An-Najah University Journal for Research – B
Humanities

مجلة جامعة النجاح للأبحاث – ب
العلوم الإنسانية

The Effectiveness of Using the Partial Credit Model in Developing Polytomous Scored Mathematics Items for Seventh Grade Students

Mohammad Dabous[1,*]

(Type: Full Article). Received: 28th May 2025, accepted: 16th Sep. 2025, Published: ××××. DOI: https://doi.

Accepted Manuscript, In Press

Abstract: Aim: This study aimed to investigate the effectiveness of the Partial Credit Model (PCM) in constructing polytomous items in mathematics assessments for seventh-grade students, and to examine the extent to which student response data align with the assumptions of this model. **Methodology:** The researcher employed a descriptive-analytical approach. To achieve the study's objectives, a mathematics achievement test was developed for seventh-grade students, consisting of 15 items varying in the number of response steps. The test was administered to a sample of 653 male and female students. The data were statistically analyzed to verify the assumptions of the Partial Credit Model. **Results:** The findings indicated that the test met the assumptions of Item Response Theory (IRT), and that all items conformed to the Partial Credit Model, reflecting its suitability for measuring diverse student abilities. The results also revealed variation in item difficulty and discrimination levels, enabling accurate differentiation among performance levels. **Conclusions:** The study affirms the validity of using the Partial Credit Model in constructing polytomous mathematics assessments for the basic education stage. It recommends adopting this model in educational institutions, training teachers and specialists in its application, and utilizing its outcomes to develop a scientifically grounded national item bank.

Keywords: Partial Credit Model, polytomous items, Item Response Theory, mathematics, seventh grade.

فاعلية نموذج التقدير الجزئي في بناء أسئلة متعددة التدريج في مبحث الرياضيات للصف السابع الأساسي

محمد طالب دبوس[1،*]

تاريخ التسليم: (2025/5/28) تاريخ القبول: (2025/9/16)، تاريخ النشر: (××××)

الملخص: الهدف: هدفت الدراسة الكشف عن فاعلية نموذج التقدير الجزئي في بناء أسئلة متعددة التدريج في مبحث الرياضيات للصف السابع الأساسي، والتحقق من مدى توافق بيانات الطلبة مع افتراضات هذا النموذج. **المنهجية**: استخدم الباحث المنهج الوصفي التحليلي منهجا للدراسة، ولتحقيق اهداف الدراسة تم إعداد اختبار تحصيلي في مادة الرياضيات للصف السابع الأساسي مؤلفا من 15 سؤالا تختلف في عدد خطوات الاستجابة، وطُبق على عينة مكونة من (653) طالبًا وطالبة. وحللت البيانات إحصائيًا للتحقق من افتراضات نموذج التقدير الجزئي. **النتائج**: أظهرت النتائج الى تحقيق الاختبار لافتراضات نظرية الاستجابة للفقرة ومطابقة جميع أسئلة الاختبار لنموذج التقدير الجزئي مما يعكس ملاءمة نموذج التقدير الجزئي لقياس القدرات المختلفة لدى الطلبة. كما أظهرت النتائج وجود تباين في مستويات الصعوبة والتمييز، الأمر الذي يتيح التفرقة بين مستويات الأداء المختلفة بدقة. **الاستنتاجات**: تؤكد الدراسة صلاحية استخدام نموذج التقدير الجزئي في بناء اختبارات الرياضيات متعددة التدريج للمرحلة الأساسية، ويوصي باعتماد هذا النموذج في المؤسسات التعليمية، وتدريب المعلمين والمختصين على تطبيقه، بالإضافة إلى توظيف نتائجه في إنشاء بنك فقرات وطني قائم على أسس علمية دقيقة.

الكلمات المفتاحية: نموذج التقدير الجزئي، الأسئلة متعددة التدريج، نظرية الاستجابة للفقرة، الرياضيات، الصف السابع الأساسي

1 Department of Psychology, Faculty of Humanities, Al-Istiqlal University, Jericho, Palestine.
* Corresponding author: m.dabous@najah.edu

1 قسم علم النفس، كلية العلوم الإنسانية، جامعة الاستقلال، أريحا، فلسطين.
* الباحث المراسل: m.dabous@najah.edu

المقدمة

يُعد التقويم التربوي حجر الزاوية في المنظومة التعليمية، فهو لا يقتصر على كونه أداة لقياس تحصيل المتعلمين، بل يمثل منظومة متكاملة تسهم في تحسين جودة التعليم وتوجيهه نحو تحقيق الأهداف المنشودة. فكل عملية تعليمية تُصبح بلا معنى إذا افتقرت إلى تقويمٍ فعّال يُنير مسارها، ويكشف مواطن القوّة والضعف لدى المتعلّمين، ويزود المعلمين وصنّاع القرار بمعلومات دقيقة وموثوقة يمكن البناء عليها.

ولقد أصبح من الضروري في ظل التحولات التربوية المعاصرة تجاوز المفهوم التقليدي للتقويم الذي ينحصر في إصدار الأحكام النهائية حول أداء المتعلمين، والانتقال إلى منظور أشمل وأكثر إنصافًا، يهدف إلى دعم عملية التعلّم وتعزيزها، من خلال تشخيص دقيق ومتكامل للاحتياجات الفردية والجماعية للطلبة، وبما يسهم في تحقيق مبدأ العدالة التربوية. وقد جاء تطور أدوات القياس استجابةً لهذه الحاجة إلى تقويم يتسم بالموضوعية والدقة، ويأخذ في اعتباره تباين قدرات الطلبة واختلاف مستوياتهم المعرفية. وفي هذا الإطار، برزت نظرية الاستجابة للفقرة (IRT) بما تتيحه من إمكانات تتفوق على النظرية الكلاسيكية في القياس، وفي مقدمتها بناء اختبارات دقيقة وموضوعية وعادلة (Hambleton, Swaminathan, & Rogers, 1991).

ويُعد نموذج التقدير الجزئي (Partial Credit Model - PCM) أحد النماذج المشتقة من هذه النظرية، حيث يتيح قياس الأداء التدريجي بصورة أكثر مرونة وعدالة، ولا سيما في المباحث التي تتضمن فقرات مفتوحة أو متعددة الخطوات، مثل مبحث الرياضيات. وتكمن أهمية هذا النموذج في منحه الطالب درجة تتناسب مع مستوى استجابته، بعيدًا عن الثنائية التقليدية "صح/خطأ"، الأمر الذي يسهم في توفير تقويم أكثر دقة وإنصافًا، فضلاً عن دوره في تحفيز المتعلمين. وتشير الأدبيات إلى فاعلية هذا النموذج في الكشف عن الكفاءة الحقيقية للطلبة، خصوصًا في المراحل الأساسية التي تتكون فيها المفاهيم الرياضية بصورة تدريجية (Masters, 1982).

وتتوافق هذه التوجهات الحديثة في مجال التقويم مع الدعوات العالمية المتزايدة إلى تطوير اختبارات قائمة على الكفايات، تراعي الفروق الفردية والمستويات المعرفية المتباينة، وتكون أكثر ارتباطًا بالأهداف التعليمية وواقع الممارسات الصفية (Schraw & Robinson, 2015) كما تؤكد الاتجاهات التربوية المعاصرة على أهمية توظيف التغذية الراجعة التحليلية المستندة إلى نتائج الاختبارات، بوصفها مدخلًا فعالًا لدعم التعلّم التكيفي، وتمكين المعلمين من تصميم أنشطة تعليمية مبنية على بيانات دقيقة تساعد الطلبة على تجاوز التحديات المعرفية المعقدة (Brookhart, 2016).

وبناءً على ما سبق، تبرز الحاجة الملحّة إلى تطوير أدوات تقويم تتوافق مع هذه الرؤية المتقدمة. وانطلاقًا من ذلك، تسعى الدراسة الحالية إلى توظيف نموذج التقدير الجزئي في بناء فقرات اختبارية متعددة التدريج في مبحث الرياضيات للصف السابع الأساسي، بما يتيح تقديم صورة أكثر واقعية عن قدرات الطلبة، ويعزّز من عدالة التقييم وجودته.

مشكلة الدراسة وأسئلتها

يمثل التقويم التربوي عنصرًا جوهريًا في العملية التعليمية، غير أن أدواته التقليدية، لا سيما في مادة الرياضيات، غالبًا ما تعجز عن قياس الأداء الحقيقي للطلبة. فهي تكتفي بتصنيف الإجابات إلى صحيحة أو خاطئة، متجاهلة مراحل التفكير والجهد المبذول في الحل، مما يُضعف من قيمتها التشخيصية ويحدّ من قدرتها على تقديم تغذية راجعة دقيقة تدعم اتخاذ قرارات تربوية فاعلة.

في ضوء هذا القصور، تبرز الحاجة إلى نماذج تقويم أكثر عدالة ومرونة، مثل نموذج التقدير الجزئي (PCM)، والذي يُراعي تعدد مستويات الأداء داخل الفقرة الواحدة، ويُقدّر كل خطوة بشكل مستقل. هذا النموذج يتناسب مع طبيعة الأسئلة التي تتطلب تدرجًا في الحل، ويُسهم في بناء أدوات تقويم دقيقة تُظهر الفروق الفردية وتُعزز الموضوعية في القياس.

وقد جاء اختيار هذا النموذج في الدراسة استجابة لحاجة ميدانية واقعية، استندت إلى خبرة الباحث العلمية في القياس والتقويم، وخبرته العملية في تدريس الرياضيات في المدارس الفلسطينية، حيث لمس عن قرب التحديات التي تواجه المعلمين في تصميم أدوات تقويم تعكس فعليًا أداء الطلبة. وقد اختير الصف السابع الأساسي مجالًا للتطبيق، نظرًا لأهمية هذه المرحلة في ترسيخ المهارات المفاهيمية والإجرائية، وما تتطلبه من أساليب تقويم تراعي تنوّع القدرات وتفاوت أنماط التفكير.

تهدف الدراسة إلى بناء اختبار رياضيات متعدد التدرج يستند إلى نموذج التقدير الجزئي، وتحليل مدى توافق استجابات الطلبة مع افتراضات هذا النموذج، بما يُسهم في تطوير أدوات تقويم حديثة تدعم جودة التعليم، وتزوّد المعلمين ببيانات دقيقة تساعدهم في اتخاذ قرارات تربوية مبنية على أسس علمية، وتُعزز من فاعلية الممارسات التعليمية بما يتناسب مع احتياجات الطلبة وتنوع مستوياتهم المعرفية.

وبالتحديد، حاولت هذه الدراسة الإجابة عن السؤال الرئيسي **ما مدى فاعلية نموذج التقدير الجزئي في تطوير أسئلة متعددة التدريج في مادة الرياضيات للصف السابع الاساسي؟**

وينبثق منه الأسئلة الفرعية التالية:

ACCEPTED

1. ما مدى توافق استجابات الطلبة على أسئلة اختبار الرياضيات متعدد التدريج لافتراضات نموذج التقدير الجزئي وفقا لنظرية الاستجابة للفقرة؟
2. ما درجة مطابقه اسئلة اختبار الرياضيات متعدد التدريج للصف السابع الأساسي لنموذج التقدير الجزئي؟
3. ما قيم معالم صعوبة كل خطوة من خطوات استجابة السؤال الواحد في اختبار الرياضيات متعدد التدريج للصف السابع الأساسي؟
4. ما تقدير قدرات الأفراد على اختبار الرياضيات متعدد التدريج للصف السابع الأساسي لنموذج التقدير الجزئي؟
5. ما قيمة دالة المعلومات لكل فقرة من فقرات اختبار الرياضيات متعدد التدريج؟

أهداف الدراسة

تهدف الدراسة إلى ما يلي:

1. تحليل فاعلية نموذج التقدير الجزئي في تطوير أسئلة متعددة التدريج في مادة الرياضيات للصف السابع الأساسي.
2. تحديد مدى مطابقة فقرات اختبار الرياضيات لنموذج التقدير الجزئي من حيث الافتراضات والمعايير الإحصائية.
3. تقييم قدرات الأفراد بناءً على استجاباتهم لفقرات الاختبار متعددة التدريج باستخدام نموذج التقدير الجزئي.
4. استخدام نتائج الدراسة لتطوير بنوك أسئلة تعتمد على النموذج في تحسين أدوات القياس والتقويم.
5. تحسين جودة التعليم والتقييم الأكاديمي من خلال بناء اختبارات موثوقة ودقيقة تتيح تقييمًا عادلًا لقدرات الطلبة.

أهمية الدراسة

الأهمية النظرية: تنبع الأهمية النظرية للدراسة من إسهامها في إثراء الأدبيات حول نموذج التقدير الجزئي وتطبيقه في تحليل الفقرات متعددة التدريج. فهي تعزز الفهم العلمي لفاعلية النموذج في قياس القدرات بدقة، واستقلال معالم الفقرات عن العينات، مما يفتح المجال لتطوير أدوات تقييم حديثة أكثر ملاءمة.

الأهمية العملية: تبرز الأهمية العملية للدراسة في تطوير اختبارات الرياضيات للصف السابع باستخدام نموذج التقدير الجزئي، بما يعزز دقة وعدالة القياس بعيدًا عن خصائص العينة. كما تتيح النتائج بناء بنوك أسئلة فعّالة، وتزويد المعلمين بأداة تشخيصية دقيقة لتحسين أساليب التدريس وتوجيه الدعم الأكاديمي للطلبة.

حدود الدراسة

- **الحد البشري**: تقتصر هذه الدراسة على عينة من طلبة الصف السابع الأساسي.
- **الحد المكاني**: المدارس التابعة لمديرية التربية والتعليم في نابلس.
- **الحد الزماني**: تم إجراء الدراسة في نهاية الفصل الدراسي الأول من العام الدراسي 2025/2024م.

مفاهيم ومصطلحات الدراسة

نموذج التقدير الجزئي Partial Credit Model (PCM): أحد النماذج المنبثقة عن نموذج راش، ويُستخدم مع الفقرات متعددة التدريج. يُمكّن هذا النموذج من تقدير معامل الصعوبة الفئوية لكل فئة من فئات الاستجابة، كما يحدد قدرة الفرد الكامنة. ويتميز بإمكانية اختلاف الفقرات في متوسط قيم العتبات الفارقة بين فئات الاستجابة (Tutz *et al.*, 2018).

ويعرّف الباحث النموذج إجرائيًا بأنه أحد نماذج نظرية الاستجابة للفقرة متعددة التدريج، ويُستخدم لتحليل فقرات اختبار الرياضيات متعدد التدرج للصف السابع، من خلال تقدير صعوبة كل خطوة في الحل، ومعلم قدرة الطالب بناءً على استجاباته، ويُعد إطارًا إحصائيًا لقياس جودة الفقرات وفعالية التمييز بين مستويات الطلبة.

الفقرات متعددة التدريج Polytomous Items: هي الفقرات المقالية والتي تتدرج فيها الإجابة في عدد من الخطوات ويعطى عليه الفرد علامة بناء على عدد الخطوات التي اجتازها بشكل صحيح وذلك حسب مستويات الاستجابة على هذه الفقرة (دبوس، 2009).

ويعرفها الباحث اجرائيا: فقرات اختبار الرياضيات التي تم بناؤها من كتاب الرياضيات للصف السابع الأساسي من المنهاج الفلسطيني لتتناسب مع نموذج التقدير الجزئي وتسمح بتقويم أداء الطلبة بشكل تدريجي.

الإطار النظري والدراسات السابقة

تُعد النظرية الكلاسيكية للاختبارات (CTT) من النماذج الأساسية في القياس النفسي والتربوي التي شكّلت مرجعًا رئيسًا في بناء الاختبارات وتحليل نتائجها لعقود طويلة (Hambleton & Jones, 1993). ورغم مساهمتها في فهم العلاقة بين أداء الأفراد ونتائجهم، إلا أن اعتمادها على مؤشرات إجمالية، مثل معامل الصعوبة والتمييز، حدّ من قدرتها على تفسير الفروق الفردية بدقة، خاصة في سياقات تطوير بنوك الأسئلة أو الاختبارات التكيفية.

استجابةً لهذه التحديات، ظهرت نظرية الاستجابة للفقرة (IRT)كبديل أكثر تقدمًا، إذ توفر إطارًا رياضيًا يربط بين استجابة المفحوص على الفقرة ومستوى السمة الكامنة لديه،

ACCEPTED

بما يتيح تحليلاً أكثر دقة ومرونة في تفسير الأداء الفردي (حبيب وعزيز، 2018).

وقد أشار دبوس (2016) بأن نظرية الاستجابة للفقرة تعتبر نظرية حديثة في القياس والتقويم التربوي يتم فيها تحديد العلاقة بين أداء المفحوص والسمة الكامنة موضع القياس وفق دالة رياضية محددة.

تستند نظرية الاستجابة للفقرة إلى افتراضات أساسية ونماذج رياضية دقيقة تمكّن من تقدير خصائص الفقرات والقدرات الكامنة للأفراد بشكل مستقل، مما يتيح بيانات أكثر دقة وموضوعية. وتمتاز هذه النظرية بمرونتها العالية في تصميم الاختبارات وتفسير النتائج، فضلاً عن قدرتها على تحسين جودة أدوات القياس من خلال تقديم تقديرات فردية دقيقة لكل من الفقرة والمفحوص، بما يعزز من صدق القرارات التربوية المستندة إليها (Hambleton & Swaminathan, 1985; Lord, 1980).

الافتراضات الأساسية لنظرية الاستجابة للفقرة

تستند نظرية الاستجابة للفقرة إلى مجموعة من الافتراضات الأساسية التي تضمن صدق النماذج واستقرارها عبر السياقات المختلفة. وقد لخصها كل من كاي واخرون (Cai *et al*, 2016) وهامبلتون وسوامنثيان (Hambelton & Swaminathan, 1985) الى أربعة محاور رئيسية وهي: أولًا، افتراض أحادية البُعد، والذي يعني أن جميع الفقرات تقيس بُعدًا كامنًا واحدًا كالسمة أو القدرة. ثانيًا، الاستقلال الموضعي، حيث تُفترض استقلالية استجابات المفحوصين للفقرات إحصائيًا عند ضبط البُعد الكامن. ثالثًا، عدم تأثير عامل السرعة، أي أن الفشل في الإجابة يُعزى إلى محدودية القدرة وليس لضيق الوقت. وأخيرًا، منحنى خصائص الفقرة(ICC)، وهو دالة غير خطية تصف العلاقة بين احتمال الإجابة الصحيحة ومستوى القدرة، وغالبًا ما يتخذ شكلًا مميزًا يعكس التدرج في الأداء مع تزايد القدرة الكامنة.

نماذج نظرية الاستجابة للفقرة

ظهرت نتيجة العمل في نظرية الاستجابة للفقرة عدة نماذج رياضية لوغاريتمية تصف العلاقة بين أداء المفحوص وسماته الكامنة باستخدام أشكال رياضية مختلفة لمنحنيات خصائص الفقرات بناءً على افتراضات محددة (Samejima, 1969؛ Embretson & Reise, 2000؛ Masters, 1982؛ Hambelton & Swaminathan, 1985؛ علام، (2012)؛ الشريفين وبني عطا (2013)) وتقسم النماذج إلى نوعين: نماذج ثنائية التدريج، ونماذج متعددة التدريج.

في النوع الأول، النماذج ثنائية التدريج، نجد نموذج راش الأحادي المعلم الذي يركّز على تقدير صعوبة الفقرات مع افتراض ثبات معلمَي التمييز والتخمين. يليه نموذج بيرنوم الثنائي المعلم، الذي يفترض تغير معاملي الصعوبة والتمييز مع إلغاء أثر التخمين. أما نموذج لورد الثلاثي المعلم، فيُدخل معلمة التخمين إلى جانب الصعوبة والتمييز، مما يجعله أكثر تعقيدًا وواقعية في تمثيل أداء المفحوصين.

أما النوع الثاني، وهو النماذج متعددة التدريج فيشمل نوعين هما: نموذج الاستجابة المتدرجة (GRM) الذي طوّره Samejima عام 1969 وهو يعد من النماذج متعددة الفئات الشائعة والمرنة في نظرية الاستجابة للفقرة (IRT) ذات المعلمَين (PL-2). ويتم في هذا النموذج تقدير خصائص الفقرة المتمثلة في معاملي التمييز والعتبات (أو الموقع) threshold (or location). وتوفر معلمات العتبة معلومات حول صعوبة أو شدة الفقرة، وتحدد موقعها على البُعد المقاس. ونموذج التقدير الجزئي(PCM) Partial Credit Model الذي طوّره Masters عام 1982 في أستراليا، امتدادًا لنموذج Rasch المخصص للفقرات ثنائية التدريج، وهو من أبسط نماذج نظرية استجابة الفقرة متعددة التدريج، إذ جاء ليمكّن من تقدير معالم الفقرات ذات الاستجابة متعددة المستويات، بحيث تكون هذه التقديرات مستقلة عن عينة الأفراد، وكذلك تقدير قدرات الأفراد بمعزل عن عينة الفقرات المستخدمة. وتمتاز هذه الآلية بمرونتها في التعامل مع مواقف اختبارية تتضمن فقرات متدرجة الإجابة، تُقاس من خلال مستويات أو خطوات متتابعة، بحيث لا يمكن للمفحوص الانتقال إلى مستوى معين إلا بعد إتمام جميع المستويات السابقة له. فعلى سبيل المثال، إذا كانت الفقرة مؤلفة من ست خطوات، فإن بلوغ المستوى الثالث يتطلب بالضرورة إتمام الخطوتين الأولى والثانية، وهكذا حتى الوصول إلى المستوى الأخير، مما يجعل هذا النموذج مثاليًا لقياس الأداء في المهام التي تتطلب تسلسلًا تدريجيًا في الحل. ويسمح هذا النموذج بتقدير احتمالية حصول الفرد على درجة معينة في فقرة معينة بناءً على قدرته وصعوبة الخطوات أو العتبات (thresholds) بين مستويات الاستجابة المختلفة (العمراني والمالكي، 2023)، و(طنطاوي، 2023).

ونظرا لأهمية استخدام نموذج التقدير الجزئي (PCM) فقد أجريت العديد من الدراسات والابحاث العلمية التي وضحت أهمية استخدام هذا النموذج في بناء الاختبارات ومنها:

دراسة **حمدان (2025)** التي هدفت إلى تحديد النموذج التحليلي الأنسب لقياس الكفاءة الذاتية في الإدارة الصفية لدى الطلبة المعلمين، بمقارنة نموذج التقدير الجزئي وسلّم التقدير. وطُبّق مقياس للكفاءة الذاتية على عينة من 741 طالبًا وطالبة بكلية التربية في جامعة أسيوط. أظهرت النتائج توافق الفقرات مع كلا النموذجين، لكن نموذج التقدير الجزئي كان الأدق والأكثر ملاءمة للبيانات التربوية.

ودراسة **الزهراني وسليمان (2023)** التي هدفت إلى تطوير مقياس للدافعية نحو التعلم وفق نموذج التقدير الجزئي. تكون المقياس من 25 فقرة موزعة على أربعة أبعاد وطُبّق

ACCEPTED

على عينة عنقودية عشوائية من 541 طالبًا ثانويًا في منطقة تبوك، حُللت البيانات ببرنامجEasy Estimation ، وأظهرت النتائج توافق الفقرات مع النموذج، وفعالية المقياس في تصنيف الطلاب وتصميم برامج تعليمية داعمة.

في حين هدفت دراسة **العنزي وسليمان (2023)** إلى تطوير مقياس لقياس التفكير ما وراء المعرفي باستخدام نموذج التقدير الجزئي. اعتمدت الدراسة المنهج الكمي، وطُبّق المقياس إلكترونيًا على عينة من 455 طالبًا وطالبة بجامعة تبوك. تم تحليل البيانات باستخدام برنامج Easy Estimation، وأظهرت النتائج توافق الفقرات مع النموذج، ووقوع مؤشرات المطابقة ضمن النطاق المقبول.

وهدفت دراسة **العمراني والمالكي (2023)** إلى فحص أثر طريقة ترتيب فئات التدريج ونموذج الاستجابة للفقرة على دقة تقدير بارامترات الفقرات وقدرات الأفراد، بمقارنة نموذج التقدير الجزئي (PCM) ونموذج الاستجابات المتدرجة (GRM). طبق الباحثان مقياس اتخاذ القرار (34 فقرة) على عينة من (2385) طالبًا وطالبة بجامعة أم القرى، وحُللت البيانات ببرنامجي SPSS و .PARSCALEأظهرت النتائج تفوق نموذج التقدير الجزئي في دقة تقدير بارامترات الصعوبة، دون فروق دالة تُعزى إلى طريقة ترتيب الفئات أو التفاعل بينها وبين النموذج. كما لم تظهر فروق في تقدير بارامترات التمييز وفق ترتيب الفئات، بينما وُجد فرق دال في تقدير قدرات الأفراد لصالح أحد النماذج.

وأجرى **فرحات (2022)** دراسة هدفت إلى إعداد مقياس للتأمل الناقد وتدريج خصائصه السيكومترية باستخدام نموذج التقدير الجزئي ونموذج الاستجابة المتدرجة. شملت الدراسة عينة مكونة من 1042 مشاركًا من طلبة كليات جامعة الأزهر. وأظهرت النتائج عدم توافق 20 فقرة مع نموذج التقدير الجزئي، و6 فقرات مع نموذج الاستجابة المتدرجة. كما بين التحليل أن نموذج التقدير الجزئي يوفر دالة معلومات أكثر دقة عند المستويات المنخفضة من القدرة، في حين يقدم نموذج الاستجابة المتدرجة دالة معلومات أفضل عند المستويين المنخفض والمتوسط من القدرة.

وهدفت دراسة **بني هاني واخرون (Bani Hani *et al,* 2021)** إلى تطوير واختبار صدق اختبار الرياضيات لطلبة الصف العاشر وفقًا لنموذج التقدير الجزئي لراش (PCM) باستخدام المنهج الوصفي الذي يتناسب مع أهداف الدراسة. تم إجراء التطبيق الأولي للاختبار للتحقق من صدق وثبات الاختبار. كما تم حساب الصدق التجريبي للاختبار، وبلغت قيمة ثبات الأفراد 0.91، فيما بلغت ثبات البنود 0.93. تم تطبيق الاختبار على عينة حجمها 250 طالبًا وطالبة. وبحسب نموذج التقدير الجزئي، أظهرت نتائج الدراسة عدة قضايا متعلقة بتحصيل مادة الرياضيات من خلال التحقق من الصدق والثبات وتحقيق فروض نظرية الاستجابة للمقياس.

وهدفت دراسة **أبو شندي وحمود والرجيبي (2021)** إلى تحليل الخصائص السيكو مترية لمقياس قلق المستقبل المهني لدى طلبة الجامعات في سلطنة عُمان باستخدام نظرية الاستجابة المتدرّجة. شملت الدراسة عينة عنقودية مكونة من 524 طالبًا من جامعتي السلطان قابوس ونزوى. أظهرت النتائج دلالات قوية على صدق وثبات المقياس، وبنية عاملية داخلية مناسبة وفق التحليل العاملي التوكيدي، كما تم تحديد درجات القطع لتصنيف مستويات القلق المهني بناءً على الدرجات الخام والقدرات المقدّرة.

في حين هدفت دراسة **الشريفين (2019)** إلى مقارنة ثلاثة نماذج من نظرية استجابة الفقرة متعددة التدريج (النموذج الإسمي، التقدير الجزئي العام، والاستجابة المتدرج) في دقة تقدير القدرات والخصائص السيكومترية. استخدمت الدراسة المنهج الكمي، وولّدت بيانات افتراضية لاختبار مكون من 50 فقرة عبر برنامجWinGen 3 ، وطبّقت على عينة من 1500 مفحوص. أظهرت النتائج تفوق نموذج التقدير الجزئي العام في تقدير القدرات، والنموذج الإسمي في التمييز، والاستجابة المتدرج في تحديد الصعوبة وكفاءة المعلومات، دون فروق دالة في معاملات الثبات بين النماذج.

أما دراسة **توسيم وآخرون (Tusiime *et al,* 2015)**، فقد سعت إلى التحقق من الخصائص السيكومترية لمقياس بيك للاكتئاب من خلال تطبيق نموذج التقدير الجزئي. وتم تنفيذ الدراسة على عينة مكونة من 123 شخصًا من المصابين بفيروس نقص المناعة المكتسبة (HIV) من ذوي الدخل المحدود في أوغندا. وقد شملت أداة القياس 21 فقرة، وتم تحليل استجابات المشاركين للتحقق من مدى توافقها مع النموذج المستخدم حيث أظهرت أن المقياس يتمتع بخصائص سيكومترية قوية وتطابق جيّد مع نموذج التقدير الجزئي، مما يعزز من موثوقيته كأداة فعالة لتقييم الاكتئاب ضمن هذه الفئة.

التعقيب على الدراسات السابقة

يتضح من مراجعة الدراسات السابقة أن هناك اهتمامًا متزايدًا لاستخدام نموذج التقدير الجزئي (PCM) في تطوير أدوات القياس، وخصوصًا في المجالات النفسية والتربوية. ومع هذا الزخم البحثي، إلا أن المتتبع لهذه الدراسات يلحظ فجوة واضحة، تتمثل في أن معظمها انصبّ على تحليل أدوات جاهزة أو تطوير مقاييس نفسية، دون التطرق الجاد لبناء اختبارات تحصيلية فعلية قائمة على النموذج. فقلّما نجد دراسات تناولت بناء أسئلة اختبارية متعددة الخطوات تُحفّز الطالب على التفكير التدريجي وتُقدّر محاولاته الجزئية بشكل فعلي، وهو أمر أساسي في تعليم مواد كالرياضيات.

ACCEPTED

إضـــافة إلى ذلك، فإن الفئة العمرية المســتهدفة في معظم هذه الدراســـات كانت من طلبة المراحل الثانوية أو الجامعية، في حين غابت المرحلة الأساسية – وتحديدًا الصف السابع – عن الاهتمـام البحثي رغم أهميتهـا في ترســـيخ المفـاهيم الرياضـية الأسـاسـية، لذا فإن هذا المسـتوى الدراسـي لا يزال بحاجة إلى المزيد من البحث والتطوير في هذا المجال.

ويبين الجدول (1) أوجـه التشـــابـه والاختلاف مـا بين الدراسة الحالية والدراسات السابقة.

جدول (1): أوجه التشابه والاختلاف ونقاط التميز للدراسة الحالية مقارنة بالدراسات السابقة.

الدراسة	أوجه تشابه الدراسات السابقة مع الدراسة الحالية	أوجه اختلاف الدراسات السابقة مع الدراسة الحالية
حمدان(2025)	– استخدام نموذج التقدير الجزئي – التحقق من توافق الفقرات مع افتراضات النموذج.	– المجال: تطبيقي على مقياس للكفاءة الذاتية في الإدارة الصفية وليس اختبار تحصيلي. – الفئة المستهدفة: طلبة جامعيون، وليس طلبة مدارس.
الزهراني وسليمان (2023)	– تطوير أداة وفق نموذج التقدير الجزئي. – التحقق من توافق الفقرات مع افتراضات النموذج.	– المجال: مقياس للدافعية نحو التعلم وليس اختبار تحصيلي. – الفئة المستهدفة: طلبة ثانوية
العنزي وسليمان (2023)	– تطوير مقياس وفق نموذج التقدير الجزئي. – التحقق من توافق الفقرات مع افتراضات النموذج.	– المجال: تطوير مقياس لقياس التفكير ما وراء المعرفي. – الفئة المستهدفة: طلبة جامعيون.
العمراني والمالكي (2023)	– استخدام نموذج التقدير الجزئي – التحقق من توافق الفقرات مع افتراضات النموذج	– المجال: مقياس اتخاذ القرار. – الفئة المستهدفة: طلبة جامعيون. – لم يتضمن تحليل دوال المعلومات. – استخدم أيضا نموذج الاستجابة المتدرجة (GRM)
فرحات(2022)	– استخدام نموذج التقدير الجزئي – التحقق من توافق الفقرات مع افتراضات النموذج	– المجال: مقياس التأمل الناقد. – الفئة المستهدفة: طلبة جامعيون.
Bani Hani *et al.* (2021)	– استخدام نموذج التقدير الجزئي – التحقق من توافق الفقرات مع افتراضات النموذج – اختبار تحصيلي في الرياضيات.	– الصف المستهدف: الصف العاشر.
أبو شندي وحمود والرجيبي (2021)	– تحليل الخصائص السيكومترية للمقياس	– استخدام نموذج الاستجابة المتدرجة – المجال: مقياس قلق المستقبل المهني – الفئة المستهدفة: طلبة جامعيون.
الشريفين(2019)	– تحليل دقة تقدير القدرات ومعالم الفقرات.	– استخدام ثلاثة نماذج من نظرية الاستجابة للفقرة. – ولّدت بيانات افتراضية لاختبار مكون من 50 فقرة عبر برنامج WinGen 3.
Tusiime *et al.* (2015)	– استخدام PCM للتحقق من الخصائص السيكومترية. – التحقق من التوافق مع النموذج.	– المجال: مقياس بيك للاكتئاب. – الفئة المستهدفة: مرضى نقص المناعة المكتسبة

الطريقة والإجراءات

منهج الدراسة

استخدم الباحث المنهج الوصفي التحليلي منهجا للدراسة، وذلـك لملاءمتـه لطبيعتهـا، حيـث يتم في هـذا المنهج جمع البيـانـات وإجراء التحليل الإحصـــائي لاســـتخراج النتـائج المطلوبة.

مجتمع الدراسة

تكون مجتمع الدراسة من جميع طلبة الصف السابع الأساسي والتابعين لمديرية التربية والتعليم في مدينة نابلس في العام الدراسي 2025/2024م والبالغ عددهم (5026) طالبا وطالبة حسب إحصاءات مديرية التربية والتعليم.

عينة الدراسة

تكونت عينة الدراسة من (653) طالبا وطالبة من طلبة الصف السابع الأساسي منهم (310) طالبا و (343) طالبة موزعين على (20) مدرسة بواقع (10) مدارس للذكور و (10) مدارس للإناث، وتم اختيار المدارس بطريقة قصدية.

أداة الدراسة

لتحقيق أهداف الدراسة، استعان الباحث بتحليل المحتوى الرسمي الصادر عن وكالة الغوث الدولية (2017) لتحديد وتصنيف الأهداف المعرفية الواردة في كتاب الرياضيات للصف السابع الأساسي – الفصل الدراسي الأول، والذي يتضمن أربع وحدات دراسية: الأعداد الصحيحة، الهندسة والقياس، التناسب، والإحصاء. وقد تم تصنيف الأهداف وفق المجالات الثلاثة: المعرفة المفاهيمية، والمعرفة الإجرائية، والاستدلال. وبناءً على هذا التصنيف، تم إعداد جدول مواصفات للاختبار المقترح، بحيث تكون في صورته النهائية من (15) سؤالًا موزعة بما يتناسب مع الوزن النسبي لكل وحدة ونوع المعرفة، وتم الاخذ بعين الاعتبار بأن الاستجابة على كل سؤال تكون ضمن خطوات متسلسلة ومضبوطة بعدد محدد من مستويات الأداء.

تم عرض تحليل المحتوى وقائمة الأهداف وجدول المواصفات وأسئلة الاختبار وإجاباتها النموذجية على مجموعة من مدرسي الرياضيات للصف السابع من أجل إبداء الرأي بها والحكم على الفقرات من حيث صياغتها ومدى ارتباط كل فقرة بالهدف الذي تقيسه، وتم الأخذ بملاحظاتهم وجرى تعديل الاسئلة التي كانت بحاجة إلى تعديل. وتم التأكد من خطوات حل كل سؤال، وكان عدد خطوات الاستجابة الكلي لجميع الأسئلة هو (45) خطوة، بحيث يعطى الطالب علامة على كل خطوة يتم الإجابة عليها بطريقة صحيحة، فتكون العلامة الكلية

ACCEPTED

للطالب مساوية لعدد خطوات استجابته على أسئلة الاختبار. وقد تراوحت خطوات الاستجابة في السؤال بين خطوتين إلى 4 خطوات. وبهذا يكون الباحث تأكد من صدق المحتوى للاختبار من خلال ملاحظات المحكمين.

تجريب الاختبار على عينة استطلاعية

جرى تجريب الاختبار على عينة استطلاعية من طلبة الصف السابع (32 طالبًا و35 طالبة) بهدف التحقق من وضوح التعليمات، وتحديد زمن التطبيق، وتقدير صعوبة الفقرات. طُلب من الطلبة كتابة خطوات الحل كاملة لتحليل أنماط الاستجابة، وأسهم التجريب في تعديل التعليمات النهائية وتحديد آلية التصحيح. تراوحت العلامات بين 4 و45 درجة، وبلغ معامل الثبات باستخدام كرونباخ ألفا (0.937)، وهي قيمة مناسبة لأغراض الدراسة. طُبق الاختبار لاحقًا على العينة الأساسية بالتعاون مع معلمي الرياضيات في المدارس، وتم تصحيحه يدويًا وفق تعليمات محددة وواضحة.

المعالجات الإحصائية

من أجل إجراء التحليلات الإحصائية قام الباحث بإدخال البيانات إلى الحاسوب، واستخدم برنامج الرزم الإحصائية في العلوم الاجتماعية (SPSS) وبرنامج MULITILOG.7 وبرنامج MINITAB 15 لتحليل البيانات حسب نموذج التقدير الجزئي وتم استخدام المعالجات الإحصائية الأتية:

- معامل ألف كرونباخ
- معاملات الصعوبة والتمييز للفقرات الاختبار حسب النظرية الكلاسيكية
- التحليل العاملي الاستكشافي Factor Analysis من الدرجة الأولى لفقرات الاختبار.
- مربع كاي (χ^2).
- معلمة الصعوبة والخطأ المعياري للفقرات حسب النظرية الحديثة
- القيمة العظمى لدالة المعلومات لكل فقرة $I(\theta)_{max}$ وقيمة القدرة المقابلة لها $(\theta)_{max}$.

نتائج الدراسة ومناقشتها

النتائج المتعلقة بسؤال الدراسة الأول: ما مدى توافق استجابات الطلبة على أسئلة اختبار الرياضيات متعدد التدريج لافتراضات نموذج التقدير الجزئي وفقا لنظرية الاستجابة للفقرة؟ للإجابة عن السؤال الأول قام الباحث بالتحقق من افتراضات نموذج التقدير الجزئي كما يلي:

افتراض أحادية البعد (Unidimensionality): تم إجراء التحليل العاملي الاستكشافي Factor Analysis من الدرجة الأولى لفقرات اختبار الرياضيات متعدد التدريج باستخدام برمجية SPSS ويظهر الجدول (2) قيمة الجذر الكامن ونسبة التباين المفسر والنسبة التراكمية للتباين المفسر.

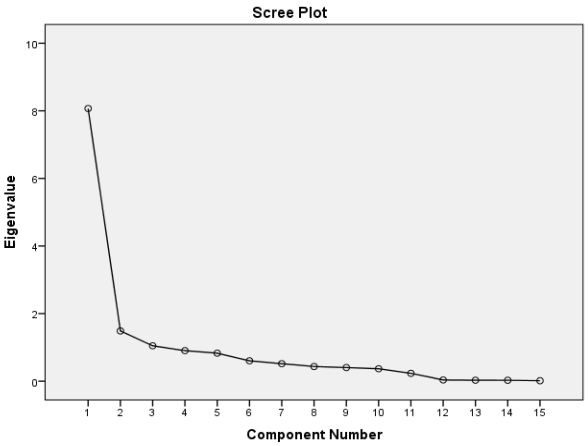
جدول (2): نتائج التحليل العاملي من الدرجة الأولى لفقرات اختبار الرياضيات متعدد التدريج.

العامل	الجذر الكامن	نسبة التباين المفسر المفسر	النسبة التراكمية للتباين المفسر%	ناتج قسمة الجذر الكامن الأول على الثاني
1	7.930	%52.869	%52.869	5.31
2	1.494	%9.962	63.708	
3	1.046	6.973	69.804	
4	.911	6.075	75.879	
5	.868	5.784	81.663	
6	.604	4.024	85.687	
7	.592	3.948	89.635	
8	.448	2.987	92.622	
9	.408	2.720	95.342	
10	.361	2.407	97.748	
11	.231	1.543	99.291	
12	.037	.243	99.534	
13	.028	.189	99.723	
14	.028	.183	99.906	
15	.014	.094	100.000	

أظهرت نتائج التحليل العاملي من الدرجة الأولى للفقرات متعددة التدريج في جدول (2) وجود عامل واحد فسّر 52.869% من التباين الكلي، يليه عامل ثانٍ بنسبة 9.962%. وبلغت قيمة الجذر الكامن للعامل الأول (7.930)، وهي مرتفعة مقارنةً بالعوامل الأخرى. ووفقًا لمعيار لورد (Lord, 1984) فإن نسبة الجذر الكامن الأول إلى الثاني (5.31) تفوق الحد المطلوب (2)، مما يدل على وجود عامل سائد يُرجّح أحادية البعد في اختبار الرياضيات متعدد التدريج.

ويوضح الشكل (1) التمثيل البياني لقيم الجذور الكامنة المكونة للعوامل المستخلصة في التحليل العاملي حيث أن قيمة الجذر الكامن للعامل الأول كبيرة جدا مقارنة بالعامل الثاني.

ACCEPTED



شكل (1): التمثيل البياني لقيم الجذور الكامنة للعوامل المكونة للاختبار.

افتراض الاستقلال الموضعي (Independence Local): يرى وورم (Warm, 1978) أن افتراض أحادية البعد يتضمن افتراض الاستقلال الموضعي. وبما أن افتراض أحادية البعد قد تحقق حسب جدول (2)، لذا يمكن القول إن افتراض الاستقلال الموضعي متحقق بالنسبة لفقرات اختبار الرياضيات متعددة التدريج.

افتراض التحرر من سرعة الأداء Seededness: تم إعطاء الطلبة الوقت الكافي للإجابة عن أسئلة الاختبار. حيث تفترض نماذج نظرية الاستجابة للفقرة أن عامل السرعة لا يلعب دورا في الإجابة عن فقرات الاختبار، بمعنى أن الذين يفشلون في الإجابة عن فقرات الاختبار يفشلون بسبب محدودية قدرتهم.

وفي إطار فحص خصائص فقرات اختبار الرياضيات متعدد التدريج وفقًا للنظرية الكلاسيكية في القياس، قام الباحث بتحليل معلمتي الصعوبة والتمييز لكل فقرة باستخدام برنامج الرزم الإحصائية للعلوم الاجتماعية (SPSS). وقد استند هذا التحليل إلى افتراض رئيس في نموذج التقدير الجزئي، وهو تساوي قدرة جميع الفقرات على التمييز بين مستويات القدرة المختلفة، مما استلزم التحقق من مدى توافر هذه الخاصية في فقرات الاختبار، فتم حساب معامل التمييز من خلال تقدير معاملات الارتباط بين أداء الطلبة على كل فقرة على حدة وأدائهم الكلي على بقية فقرات الاختبار، حيث يُعد هذا الأسلوب من الأساليب المعتمدة في النظرية الكلاسيكية لتقدير فعالية الفقرات في التمييز بين الطلبة ذوي المستويات المختلفة من التحصيل. كما تم احتساب معامل الصعوبة لكل سؤال من خلال قسمة مجموع العلامات المحصلة على السؤال على النهاية العظمى للفقرة مضروبة في العدد الكلي للأفراد المفحوصين، ويُعد مؤشرًا على مدى سهولة أو صعوبة الفقرة. والجدول (3) يبين نتائج التحليل.

جدول (3): معلمتي الصعوبة والتمييز لكل فقرة من فقرات اختبار الرياضيات باستخدام برنامج الرزم الإحصائية للعلوم الاجتماعية (SPSS)

السؤال	معامل الصعوبة	معامل التمييز	السؤال	معامل الصعوبة	معامل التمييز
1	0.333	0.605	9	0.343	0.703
2	0.520	0.708	10	0.702	0.493
3	0.255	0.658	11	0.335	0.632
4	0.580	0.549	12	0.495	0.661
5	0.508	0.573	13	0.258	0.623
6	0.474	0.618	14	0.595	0.631
7	0.346	0.722	15	0.629	0.706
8	0.670	0.597			

تشير بيانات الجدول (3) إلى أن قيم معلم الصعوبة للفقرات تراوحت القيم بين (0.255) للفقرة (3) و(0.702) للفقرة (10)، ما يشير إلى أن الفقرات تمثل مستويات متفاوتة من الصعوبة مع ميل عام نحو الصعوبة المتوسطة، وهو أمر مرغوب فيه في بناء الاختبارات التربوية لقياس مدى واسع من القدرات.

وفيما يخص معاملات التمييز للفقرات فكانت مرتفعة بوجه عام، حيث تراوحت بين (0.493) للفقرة (10) و(0.722) للفقرة (7)، مما يدل على قدرة الفقرات على التمييز بدرجة جيدة بين الطلبة ذوي التحصيل المرتفع والمنخفض. كما أن جميع الفقرات أظهرت معاملات تمييز ضمن النطاق المقبول تربويًا (0.4 فأعلى)، وهو ما يُعزز من صلاحية الفقرات للاستخدام في القياس التربوي.

يفسّر الباحث نتائج السؤال الأول في ضوء نظرية الاستجابة للفقرة، حيث يُعد تحقق افتراض أحادية البعد مؤشرًا بنيويًا على تجانس البنية المفاهيمية للاختبار، مما يدل على قياس بُعد معرفي موحّد، يُحتمل أن يكون "القدرة الرياضية العامة" أو "حل المشكلات متعددة الخطوات". ويعزز ذلك التصميم المنهجي المتماسك للاختبار، كما تؤكده نسبة التباين المفسَّر للعامل الأول (52.869%) والفجوة الكبيرة بين الجذرين الكامنين للعاملين الأول والثاني (5.31)، مما يبرر اعتماد نموذج التقدير الجزئي. وبما أن أحادية البعد متحققة، فإن افتراض الاستقلال الموضعي يُعد متحققًا أيضًا، مما يعزز موثوقية البيانات وصلاحيتها للتحليل.

أما افتراض "التحرر من سرعة الأداء"، فقد تحقق من خلال منح الطلبة وقتًا كافيًا للإجابة، مما ضمن أن الاستجابات تعكس القدرة الحقيقية لا عامل السرعة.

ويُظهر تحليل الخصائص الكلاسيكية للفقرات ارتفاعًا في معامل التمييز، مما يدل على قدرة الفقرات المتدرجة على التفرقة الدقيقة بين مستويات الطلبة، إذ تقيس نمط التفكير الرياضي وليس مجرد الإجابة، وهو ما يعزز دقة القياس.

تتوافق النتائج مع دراسة بني هاني وآخرين (2021) التي أثبتت تحقق أحادية البعد في اختبار رياضيات وفق نموذج PCM، مما يدعم تطبيقه في الصفوف الأدنى. كما تؤيدها

ACCEPTED

دراستا الزهراني وسليمان (2023) والعنزي وسليمان (2023)، حيث تحقق النموذج في أدوات متعددة التدريج رغم اختلاف المجالات.

بناءً عليه، يرى الباحث أن تحقق افتراضات نموذج التقدير الجزئي في اختبار الرياضيات الحالي يُعد مؤشرًا على جودة البناء المفاهيمي والدقة التقنية، ويؤكد صلاحية النموذج في تطوير اختبارات تحصيلية تقيس مستويات معرفية متقدمة، وتوفر تقييمات أكثر عدالة وعمقًا، بما يتجاوز أنماط التقييم التقليدية ويعزز موضوعية قياس تحصيل طلبة الصف السابع في الرياضيات.

النتائج المتعلقة بسؤال الدراسة الثاني: ما درجة مطابقه اسئلة اختبار الرياضيات متعدد التدريج للصف السابع الأساسي لنموذج التقدير الجزئي؟ للإجابة على السؤال الثاني قام الباحث بالتحقق من درجة مطابقة أسئلة اختبار الرياضيات متعدد التدريج لنموذج التقدير الجزئي وذلك باستخدام البرمجيتين MULITILOG.7 وMINITAB 15 عند مستوى الدلالة $\alpha= 0.05$. ويظهر الجدول (4) قيمة مربع كاي ومستوى الدلالة لفحص المطابقة لكل فقرة من فقرات الاختبار مع النموذج ثنائي المعلم عند مستوى الدلالة $\alpha = 0.05$.

جدول (4): مربع كاي ومستوى الدلالة لفحص مطابقة الفقرات متعددة التدريج مع نموذج الاستجابة المتدرجة.

الفقرة	قيمة χ^2	مستوى الدلالة	الفقرة	قيمة χ^2	مستوى الدلالة
1	0.03252259	0.995	9	0.00703439	0.999
2	0.00857469	1	10	0.00096889	1
3	0.00156509	0.999	11	0.02632869	0.996
4	0.00076719	1	12	0.00848569	1
5	0.00649419	0.998	13	0.01207499	1
6	0.00151659	1	14	0.00973659	0.997
7	0.01075319	1	15	0.00414409	1
8	0.00245769	1			

تشير نتائج تحليل مطابقة الفقرات إلى توافق تام بين بيانات استجابات الطلبة ونموذج التقدير الجزئي (PCM). فقد أظهرت القيم العالية لمستوى الدلالة لجميع الفقرات والتي تجاوزت بكثير حد الدلالة الإحصائية 0.05 أن فروق مربع كاي (χ^2) بين القيم المرصودة والمتوقعة ليست ذات دلالة إحصائية. وهذا يعني أن الفقرات تقيس المتغير الكامن المرتبط بالقدرة الرياضية بدقة وفق النموذج المستخدم، مما يعكس جودة بناء الأسئلة وملاءمتها لنموذج التقدير الجزئي.

من منظور الباحث، يُعد هذا الانسجام بين البيانات والنموذج إشارة واضحة إلى أن الفقرات لا تقتصر على التمييز بين النجاح والفشل فحسب، بل تقيس مستويات متفاوتة من أداء الطلبة، وهو أمر جوهري في تقييم المهارات الرياضية المتعددة التدريج. كما أن قدرة النموذج على التعامل مع هذه الدرجات المتعددة تمثل تطورًا مهمًا مقارنة بالتصحيح الثنائي التقليدي، حيث يمكن تقدير الأداء بدقة أعلى وشمولية أكبر. وهذا التوافق يعزز مصداقية الاختبار كأداة تقييمية عادلة وموضوعية، ويؤكد صلاحية استخدام نموذج التقدير الجزئي في بناء اختبارات ذات درجات متدرجة، لا سيما في المواد التي تعتمد على التفكير التحليلي والخطوات الحسابية المتعددة مثل الرياضيات.

علاوة على ذلك، فإن هذه النتيجة تشير إلى أن تصميم الفقرات أخذ في الاعتبار المراحل المختلفة من صعوبة الأسئلة ومستويات الأداء، ما سمح للنموذج بالتعرف بدقة على الفروق الفردية بين الطلبة. وهذا يعد دلالة إيجابية على أن الاختبار قادر على التمييز بين الطلاب ذوي القدرات المختلفة بطريقة دقيقة وعادلة.

تتفق النتائج مع دراسات سابقة، مثل بني هاني وآخرين (2021) والزهراني وسليمان، والعنزي وسليمان (2023)، التي أثبتت صلاحية نموذج التقدير الجزئي في تحليل الفقرات متعددة التدريج في السياقات التربوية والنفسية. إلا أنها تختلف جزئيًا عن دراسة فرحات (2022)، التي وجدت ضعف التوافق في بعض الفقرات، ويرجّح أن ذلك يعود إلى طبيعة الأداة أو ضعف التصميم، مما يؤكد أهمية الإعداد الجيد قبل تطبيق النموذج.

النتائج المتعلقة بسؤال الدراسة الثالث: ما قيم معالم صعوبة وتمييز كل خطوة من خطوات استجابة السؤال الواحد في اختبار الرياضيات متعدد التدريج للصف السابع الأساسي؟ للإجابة على السؤال الثاني قام الباحث باستخدام برمجية MULTILOG.7 لحساب معلمة الصعوبة والخطأ المعياري لمعلمة الصعوبة لكل خطوة من خطوات أسئلة الاختبار، كذلك تم حساب معلمة التمييز والخطأ المعياري لمعلمة التمييز لكل فقرة من فقرات الاختبار. ويوضح الجدول (5) نتائج التحليل.

ACCEPTED

جدول (5): معالم الصعوبة والخطأ المعياري لمعالم الصعوبة لكل خطوة من خطوات فقرات الاختبار متعدد التدريج ومعلم التمييز والخطأ المعياري لمعلم التمييز لكل فقرة باستخدام برمجية MULTILOG.7

معالم الصعوبة للفقرات متعددة التدريج									معالم التمييز	
رقم السؤال	صعوبة الخطوة الاولى	الخطأ المعياري لصعوبة الخطوة الأولى	صعوبة الخطوة الثانية	الخطأ المعياري لصعوبة الخطوة الثانية	صعوبة الخطوة الثالثة	الخطأ المعياري لصعوبة الخطوة الثالثة	صعوبة الخطوة الرابعة	الخطأ المعياري لصعوبة الخطوة الرابعة	التمييز	الخطأ المعياري للتمييز
1	-0.62	0.15	2.07	0.22	2.13	0.23			1.37	0.20
2	-1.41	0.15	0.40	0.08	0.45	0.09			1.66	0.19
3	0.8-	0.14	1.14	0.15	2.52	0.28	2.69	0.29	1.15	0.13
4	1.41-	0.22	0.48	0.11					1.01	0.14
5	1.18-	0.31	0.02-	0.26					0.73	0.13
6	0.92-	0.12	0.12	0.09	1.11	0.12			1.55	0.16
7	0.87-	0.13	0.37	0.08	1.26	0.13	1.97	0.18	1.51	0.15
8	1.45-	0.15	0.43-	0.09					1.68	0.20
9	0.65-	0.09	0.16	0.06	1.36	0.12	1.75	0.15	1.92	0.17
10	5.30-	1.15	2.92-	0.89	1.64-	0.59			0.44	0.12
11	0.12	0.09	10.5	0.11	1.34	0.13			2.07	0.23
12	1.34-	0.16	0.30-	0.09	0.62	0.10			1.71	0.18
13	1.89-	0.32	0.14-	0.14	1.82	0.24	2.05	0.26	1.12	0.15
14	1.60-	0.27	0.85-	0.17					1.40	0.17
15	1.44-	0.15	0.30-	0.08	0.53	0.09			1.87	0.18

تشير نتائج جدول (5) إلى وجود تفاوت واضح في معلمات الصعوبة عبر خطوات الفقرات المختلفة، حيث تراوحت قيم الصعوبة بين -1.64 لوجيت في الفقرة رقم (10) إلى 2.69 لوجيت في الفقرة رقم (3)، بمتوسط قدره 0.925، مما يعكس مدى تنوع مستوى التحدي الذي تواجهه خطوات الأسئلة. ويُظهر هذا التوزيع أن بعض الخطوات في الفقرات تعد سهلة نسبيًا (قيم سلبية للصعوبة)، في حين أن أخرى صعبة للغاية (قيم إيجابية مرتفعة)، وهو ما يتوافق مع طبيعة الأسئلة متعددة التدريج التي تستهدف قياس مستويات متفاوتة من المهارة.

ويرى الباحث أن هذا التفاوت يعكس قدرة الاختبار على استهداف طيف واسع من مستويات الطلبة، مما يسهم في تقديم صورة أدق عن قدراتهم، خاصة عند استخدام نموذج التقدير الجزئي الذي يسمح بتقييم كل خطوة بشكل مستقل. هذه النتائج تؤكد جودة بناء الاختبار من حيث التنوع والعدالة في تقدير الأداء، حيث لا تعتمد على تقييم موحد وإنما تأخذ بعين الاعتبار الفروق الدقيقة بين خطوات الحل.

أما فيما يتعلق بمعاملات التمييز، فتراوحت بين 0.44 و2.07 وهو ما يدل على أن جميع الفقرات قادرة على التمييز بين مستويات أداء الطلبة بدرجة جيدة، مع تفوق بعض الفقرات مثل الفقرة رقم (11) التي حققت أعلى قيمة تمييز، مما يعني أن تلك الفقرة كانت أكثر قدرة على فصل الطلبة ذوي المستويات المختلفة بدقة.

بالمقارنة مع نتائج النظرية الكلاسيكية، التي أظهرت نطاقًا أضيق لقيم الصعوبة والتمييز، تبرز نظرية استجابة الفقرة (IRT) كأداة أكثر دقة وحساسية، خصوصًا في سياق الاختبارات متعددة التدريج. فقد وفرت IRT تقديرات تفصيلية لكل خطوة على حدة، وهو ما لم تستطع النظرية الكلاسيكية تحقيقه، مما يعزز من الاعتماد على IRT في تطوير وتحليل الاختبارات التعليمية المعقدة.

تتوافق نتائج الدراسة مع ما ورد في دراسة بني هاني وآخرين (2021)، التي أكدت كفاءة نموذج التقدير الجزئي في تقدير معالم الصعوبة والتمييز لكل خطوة، مما يعزز دقة التقييم في الاختبارات متعددة التدريج. كما تدعمها نتائج الزهراني وسليمان (2023)، التي أبرزت قدرة النموذج على الكشف عن الفروق الدقيقة بين مستويات الطلبة، وتحسين جودة أدوات القياس. وتتسق أيضًا مع العمراني والمالكي (2023)، اللذين أظهرا تفوق النموذج على نماذج أخرى في تقليل الخطأ المعياري وتوفير تقديرات أكثر دقة.

ورغم هذا الاتفاق، أظهرت الدراسة الحالية تفاوتًا أكبر في قيم الصعوبة مقارنة بدراسة فرحات (2022)، التي أشارت إلى دقة النموذج عند المستويات المنخفضة من القدرة، بينما أظهر اختبار هذه الدراسة توازنًا في استهداف المستويات السهلة والمتوسطة والصعبة، مما يعكس تصميمًا شاملًا.

ويرى الباحث أن هذه النتائج تعزز الثقة في نموذج التقدير الجزئي كإطار تحليلي متقدم لبناء اختبارات رياضيات متعددة التدريج.

النتائج المتعلقة بسؤال الدراسة الرابع: ما تقدير قدرات الأفراد على اختبار الرياضيات متعدد التدريج للصف السابع الأساسي لنموذج التقدير الجزئي؟ تم استخدام البرنامج MULTILOG.7 لتقدير قدرات الأفراد المستجيبين عن فقرات المتعددة التدريج. ويبين جدول (6) القيم الدنيا والعليا

ACCEPTED

والمتوسط الحسابي لتقديرات القدرة في عينة الدراسة لنموذج الاختبار.

جدول (6): القيم الدنيا والعليا والمتوسط الحسابي لتقديرات القدرة في عينة الدراسة لاختبار الرياضيات متعدد التدريج.

عدد افراد العينة	القيمة العليا للقدرة	القيمة الدنيا للقدرة	المتوسط الحسابي
653	2.296	2.083-	0.282

تشير نتائج الجدول (6) أن قدرات الطلبة الذين خضعوا للاختبار تراوحت ما بين -2.083 إلى 2.296 لوجيت، بينما بلغ المتوسط الحسابي لتقديرات القدرة (0.282 لوجيت)، وهو يشير إلى أن مستوى أداء الطلبة بشكل عام كان قريبًا من المتوسط المعياري في سلم التقدير المستخدم الذي يُمثّل عادة بالقيمة 0 في نماذج(IRT).

من وجهة نظر الباحث، يعكس التوزيع المتوازن لقدرات الطلبة في هذه الدراسة فعالية تصميم الفقرات متعددة التدريج، ومدى ملاءمتها لنموذج التقدير الجزئي، حيث أتاح الاختبار تمييزًا دقيقًا بين مستويات الأداء المختلفة ضمن عينة متنوعة في الكفاءة الرياضية. ويُظهر هذا التوازن أن الاختبار لم يكن منحازًا لفئة محددة، بل مكّن جميع الطلبة من التعبير عن قدراتهم، مما يعزز العدالة في التقييم ويؤكد التوجه نحو قياس تربوي إنساني يحترم الفروق الفردية.

وتدعم هذه النتائج صلاحية نموذج التقدير الجزئي في تقديم تقديرات دقيقة وغير متحيزة، كما تؤكد أن بنية الاختبار وتنوع فقراته مكّنت من قياس الفروق الفعلية بين الطلبة، مما يعزز مصداقية نتائجه لأغراض التقويم والتشخيص. وتتفق هذه النتيجة مع دراسات سابقة، منها حمدان (2025)، الزهراني وسليمان (2023)، العنزي وسليمان (2023)، وبني هاني وآخرون (2021)، التي أثبتت قدرة النموذج على تمثيل البيانات بدقة، وتمييز مستويات الأداء، وتحقيق ثبات وصدق مرتفعين.

في المقابل، تختلف هذه الدراسة جزئيًا عن نتائج فرحات (2022)، التي أشارت إلى تفوق النموذج في تمثيل المستويات المنخفضة فقط، بينما أظهر اختبار هذه الدراسة قدرة متوازنة عبر جميع المستويات.

النتائج المتعلقة بسؤال الدراسة الخامس: ما قيمة دالة المعلومات لكل فقرة من فقرات اختبار الرياضيات متعدد التدريج؟ دالة معلومات الفقرة هي دالة تصف مقدار "المعلومات" التي توفرها الفقرة عن قدرة المفحوص عند كل قيمة محتملة من القدرة .(θ) بمعنى آخر، تُظهر هذه الدالة عند أي مستوى من القدرة تكون الفقرة أكثر فعالية في التمييز بين الأفراد.

وقد قام الباحث باستخدام برمجية MULITILOG.7 لحساب القيمة العظمى لدالة المعلومات لكل فقرة $I(\theta)_{max}$ وقيمة القدرة المقابلة للقيمة العظمى لدالة المعومات $(\theta)_{max}$ ويبين الجدول (7) النتائج.

جدول (7): القيمة العظمى لدالة المعلومات والقدرة المقابلة للقيمة العظمى لدالة المعلومات لكل فقرة.

الفقرة	القيمة العظمى لدالة المعلومات $I(\theta)_{max}$	القدرة المقابلة للقيمة العظمى لدالة المعلومات $(\theta)_{max}$	الفقرة	القيمة العظمى لدالة المعلومات $I(\theta)_{max}$	القدرة المقابلة للقيمة العظمى لدالة المعلومات $(\theta)_{max}$
1	0.498	2.0	9	1.122	0.1
2	0.739	0.3	10	0.062	3.0-
3	0.415	1.6	11	1.294	1.0
4	0.302	0.3-	12	0.897	0.2-
5	0.155	0.5-	13	0.375	0.3-
6	0.743	0.2	14	0.516	1.1-
7	0.732	0.1	15	1.067	0.0
8	0.834	1.0-			

أظهرت نتائج السؤال الرابع أن فقرات اختبار الرياضيات متعدد التدريج اختلفت في كفاءتها المعلوماتية تبعًا لمستوى القدرة المستهدف. فقد برزت الفقرة (11) كأعلى الفقرات كفاءة، حيث بلغت قيمة دالتها المعلوماتية (1.294) عند مستوى قدرة متوسط مرتفع (θ = 1.0)، تلتها الفقرة (9) بقيمة (1.122) عند قدرة قريبة من المتوسط (θ = 0.1)، ثم الفقرة (15) بقيمة (1.067) عند قدرة متوسطة تمامًا (θ = 0.0). يشير ذلك إلى أن هذه الفقرات توفر أكبر قدر من الدقة في تقدير القدرات لدى الطلبة الذين يقعون في النطاق المتوسط وما فوقه بقليل، مما يجعلها أدوات قياس موثوقة لهذه الفئة.

في المقابل، أظهرت بعض الفقرات كفاءة منخفضة، مثل الفقرة (10) التي سجلت (0.062) عند قدرة منخفضة جدًا (θ = -3.0)، والفقرة (5) التي سجلت (0.155) عند قدرة منخفضة نسبيًا (θ = -0.5)، وهو ما يعكس ضعف قدرتها على تقديم معلومات دقيقة عن الطلبة ذوي الأداء المنخفض. ويمكن تفسير ذلك بأن هذه الفقرات قد لا تمنح الطلبة ذوي القدرات المحدودة فرصة كافية لإظهار أي أداء جزئي، أو أنها لا تعكس بوضوح المهارات الأساسية التي تناسب مستواهم، الأمر الذي قد يؤدي إلى إحباطهم أو شعورهم بالعجز، مما يستدعي مراجعتها أو إعادة صياغتها لتكون أكثر حساسية لهذه الفئة.

تتوافق نتائج هذه الدراسة مع ما أشار إليه حمدان (2025) من أن فقرات المقياس قد تتلاءم مع أكثر من نموذج تحليلي،

ACCEPTED

إلا أن بعضها يتفوق في دقة تمثيل البيانات، وهو ما انعكس في القيم العظمى لدالة المعلومات لبعض الفقرات هنا. كما تتسق مع الزهراني وسليمان (2023)، اللذين أكدا أن توافق الفقرات مع نموذج التقدير الجزئي لا يعني تجانسها في الكفاءة المعلوماتية، إذ يظل هناك تفاوت في قدرتها على التمييز بين مستويات الأداء.

وتتشابه النتائج أيضًا مع ما ورد في دراسة فرحات (2022)، التي بيّنت أن دقة النموذج تختلف باختلاف مستويات القدرة وطبيعة المحتوى؛ حيث ركزت الفقرات عالية الكفاءة في هذه الدراسة على المستويات المتوسطة والعليا، بينما أظهرت دراسة فرحات تفوقًا عند المستويات المنخفضة في مقياس التأمل الناقد.

في المقابل، تختلف هذه النتائج عن دراسة العنزي وسليمان (2023)، التي أظهرت تجانسًا في الكفاءة المعلوماتية لجميع فقرات مقياس التفكير ما وراء المعرفي، وهو ما لم يتحقق في هذه الدراسة، ويُحتمل أن يعود ذلك إلى طبيعة مادة الرياضيات التي تتسم بتدرج أكبر في الصعوبة وتعقيد المحتوى.

وبناءً على ما سبق، يتضح أن نموذج التقدير الجزئي لا يقتصر دوره على التحقق من ملاءمة الفقرات إحصائيًا، بل يكشف أيضًا عن الفروق في دقتها المعلوماتية، مما يتيح تحسين جودة الاختبارات من خلال الاحتفاظ بالفقرات عالية الكفاءة وتعديل أو استبعاد الفقرات ضعيفة الأداء. وتبرز النتيجة الحالية الحاجة إلى تعزيز شمولية الاختبار بإضافة فقرات تستهدف المستويات الدنيا من القدرة، لضمان عدالة القياس وفعاليته عبر مختلف فئات الطلبة.

ويخلص الباحث إلى أن نتائج هذه الدراسة تعزز كفاءة نموذج التقدير الجزئي كأداة تحليلية متقدمة في بناء اختبارات الرياضيات المدرسية، بما يتماشى مع فلسفة القياس الإنساني التي تثمّن كل مستوى من الأداء، وتبتعد عن الأحكام الثنائية، مما يجعله مناسبًا لدعم التشخيص التربوي وتحسين جودة التعليم.

التوصيات

في ضوء النتائج يوصي الباحث بما يلي:

1. اعتماد نموذج التقدير الجزئي (PCM) في تصميم وبناء الاختبارات التحصيلية في الرياضيات والمجالات المشابهة، نظرًا لما أثبته من قدرة على قياس الأداء المتدرج وتمييز الفروق الفردية بدقة وعدالة.
2. تضمين فقرات متعددة التدريج في الاختبارات، بحيث تقيس خطوات الحل ومراحل التفكير، بدلًا من الاقتصار على التصحيح الثنائي (صحيح/خطأ)، مما يعكس مستوى التفكير التحليلي لدى الطلبة.
3. منح الطلبة وقتًا كافيًا في الاختبارات لتقليل أثر عامل السرعة، بحيث تعكس النتائج قدراتهم الحقيقية.
4. الاستفادة من تقديرات القدرة الفردية الناتجة عن نموذج التقدير الجزئي في التشخيص الأكاديمي ووضع خطط علاجية أو إثرائية لكل طالب وفق مستواه.
5. إجراء دراسات مقارنة بين نموذج التقدير الجزئي (PCM) ونماذج أخرى من نظرية الاستجابة للفقرة، مثل نموذج راش للدرجات المتعددة (RSM) أو نموذج التدرج الجزئي المعدل، لبحث فروق الدقة والملاءمة في مجالات مختلفة.

بيانات الإفصاح

- **الموافقة الأخلاقية:** لا ينطبق
- **توافر البيانات والمواد:** متوافرة لدى الباحث m.dabous@najah.edu
- **مساهمة الباحثين:** البحث كله جهد خالص للباحث.
- **تضارب المصالح:** لا يوجد
- **التمويل**: لا يوجد
- **شكر وتقدير:** يتقدم الباحث بالشكر والتقدير إلى مجلة جامعة النجاح للأبحاث – ب (العلوم الإنسانية) جامعة النجاح الوطنية. نابلس. فلسطين. (www.najah.edu).

Open Access

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third-party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit https://creativecommons.org/licenses/by-nc/4.0/

ACCEPTED

المراجع

- ياسر، حمود، محمود. والرجيبي، يحيى. (2021). استخدام نموذج استجابة الفقرة المتدرجة في تقدير الخصائص السيكومترية لمقياس قلق المستقبل المهني لدى طلبة الجامعات في سلطنة عُمان. *مجلة جامعة النجاح للأبحاث -ب (العلوم الإنسانية)*. *35*(6)، 901-926. https://doi.org/10.35552/0247-035-006-002
- حبيب، صالح. وعزيز، علي. (2018). *بناء المقياس النفسية وفقا لنظرية الاستجابة للفقرة باستعمال الأنموذج الكشفي المعمم*. نابلس: الدار المنهجية للنشر والتوزيع.
- حمدان، عادل سمير محمد. (2025). مقارنة بين نموذجي "التقدير الجزئي" و "سلم التقدير" لمقياس الكفاءة الذاتية في الإدارة الصفية للطالب المعلم. *مجلة العلوم التربوية*، *25*(1)، 151 – 176.
- دبوس، مهند طاهر. (2016). استخدام نظرية الاستجابة للفقرة في بناء فقرات اختبار محكي المرجع في الرياضيات بفقرات ثنائية التدريج ومتعددة التدريج وفق النموذج اللوجستي ثنائي المعلم. *مجلة جامعة النجاح للأبحاث – ب (العلوم الإنسانية)* 30 (7)، 1453 -1480. https://doi.org/10.35552/0247-030-007-007
- دبوس، مهند طاهر. (2009). *تطوير نظام بنك محوسب لفقرات اختبار في الرياضيات متعدد الأبعاد باستخدام فقرات ثنائية التدريج ومتعددة التدريج*. رسالة دكتورة غير منشورة، جامعة عمان العربية للدراسات العليا، الأردن.
- الزهراني، منى بنت محمد. وسليمان، شوق خليل. (2023). تطوير مقياس الدافعية نحو التعلم وفق نموذج التقدير الجزئي: دراسة على عينة من طالبات المرحلة الثانوية في منطقة تبوك. *المجلة الدولية للدراسات التربوية والنفسية*، *12*(1)، 208- 223.
- الشريفين، نعمة كاظم. (2019). أثر نموذج نظرية الاستجابة للفقرة ذات الاستجابة المتعددة التدريج على دقة تقدير القدرات للأفراد والخصائص السيكومترية للفقرات والاختبار. *المجلة التربوية*، *33* (130)، 241- 295.
- الشريفين، نعمة كاظم. وبني عطا، زيد صادق. (2013). تقصي أثر عدد خطوات الأسئلة متعددة التدريج وشكل التوزيع لصعوبتها على تقديرات القدرة للأفراد والصعوبة للأسئلة ودالة المعلومات للاختبار وفق نموذج التقدير الجزئي. *المجلة التربوية، جامعة الكويت- مجلس النشر العلمي*، *28* (109)، 275 – 213.
- طنطاوي، شوق فخر. (2022). تأثيرات طول الاختبار وطريقة التصحيح ونموذج تحليل المفردة على تقديرات القدرة ومعايير التحسين للاختبارات ذات المفردات المختلطة. *مجلة الدراسات النفسية المعاصرة*، *(4)* 1، 38 -97.
- علام، صالح. (2012). *نماذج الاستجابة للمفردة الاختبارية*. القاهرة: دار الفكر.
- العمراني، عادل غازي، والمالكي، ذياب عيسى. (2023). أثر ترتيب فئات التدريج ونوع النموذج المستخدم على دقة تقدير بارامترات الفقرات والأفراد. *مجلة كلية التربية - جامعة طنطا*، *89* (3)، 1120 – 1166.
- العنزي، تثليث ناصر. وسليمان، شوق خليل. (2023). تطوير مقياس التفكير ما وراء المعرفي وفق نموذج التقدير الجزئي: دراسة على عينة من طلبة جامعة تبوك. *دراسات عربية في التربية وعلم النفس*، *147* (2)، 23 -52.
- فرحات، رمزي سعيد. (2022). بناء وتخريج مقياس التأمل الناقد في ضوء نموذج التقدير الجزئي والاستجابة المتدرجة لدى طلاب جامعة الأزهر. *المجلة المصرية للدراسات النفسية 32* (114)، 135 - 208.
- وكالة الغوث الدولية. (2017). *تحليل أهداف منهاج الرياضيات للصف السابع - الفصل الدراسي الأول*. مركز التطوير التربوي، وحدة التطوير المهني والمنهاج.

References

- Abu Shindi, Y., Hamoud, M. S., & Alrojaibi, Y. (2021). Using the graded response model in estimating the psychometric properties of the career future anxiety scale among universities students in the Sultanate of Oman. *An-Najah University Journal for Research - B (Humanities)*, *35*(6), 901–926. https://doi.org/10.35552/0247-035-006-002
- Al-Enezi, T. N., & Suleiman, S. K. (2023). Developing a metacognitive thinking scale according to the partial credit model: A study on a sample of University of Tabuk students. *Arab Studies in Education and Psychology*, *147*(2), 23–52.
- Allam, S. M. (2012). *Item response models*. Cairo: Dar Al-Fikr.
- Al-Omrani, A. G., & Al-Maliki, D. A. (2023). The effect of the order of grading categories and the type of model used on the accuracy of estimating item and individual parameters. *Journal of the Faculty of*

ACCEPTED

Review of Statistics and Its Application, 3, 297–321. https://doi.org/10.1146/annurev-statistics-041715-033702

– Dabbous, M. (2016). Using Item Response Theory in Constructing the Item Pool in Criterion- Referenced Testing with Dichotomous and Polytomous Items according to Two – Parameter Logistic Model. *An-Najah University Journal for Research - B (Humanities), 30*(7), 1453–1480. https://doi.org/10.35552/0247-030-007-007

– Dabous, M, T, (2009*). Developing a Computerized Banking System for Multidimensional Mathematics Test Items Utilizing Dichotomous and Polytomous Items*. Unpublished doctoral dissertation, Amman Arab University for Graduate Studies, Jordan.

– Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Psychology Press. New York, United States of America.

– Farhat, R. S. (2022). Constructing and validating a critical reflection scale in light of the partial credit model and the graded response model among Al-Azhar University students. *The Egyptian Journal of Psychological Studies, 32*(114), 135–208.

– Habeeb, S., & Aziz, S. (2018). Constructing psychological scales according to item response theory using the generalized diagnostic model. Nablus: Al-Manhajiah Publishing and Distribution House.

– Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Kluwer. Nijhoff Publishing, Boston.

– Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Sage Publications, Inc.

– Hamdan, A. S. M. (2025). A comparison between the "partial credit" and "rating *Education - Tanta University, 89*(3), 1120–1166.

– Alsharaifin, N., & Bani – Ata, Z. (2013). An investigation of the effect of number of steps and difficulty distribution of Polytomous Test Question on Persons Ability, Questions Difficulty Estimations and Information Function of the Test According to Partial Credit Model. *The Educational Journal, Kuwait University – Academic Publication Council*, 28(109), 213–275.

– Al-Shurayfeen, N. K. (2019). The effect of the polytomous item response theory model on the accuracy of estimating individuals' abilities and the psychometric properties of items and tests. *The Educational Journal, 33*(130), 241–295.

– Al-Zahrani, M. M., & Suleiman, S. K. (2023). Developing a motivation-to-learn scale according to the partial credit model: A study on a sample of female secondary school students in the Tabuk region. *International Journal of Educational and Psychological Studies, 12*(1), 208–223.

– Ayanwale, M. A. (2021). Calibration of Polytomous Response Mathematics Achievement Test Using Generalized Partial Credit Model of Item Response Theory. *Journal of Science, Mathematics and Technology, 8*(1), 57-69.

– Bani Hani, A., Talib, R., Zrekat, M., Nasir, M. A., Al_Ahmad, A., & Wedian, N. (2021). Development and validation of the mathematics test for tenth grade Jordanian students, according to the partial credit model. *Turkish Journal of Computer and Mathematics Education, 12*(6), 1527–1536.

– Brookhart, S. M. (2016). *How to make decisions with different kinds of student assessment data*. ASCD, United States of America.

– Cai, L., Choi, K., Hansen, M., & Harrell, L. (2016). Item response theory. *Annual*

ACCEPTED

model on ability estimates and improvement criteria for tests with mixed items. *Journal of Contemporary Psychological Studies, 4*(1), 38-97.

– Tusiime, J., Bangsberg, D., & Mark, W. (2015). Examining the psychometric of the Beck depression inventory – II using an item response modeling approach in an HIV infected population in Kampala, Uganda. *Journal of Depression and Anxiety, 4*(2), 1–8.

– Tutz, L. (2018). Response styles in the partial credit model. *Applied Psychological Measurement, 42*(6), 407–427.

– United Nations Relief and Works Agency (UNRWA). (2017). *Analysis of Mathematics Curriculum Objectives for Grade Seven – First Semester*. Educational Development Center, Professional Development and Curriculum Unit.

– Warm, T.A. (1978). *A Primer of item Response Theory*. Oklahoma: U .S. Coast Guard Institute 73/69.

scale" models for a classroom management self-efficacy scale for student teachers. *Journal of Educational Sciences, 25*(1), 151–176.

– Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Lawrence Erlbaum Associates, New York. United States of America.

– Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*(2), 149–174. https://doi.org/10.1007/BF02296272

– Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph*, 17.

– Schraw, G., & Robinson, D. H. (Eds.). (2015). *Assessment of Higher Order Thinking Skills. Current Perspectives on Cognition, Learning and Instruction*. IAP - Information Age Publishing, Inc. United States of America.

– Tantawi, Sh. F. (2022). Effects of test length, scoring method, and item analysis

ACCEPTED