد على مفردات هذا الاختبار الذي يقيس هذه السمة. وتعتمد هذه النظرية على فرضية أساسية وهي أن القيمة الاحتمالية لاستجابة المفحوص على المفردة الاختبارية تكون دالة في متغيرين رئيسين هما: قدرة الفرد في المتغير المراد قياسه، وخصائص المفردة التي يحاول الفرد الإجابة عنها.

ويمكن تعريفها اجرائياً بأنها نظرية من نظريات القياس الحديثة التي يتم من خلالها تحديد العلاقة بين قدرة الفرد واحتمال اجابته إجابة صحيحة على المفردات، والتي تتسم

بدرجة عالية من موضوعية القياس، حيث يمكن من خلالها تقدير قدرة الفرد بدقة عالية في سمةٍ ما ومن دون الاعتماد على اختبار محدد في قياس هذه السمة، وكذلك تقدير الخصائص السيكومترية للمفردات بدقة عالية ومن دون الاعتماد على أفراد محددين.

النموذج الثلاثي البارامتر: يعرفه (Hambleton & Swaminathan, 1985) بأنه أحد نماذج نظرية الاستجابة للمفردة أحادية البعد، ويقوم على افتراض اختلاف المفردات في بارامترات صعوبتها وتمييزها والحد التقاربي الأدنى.

ويعرف اجرائياً بأنه: أحد نماذج نظرية الاستجابة للمفردة أحادية البعد، ويتم فيه تقدير ثلاثة بارامترات للمفردة وهي بارامتر الصعوبة وبارامتر التمييز، ويتميز بإضافة تقدير لبارامتر التخمين الذي يشير إلى احتمال توصل الأفراد ذوي القدرة المنخفضة جداً إلى الإجابة الصحيحة على المفردة عن طريق التخمين.

الإطار النظري

نظرية الاستجابة للمفردة

يقوم القياس النفسي والتربوي في الأساس على نظريتين رئيسيتين تمثل الأسس النظرية، والخطوات العملية المستخدمة في بناء وتصميم الاختبارات والمقاييس، وطرق تحليل بياناتها، وهما: النظرية التقليدية للقياس (Classical Test Theory) ونظرية الاستجابة للمفردة (Item Response Theory).

ولقد شاع استخدام النظرية التقليدية في القياس لسنوات طويلة خلال القرن الماضي وكان يتم من خلالها التحقق من الخصائص السيكومترية على مستوى المفردات والاختبار. إلا أن الكثير من العلماء في مجال القياس والتقويم وكذلك الباحثين وجدوا أن هناك جوانب كثيرة من القصور في هذه النظرية تحد من التطبيقات الواقعية التي من المفترض أن تتسم بدرجة عالية من الدقة وبأقل نسبة من الأخطاء فقد أشار علام (2001) إلى أن جوانب القصور في نظرية القياس التقليدية تتمثل في أن الخصائص السيكومترية للاختبار تعتمد على كل من خصائص عينة المفحوصين التي يجرى عليها الاختبار، وخصائص عينة المفردات التي تمثل الاختبار، مما يجعل نظام القياس غير مستقر. كما أنها تعتمد على مقارنة أداء المفحوص بأداء أقرانه وفقاً للمجموع الكلي لدرجاتهم في الاختبار، وسواء أجريت هذه المقارنة على أساس الدرجات الخام أو الدرجات المعيارية بأنواعها المختلفة، فإن هذه

الدرجات تخضع لخصائص عينة الأسئلة أو المفردات الاختبارية. ولعل اختلاف هذه الخصائص باختلاف عينة الأفراد المختبرين وعينة المفردات يؤثر في صدق إجراء هذه المقارنات، وبذلك يصعب تعميم نتائج هذه الاختبارات أو الاستفادة العملية منها. (عودة، 2016)

ونظراً لقصور هذه النظرية في بعض الجوانب فقد ظهرت نظرية الاستجابة للمفردة وحاولت التغلب على أوجه القصور فيها، فمنذ بداية السبعينيات من القرن الماضي بدأت نظرية الاستجابة للمفردة تحل محلها في بناء وتطوير الاختبارات، وقد أصبحت تمثل الإطار النظري الأكثر استخداماً في مجال القياس. (علام، 2001؛ Crocker & , 1986 Algina)، كما أنها تحقق موضوعية القياس (علام، 2015؛ Kaplan & Saccuzzo, 2008).

ويمكن الحكم على القياس بالموضوعية إذا كانت نتائج القياس مستقلة عن أداة القياس المستخدمة في التوصل إلى هذه النتائج. وتعد استقلالية القياس بمثابة النقطة المفصلية بين النظرية التقليدية ونظرية الاستجابة للمفردة، وهو ما عبر عنه لورد (1980) بخاصية اللا تغاير(Invariance) وباختصار يمكن القول بأن أحد أهم ميزات نظرية الاستجابة للمفردة هي استقلالية القياس، حيث أن تقدير بارامترات المفردات لا يتغير بتغير الأفراد، ومن ثم يكون ذلك التقدير مستقلاً عن خصائص مجموعة الأفراد التي استخدمت في تقدير تلك البارامترات، كما أن تقدير القدرة لا يتغير بتغير مجموعة المفردات المستخدمة في القياس ما دام أنها مفردات مناسبة لقياس السمة، ويكون تقدير قدرة الفرد مستقلاً عن مجموعة المفردات التي تطبق عليه، وبالتالي يمكن مقارنة النتائج على الرغم من اختلاف الاختبارات التي تقيس تلك القدرة، والحصول على خطأ معياري في تقدير القدرة لتقدير درجة الدقة في قياس القدرة لكل فرد على حدة. (Hambleton & Swaminathan, 1985)

وتعتمد هذه النظرية على فرضية أساسية وهي أن القيمة الاحتمالية لاستجابة المفحوص على المفردة الاختبارية تكون دالة في متغيرين رئيسين هما: المتغير المراد قياسه، وخصائص المفردة التي يحاول المفحوص الإجابة عنها (علام، 2001). ويتطلب تحديد هذه العلاقة تحديد بعض المعلومات عن كل من الفرد والمفردة الاختبارية، فبالنسبة للفرد نحتاج إلى قيمة عددية واحدة هي بارامتر القدرة، وبالنسبة للمفردة نحتاج إلى قيمة عددية واحدة أو

أكثر تمثل البارامترات المطلوبة للمفردة والمتمثلة في الصعوبة والتمييز وأحياناً التخمين أو ما يسمى ببارامتر شبه الصدفة. (الوليلي والدوسري، 2019)

وقد تعددت النماذج المتعلقة بهذه النظرية وتباينت أشكال المنحنيات المميزة للمفردات تبعاً لاختلاف عدد البارامترات المتطلبة لكل نموذج (Hambleton, Swaminathan & Rogers, 1991). إلا أن استخدام هذه النماذج يتطلب أولاً التحقق من عدة افتراضات تتطلبها النظرية ومن دون تحققها لا يمكن استخدام نماذج نظرية الاستجابة للمفردة، وتتمثل هذه الافتراضات في:

أحادية البعد (Uni-dimensionality) وتعني أحادية البعد أن هناك قدرة واحدة تفسر أداء المفحوص في الاختبار، كما تعني وجود عامل واحد أو مكون واحد مسيطر يكمن وراء الأداء في الاختبار وهذا المكون هو القدرة التي يسعى الاختبار إلى قياسها. والاستقلال الموضعي (Local Independence) ويعني أن إجابة الفرد على إحدى مفردات المقياس مستقلة إحصائياً عن إجابته على بقية المفردات أو على استجابته على أي مفردة أخرى من مفردات المقياس ولا يتأثر أداؤه سلباً أو إيجاباً بأدائه على مفردات أخرى (Hambleton, Swaminathan & Rogers, 1991). وتحقق منحنى خصائص المفردة (Item Characteristic Curve) وهو دالة انحدار غير خطية تربط ما بين احتمالية الاستجابة على المفردة استجابة صحيحة وقدرة المفحوص التي تقيسها مجموعة من المفردات. ومقدار الاحتمال يكون دالة متزايدة لموقع الأفراد على متصل السمة، وهذا يعني أن احتمال الاستجابة الصحيحة للفرد يزداد بازدياد مقدار السمة أو السمات الكامنة لديه (علام، 2005). بالإضافة إلى افتراض التحرر من السرعة ويعني أن الاختبار لم يتم تطبيقه تحت عامل السرعة بمعنى أن إخفاق المفحوص في الإجابة على المفردة يعود لانخفاض قدرة المفحوص، وليس إلى عدم قدرته على الوصول إلى المفردة بسبب الوقت غير الكافي للاختبار. (Hambleton, Swaminathan & Rogers, 1991)

وبعد التحقق من توفر هذه الافتراضات يمكن استخدام أيٍ من نماذج نظرية الاستجابة للمفردة، ويوجد ثلاثة نماذج أساسية شائعة الاستخدام وهي: النموذج الأحادي البارامتر الذي يفترض أن تختلف المفردات في بارامترات صعوبتها فقط، وأن بارامترات التمييز

ثابته لجميع المفردات، وأن الفرد لا يلجأ إلى التخمين العشوائي في إجابته. وتتمثل الصيغة الرياضية لهذا النموذج في التالي:

$$P_i(\Theta) = \frac{e^{\theta j - bi}}{1 + e^{\theta j - bi}} \quad \ldots\ldots\ldots\ldots$$

حيث: Pi (Θ): هو احتمال أن يجيب فرد قدرته Θ إجابة صحيحة عن المفردة i.

bi: بارامتر صعوبة المفردة. Θ: قدرة الفرد.

e: الأساس اللوغاريتمي الطبيعي، ويساوي 2.718.

والنموذج الثنائي البارامتر الذي يختلف عن النموذج الأحادي في أنه يسمح للمفردات بأن تختلف في كلٍ من بارامتري الصعوبة والتمييز. وتتمثل الصيغة الرياضية لهذا النموذج في التالي:

$$Pi(\Theta) = \frac{e^{ai(\theta j - bi)}}{1 + e^{ai(\theta j - bi)}}$$

حيث: ai: بارامتر تمييز المفردة.

والنموذج الثلاثي البارامتر الذي أضاف هذا النموذج بارامتراً ثالثاً وأطلق عليه بارامتر الخط التقاربي الأدنى أو بارامتر التخمين فيما يتعلق بالمفردات الاختبارية التي تتطلب الاختيار بين بدائل معطاة. وهذا البارامتر يحدد احتمال أن يجيب فرد يفترض أن مستوى قدرته منخفضاً انخفاضاً لا نهائياً ومع هذا يجيب إجابة صحيحة عن غالبية مفردات الاختبار عن طريق التخمين، لذلك تبدو أهمية هذا البارامتر عند مطابقة بيانات مستمدة من مفردات اختبارية من نوع الاختيار من متعدد لهذا النموذج. والصيغة الرياضية للنموذج ثلاثي البارامتر كالتالي:

$$Pi(\Theta) = ci + (1 - ci)\frac{e^{ai(\theta j - bi)}}{1 + e^{ai(\theta j - bi)}}$$

حيث: ci: بارامتر شبه الصدفة (التخمين). (علام، 2005؛ التقي، 2013؛ الكيلاني والبرصان، 2017)

التفكير الاستدلالي

إن التوجهات الحديثة في المجالات التربوية تؤكد على ضرورة الاهتمام بمهارات التفكير المختلفة كالتفكير الناقد والتفكير المنظومي والتفكير الاستدلالي وذلك لما لها من أثر كبير على المخرجات التعليمية حيث إنها تدرب الأفراد على أسلوب حل المشكلات من خلال

مهارات التفكير الاستدلالي وذلك بتحليل المواقف واستخدام المهارات المكتسبة والمعارف السابقة.

ويعرف جروان (2016) التفكير الاستدلالي بأنه عملية تفكيرية تتضمن وضع الحقائق أو المعلومات بطريقة منظمة بحيث تؤدي إلى استنتاج أو اتخاذ قرار أو حل المشكلة. ويرى رزوقي وعبدالكريم (2015) أن التفكير الاستدلالي يعبر عن قدرة الفرد على تطبيق عمليات عقلية تعتمد على الاستقراء والاستنباط والاستنتاج من أجل كشف الغموض عن شيء محدد والتعرف عليه للوصول إلى حل المشكلات المتعلقة به. وأشار العبيدي والعبيدي وولي (2015) إلى أن التفكير الاستدلالي حظي باهتمام كبير من بين أنماط التفكير الأخرى وذلك لكونه أرقى الأنشطة العقلية للإنسان الذي يتميز بإدراكه للعلاقات القائمة بين الأشياء، ويميز الاختلافات التي بينها باستعمال الرموز الذهنية والمعاني والخبرات السابقة الموجودة لدى الفرد وبذلك يتم فيه الانتقال من المعلوم إلى المجهول. ومن خلال التعريفات السابقة يمكن القول بأن التفكير الاستدلالي يعبر عن قدرة الفرد على تحليل المواقف المختلفة وحل المشكلات المتعلقة بها من خلال الاعتماد على معلومات وتجارب وخبرات سابقة لديه وذلك بإعمال الحجج والبراهين. والتفكير الاستدلالي يقوم على مقدمة يستدل بها من أجل التوصل إلى نتيجة بشرط أن تكون هناك علاقة منطقية بين المقدمة والنتيجة من أجل اتخاذ القرارات المناسبة لحل المشكلات وتوليد معارف جديدة. (مراد، 2013)

وذكر كل من رزوقي وعبدالكريم (2015) وأبو زيد (2002) أن للتفكير الاستدلالي العديد من الخصائص التي تميزه عن غيره من أنماط التفكير ويمكن تلخيصها في أن التفكير الاستدلالي يعد من العمليات العقلية التي تعتمد على المنطق للتوصل من المقدمات أو المعلومات السابقة إلى نتائج وعلاقات منطقية لاستنباط الفروض والحلول من أجل تكوين مفاهيم معينه مع ربط السبب بالنتيجة بطريقة ما. وذلك من أجل الوصول إلى التعميمات التي توضح العلاقات بين المدخلات والمخرجات، كما أنه يعتبر من العمليات العقلية العليا التي ترتبط بشكل كبير بالذكاء، فهي تقوم على الانتقال من المعلوم إلى المجهول، بالإضافة إلى أن التفكير الاستدلالي يعتبر تفكير متطور ونمائي يختلف في درجته ومستوياته حسب اختلاف المراحل العمرية وقابل للتغير تبعاً لمراحل النمو والخبرات السابقة للأفراد.

وهناك ثلاثة أنماط رئيسية للتفكير الاستدلالي وهي الاستدلال الاستنباطي: ويعبر عن العمليات العقلية التي يستخلص من خلالها الفرد الحالات الخاصة من الحالات العامة المسلم بها. وتركز هذه الطريقة على عرض القاعدة، أو المفهوم أولاً ثم تضع بعدها الشروح والأمثلة والتطبيقات. بمعنى أنه يبدأ من قاعدة كلية ليصل إلى نتيجة تنطبق على الأمثلة الجديدة فهو انتقال العقل من قواعد وأحكام عامة مسلم بصحتها إلى حكم خاص (المهداوي، 2013). والاستدلال الاستقرائي وهذه الطريقة تتبع شرح الدرس تفصيلياً وتطبيق أمثلته وفهم المعلومات التي يحتويها ثم تستخرج منه ملخص لهذه المعلومات وتتوصل إلى القاعدة العامة، فهو بذلك نشاط فكري وعقلي يقوم على معالجة البيانات والمعارف والمعلومات وتطبيقها وتجريبها ومن ثم استخراج القاعدة العامة منها (المنصور، 2010). والاستدلال الاستنتاجي: هو الأداء المعرفي العقلي المنطقي الذي ينتقل فيه التفكير من المعرفة المكتسبة إلى معارف مجهولة مع وجود بيانات متوفرة مسبقاً، ويقوم بربط المعارف والمعلومات السابقة التي اكتسبها مع المواقف الجديدة لفهمها وتفسيرها. (محمد، 2013)

دراسات سابقة

تم الاطلاع على عدد من الدراسات السابقة ذات الصلة بموضوع الدراسة الحالية ومن أهمها ما يلي:

دراسة المصاروة (2008) التي هدفت إلى التعرف على دلالات صدق وثبات اختبار التفكير الاستدلالي لطلبة المرحلة الثانوية في دولة الإمارات العربية المتحدة. وتكونت العينة من (1149) طالبا وطالبة 589 طالباً و560 طالبة تم اختيارهم بطريقة عشوائية من مدارس منطقة الشارقة التعليمية. وقد أعد الباحث اختباراً للتفكير الاستدلالي يتكون من خمسين مفردة من نوع الاختيار من متعدد تتوزع على خمس مجالات رئيسة هي: الاستدلال العددي؛ ويشتمل على (13) سؤال تم توزيعها على بعدين فرعيين هما: المتتاليات أو الأنساق العددية والمتقابلات العددية أو الأشكال. الاستدلال اللفظي: ويشتمل على (15) سؤال تم توزيعها على ثلاثة أبعاد فرعية هي: علاقات بين مفاهيم الألفاظ ومعانيها المرادفة والمتقابلات اللفظية وعلاقات أفعل التفضيل (تقع ضمن الاستدلال الاستنباطي). الاستدلال العددي اللفظي: ويشتمل على (9) أسئلة. الاستنباط العام ويشتمل على (11) سؤال. والاستقراء العام ويشتمل على سؤالين. وقد تم حساب ثبات الاختبار من خلال معامل ألفا

كرونباخ حيث بلغت قيمته 0.87. وكذلك من خلال التجزئة النصفية وبلغت قيمته 0.93. وقد أظهرت نتائج الدراسة أن معاملات الصعوبة تراوحت بين (0.34 – 0.79) وهي معاملات صعوبة مقبولة. كما تراوحت معاملات التمييز بين (0.24 – 0.79) وهي معاملات تمييز مقبولة. كما أن معاملات الارتباط بين أداء الطلبة على الاختبار وتحصيلهم الدراسي في عدد من المقررات تراوحت بين (0.74 – 0.96) وجميعها دالة إحصائيا عند مستوى الدلالة (0.01). وكان أبرزها للطلبة الذكور من الصف الحادي عشر العلمي، وأدناها للطالبات الإناث من الصف الحادي عشر الأدبي.

ودراسة المنشدي (2019) التي قام فيها ببناء اختبار لقياس القدرة الاستدلالية لدى الطلبة تكون في صورته النهائية من (40) مفردة، تم تطبيق الاختبار على عينة مكونة من (200) طالب وطالبة، تم اختيارهم بطريقة عشوائية. وقد توصلت الدراسة إلى أن اختبار القدرة الاستدلالية يتمتع بخصائص سيكومترية جيدة، حيث تم التحقق من صدق وثبات الاختبار باستخدام الأساليب الإحصائية وفقاً لنظرية القياس التقليدية كما تم استخراج صدق وتمييز المفردات. وأشارت النتائج إلى أنه يوجد مستوى عالي من القدرة الاستدلالية لدى عينة الدراسة. وقد أوصت الدراسة بإمكانية استخدام الاختبار من قبل المرشدين التربويين في تشخيص مستوى القدرة الاستدلالية لدى الطلبة بالمدارس الإعدادية نظراً لجودته من الناحية السيكومترية.

وسعت دراسة عيال وجاسم (2019) إلى تطوير اختبار مهارات الاستدلال الاستقرائي لدى طلبة الجامعة في العراق. ولتحقيق هذا الهدف، استخدم الباحثان اختبار لويولا الذي وضعه Charles Doyle والذي يعد اختباراً مناسباً لقياس القدرة على الاستدلال الاستقرائي لدى طلبة جامعة. يتكون الاختبار من 50 مفردة، وقد تم عرض الاختبار على الخبراء والمحكمين في العلوم التربوية والنفسية، للتأكد من توافر الخصائص المناسبة له ولبدائله من حيث الشكل والمضمون، وأخذ الباحثان بآرائهم. كانت مفردات الاختبار جميعها صالحة منطقياً لقياس ما وضعت لقياسه، وقد تم تطبيق الاختبار على عينة مكونة من (1000) طالبا وطالبة من الجامعات في (بغداد) اختيروا بأسلوب العينة العشوائية متعددة المراحل. واعتمد الباحثان على النموذج ثلاثي البارامتر، وهو أحد نماذج نظرية الاستجابة للمفردة. وقد تأكد الباحثان من تحقق افتراضات النموذج المتمثلة في أحادية البعد من خلال إجراء التحليل

العاملي بطريقة المكونات الأساسية وعلاقة المفردة بالدرجة الكلية، فضلاً عن مطابقة المفردات للنموذج ثلاثي البارامتر. وقد تم قبول صعوبة المفردة ضمن المعيار (2.5, -2.5)، وكان معيار قبول تمييز المفردة بين (0.5، 2.5)، ومعيار التخمين لا يزيد عن (0.25) وتبين أن جميع المفردات كانت ضمن الحدود المقبولة، وعليه لم تحذف أي مفردة. كما تم التحقق من صدق الاختبار وثباته عن طريق مؤشرين هما: نسبة التباين لتقدير ثبات الاختبار ودالة المعلومات لتقدير ثبات المفردة.

وهدفت الدراسة التي أجراها باسيرير BASERER (2020) إلى تحديد مستوى التفكير المنطقي لدى الطلاب المتدربين على مهنة التعليم من التخصصات المختلفة. ولهذا الغرض، تم تطبيق مقياس التفكير المنطقي الذي طوره الباحث على 525 طالبًا في كلية التربية بجامعة أتاتورك. تكون الاختبار من 25مفردة وقد تم التحقق من خصائصه السيكومترية وفق النظرية التقليدية للقياس حيث تم التحقق من الصدق من خلال إجراء التحليل العاملي الاستكشافي والتوكيدي وقد أكدت النتائج على أن المفردات تشبعت وبقيم مرتفعة على عامل واحد وهو التفكير المنطقي. كما تم حساب معامل ثبات ألفا كرونباخ وقد بلغ 0.83. تم تحديد مجموعة من المعلومات الشخصية للمشاركين كالنوع والتخصص الأكاديمي (نظري أو تطبيقي) وتكرار قراءة الكتب (لمعرفة ما إذا كانت مهارات التفكير تتحسن مع تكرار القراءة) وتم إيجاد العلاقة بينها وبين مجموع الدرجات التي حصلوا عليها في اختبار التفكير المنطقي. أشارت النتائج إلى أن الطلاب الذين يقرؤون الكتب كل يوم لديهم مستويات أعلى من التفكير المنطقي. كما أشارت النتائج إلى أن الطلاب الذين يدرسون في العلوم التربوية وتعليم العلوم التركية والاجتماعية يتمتعون بمستويات تفكير منطقي أعلى من الطلاب الذين يدرسون في الأقسام الأخرى.

كما هدفت دراسة كل من فيرولا، تالمان، هوبلي، لاكونين، انجبلوم وهافيستو Vierula, Talman, Hupli, Laakkonen, Engblom and Haavisto (2021) إلى تطوير اختبار يقيس مهارات التفكير الاستدلالي، واختباره من الناحية السيكومترية لتقييم مهارات التفكير لدى المتقدمين لكلية التمريض في مرحلة البكالوريوس وذلك لاختيار الأفضل منهم. وقد تم تطوير الاختبار كجزء من اختبارات القبول الإلكترونية المتعددة. تضمن الاختبار وصف حالة تبعها ثلاثة أقسام من الأسئلة لتقييم مهارات التفكير الاستدلالي لدى المتقدمين.

تم استخدام نظرية الاستجابة للمفردة في دراسة الخصائص السيكومترية للاختبار وذلك لتحديد بارامترات الصعوبة والتمييز. وقد تم تطبيق الاختبار على عينة تكونت من 1056 طالباً من المتقدمين لكليات التمريض في ست جامعات فنلندية (28 مايو 2019). تم التحقق من صدق محتوى الاختبار من خلال عرضه على مجموعة من المحكمين الذين أبدو موافقتهم عليه. وقد أشارت النتائج إلى أن معاملات التمييز كانت مقبولة وأن صعوبة المفردات كانت منخفضة في مجملها مما يعني أن الاختبار كان سهلاً نوعاً ما. وأوصت الدراسة بإمكانية إعادة النظر في كتابة المشتتات مما قد يحسن من معاملات صعوبة المفردات ويحدث تباين بينها. إلا أن الباحثين أشاروا أخيراً بأن الاختبار يعد موضوعياً ويقيس بشكل جيد مهارات التفكير الاستدلالي لدى المتقدمين على كلية التمريض ويمكن الوثوق في النتائج المستمدة منه.

كما أجرى بولات وامير Polat and Emre (2021) دراسة هدفت إلى تطوير اختبار مهارات التفكير المنطقي الاستدلالي لمعلمي العلوم المتدربين. وقد اعتمدت الدراسة على المنهج المختلط في إجرائها. تم بناء الاختبار بالاعتماد على اختبار لوسون للتفكير المنطقي الاستدلالي في العلوم (1978). وقد تم بدايةً بناء عدد كبير من الأسئلة في موضوعات مقرر تطبيقات الكيمياء العامة في المختبرات وقد تم التركيز على ست موضوعات فرعية من المقرر. قام الباحثان بالتحقق من صدق محتوى الاختبار من خلال عرضه على مجموعة من الخبراء والمتخصصين في المجال، وتم إجراء التعديلات اللازمة على الاختبار. ثم تم إعداد نسخة تجريبية للاختبار تكونت من 13 سؤال من نوع الاختيار من متعدد وسؤالين مفتوحين. وتتطلب الإجابة على أسئلة الاختيار من متعدد خطوتين؛ الخطوة الأولى يقوم الطالب باختيار الإجابة الصحيحة على المفردة (أسئلة المجموعة أ). وفي الخطوة الثانية يقوم الطالب باختيار سبب الإجابة التي اختارها أولاً (أسئلة المجموعة ب). تم تطبيق نسخة الاختبار التجريبية على 142 متدرباً من معلمي العلوم. تم تحليل أسئلة الاختبار، وبعد نهاية التحليل تم حذف المفردة رقم 12 من الاختبار لظهور مشكلات سيكومترية بها حيث إن معاملات الصعوبة والتمييز لها لم تكن ضمن المدى المقبول. أشارت نتائج التحليل أن متوسط مؤشرات التمييز لباقي المفردات في المجموعتين أ و ب كانت مرتفعة ومعاملات الصعوبة كانت جميعها ضمن المدى المقبول. وقد بلغ معامل ثبات ألفا كرونباخ للاختبار0.96. وبناء على هذه النتائج

أوصى الباحثان بإمكانية استخدام اختبار قدرات التفكير المنطقي الاستدلالي وتطبيقه على معلمي العلوم المتدربين حيث إنه يتمتع بخصائص سيكومترية جيدة من صدق وثبات.

في حين هدف بحث الحجيلي (2022) إلى الكشف عن البناء العاملي لاختبار الاستدلال المنطقي لدى طالبات المرحلة الثانوية بمنطقة المدينة المنورة، وحساب خصائصه السيكومترية. وتحقيقاً لذلك تم بناء اختبار لقياس الاستدلال المنطقي يتكون من (30) مفردة، وتم تطبيقه على عينة استطلاعية تكوّنت من (100) طالبة؛ بهدف التحقق من صلاحية الاختبار الذي تم إعداده، ومن ثم تم تطبيقه على أفراد عينة البحث الأساسية البالغ عددها (728) طالبة من طالبات المرحلة الثانوية بمنطقة المدينة المنورة في الفصل الدراسي الثاني للعام الدراسي 1442هـ تم اختيارهن عشوائيًّا. واتبع البحث المنهج الوصفي التحليلي، كما تم التحقق من الخصائص السيكومترية للاختبار باستخدام بعض الأساليب الإحصائية، منها: التحليل العاملي الاستكشافي باستخدام طريقة المكوّنات الرئيسية مع التدوير المتعامد للمحاور؛ للكشف عن البناء العاملي للاختبار، ومعامل ارتباط (بيرسون)، ومعامل ألفا كرونباخ، ومعامل كيودر ريتشاردسون (20). وأشارت النتائج إلى أن الاختبار ذو بنية عاملية ثلاثية تكونت من العوامل الآتية: (العمليات الحسابية– الاستعداد الحسابي– الاستعداد الرياضي)، وقد فسّرت ما مقداره (40%) من التباين الكلي للأداء على الاختبار، كما أشارت النتائج إلى تمتع الاختبار بخصائص سيكومترية مناسبة تسمح باستخدامه كأداة لقياس الاستدلال المنطقي لدى الفئة المقصودة بالبحث. وقد أوصى البحث باعتماد هذا الاختبار للحكم على مدى امتلاك الطلاب في المرحلة الثانوية لمهارات الاستدلال المنطقي.

وقد أجرى كل من باسزتور، ماجيار، باسزتوركوفاش وراوش Pásztor, Magyar, Pásztor–Kovács and Rausch (2022) دراسة هدفت إلى تطوير أداة لتقييم مجال الاستدلال الاستقرائي من خلال الحاسب الآلي. كما هدفت إلى تطوير لعبة على الإنترنت لتعزيز الاستدلال الاستقرائي من خلال المحتوى الرياضي فيها والتحقق من فعاليتها. تم التطبيق على عينة من طلاب الصف الخامس تكونت من 267 طالباً في المجموعتين الضابطة والتجريبية، وقد أجريت الدراسة التجريبية على مجموعة تكونت من 122 طالباً. وقد قام الباحثون ببناء اختبار الكتروني تكون من 54 مفردة تمثل مهارات الاستدلال الاستقرائي. تم تحليل الاختبار في ضوء نظرية القياس التقليدية، وقد أشارت النتائج إلى أن الاختبار

يتمتع بخصائص سيكومترية جيدة فقد بلغ معامل ثبات ألفا كرونباخ 0.90، كما أن معاملات التمييز والصعوبة لجميع المفردات كانت ضمن المدى المقبول. وقد كان ارتباط جميع المفردات بالدرجة الكلية على الاختبار ارتباطاً موجباً. كما كشف التحليل العاملي التوكيدي إلى أن الاختبار يتكون من بعد واحد وذلك لوجود عامل واحد مسيطر تشبعت عليه جميع المفردات. وأكد الباحثون على جودة الاختبار من الناحية السيكومترية وإمكانية استخدامه والوثوق في نتائجه لتقييم مهارة الاستدلال الاستقرائي للطلاب.

كما قام قومريه ودارمايانتي Qomariyah and Darmayanti (2023) بإجراء دراسة هدفت إلى تطوير أداة لقياس مهارة التفكير الرياضي الاستدلالي في تعلم الرياضيات لطلاب المدارس الثانوية. تكون مجتمع الدراسة من أربع مدارس في مدينة باسوروان بأندونيسيا. وقد تم التطبيق على عينة من طلبة الصف الثاني عشر. تم جمع البيانات باستخدام اختبار التفكير الرياضي الاستدلالي المكون من ثلاثة أبعاد من خلال ثلاثة أسئلة وصفية. وقد تم تحليل البيانات والتحقق من صدق محتوى الاختبار من خلال عرضه على مجموعة من المحكمين الذين كانت نسبة اتفاقهم مرتفعة على جودة محتوى الاختبار. كما تم حساب معامل ثبات ألفا كرونباخ للاختبار وقد بلغ 0.743، وكذلك تم استخراج معاملات الصعوبة والتمييز لأبعاد الاختبار. أشارت نتائج الدراسة إلى أن أداة التقييم تقيس قدرات التفكير الرياضي الاستدلالي لطلاب المدارس الثانوية وقد تم تطويرها بشكل جيد، وقد أوصت الدراسة باستخدامها وتطبيقها مستقبلاً لقياس مهارة التفكير الرياضي الاستدلالي للطلاب.

من خلال استعراض الدراسات السابقة يتضح اهتمام الباحثين بدراسة مهارة التفكير الاستدلالي من عدة جوانب، فقد سعى بعضهم إلى بناء اختبارات تقيس هذه المهارة مثل دراسة كل من (المصاروة، 2008؛ المنشدي، 2019؛ فيرولا وآخرون، 2021؛ بولات وامير، 2021؛ الحجيلي،2022؛ باسزتور وآخرون، 2022 و قومريه ودارمايانتي، 2023)، وركز البعض الآخر على معرفة مستواها لدى الطلبة أو مدى توفرها في موضوعات المقررات الدراسية وعلاقتها بعدد من المتغيرات كدراسة (باسيرير، 2020). ويلاحظ أن أغلب هذه الدراسات عالجت الأدوات المستخدمة وقامت بتحليلها والتحقق من الخصائص السيكومترية لها في ضوء نظرية القياس التقليدية ماعدا دراسة (فيرولا وآخرون، 2021) التي استخدمت نظرية الاستجابة للمفردة التي أثبتت جودتها وتفوقها على النظرية التقليدية

بالإضافة إلى دراسة (عيال وجاسم، 2019) التي تم تدريج مفردات الاختبار بها باستخدام النموذج الثلاثي البارامتر، إلا أنها تختلف مع الدراسة الحالية في تركيزها فقط على الاستدلال الاستقرائي كنوع من مهارات التفكير الاستدلالي كما أنها اعتمدت طلبة المرحلة الجامعية كعينة لإجراء الدراسة. في حين تستخدم الدراسة الحالية اختباراً مختلفاً في قياس التفكير الاستدلالي ويشمل عدة أنواع من الاستدلال، كما أنها تهتم بدراسة هذه المهارة لدى طلبة المرحلة الثانوية والذين قد يفيدهم الكشف عن هذه المهارة لديهم وتطويرها في توجهاتهم المستقبلية.

منهجية الدراسة واجراءاتها:

منهج الدراسة: **تم استخدام المنهج الوصفي وذلك لملائمته للدراسة وللإجابة على أسئلتها.**

عينة الدراسة: **اختارت الباحثة عينة الدراسة من طلبة المرحلة الثانوية شملت التخصصات العلمية والأدبية والمسار المشترك، وقد تم التطبيق في العام الدراسي 1443هـ وذلك باستخدام الطريقة العشوائية الطبقية وقد بلغ حجم العينة (2023) منهم (1081) طالبة و(942) طالباً والجدول (1) يبين توزيع أفراد العينة وفقاً لمتغيري النوع والتخصص الدراسي.**

جدول1: توزيع أفراد العينة حسب النوع والتخصص

التخصص	النوع		المجموع
	ذكور	إناث	
علمي	613	513	1126
أدبي	290	473	763
مسار مشترك	39	95	134
المجموع	942	1081	2023

أداة الدراسة:

الأداة المستخدمة في هذه الدراسة هي اختبار أيسر للقدرة الاستدلالية (ACER TEST OF REASONING_ABILITY, 1990) ويتكون الاختبار من (75) مفردة، وأشارت نتائج تقنين هذا الاختبار في أستراليا عام (1978) إلى أن الصدق التلازمي للاختبار تراوح بين (0.82 و 0.9). وأن ثبات الاختبار KR – 20 كان أعلى من (0.91) لعينة سن (15) سنة و (0.88) لعينة سنة أولى كليات (ACER, 1983). وعند تحويل المفردات الـ (75) من صورة أسئلة مفتوحة إلى صورة اختيار من متعدد عام (1986) تم تحويل (70) مفردة فقط وتعذر تحويل المفردات الخمس الباقية، وبذلك أصبحت المفردات

السبعين تمثل اختبار أيسر للقدرة الاستدلالية بصورته النهائية. وقد تم ترجمة وتقنين هذا الاختبار على عينة من طلبة الصف الأول الثانوي في الأردن من قِبل حجازين (2003).

تصحيح الاختبار

تصحح الإجابات على الاختبار آلياً أو يدوياً وذلك باستخدام مفتاح للتصحيح بحيث تعطى درجة واحدة للإجابة الصحيحة، وتعطى درجة صفر للإجابة الخاطئة، وبما أن الاختبار يتكون من (70) مفردة من نوع الاختيار من متعدد، فإن النهاية العظمى للاختبار تكون (70) والنهاية الصغرى صفر.

الخصائص السيكومترية للاختبار

صدق وثبات الاختبار في البيئة الأسترالية

الصدق: **تم تحديد الصدق التلازمي لاختبار أيسر للقدرة الاستدلالية في نسخته الأصلية بإيجاد الارتباطات بين أداء الأفراد على اختبار أيسر بأدائهم على اختبارات أخرى وهي: اختبار الرياضيات واختبار اللغة الإنجليزية ومقياس الاتجاه نحو المدرسة، وقد جاءت معاملات الارتباطات متوسطة وتراوحت ما بين 0.50 و 0.58 و 0.61.**

الثبات: **تم تقدير الثبات لاختبار أيسر للقدرة الاستدلالية باستخدام معامل الثبات لكيودر ريتشاردسون (KR – 21) للصفوف التاسع والعاشر والحادي عشر. وقد تراوحت معاملات الثبات للصفوف الثلاث بين (0.85) و (0.89) وتشير إلى معاملات ثبات مقبولة.**

صدق وثبات الاختبار في البيئة الأردنية

الصدق: **تم التحقق من صدق الاختبار في النسخة العربية من خلال صدق المحكمين وكذلك من خلال الصدق المرتبط بالمحك. وقد تم إيجاد معاملات الارتباط بين درجات الطلبة على الاختبار مع درجاتهم في كلٍ من مبحث الرياضيات ومبحث اللغة العربية ومبحث الرياضيات واللغة العربية معاً، وقد أشارت النتائج إلى أن قيم معاملات ارتباط بيرسون كانت (0.66 ، 0.62 ، 0.65) على التوالي وهي قيم متوسطة وتعني أن اختبار أيسر للقدرة الاستدلالية يقيس القدرة الاستدلالية اللفظية والقدرة الاستدلالية العددية.**

الثبات: **تم حساب معامل ثبات الانساق الداخلي باستخدام معادلة KR – 20 حيث بلغت قيمته (0.79)، وكذلك تم حساب معامل الثبات بطريقة التجزئة النصفية، وبلغت قيمة المعامل المصحح بمعادلة سبيرمان – براون (0.7)**

صدق وثبات الاختبار في الدراسة الحالية

الصدق: **تم التحقق من الصدق في الدراسة الحالية من خلال صدق المحكمين وذلك من خلال عرض الاختبار على مجموعة من أعضاء هيئة التدريس المتخصصين في علم النفس التربوي والمناهج وطرق التدريس لإبداء آراءهم في ملائمة المفردات لقياس السمة موضع الدراسة وكذلك وضوح المفردات وصحة الصياغة اللغوية لها. وقد تم إجراء بعض التعديلات بناءً على ملاحظاتهم. وقد أبدى المحكمون موافقتهم على ملائمة الاختبار للتطبيق على الطلبة في السعودية ووضوح مفرداته واتساقها وسلامتها اللغوية. كما تم حساب الصدق التلازمي للاختبار مع درجات الطلاب في اختبار القدرات العامة والذي تشير هيئة تقويم التعليم على أنه يقيس القدرة التحليلية والاستدلالية لدى الطالب ومعرفة قابليته للتعلم (موقع هيئة تقويم التعليم الالكتروني، 2021) من خلال حساب معامل ارتباط بيرسون بين الدرجة الكلية للطلبة في اختبار التفكير الاستدلالي ودرجاتهم في اختبار القدرات العامة، وقد كانت قيمة معامل الارتباط **0.726 وهي قيمة دالة احصائياً عند مستوى 0.001، وتدل على صدق المحك التلازمي بين درجات الطلبة في اختبار التفكير الاستدلالي ودرجاتهم في اختبار القدرات العامة.**

الثبات: **تم حساب معامل الثبات مبدئياً باستخدام معامل كيودر ريتشاردسون كخطوة أولية للتحقق من ثبات الاختبار، وقد أشارت النتائج إلى أن معامل ثبات كيودر ريتشاردسون يعتبر مقبولاً حيث بلغ 0.903.**

إجراءات الدراسة

- **تم تطبيق الاختبار على الطلبة في الفصل الدراسي الثالث للعام الدراسي 1443هـ.**
- **تم تصحيح إجابة الطلبة بإعطاء درجة (1) للإجابة الصحيحة والدرجة (0) للإجابة الخاطئة.**
- **تم تفريغ استجابات المفحوصين في ملفات خاصة ببرنامج الـ SPSS وبرنامج (BILOG – MG) وذلك لعمل التحليلات الإحصائية اللازمة للإجابة على أسئلة الدراسة.**

نتائج الدراسة ومناقشتها

أولاً: نتائج السؤال الأول ونصه: هل تتحقق افتراضات نماذج نظرية الاستجابة للمفردة أحادية البعد في بيانات الدراسة الحالية؟

تم التأكد من تحقق افتراضات نماذج نظرية الاستجابة للمفردة أحادية البعد في بيانات الدراسة الحالية من خلال الخطوات التالية:

التحقق من افتراض أحادية البعد

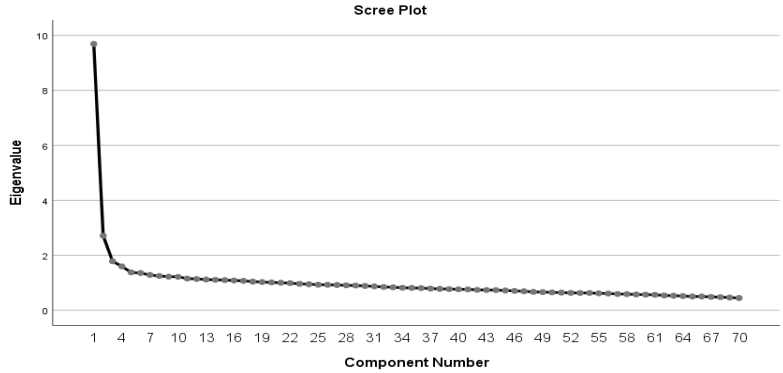
قبل إجراء التحليل العاملي على بيانات الدراسة تم التحقق من المؤشرات التي تسمح باستخدام أسلوب التحليل العاملي والاعتماد على نتائجه، فقد تم التحقق من قابلية مصفوفة معاملات الارتباط للتحليل العاملي حيث يجب أن تكون القيمة المطلقة لها لا تساوي الصفر ولابد أن تكون أكبر من (0.00001) وأشارت نتائج التحليل العاملي للاختبار أن قيمة محدد مصفوفة الارتباط بلغ (0.00106) وهي قيمة أكبر من القيمة المطلقة لمحدد مصفوفة الارتباط. كما تم التحقق من ملاءمة وكفاية حجم العينة وذلك من خلال قيمة اختبار Kaiser–Meyer–Olkin (KMO–Test) لكفاية حجم العينة والتي بلغت (0.917) وهي بحسب محك كايزر تعد قيمة مقبولة حيث إنها أكبر من (0.5). كما تم أيضاً إجراء اختبار بارتليت للتحقق من تجانس العينة واختبار فرضية عدم تماثل مصفوفة الارتباط وقد بلغت قيمته (22989.408) وهي دالة احصائياً عند مستوى أقل من (0.01) وتشير هذه الدلالة إلى أن مصفوفة الارتباطات ليست مصفوفة الوحدة وبالتالي يمكن إجراء التحليل العاملي.

بعد التحقق من ملاءمة البيانات، تم التحقق من افتراض أحادية البعد من خلال إجراء التحليل العاملي لبيانات الدراسة الحالية وذلك لمعرفة عدد العوامل التي يتكون منها الاختبار. ويوضح الجدول (2) نتائج التحليل العاملي لبيانات الدراسة.

جدول 2: قيم الجذور الكامنة ونسبة التباين المفسر والنسبة التراكمية للتباين المفسر لها

رقم العامل	قيمة الجذر الكامن	نسبة التباين المفسر	النسبة التراكمية للتباين المفسر
1	9.692	13.845	13.845
2	2.713	3.875	17.721
3	1.784	2.549	20.270
النسبة بين قيمة الجذر الكامن للعامل الأول إلى قيمة الجذر الكامن للعامل الثاني		3.573	

يتضح من الجدول (2) تحقق افتراض أحادية البعد وفقاً لمحك ريكاس (Reckase, 1979) حيث إن قيمة الجذر الكامن للعامل الأول أكبر من ضعف قيمة الجذر الكامن للعامل الثاني، وأن النسبة بينهما أكبر من 2 حيث بلغت 3.573 وهذا يعني أن هناك عامل واحد مسيطر يعود إليه التباين بين إجابات المختبرين وأن الاختبار يقيس سمة واحدة. ويتضح ذلك أيضاً من خلال التمثيل البياني Scree Plot لقيم الجذور الكامنة لبيانات الاختبار كما هو موضح في الشكل (1).



شكل1: التمثيل البياني لقيم الجذور الكامنة

ويتضح من الرسم البياني وجود انحدار واضح في التمثيل البياني لقيم الجذور الكامنة التي تزيد عن الواحد الصحيح في بيانات الاختبار وهذا يشير إلى تحقق أحادية البعد فيها أيضاً.

التحقق من افتراض الاستقلال الموضعي

تم التحقق من افتراض الاستقلال الموضعي لمفردات الاختبار من خلال حساب متوسط معاملات الارتباط البينية بين المفردات لمجموعة الأفراد المختبرين جميعهم، وللمجموعة المرتفعة، وللمجموعة الدنيا. توضح بيانات الجدول (3) متوسط معاملات الارتباط لكل مجموعة.

جدول 3: متوسطات معاملات الارتباطات البينية بين المفردات

	المجموعة العليا	المجموعة الدنيا	العينة ككل
متوسط معاملات الارتباط	.020	.024	.117

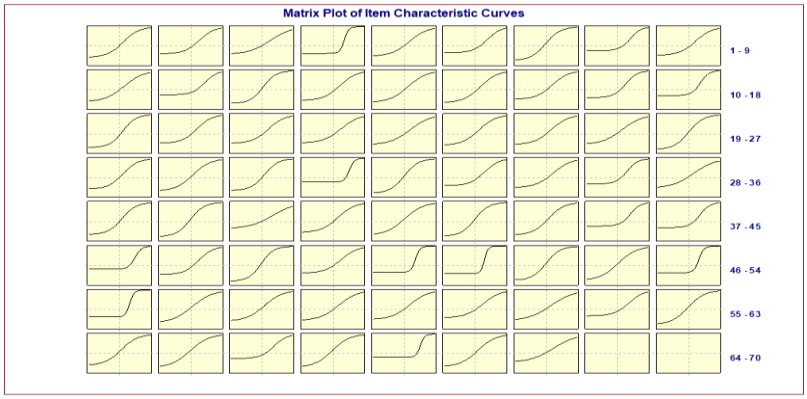
يتضح من متوسطات معاملات الارتباط تحقق افتراض الاستقلال الموضعي، حيث بلغ متوسطا معاملات الارتباط في المجموعة المرتفعة والمنخفضة (0.024 – 0.020) على التوالي وهما قريبين من الصفر وأقل من متوسط معاملات ارتباط المفردات في المجموعة الكلية للمختبرين التي بلغت (0.117) وبذلك يكون الاستقلال الموضعي لمفردات الاختبار متحققاً (Onder, 2007).

التحقق من افتراض التحرر من السرعة

تم الحرص على تحقيق هذا الافتراض من خلال إعطاء المختبرين الوقت الكافي للإجابة على الاختبار بحيث لم يؤثر عامل السرعة في أدائهم، وإنما كانت نتائج تحصيلهم تعكس قدرتهم الحقيقية في الإجابة على مفردات الاختبار. كما أن هذا الافتراض يعتبر من الافتراضات الضمنية وخاصة بعد التحقق من افتراض أحادية البعد والتي تعني وجود عامل واحد وهو السمة الكامنة للأفراد (الناغي،2011)، وبالتالي يمكن القول أنه ليس هناك أي تأثير لعامل السرعة. وقد تم التأكد من أن جميع المختبرين أجابوا على جميع أسئلة الاختبار، وبناءً على ما ذكره هامبلتون وسوامينثان (Hambleton & Swaminathan, 1985) أنه إذا أكمل (75%) من المختبرين الاختبار، و(80%) من المفردات تمت الإجابة عليها فإن السرعة لن تعتبر من العوامل المؤثرة في الأداء على الاختبار (الشريف وفراج، 2013)

تحقق منحنى خصائص المفردة

تم التحقق من هذا الافتراض من خلال إيجاد العلاقة الوتيرية بين قدرة الأفراد واستجاباتهم على المفردات فقد تم رسم منحنى خصائص المفردة (ICC) لكل مفردة من مفردات الاختبار. وتشير الحموري (2011) إلى أنه يمكن التحقق من هذا الافتراض من خلال إيجاد بارامترات المفردات التي تحدد منحنى خصائص كلٍ منها ويوضح الشكل (2) تحقق وتيرية السمة في المفردات.



شكل 2: الرسم البياني لمنحنى خصائص المفردة لجميع مفردات الاختبار

وبناءً على ما سبق يمكن القول بأن افتراضات نماذج نظرية الاستجابة للمفردة أحادية البعد متحققة في بيانات الدراسة الحالية مما يمكّن من استكمال تحليل هذه البيانات باستخدام أحد نماذج النظرية لمطابقة البيانات.

ثانياً: نتائج السؤال الثاني ونصه: هل تتطابق مفردات اختبار التفكير الاستدلالي مع النموذج الثلاثي البارامتر؟

للإجابة على هذا السؤال تم تحليل البيانات باستخدام برنامج Bilog-Mg، وتم التحقق من مطابقة مفردات اختبار التفكير الاستدلالي للنموذج الثلاثي البارامتر في بيانات الدراسة الحالية، حيث تم حساب قيمة اختبار مربع كاي لحسن المطابقة والدلالة الإحصائية للمطابقة لجميع المفردات وذلك عند مستوى دلالة (0.01) كما في دراسات كل من (عليمات، 2021 ؛ عالم وأبو هادي، 2019 ؛ Aksu, Reyhanhoglu & Eser, 2017 ؛ أبو فودة، 2016 ؛ Hamadneh, 2015 ؛ بني عطا والرباعي، 2013) ويظهر الجدول (4) النتائج الخاصة بمطابقة المفردات للنموذج الثلاثي البارامتر.

جدول 4: إحصائيات مربع كاي لمطابقة المفردات للنموذج الثلاثي البارامتر

رقم المفردة	كاي تربيع	الدلالة	رقم المفردة	كاي تربيع	الدلالة	رقم المفردة	كاي تربيع	الدلالة
1	17.9	0.0358	25	11.0	0.2741	49	12.5	0.1877
2	10.5	0.3131	26	404	0.8857	50	62.9	0.000**
3	8.9	0.4459	27	20.4	0.0154	51	88.3	0.000**
4	67.2	0.000**	28	18.8	0.0272	52	13.6	0.1372
5	8.2	0.5110	29	12.8	0.1741	53	37.3	0.000**
6	16.5	0.0576	30	604	0.7014	54	60.4	0.000**
7	13.7	0.1326	31	82.0	0.000**	55	55.4	0.000**
8	18.7	0.0283	32	26.3	0.0018**	56	28.9	0.0007**
9	11.9	0.2203	33	14.9	0.0927	57	8.9	0.4453
10	17.7	0.0390	34	16.5	0.0579	58	21.3	0.0114
11	4.3	0.8908	35	23.3	0.0056**	59	7.0	0.6417
12	11.3	0.2557	36	16.9	0.0510	60	7.6	0.5761
13	8.3	0.5067	37	11.5	0.2454	61	19.1	0.0242
14	15.8	0.0719	38	41.9	0.000**	62	8.9	0.4477
15	20.3	0.0160	39	21.0	0.0126	63	19.4	0.0220
16	8.0	0.5340	40	12.1	0.2052	64	10.9	0.2831
17	10.6	0.3012	41	40.4	0.000**	65	31.7	0.0002**
18	15.5	0.0773	42	14.4	0.1090	66	30.2	0.0004**
19	20.9	0.0133	43	9.4	0.3984	67	18.6	0.0289
20	8.7	0.4677	44	10.2	0.3331	68	114.4	0.000**
21	13.9	0.1254	45	19.4	0.0218	79	30.9	0.0003**
22	5.8	0.7631	46	23.4	0.0054**	70	10.6	0.3065
23	17.3	0.0439	47	9.4	0.3982			
24	9.1	0.4323	48	15.0	0.0905			

يلاحظ من الجدول (4) تطابق 53 مفردة مع النموذج الثلاثي البارامتر، وعدم مطابقة 17 مفردة للنموذج وذلك لكونها دالة احصائياً عند مستوى (0.01) وبالتالي تم استبعاد المفردات غير المطابقة والإبقاء على المفردات المطابقة وعددها 53 مفردة. وقد تم التوصل إلى هذه النتيجة بعد محاولة مطابقة البيانات للنماذج الثلاثة إلا أن أفضلها كان النموذج الثلاثي حيث تطابقت معه عدد مفردات أكثر مما تطابق مع النموذجين الآخرين. ولكون الاختبار من نوع الاختيار من متعدد فترى الباحثة أن النموذج الثلاثي البارامتر هو الأكثر ملائمة بالفعل له وذلك لأنه أكثر واقعية، وكما ذكر (Hambleton & Traub, 1971) فهو النموذج الأمثل في تقدير قدرة الفرد نظراً لاحتوائه على بارامترات الصعوبة والتمييز بالإضافة إلى أنه يعطي تقديراً لبارامتر التخمين الذي تزيد فرصة الوقوع به في أسئلة الاختيار من متعدد، وبالتالي يمكن الوصول من خلاله إلى نتائج أكثر دقة وموضوعية.

ثالثاً: نتائج السؤال الثالث ونصه: ما تقدير بارامترات الصعوبة والتمييز والتخمين لمفردات الاختبار؟

تم حساب بارامترات المفردات الصعوبة والتمييز والتخمين والأخطاء المعيارية لكلٍ منها. ويتضح ذلك في جدول (5).

جدول 5: معاملات الصعوبة والتمييز والتخمين والأخطاء المعيارية لها لكل مفردة من مفردات الاختبار

تسلسل المفردة	الصعوبة	الخطأ المعياري	التمييز	الخطأ المعياري	التخمين	الخطأ المعياري
1	0.277	0.126*	1.152	0.134*	0.194	0.043*
2	0.734	0.131*	1.160	0.182*	0.266	0.041*
3	0.750	0.209*	0.858	0.152*	0.267	0.055*
4	0.439	0.139*	1.102	0.145*	0.220	0.044*
5	1.144	0.126*	1.212	0.232*	0.305	0.036*
6	-0.080	0.088*	1.394	0.116*	0.134	0.034*
7	1.297	0.108*	1.637	0.346*	0.375	0.026*
8	0.630	0.127*	1.178	0.175*	0.240	0.042*
9	0.418	0.135*	1.093	0.141*	0.202	0.044*
10	1.399	0.131*	1.339	0.288*	0.360	0.030*
11	0.041	0.073*	1.618	0.138*	0.134	0.030*
12	0.610	0.120*	1.168	0.159*	0.213	0.040*
13	0.339	0.117*	1.131	0.126*	0.169	0.040*
14	0.522	0.110*	1.198	0.151*	0.185	0.038*
15	0.531	0.126*	1.204	0.163*	0.246	0.041*
16	0.750	0.090*	1.571	0.218*	0.289	0.031*
17	1.159	0.082*	1.994	0.370*	0.349	0.023*
18	0.254	0.075*	1.556	0.149*	0.141	0.031*
19	0.570	0.120*	1.256	0.176*	0.253	0.040*
20	0.567	0.120*	1.224	0.166*	0.240	0.040*
21	0.720	0.137*	1.113	0.176*	0.254	0.043*
22	0.653	0.156*	1.012	0.157*	0.237	0.047*
23	0.318	0.125*	1.152	0.132*	0.195	0.041*
24	0.462	0.139*	1.088	0.146*	0.213	0.045*
25	0.436	0.167*	0.957	0.133*	0.218	0.049*
26	0.008	0.075*	1.382	0.110*	0.103	0.028*
27	0.316	0.101*	1.419	0.165*	0.217	0.038*
28	0.231	0.118*	1.153	0.126*	0.171	0.041*
29	0.216	0.089*	1.439	0.144*	0.168	0.034*
30	0.805	0.130*	1.145	0.185*	0.260	0.041*
31	0.600	0.137*	1.098	0.160*	0.229	0.044*
32	0.462	0.172*	0.961	0.140*	0.229	0.051*
33	0.165	0.086*	1.454	0.142*	0.155	0.034*

تسلسل المفردة	الصعوبة	الخطأ المعياري	التمييز	الخطأ المعياري	التخمين	الخطأ المعياري
34	1.059	0.210*	0.877	0.190*	0.326	0.052*
35	0.473	0.131*	1.140	0.152*	0.220	0.043*
36	-0.052	0.080*	1.504	0.124*	0.131	0.032*
37	0.153	0.099*	1.280	0.122*	0.152	0.036*
38	1.281	0.098*	1.835	0.347*	0.386	0.023*
39	1.063	0.088*	1.719	0.288*	0.333	0.026*
40	0.703	0.121*	1.202	0.173*	0.248	0.039*
41	-0.100	0.067*	1.632	0.125*	0.104	0.028*
42	0.669	0.158*	1.048	0.172*	0.271	0.047*
43	-0.019	0.078*	1.553	0.136*	0.134	0.032*
44	0.161	0.137*	1.085	0.125*	0.184	0.045*
45	0.597	0.157*	1.002	0.153*	0.231	0.048*
46	0.452	0.156*	1.018	0.140*	0.221	0.047*
47	0.674	0.149*	1.080	0.168*	0.270	0.045*
48	0.423	0.175*	0.914	0.129*	0.209	0.051*
49	1.125	0.141*	1.139	0.225*	0.323	0.039*
50	-0.029	0.100*	1.207	0.106*	0.131	0.035*
51	0.114	0.117*	1.264	0.138*	0.191	0.042*
52	-0.082	0.114*	1.192	0.109*	0.157	0.040*
53	0.605	0.229*	0.792	0.132*	0.246	0.057*

يلاحظ من جدول (5) أن قيم بارامتر الصعوبة تراوحت بين (0.1- و1.399) وهذا يدل على أن اختبار التفكير الاستدلالي يميل إلى أن يكون متوسط الصعوبة نوعاً ما، وكانت المفردة (10) هي أصعب مفردة في الاختبار، والمفردة (41) هي الأسهل. وأن قيم بارامتر التمييز تراوحت بين (0.792 و1.994) وأن المفردة (17) هي الأعلى تمييزاً والمفردة (53) هي الأقل تمييزاً. كما يلاحظ أن بارامترات التخمينين تراوحت بين (0.103 و0.386) وكانت المفردة (38) هي الأكثر تخميناً، والمفردة (26) هي الأقل تخميناً. ويلاحظ من جميع القيم السابقة أنها جاءت ضمن المدى العلمي المقبول في القياس التربوي والنفسي حيث يشير هامبلتون (Hambleton & Swaminathan, 1985) إلى أنه عندما تتراوح قيم بارامتر الصعوبة بين (2- ,2) والتمييز بين (0.4, 2) فإن البارامترات تتسم بالاتساق وعند المقارنة بين هذه القيم بنتائج الدراسة الحالية في الجدول (5) نجد بأن البارامترات متسقة، وجميع القيم مقبولة. وبالنسبة للتخمين فجميع المفردات كانت أقل من (0.4) إلا أن هناك بعض المفردات كانت قيم التخمين لها مرتفعة نوعاً ما وتجاوزت (0.3) وترجع الباحثة ارتفاع هذه

القيم إلى طبيعة الأسئلة ونوعها الذي يعطي فرصة أكبر للتخمين كونها من نوع الاختيار من متعدد، ومع ذلك تعتبر جميع القيم ضمن المدى المقبول أيضاً.

رابعاً: نتائج السؤال الرابع ونصه: ما تقدير دالة المعلومات للمفردات وللاختبار؟

تم حساب أقصى قيمة للدالة المعلوماتية لكل مفردة من المفردات وكذلك القدرة المقابلة لها ويتضح ذلك في جدول (6).

جدول 6: الدوال المعلوماتية لمفردات الاختبار والقدرات المقابلة لها

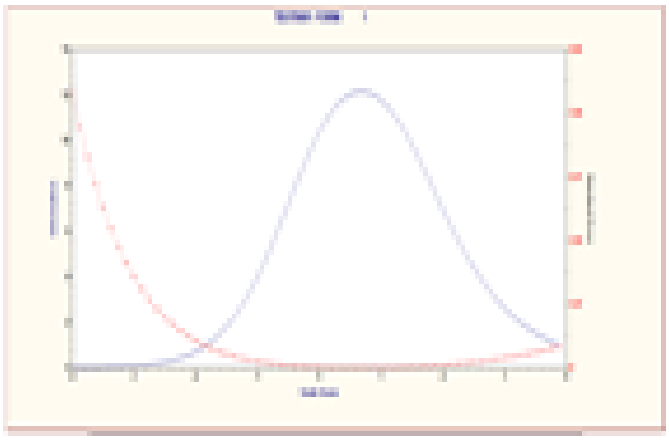
رقم المفردة	أقصى قيمة للدالة المعلوماتية	الخطأ المعياري	القدرة المقابلة	الخطأ المعياري	رقم المفردة	أقصى قيمة للدالة المعلوماتية	الخطأ المعياري	القدرة المقابلة	الخطأ المعياري
1	0.2286	0.0362*	0.5042	0.1118*	28	0.2394	0.0370*	0.4376	0.1059*
2	0.2019	0.0419*	1.0146	0.1155*	29	0.3753	0.0543*	0.3795	0.0804*
3	0.1101	0.0242*	1.1304	0.1795*	30	0.1990	0.0429*	1.0851	0.1167*
4	0.1989	0.0343*	0.6984	0.1211*	31	0.1943	0.0375*	0.8676	0.1202*
5	0.2038	0.0534*	1.4377	0.1288*	32	0.1486	0.0276*	0.7681	0.1489*
6	0.3753	0.0481*	0.0628	0.0817*	33	0.3924	0.0563*	0.3175	0.0785*
7	0.3215	0.0954*	1.5449	0.1256*	34	0.1023	0.0276*	1.4822	0.1888*
8	0.2187	0.0432*	0.8879	0.1108*	35	0.2131	0.0377*	0.7240	0.1147*
9	0.2025	0.0350*	0.6643	0.1183*	36	0.4396	0.0564*	0.0773	0.0742*
10	0.2221	0.0663*	1.6945	0.1522*	37	0.3062	0.0428*	0.3240	0.0901*
11	0.5054	0.0666*	0.1638	0.0675*	38	0.3949	0.1052*	1.5056	0.1104*
12	0.2269	0.0421*	0.8492	0.1064*	39	0.3871	0.0919*	1.2819	0.0911*
13	0.2312	0.0364*	0.5487	0.1056*	40	0.2245	0.0433*	0.9610	0.1077*
14	0.2516	0.0444*	0.7332	0.0986*	41	0.5442	0.0685*	0.0001	0.0630*
15	0.2261	0.0404*	0.7876	0.1092*	42	0.1631	0.0340*	0.9825	0.1357*
16	0.3536	0.0679*	0.9684	0.0821*	43	0.4660	0.0628*	0.1090	0.0722*
17	0.5038	0.1357*	1.3540	0.0914*	44	0.2068	0.0325*	0.3929	0.1214*
18	0.4612	0.0671*	0.3873	0.0688*	45	0.1612	0.0321*	0.8922	0.1366*
19	0.2427	0.0449*	0.8205	0.1042*	46	0.1695	0.0302*	0.7337	0.1353*
20	0.2366	0.0426*	0.8144	0.1052*	47	0.1736	0.0346*	0.9782	0.1288*
21	0.1902	0.0397*	1.0030	0.1206*	48	0.1399	0.0260*	0.7249	0.1526*
22	0.1625	0.0328*	0.9502	0.1356*	49	0.1735	0.0457*	1.4495	0.1403*
23	0.2282	0.0355*	0.5455	0.1104*	50	0.2831	0.0387*	0.1328	0.0921*
24	0.1967	0.0349*	0.7185	0.1214*	51	0.2773	0.0414*	0.3181	0.1032*
25	0.1510	0.0273*	0.7325	0.1457*	52	0.2627	0.0355*	0.1063	0.1041*
26	0.3915	0.0512*	0.1250	0.0705*	53	0.0976	0.0203*	0.9952	0.1971*
27	0.3322	0.0527*	0.5152	0.0886*					

يتضح من القيم الموجودة في الجدول (6) تباين المفردات في تقديم معلومات عن الاختبار والمختبرين والذي وضحه ارتفاع المنحنيات على متصل القدرة، فقد كانت المفردة (41) أعلى المفردات ارتفاعاً في منحناها وبلغ حجم مساهمتها بالمعلومات (0.5442) عند

مستوى القدرة (0.0001) بمعنى أنها تعطي معلومات أكثر عند المستوى المتوسط من القدرة. وقد تدرجت المفردات في ارتفاع منحنياتها واعطاءها للمعلومات، إلا أن المفردة (53) كانت هي الأدنى والأقل في كمية المعلومات التي تعطيها عن المختبرين فقد بلغت أقصى قيمة للمعلومات التي تعطيها (0.0976) عند مستوى القدرة (0.9952).

كما تم حساب الدالة المعلوماتية للاختبار ككل والخطأ المعياري لها، وتعكس دالة معلومات الاختبار مدى جودة الاختبار في قياس السمة المراد قياسها، وهي تتناسب عكسياً مع مقدار الخطأ في القياس (Momani,2018)، وقد تم تمثيل دالة معلومات الاختبار في الدراسة الحالية كما هو موضح في شكل (3) وتتضح أهم الإحصاءات الخاصة بها في جدول (7).

شكل 3: دالة معلومات الاختبار والخطأ المعياري لها



جدول 7: إحصاءات الدالة المعلوماتية للاختبار

الدالة المعلوماتية للاختبار	
أقصى قيمة للدالة المعلوماتية	12.130
القدرة المقابلة	0.6250
متوسط الدالة المعلوماتية	8.356
مؤشر الثبات	0.893

يلاحظ من الشكل (3) أن أول قيمة لدالة المعلومات التي يقدمها الاختبار هي (0.2) عند مستوى القدرة (2.8−)، حيث بدأت هذه القيمة بالارتفاع التدريجي إلى أن وصلت إلى أعلى قيمة لها (12.13) والمقابلة لمستوى القدرة (0.6250)، ثم بدأت بعد ذلك بالانخفاض التدريجي مع ارتفاع مستويات القدرة. كما يبين منحنى دالة معلومات الاختبار ومنحنى الخطأ المعياري في تقدير القدرة للأفراد عند مستويات القدرة المختلفة أن القيمة القصوى لدالة المعلومات كانت قريبة نوعاً ما من منتصف مستوى القدرة، وتتناقص تدريجياً عند الابتعاد عن تلك النقطة وهذا يعني دقة تقدير القدرة عند القدرة الأعلى قليلاً من منتصف مستوى القدرة التي يقيسها الاختبار على مقياس متصل القدرة، أما عند أطراف مستوى القدرة فيلاحظ انخفاض قيمة دالة المعلومات للاختبار، كما يظهر في الشكل أن الخطأ المعياري المقابل للقيمة القصوى للمعلومات كان أقل ما يمكن ويقترب من الصفر وهذا يعني أنه كلما زادت كمية المعلومات قل الخطأ المعياري لتقدير القدرة والعكس صحيح.

وبشكل أكثر توضيحاً، يُلاحظ من شكل (3) أن الخطأ المعياري والذي يمثله الخط المتقطع يتناقص تدريجياً وبشكل عكسي مع معلومات الاختبار كلما انتقلنا من المستويات الدنيا للقدرة إلى المستويات العليا، وقد بلغت قيمة الخطأ المعياري أقل قيمة لها وهي (0.38) عند مستوى قدرة (0.6250) وهي القدرة المقابلة لأعلى قيمة لدالة معلومات الاختبار (12.13). كما عادت قيمة الخطأ المعياري بالتزايد التدريجي وكمية المعلومات بالتناقص لمدى القدرة (0.6250 إلى 3.047). ويشير الشكل إلى أن الاختبار يقدم معلومات عند هذا المستوى من القدرة أكبر مما يقدمه عند المستويات المنخفضة من القدرة (1.5− إلى − 3.992) وبمقدار خطأ أقل. وهذا يشير بشكل عام إلى أن الاختبار المستخدم في الدراسة الحالية يميل إلى قياس القدرة المتوسطة – المرتفعة، كما يقدم معلومات أكبر للمختبرين الذين يمتلكون مستوى متوسط – مرتفع من السمة.

خامساً: نتائج السؤال الخامس ونصه: ما تقدير بارامتر القدرة للأفراد؟

باستخدام طريقة الأرجحية العظمى؛ تم تقدير قدرات المختبرين بوحدة اللوجيت، ويوضح الجدول (8) الإحصاءات الوصفية لتقديرات القدرة.

جدول 8: إحصاءات تقدير القدرة للمختبرين

الاحصائي	تقدير القدرة باللوجيت
أعلى قيمة	3.0473
أقل قيمة	-3.992
المتوسط	-0.0846
الانحراف المعياري	1.2867

تشير النتائج الواردة في جدول (8) إلى أن أعلى تقدير للقدرة في بيانات الدراسة الحالية كان (3.0473) لوجيت، وأقل تقدير للقدرة كان (3.992−) لوجيت، في حين كان متوسط تقدير القدرة (0.0846−) لوجيت بانحراف معياري بلغ (1.2867). وتشير هذه النتائج إلى أن اختبار التفكير الاستدلالي يقيس مدى واسع من قدرات المفحوصين وأن قدرات الأفراد تتوزع بشكل اعتدالي على هذا الاختبار كما أنه يقيس المستوى من المتوسط إلى المرتفع من القدرة بصورة أفضل من باقي مستويات القدرة.

سادساً: نتائج السؤال السادس ونصه: ما تقدير قيمة معامل الثبات التجريبي للاختبار؟

تم حساب معامل الثبات التجريبي للاختبار وقد بلغت قيمته 0.929 وتعد هذه القيمة مقبولة، كما تعد مؤشراً جيداً يمكن الأخذ به للحكم على جودة الاختبار المستخدم في الدراسة الحالية.

التوصيات والمقترحات

في ضوء النتائج التي تم الوصول إليها توصي الدراسة بما يلي:

- **استخدام النموذج الثلاثي البارامتر كأحد نماذج نظرية الاستجابة للمفردة في تدريج مفردات الاختبارات والمقاييس النفسية والتربوية المختلفة.**
- **الاهتمام بإجراء المزيد من الدراسات السيكومترية على أنواع أخرى من الاختبارات التي تقيس أنماط تفكير مختلفة كالتفكير الناقد والتفكير المنظومي والتفكير المجرد، باستخدام نظرية الاستجابة للمفردة.**
- **إعادة إجراء الدراسة على فئات عمرية مختلفة ومستويات تعليمية أخرى كطلبة المرحلة الجامعية.**

قائمة المراجع:

- أبو زيد، عبدالهادي السيد (2002). *الخصائص النفسية للحدسيين والاستدلاليين من طلاب المرحلة الثانوية*. [رسالة ماجستير]. كلية التربية: جامعة عين شمس.
- أبو فودة، باسل خميس. (2016). التوافق بين النظرية التقليدية في القياس ونظرية استجابة الفقرة في مطابقة فقرات اختبار محكي المرجع في وحدة الهندسة التحليلية. دراسات عربية في التربية وعلم النفس، (73)، 189–2015.
- بني عطا، زايد صالح و الرباعي، إبراهيم محمد. (2013). أثر عدد البدائل وتغيير موقع المموه القوي في فقرات اختبار الاختيار من متعدد على معالم الفقرات وقدرة الفرد ودالة المعلومات. *المجلة الأردنية في العلوم التربوية، 9* (3)، 319 – 333.
- التقي، أحمد محمد. (2013). *النظرية الحديثة في القياس*. دار المسيرة للطباعة والنشر. عمان: الأردن.
- جروان، فتحي عبدالرحمن. (2016). *تعليم التفكير مفاهيم وتطبيقات*. دار الفكر ناشرون وموزعون. عمان: الأردن.
- الحجازين، نايل عيد. (2014). *التفكير الاستدلالي*. عمان: دار جليس الزمان.
- الحجيلي، غدير أحمد. (2022). البناء العاملي لاختبار الاستدلال المنطقي لدى طالبات المرحلة الثانوية بمنطقة المدينة المنورة. *المجلة العربية للعلوم التربوية والنفسية. 6* (25)، 351 – 394.
- الحموري، هند عبدالمجيد. (2011). دراسة استكشافية لملائمة نماذج نظرية الاستجابة للمفردة في بناء اختبار تحصيلي من إعداد المعلم. *مجلة العلوم التربوية والنفسية، جامعة البحرين، 2* (12)، 47 – 81.
- خوالده، أكرم صالح محمود. (2016). *اللغة والتفكير الاستدلالي*. دار الحامد للنشر والتوزيع، عمان: الأردن.
- دي ايالا، ر. ج. (2017). *النظرية والتطبيق في نظرية الاستجابة للفقرة*. (عبد الله الكيلاني واسماعيل البرصان، مترجم). الرياض: دار جامعة الملك سعود للنشر.
- راجح، أحمد عزت. (1977). *أصول علم النفس*. الإسكندرية: دار المعارف.
- رزوقي، رعد مهدي وعبد الكريم، سهى إبراهيم. (2015). *التفكير وأنماطه: التفكير الاستدلالي – التفكير الإبداعي – التفكير المنظومي – التفكير البصري*. عمان: دار المسيرة للنشر والتوزيع.

- الشرقاوي، أنور؛ الشيخ، سليمان؛ كاظم، أمينة وعبدالسلام، نادية. (١٩٩٦). *اتجاهات معاصرة في القياس والتقويم النفسي والتربوي*. القاهرة: مكتبة الأنجلو المصرية.
- الشريف، خالد وفراج، محمد. (2013). فاعلية استخدام كل من نظرية السمات الكامنة والنظرية الكلاسيكية في قياس التحصيل الدراسي في الجبر بالمرحلة الإعدادية: دراسة مقارنة. *مجلة العلوم الإنسانية والاجتماعية، جامعة الإمام محمد بن سعود الإسلامية*، (28)، 89 – 147.
- عالم، توفيق علي وأبو هادي، سوسن حسن. (2019). بناء بنك أسئلة في مادة الرياضيات للصف التاسع الأساسي في ضوء نظرية استجابة الفقرة والتحقق من جودة توظيف برنامج حاسوبي لإدارته. *مجلة أبحاث، 1* (15)، 160–197.
- عبد الحافظ، شحته عبد الولي. (2016). الدقة الإحصائية لتقدير بارامترات النماذج الرياضية للاستجابة للمفردة. *مجلة رسالة التربية وعلم النفس*. 52، 129–161.
- عبد الوهاب، محمد محمود. (2010). *استخدام نماذج الاستجابة للمفردة الاختبارية في تدريج مفردات بعض الاختبارات المعرفية*. [رسالة دكتوراه]، جامعة المنيا، جمهورية مصر العربية.
- العبيدي، باسم؛ العبيدي، آلاء وولي، محمد. (2015). *اكتساب المعرفة وتعليم التفكير الاستدلالي والتفكير التماثلي*. مركز ديبونو لتعليم التفكير.
- العكايلة، عبد الناصر سند. (2017). استخدام النموذج ثنائي المعلم كأحد نماذج نظرية الاستجابة للمفردة في تدريج مفردات التفكير الاستدلالي المعرفي على طلبة المرحلة الثانوية في مديرية لواء الجامعة بالأردن. *المجلة الدولية التربوية المتخصصة. 6* (8)، 159 – 174.
- علام، صلاح الدين محمود. (2001). *الاختبارات التشخيصية مرجعية المحك في المجالات النفسية والتربوية والتدريبية*. ط2. القاهرة: دار الفكر العربي.
- علام، صلاح الدين محمود (2005). *نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي*. القاهرة: دار الفكر العربي.
- علام، صلاح الدين محمود. (2015). *القياس والتقويم التربوي والنفسي، أساسياته وتطبيقاته وتوجهاته المعاصرة*. الطبعة السادسة. القاهرة: دار الفكر العربي.
- عليمات، مهدي. (2021). بناء اختبار تحصيلي في الرياضيات وفق نظرية الايتجابة للفقرة (النموذج ثنائي المعلمة). *مجلة جامعة النجاح للأبحاث (العلوم الإنسانية)، 35* (7)، 1080 – 1104.

- عودة، موسى عزت. (2016). *تقدير معالم الفقرة والقدرة بالطرق التي تستند إلى نظرية الاستجابة للفقرة تبعاً لمتغيري طول الاختبار وعدد البدائل*. [رسالة دكتوراه]. كلية الدراسات العليا: الجامعة الأردنية.
- عيال، ياسين وجاسم، خالد. (2019). الخصائص القياسية لاختبار الاستدلال الاستقرائي لدى طلبة الجامعة على وفق نظرية الاستجابة للمفردة الاختبارية. *مجلة الأستاذ للعلوم الإنسانية والاجتماعية*. *58* (4)، 273 – 298.
- محمد، محمد. (2013). *فاعلية استخدام خرائط التفكير في تنمية مهارات الاستدلال وتحصيل مادة التربية في الدراسات الاجتماعية لدى تلاميذ المرحلة الاعدادية*. [رسالة ماجستير]. جامعة المنصورة.
- مراد، عودة. (2013). معايير أداء طلبة الصف العاشر الأساسي في الأردن على اختبار القدرة الاستدلالية. *مجلة العلوم التربوية والنفسية*. *16* (1)، 321 – 339.
- المصاروة، سامي سلامة. (2008). *دلالات الصدق والثبات لاختبار التفكير الاستدلالي لطلبة المرحلة الثانوية في دولة الامارات العربية المتحدة*. [رسالة ماجستير]. جامعة عمّان العربية، الأردن.
- المنشدي، محمد عبد الكريم. (2019). القدرة الاستدلالية لدى طلبة المرحلة الإعدادية. *مجلة أبحاث الذكاء: وقائع المؤتمر العلمي الدولي الأول للدراسات الإنسانية (الذكاء والقدرات العقلية)*. 144 – 172.
- المنصور، غسان. (2010). الاستدلال المنطقي وحل المشكلات. *مجلة جامعة دمشق*، *28* (1)، 107 – 143.
- المهداوي، وفاء. (2013). *أثر تدريس الرياضيات وفقاً للتعلم النشط في تحصيل طالبات المدارس المهنية في محافظة ديالي وتنمية تفكيرهن الاستدلالي*. [أطروحة دكتوراه]، جامعة سانت كليمنت العالمية.
- الناغي، هبة إبراهيم. (2011). أثر عدد مفردات الاختبار على تقدير قدرات الأفراد وذالة المعلومات لاختبار تحصيلي مرجعي المحك في الرياضيات باستخدام نماذج نظرية الاستجابة للمفردة. *مجلة كلية التربية، جامعة بور سعيد*، (10)، 676 – 721.
- هيئة تقويم التعليم والتدريب، اختبار القدرة المعرفية العامة. (5 – 9 – 202) تم الاسترداد من: https://etec.gov.sa/ar/productsandservices/Qiyas/profession/knowledge/Pages/default.a

- الوليلي، إسماعيل حسن والدوسري سعيد عبد الله. (2019). استخدام بعض طرق نظرية الاستجابة للمفردة في الكشف عن الأداء التفاضلي لمتغير الجنس في مفردات اختبار القدرات العامة وأثرها على دقة تقدير القدرة والخصائص السيكومترية للاختبار. *مجلة كلية التربية، جامعة بني سويف*. عدد أكتوبر، الجزء 2. 156 – 201.
- Aksu, G., Reyhanhoglu, C. & Eser, M. (2017). Examining the Two Categorical Datas by Jmetrik, Bilog-Mg and Irtpro with Application of Mathematics Exam. European Scientific Journal, 13 (33), 20 – 43.
- Baker, F. (2001). The Basics of Item Response Theory. The ERIC Clearinghouse on Assessment and Evaluation, USA.
- Baserer, D. (2020). Logical Thinking Levels of Teacher Candidates. Educational Policy Analysis and Strategic Research. 15 (4), 176 – 190.
- Crocker, L.& Algina, J. (1986). Introduction to Classical and Modern Test Theory. California: Cole Publishing Company.
- Hamadneh, I. M. (2015). The Impact of Escape Alternative Position Change in Multiple-Choice Test on the Psychometric Properties of a Test and it's Items Parameters. Journal of Education and Practice, 6 (9), 23 – 33.
- Hambleton, R. K., & Traub, R. (1971). Information Curves and Efficiency of Three Logistic Test Model. British Journal of Mathematical and Statistical Psychology, 24, 273-281.
- Hambleton, R. & Swaminathan, H. (1985). Item Response Theory: principles and applications. Boston: Kluwer Neuhoff publishing.
- Hambleton, R.; Swaminathan, H. & Rogers, H. (1991). Fundamentals of Item Response Theory. California, New bury Park SAGE Publications Inc.
- Hüseyin Polat, Fatma Bilge Emre. (2021). Logical Thinking Abilities Scale: Reliability and Validity Studies. Universal Journal of Educational Research. 9(1), 180 - 185. DOI: 10.13189/ujer.2021.090120
- Kaplan, R. & Saccuzzo, D. (2008). Psychological Testing: Principles, Applications, and Issues. Pacific Grove Wads Worth Pubco.
- Lord, F. (1980). Application of Item Response Theory to Practical Testing Problems. New Jersey: Lawrence Erlbaum Associates, Publishers.
- Momani, Rana TH. (2018). Using Item Response Theory to Evaluate Self-Directed Learning Readiness Scale. Journal of Educational and Developmental Psychology, 8 (1), 14-27.
- Onder, I. (2007). An Investigation of Goodness of Model Data Fit. Hacettepe Universitesi Egitim Fakultesi Dergisi, 32,210-220.

- Pásztor, Attila, Andrea Magyar, Anita Pásztor-Kovács, and Attila Rausch. (2022). Online Assessment and Game-Based Development of Inductive Reasoning. Journal of Intelligence 10(3): 59. https://doi.org/10.3390/ jintelligence10030059.
- Qomariyah, S. and Darmayanti, R (2023). Development of High School Students' Mathematical Reasoning Ability Instruments on Three Dimension Material. JEMS (Journal of Mathematics and Science Education), 11(1), 249-260 DOIs: 10.25273/jems.v11i1.14705.
- Reeve, B. (2004). An Introduction To Modern Measurement Theory. Division Of Cancer Control And Population Science, National Cancer Institute.
- Vierula J, Talman K, Hupli M, Laakkonen E, Engblom J, Haavisto E. (2021). Development and psychometric testing of Reasoning Skills test for nursing student selection: An item response theory approach. Journal of Advanced Nursing. 77 (2), 2549–2560. https://doi.org/10.1111/jan.14799