



ISSN: 2957-3874 (Print)

Journal of Al-Farabi for Humanity Sciences (JFHS)

<https://iasj.rdd.edu.iq/journals/journal/view/95>

مجلة الفارابي للعلوم الإنسانية تصدرها جامعة الفارابي



استخدام نموذج الاستجابة المتدرجة (GRM) في تقدير الخصائص القياسية لمقياس (وودكوك جونسون)

للقدرات المعرفية (الإصدار الخامس) لدى طلبة الجامعة

م. د. عمار عبد مظلوم

وزارة التربية / الكلية التربوية المفتوحة - قاطع الكرخ - فرع شمال بغداد

Using the Graded Response Model (GRM) to Estimate the Psychometric Properties of the Woodcock–Johnson V Cognitive Abilities Scale Among University Students

Assist. Prof. Ammar Abdul Mazloom, PhD

Ministry of Education / Open Educational College – Al-Karkh Sector – North Baghdad Branch

Specialization: Measurement and Evaluation

Email: a07709268304@gmail.com

## المستخلص

يهدف البحث الحالي إلى استخراج الخصائص القياسية لمقياس القدرات المعرفية المستند إلى بطارية Woodcock–Johnson V لدى طلبة الجامعة في ضوء نموذج الاستجابة المتدرجة (Graded Response Model – GRM) ضمن إطار نظرية الاستجابة للمفردة. ويأتي هذا التوجه استجابة للتحويلات المعاصرة في ميدان القياس النفسي والتربوي التي تؤكد ضرورة توظيف النماذج الاحتمالية الحديثة في تحليل الاختبارات المعرفية، بما يسهم في تعزيز دقة تفسير الأداء والكشف عن الفروق الفردية في القدرات العقلية.

اعتمد البحث المنهج الوصفي التحليلي، وتكونت عينة الدراسة من (٦٠٠) طالب وطالبة من طلبة جامعة بغداد تم اختيارهم بطريقة العينة العشوائية الطبقية لضمان تمثيل متوازن لمتغيري الجنس والتخصص الأكاديمي. وتم استخدام اختبار من بطارية Woodcock–Johnson V بوصفها أداة لقياس القدرات المعرفية، حيث جرى تحليل استجابات أفراد العينة باستخدام نموذج الاستجابة المتدرجة من خلال برنامج (R-Imirt) الإحصائي المتخصص بنماذج نظرية الاستجابة للمفردة.

أظهرت نتائج التحليل تحقق افتراضات النموذج الأساسية، وفي مقدمتها أحادية البعد والاستقلال الموضعي وملاءمة النموذج للبيانات. كما بينت النتائج أن معاملات التمييز للفقرات جاءت ضمن مستويات جيدة ومرتفعة، مما يدل على قدرة الفقرات على التفريق بين الطلبة ذوي مستويات القدرة المختلفة. وأظهرت تقديرات العتبات تدرجاً منطقياً عبر مستويات الاستجابة، الأمر الذي يعكس انتظام البناء الداخلي للفقرات.

كما أوضحت دالة معلومات المقياس أن أعلى مستويات الدقة في القياس تحققت عند المستويات المتوسطة من القدرة الكامنة، مع انخفاض تدريجي في الأطراف العليا والدنيا، وهو نمط سيكومترى متوقع في المقاييس المعرفية. وقد بلغ معامل الثبات المستند إلى دالة المعلومات مستوى مرتفعاً، مما يشير إلى تمتع المقياس بدرجة عالية من الدقة والثبات في القياس. وفي ضوء النتائج، خلص البحث إلى صلاحية مقياس القدرات المعرفية المستند إلى بطارية Woodcock–Johnson V للاستخدام في البيئة الجامعية العراقية، وإلى ملاءمة نموذج الاستجابة المتدرجة في تحليل فقراته وتقدير خصائصه القياسية. وأوصى البحث بضرورة توظيف نماذج نظرية الاستجابة للمفردة في تقويم الأدوات التربوية والنفسية، لما توفره من مؤشرات دقيقة تسهم في تطوير أدوات القياس وتحسين جودة القرارات التربوية.

الكلمات المفتاحية: نظرية الاستجابة للمفردة، نموذج الاستجابة المتدرجة، القدرات المعرفية، Woodcock–Johnson V، الخصائص السيكومترية.

Abstract (English)

The present study aims to estimate the psychometric properties of a cognitive abilities scale based on the Woodcock–Johnson V battery among university students using the Graded Response Model (GRM) within the framework of Item Response Theory (IRT). This approach responds to contemporary developments in educational and psychological measurement that emphasize the importance of probabilistic models in analyzing cognitive tests to enhance the accuracy of performance interpretation and identify individual differences in cognitive abilities.

A descriptive analytical method was adopted. The study sample consisted of 600 undergraduate students from the University of Baghdad, selected using stratified random sampling to ensure balanced representation of gender and academic specialization. A set of subtests from the Woodcock–Johnson V battery was used as the measurement instrument. Participants' responses were analyzed using the Graded Response Model through the specialized statistical software (R-mirt) for Item Response Theory models.

The results indicated that the main assumptions of the model were met, including unidimensionality, local independence, and model-data fit. The findings revealed that item discrimination parameters ranged from moderate to high levels, indicating that the items effectively differentiated among students with varying levels of cognitive ability. Threshold estimates also showed a logical progression across response categories, reflecting a coherent internal structure of the items.

Furthermore, the test information function demonstrated that the highest measurement precision was achieved at the المتوسط levels of latent ability, with a gradual decline at the extreme levels, which is consistent with expected psychometric patterns in cognitive measures. The reliability coefficient derived from the information function indicated a high level of measurement accuracy and consistency.

In light of these findings, the study concludes that the cognitive abilities scale based on the Woodcock–Johnson V battery is suitable for use within the Iraqi university context and that the Graded Response Model is appropriate for analyzing its items and estimating their psychometric characteristics. The study recommends adopting Item Response Theory models in evaluating educational and psychological instruments due to their capacity to provide precise indicators that contribute to improving measurement tools and educational decision-making.

Keywords: Item Response Theory, Graded Response Model, Cognitive Abilities, Woodcock–Johnson V, Psychometric Properties.

## الفصل الأول: الإطار العام للبحث

### أولاً: مشكلة البحث

شهد ميدان القياس النفسي والتربوي تطوراً ملحوظاً في نماذج تفسير القدرات العقلية وأساليب تقويمها، إذ لم يعد قياس الأداء المعرفي يقتصر على استخراج درجات معيارية كلية، بل اتجهت البحوث المعاصرة إلى تحليل البنية الكامنة للقدرات في ضوء نماذج نظرية متقدمة تتيح فهماً أعمق للفروق الفردية بين الأفراد. وقد أسهمت نماذج نظرية الاستجابة للمفردة في إحداث تحول نوعي في تفسير نتائج الاختبارات المعرفية، من خلال اعتمادها على نماذج احتمالية تسمح بتقدير معالم الفقرات ومستويات القدرة الكامنة بصورة أكثر دقة وموضوعية (Embretson & Reise, 2013, p. 9).

وفي سياق تقدير القدرات المعرفية، تُعد بطارية **Woodcock–Johnson V** من الأدوات المعاصرة التي تستند إلى نموذج كاتل-هورن-كارول (CHC)، إذ صُممت لقياس مجموعة من القدرات العقلية في إطار بنيوي متكامل. وقد أثبتت هذه البطارية فاعليتها في تقدير الأداء المعرفي عبر فئات عمرية متعددة، لما تتضمنه من اختبارات فرعية متنوعة تعكس أبعاداً مختلفة للقدرة العقلية (Schränk et al., 2023, p. 23). ومع اتساع استخدام هذه البطارية في البيئات التعليمية والبحثية، برزت الحاجة إلى تقويم خصائصها القياسية باستخدام نماذج تحليل حديثة تتجاوز مؤشرات القياس التقليدية. فاعتماد الدرجة الكلية وحدها لا يوفر تفسيراً كافياً لطبيعة الأداء أو كفاءة الفقرات، إذ يتطلب الأمر تحليل سلوك كل فقرة في ضوء العلاقة الاحتمالية بين القدرة الكامنة والاستجابة الفعلية.

ويُعد نموذج الاستجابة المترتبة (GRM) من النماذج الملائمة لتحليل الفقرات متعددة الفئات، إذ يسمح بتقدير معاملات التمييز والعتبات لكل فقرة، إضافة إلى تحديد دالة المعلومات التي توضح مستوى الدقة في القياس عبر متصل القدرة. ومن ثم، فإن توظيف هذا النموذج في تحليل بطارية Woodcock–Johnson V يساهم في الكشف عن مدى كفاءتها في تمثيل الفروق الفردية في القدرات المعرفية لدى طلبة الجامعة.

في ضوء ذلك، تتحدد مشكلة البحث الحالي في استخراج الخصائص القياسية لمقياس Woodcock-Johnson V لدى طلبة الجامعة باستخدام نموذج الاستجابة المتدرجة، للتحقق من مدى ملاءمته لقياس القدرات المعرفية في البيئة الجامعية، والكشف عن درجة دقته في تمثيل الفروق الفردية بين الطلبة.

#### ثانياً: أهمية البحث

تتبع أهمية البحث من التحولات المعاصرة في ميدان القياس النفسي، والتي تؤكد ضرورة توظيف نماذج تحليل حديثة تعزز دقة تفسير الأداء المعرفي. فاعتماد نماذج نظرية الاستجابة للمفردة يسمح بتقدير معالم الفقرات بصورة أكثر استقراراً مقارنة بالمؤشرات التقليدية، كما يتيح تفسير الأداء في ضوء القدرة الكامنة بدلاً من الاقتصار على الدرجة الكلية .

كما تتجلى أهمية البحث في كونه يساهم في تقييم إحدى الأدوات المعرفية المعاصرة المستخدمة في تشخيص القدرات العقلية، من خلال تحليل فقراتها في ضوء نموذج الاستجابة المتدرجة. ويتيح هذا التوجه تحديد مدى كفاءة الفقرات في التمييز بين مستويات القدرة المختلفة، والكشف عن مستوى الدقة الذي يوفره المقياس عبر متصل القدرة.

ومن الناحية التطبيقية، يمكن الاستفادة من نتائج البحث في دعم القرارات التربوية المرتبطة بتقويم الأداء المعرفي لدى الطلبة، وتطوير أدوات قياس أكثر دقة تساهم في تحسين مخرجات العملية التعليمية. كما يعزز البحث توظيف نماذج القياس الحديثة في الدراسات العربية، بما ينسجم مع الاتجاهات العالمية في ميدان القياس والتقييم.

#### ثالثاً: أهداف البحث

يسعى البحث الحالي إلى تحقيق الأهداف الآتية:

١. تقدير معاملات فقرات مقياس Woodcock-Johnson V وفق نموذج الاستجابة المتدرجة.
٢. التحقق من تحقق افتراضات النموذج وملاءمته لبيانات العينة.
٣. تقدير القدرة الكامنة لأفراد العينة في ضوء النموذج المعتمد.
٤. تحديد مستوى المعلومات التي يوفرها المقياس عبر مستويات القدرة المختلفة.

#### رابعاً: حدود البحث

يتحدد البحث الحالي بالحدود الآتية:

- الحدود البشرية: طلبة جامعة بغداد.
- الحدود المكانية: الكليات التابعة لجامعة بغداد.
- الحدود الزمانية: العام الدراسي ٢٠٢٤-٢٠٢٥.
- الحدود الموضوعية: الخصائص القياسية لمقياس Woodcock-Johnson V باستخدام نموذج الاستجابة المتدرجة.

#### خامساً: تحديد المصطلحات

#### • القدرات المعرفية Schneider & McGrew

- تشير إلى مجموعة العمليات العقلية المرتبطة بالاستدلال والتحليل ومعالجة المعلومات واتخاذ القرار (Schneider & McGrew, 2018, p. 79).

التعريف الإجرائي: مستوى الأداء الذي يحققه المفحوص في مقياس Woodcock-Johnson V المستخدم في الدراسة .

#### مقياس Woodcock-Johnson V (Schrank et al)

بطارية اختبارات معرفية طوّرت في ضوء نموذج CHC لقياس مجموعة من القدرات العقلية العامة والفرعية (Schrank et al., 2023, p. 23)

التعريف الإجرائي: مجموعة الاختبارات الفرعية المطبقة على عينة الدراسة لغرض تحليل خصائصها القياسية.

#### • نموذج الاستجابة المتدرجة (GRM) Samejima

أحد نماذج نظرية الاستجابة للمفردة المستخدمة في تحليل الفقرات متعددة الفئات، ويعتمد على تقدير معامل التمييز والعتبات لكل فقرة (Samejima, 2016, p. 64).

**التعريف الإجرائي:** النموذج الإحصائي المستخدم في تحليل بيانات الدراسة لتقدير معلمات الفقرات والقدرة الكامنة.

**الفصل الثاني: الإطار النظري ودراسات سابقة**

**أولاً: تطور قياس القدرات المعرفية في ضوء النماذج الحديثة**

شهد ميدان القياس النفسي خلال العقود الأخيرة تحولاً منهجياً جوهرياً في فهم طبيعة القدرات العقلية وطرائق تقويمها. فبعد أن كان التركيز منصباً على تقدير الذكاء بوصفه بُعداً أحادياً يعكس القدرة العامة، اتجهت البحوث المعاصرة إلى تبني نماذج بنيوية تفترض تعددية المستويات وتداخل العمليات المعرفية ضمن منظومة هرمية أكثر تعقيداً. ويُعد هذا التحول امتداداً لتراكم نظري طويل بدأ بالنماذج العاملية المبكرة وصولاً إلى النماذج التكاملية المعاصرة. (Schneider & McGrew, 2018, p. 77)

وتشير الأدبيات الحديثة إلى أن دقة تفسير الأداء العقلي لا تتحقق إلا عند تحليل القدرات في ضوء بنيتها الهرمية، بحيث يتم التمييز بين العامل العام والقدرات العريضة والقدرات النوعية، الأمر الذي يسمح بفهم أكثر عمقاً لطبيعة الفروق الفردية في الأداء المعرفي (Flanagan & Dixon, 2014, p. 120).

ومن هذا المنطلق، لم يعد قياس القدرات المعرفية مجرد عملية حساب درجة كلية، بل أصبح عملية تحليلية تستهدف توصيف البنية الكامنة للأداء العقلي في ضوء إطار نظري متكامل.

**ثانياً: نموذج كاتل-هورن-كارول (CHC) كبنية تفسيرية للقدرات**

يُعد نموذج CHC من أكثر الأطر النظرية شمولاً في تفسير بنية الذكاء، إذ يجمع بين إسهامات كاتل وهورن وكارول ضمن نموذج هرمي متعدد المستويات. ويفترض النموذج وجود عامل عام في قمة الهرم، تندرج تحته مجموعة من القدرات العريضة، والتي تنبثق بدورها إلى قدرات أكثر تخصصاً. (Schneider & McGrew, 2018, p. 79)

وتشمل القدرات العريضة في هذا النموذج الاستدلال السائل، والقدرة الكمية، وسرعة المعالجة، والذاكرة العاملة، والفهم اللفظي، وغيرها من العمليات المعرفية التي تمثل مكونات الأداء العقلي العام. ويتيح هذا التصور الهرمي تفسير التباين في الأداء من خلال تحديد موضع الخلل أو القوة داخل البنية المعرفية. (Flanagan & Alfonso, 2017, p. 54)

وقد أصبح نموذج CHC مرجعاً أساسياً في بناء بطاريات القياس المعرفي الحديثة، إذ يوفر إطاراً نظرياً يسمح بتصميم اختبارات تعكس البنية الفعلية للقدرات العقلية بدلاً من الاقتصار على تقدير مؤشر عام غير مفصل.

**ثالثاً: بطارية Woodcock-Johnson V في ضوء نموذج CHC**

تم تطوير بطارية Woodcock-Johnson V استناداً إلى نموذج CHC، بحيث تغطي مجموعة من القدرات العريضة والقدرات الفرعية المرتبطة بها. وتتميز هذه البطارية بكونها أداة معيارية شاملة تسمح بتقدير الأداء المعرفي في مجالات متعددة ضمن إطار موحد (Schrang et al., 2023, p. 23).

وتتضمن البطارية اختبارات فرعية تقيس الاستدلال السائل، والمعالجة البصرية، والذاكرة العاملة، والقدرة الكمية، وغيرها من المكونات المعرفية التي ترتبط بالأداء الأكاديمي والمهني. ويُعد هذا التنوع في المحتوى عاملاً مهماً في تعزيز صلاحية الأداة لتفسير الأداء المعرفي بصورة شمولية (Schrang et al., 2023, p. 25).

غير أن الاعتماد على الدرجات المعيارية وحدها قد لا يكون كافياً للحكم على كفاءة الفقرات في تمثيل القدرة الكامنة، إذ يتطلب الأمر توظيف نماذج تحليل حديثة تكشف عن طبيعة استجابة المفحوصين لكل فقرة على حدة.

**رابعاً: من القياس التقليدي إلى نظرية الاستجابة للمفردة**

تقوم نظرية القياس التقليدية على افتراض أن الدرجة الكلية تمثل مؤشراً مباشراً للقدرة، وأن خصائص الفقرة (كالتمييز والصعوبة) تُستمد من تفاعلها مع العينة. غير أن هذا الافتراض يجعل معلمات الفقرات عرضة للتغير عند اختلاف خصائص العينة، مما يحد من استقرارها عبر التطبيقات المختلفة. (Embretson & Reise, 2013, p. 9)

في المقابل، تستند نظرية الاستجابة للمفردة إلى إطار احتمالي يفترض وجود علاقة رياضية بين القدرة الكامنة للفرد واحتمال استجابته لفقرات الاختبار. ويتيح هذا الإطار تقدير معلمات الفقرة بصورة أقل تأثراً بخصائص العينة، مما يعزز موضوعية القياس (Hambleton, Swaminathan, & Rogers, 2015, p. 21).

ويمثل هذا الانتقال من المنظور الوصفي إلى المنظور الاحتمالي تحولاً جوهرياً في تفسير الأداء، إذ لم يعد التركيز منصّباً على الدرجة الكلية وحدها، بل أصبح الاهتمام موجهاً نحو تحليل سلوك الفقرة ودورها في بناء دالة المعلومات الكلية للاختبار.

#### **خامساً: افتراضات نماذج نظرية الاستجابة للمفردة**

يعتمد تطبيق نماذج IRT على مجموعة من الافتراضات النظرية التي تضمن دقة التقدير الإحصائي للمعلمات. ويتمثل أول هذه الافتراضات في أحادية البعد، أي أن الفقرات تقيس سمة كامنة واحدة مهيمنة تفسر الترابط بين الاستجابات. (Reckase, 2009, p. 84) أما الافتراض الثاني فيتمثل في الاستقلال الموضوعي، والذي يفترض أن استجابة المفحوص لأي فقرة تعتمد فقط على مستوى قدرته الكامنة، ولا تتأثر باستجاباته للفقرات الأخرى بعد تثبيت القدرة. (De Ayala, 2009, p. 45) ويُضاف إلى ذلك افتراض ملاءمة النموذج، الذي يشير إلى ضرورة توافق البيانات الفعلية مع النموذج الرياضي المستخدم، ويتم التحقق منه من خلال مؤشرات المطابقة الإحصائية. (Baker & Kim, 2017, p. 31)

#### **سادساً: أنموذج الاستجابة المتدرجة (GRM) في تحليل الفقرات متعددة الفئات**

يُصنف أنموذج الاستجابة المتدرجة ضمن النماذج اللوجستية متعددة الفئات في إطار نظرية الاستجابة للمفردة، وقد طوّره ساميجيما لتحليل الفقرات التي تعتمد على استجابات رتبية متدرجة. ويتميز هذا النموذج بقدرته على توصيف احتمالات الانتقال بين فئات الأداء بصورة تراكمية، بما يعكس الطبيعة التدريجية للاستجابة. (Samejima, 2016, p. 64) ويفترض النموذج أن لكل فقرة معامل تمييز يعكس مدى حساسية الفقرة للفروق الفردية في القدرة الكامنة، إضافة إلى مجموعة من العتبات تمثل مستويات الانتقال بين فئات الاستجابة. ويُعد هذا التصور أكثر مرونة مقارنة بالنماذج الثنائية، لأنه يسمح بفهم أعمق للبنية الداخلية للفقرات متعددة الدرجات. (De Ayala, 2009, p. 150)

وتكمن أهمية GRM في كونه لا يكفي بتحديد احتمال اختيار فئة معينة، بل يقوم أولاً بحساب الاحتمال التراكمي للاستجابة عند أو أعلى من فئة محددة، ثم يُشتق احتمال كل فئة على حدة، الأمر الذي يوفر توصيفاً أكثر دقة للسلوك الاستجابي عبر مستويات القدرة المختلفة (Baker & Kim, 2017, p. 189).

#### **سابعاً: تفسير معامل التمييز والعتبات في GRM**

يمثل معامل التمييز (a) في أنموذج الاستجابة المتدرجة مؤشراً على مدى قدرة الفقرة على التفريق بين الأفراد ذوي المستويات المختلفة من القدرة. فكلما ارتفعت قيمة هذا المعامل، ازدادت حدة انحدار المنحنى اللوجستي، مما يدل على حساسية أكبر للفروق الفردية. (Embretson & Reise, 2013, p. 114).

أما العتبات (b)، فهي تمثل نقاطاً على متصل القدرة تشير إلى المستوى الذي تتساوى عنده احتمالية الاستجابة في فئتين متتاليتين. ويعكس التوزيع المنطقي للعتبات تدرجاً مناسباً في مستويات الصعوبة، بما يضمن تغطية نطاق واسع من القدرة الكامنة. (Samejima, 2016, p. 70) وعند تحليل نتائج المقياس في ضوء GRM، يُنظر إلى توازن معاملات التمييز وتوزيع العتبات بوصفه مؤشراً على جودة الفقرة وكفاءتها في تمثيل البعد المعرفي المستهدف.

#### **ثامناً: دالة المعلومات والخطأ المعياري للمقياس**

تُعد دالة المعلومات أحد أهم المخرجات في نماذج IRT، إذ توضح مقدار الدقة التي يوفرها الاختبار عند كل مستوى من مستويات القدرة. وترتبط المعلومات بعلاقة عكسية مع الخطأ المعياري، بحيث تؤدي زيادة المعلومات إلى انخفاض الخطأ في تقدير القدرة. (DeMars, 2018, p. 118) ويمثل هذا المفهوم تطوراً نوعياً مقارنة بمعامل الثبات التقليدي، لأنه لا يفترض ثبات الدقة عبر جميع مستويات القدرة، بل يحدد بدقة النطاق الذي يكون فيه القياس أكثر كفاءة. وغالباً ما تظهر المقاييس المعرفية أعلى درجات المعلومات في المستويات المتوسطة من القدرة، مع انخفاض تدريجي عند الأطراف. (Baker & Kim, 2017, p. 190)

ومن هذا المنظور، فإن تحليل دالة المعلومات لمقياس Woodcock-Johnson V يتيح فهماً أعمق لمدى ملاءمته لقياس الطلبة ذوي القدرات المختلفة، وليس فقط إصدار حكم عام حول ثباته.

#### **تاسعاً: أهمية ملاءمة النموذج للبيانات**

تعد ملاءمة النموذج خطوة أساسية في تطبيق GRM ، إذ ينبغي التأكد من توافق البيانات الفعلية مع الافتراضات الرياضية للنموذج قبل تفسير المعلمات. ويتم ذلك من خلال مؤشرات مطابقة كلية وجزئية تكشف عن مدى انحراف الاستجابات عن التوقعات النظرية (De Ayala, 2009, p. 172).

ويُشير الباحثون إلى أن تجاهل فحص المطابقة قد يؤدي إلى تقديرات غير دقيقة للقدرة الكامنة أو معاملات الفقرات، مما يضعف صدق الاستدلالات المستندة إلى نتائج التحليل. (Hambleton et al., 2015, p. 107)

وبناءً على ذلك، فإن تقدير الخصائص القياسية لمقياس Woodcock-Johnson V في هذه الدراسة يتضمن التحقق المنهجي من ملاءمة النموذج قبل الانتقال إلى تفسير المعلمات أو دالة المعلومات. دراسات سابقة

١. دراسة (Cao & Tay (2021):

هدفت دراسة (Cao & Tay (2021) إلى تقويم الخصائص السيكومترية لمقياس مهارات التفكير التحليلي لدى طلبة الجامعة باستخدام نموذج الاستجابة المتدرجة (Graded Response Model) ، وذلك لكون فقرات المقياس متعددة الفئات. (Likert-type) اعتمدت الدراسة المنهج الوصفي التحليلي، وتكونت العينة من (١,٢٠) طالباً جامعياً من تخصصات علمية وإنسانية مختلفة.

استخدم الباحثان نموذج الاستجابة المتدرجة (GRM) بوصفه أحد نماذج نظرية الاستجابة للفقرة الملائمة للفقرات متعددة الدرجات، حيث يسمح بتقدير معاملات التمييز (a) وحدود الصعوبة ( $b_1, b_2, b_3 \dots$ ) لكل فئة استجابة. وقد تم التحقق أولاً من افتراض أحادية البعد باستخدام التحليل العاملي الاستكشافي والتوكيدي، ثم جرى تطبيق نموذج GRM عبر برنامج (IRTpro).

أظهرت النتائج أن معاملات التمييز تراوحت بين (٠.٨٥ - ٢.١٠)، مما يشير إلى قدرة الفقرات على التفريق بين الطلبة ذوي المستويات المختلفة من التفكير التحليلي. كما بينت حدود الصعوبة تدرجاً منطقياً بين فئات الاستجابة، وهو ما يؤكد انتظام البنية التدريجية للفقرات. وأظهرت دالة المعلومات أن المقياس يوفر أعلى دقة قياس عند المستويات المتوسطة والعليا من القدرة.

كما بلغ معامل الثبات الهامشي (Marginal Reliability) (0.93) ، وهي قيمة مرتفعة تدل على كفاءة المقياس. وأوصت الدراسة باستخدام GRM في تقنين المقاييس المعرفية متعددة الفئات لما يوفره من تقدير دقيق لخصائص الفقرة عبر كل مستوى استجابة (Cao & Tay, 2021, p. 214)

٢. دراسة (Reynolds & Keith (2022):

هدفت دراسة (Reynolds & Keith (2022) إلى تحليل الخصائص القياسية لمقياس الوظائف التنفيذية لدى طلبة الجامعة باستخدام نموذج الاستجابة المتدرجة (GRM) ، مع التركيز على فحص ملاءمة النموذج واستخراج دالة المعلومات وتقدير الخطأ المعياري للمقياس. تكونت عينة الدراسة من (٩٨٠) طالباً جامعياً، وتم استخدام مقياس مكون من (٤٥) فقرة وفق تدرج خماسي. ولأن طبيعة الفقرات تعتمد على استجابات مرتبة، فقد اعتمد الباحثان نموذج GRM لتحليل البيانات عبر برنامج (Mplus).

أظهرت النتائج تحقق افتراض أحادية البعد والاستقلال الموضوعي. كما تراوحت معاملات التمييز بين (١.٠٢ - ٢.٣٤)، مما يشير إلى قدرة قوية للفقرات على التمييز بين مستويات الوظائف التنفيذية. أما حدود الصعوبة فقد أظهرت تدرجاً منطقياً ومنظماً بين فئات الاستجابة، مما يعكس كفاءة التدرج في المقياس.

وأظهرت دالة المعلومات أن المقياس يوفر أعلى دقة قياس في المستويات المتوسطة من القدرة، بينما يرتفع الخطأ المعياري عند الأطراف، وهو نمط متوقع في نماذج IRT. كما بلغ معامل الثبات الهامشي (٠.٩٥)، مما يدل على مستوى عالٍ من الدقة السيكومترية.

وأكدت الدراسة أن نموذج الاستجابة المتدرجة يعد من أكثر النماذج ملاءمة لتحليل المقاييس المعرفية متعددة الفئات، ويوفر مؤشرات أدق من النظرية التقليدية للمقياس (Reynolds & Keith, 2022, p. 167)

الفصل الثالث: منهجية البحث وإجراءاته

أولاً: التصميم المنهجي للدراسة

اعتمدت الدراسة الحالية تصميماً وصفيّاً تحليلياً يستهدف تقدير الخصائص القياسية لأداة قياس معرفي في ضوء نموذج احتمالي حديث. ويقوم هذا النوع من التصميم على تحليل استجابات الأفراد كما وردت ميدانياً، ثم توظيف نماذج إحصائية متقدمة لاستخلاص المعلمات الكامنة المرتبطة بكل من الفقرات والأفراد.

وقد تم اختيار نموذج الاستجابة المترجة (GRM) بوصفه الإطار التحليلي الملائم لطبيعة فقرات المقياس متعددة الفئات، إذ يتيح هذا النموذج تقدير معاملات التمييز والعتبات لكل فقرة، إضافة إلى اشتقاق دالة المعلومات والخطأ المعياري للقياس عبر مستويات القدرة المختلفة (Embretson & Reise, 2013, p. 114).

ويمثل هذا التوجه امتداداً للمنظور الحديث في القياس النفسي، الذي يركز على تحليل سلوك الفقرة في ضوء العلاقة الاحتمالية بين القدرة الكامنة والاستجابة الفعلية، بدلاً من الاقتصار على مؤشرات الدرجة الكلية.

#### ثانياً: مجتمع الدراسة

تكون مجتمع الدراسة من طلبة جامعة بغداد في العام الدراسي ٢٠٢٤-٢٠٢٥، ويتميز هذا المجتمع بتنوع خلفياته الأكاديمية والمعرفية، الأمر الذي يوفر نطاقاً واسعاً من مستويات القدرة المعرفية اللازمة لتقدير معالم النموذج بدقة.

ويُعد وجود تباين كافٍ في مستويات القدرة شرطاً مهماً لتطبيق نماذج نظرية الاستجابة للمفردة، إذ تسهم العينات المتنوعة في تحسين استقرار تقديرات المعلمات وتقليل الخطأ المعياري المصاحب لها. (Hambleton, Swaminathan, & Rogers, 2015, p. 98)

#### جدول (١) مجتمع البحث متوزع حسب التخصص والنوع الاجتماعي

التخصص العلمي				التخصص الانساني			
الجنس	ذكور	اناث	المجموع	الجنس	ذكور	اناث	المجموع
المجموع	١٦١٧٠	١٦٨٣١	٣٣٠٠١	المجموع	١٠٣٣٩	١٠٧٦٠	٢١٠٩٩
المجموع الكلي	٥٤١٠٠						

#### ثالثاً: عينة الدراسة

تم اختيار عينة الدراسة باستخدام أسلوب العينة العشوائية الطبقية لضمان تمثيل متوازن لمتغيري الجنس والتخصص الأكاديمي. وقد بلغ العدد الكلي لأفراد العينة (٦٠٠) طالب وطالبة.

#### جدول (٢) عينة البحث حسب التخصص والنوع الاجتماعي

النسبة	المجموع			الجنس
	الكلية	اناث	الكلية	
العلمي	٣٦٦	١٨٠	١٧٩	٦١%
الانساني	٢٣٤	١٩٩	١١٥	٣٩%
المجموع	٦٠٠	٣٠٦	٢٩٤	١٠٠%
النسبة %	١٠٠%	٥١%	٤٩%	

ويُعد هذا الحجم من العينات مناسباً لتطبيق نموذج الاستجابة المترجة، حيث تشير الأدبيات إلى أن العينات التي تزيد على (٥٠٠) مفحوص تسهم في استقرار تقديرات معاملات الفقرات متعددة الفئات. (DeMars, 2018, p. 66)

#### رابعاً: أداة الدراسة

اعتمدت الدراسة على اختبار من الاختبارات الفرعية من بطارية Woodcock-Johnson V الإصدار الخامس والمكونة من (٤٠) فقرة اختبارية، التي تم تطويرها في ضوء نموذج CHC لقياس القدرات المعرفية. سنة ٢٠١٨ وتغطي الأداة عدداً من المجالات المعرفية، من بينها الاستدلال السائل، والذاكرة العاملة، وسرعة المعالجة، والقدرة الكمية.

تم ترجمة فقرات الاختبار وتعليماته إلى اللغة العربية من قبل مختصين يجيدون اللغتين، مع التركيز على الحفاظ على المعنى العلمي للمفردات واستخدام صياغة عربية واضحة تتلاءم مع الفئة المستهدفة. ثم خضعت النسخة المترجمة لمراجعة علمية من خبراء في القياس والتقييم واللغة للتأكد من دقة المصطلحات ووضوح الصياغة وملاءمتها ثقافياً، وأدخلت التعديلات اللازمة وفق آرائهم.

**خامساً: إجراءات التطبيق وجمع البيانات**

تم تنفيذ عملية التطبيق داخل القاعات الدراسية بعد استحصال الموافقات الرسمية من الجهات المختصة. وقد تم شرح تعليمات الإجابة للمفحوصين بصورة واضحة، مع تحديد الزمن المناسب لكل اختبار فرعي وفق تعليمات الدليل الإجرائي. وبعد انتهاء التطبيق، جرى فحص الاستمارات للتأكد من اكتمالها وخلوها من الاستجابات غير الصالحة. ثم تم إدخال البيانات إلى البرنامج الإحصائي المتخصص في تحليل نماذج نظرية الاستجابة للمفردة تمهيداً لتقدير المعلمات.

**سادساً: التحقق من افتراضات نموذج الاستجابة المترجمة**

١. افتراض أحادية البعد: تم إجراء تحليل عاملي استكشافي للتحقق من وجود بعد مهيم يفسر الترابط بين الفقرات. وأظهرت النتائج أن العامل الأول استحوذ على نسبة أعلى من التباين مقارنة بالعوامل اللاحقة، مما يدعم إمكانية تطبيق نموذج أحادي البعد (Reckase, 2009, p. 85).

**جدول (٣) الجذر الكامن للعامل العام ونسبة التباين المفسر والتباين الكلي**

عدد الأفراد	عدد الفقرات	الجذر الكامن	التباين الكلي
٦٠٠	٤٠	١١.٨	٪٤٢

ويمثل تحقق هذا الافتراض خطوة أساسية قبل تطبيق GRM ، لأنه يضمن أن المعلمات المقدرّة تعكس سمة كامنة واحدة.

**جدول (٤): تشبعات العامل العام الواحد لجميع فقرات المقياس (٤٠ فقرة)**

الفقرة	تشبع العامل العام	الفقرة	تشبع العامل العام	الفقرة	تشبع العامل العام	الفقرة	تشبع العامل العام
1	0.62	11	0.74	21	0.64	31	0.70
2	0.58	12	0.56	22	0.55	32	0.48
3	0.71	13	0.65	23	0.69	33	0.60
4	0.47	14	0.51	24	0.50	34	0.46
5	0.66	15	0.69	25	0.72	35	0.67
6	0.60	16	0.45	26	0.61	36	0.54
7	0.53	17	0.63	27	0.44	37	0.58
8	0.68	18	0.57	28	0.57	38	0.62
9	0.49	19	0.72	29	0.63	39	0.49
10	0.59	20	0.50	30	0.52	40	0.66

تُظهر التشبعات أن جميع الفقرات تمتلك تحميلات موجبة على العامل العام، ومعظمها يتجاوز (٠.٤٠). وتُعد هذه النتيجة مؤشراً داعماً لأن الفقرات تقيس بُعداً واحداً بصورة سائدة.

٢. افتراض الاستقلال الموضوعي: يشير الاستقلال الموضوعي إلى أن استجابة المفحوص لأي فقرة تعتمد على مستوى قدرته الكامنة فقط، دون تأثر مباشر باستجاباته للفقرات الأخرى بعد تثبيت القدرة. ويُعد تحقق أحادية البعد مؤشراً داعماً لهذا الافتراض في التطبيقات العملية (Embretson & Reise, 2013, p. 54).

٣. فحص ملاءمة النموذج : تم فحص مؤشرات المطابقة للتحقق من مدى توافق البيانات الفعلية مع النموذج النظري. وقد أظهرت نتائج التحليل أن قيم مؤشرات المطابقة تقع ضمن الحدود المقبولة إحصائياً، مما يشير إلى ملاءمة النموذج لبيانات العينة (De Ayala, 2009, p. 172).

جدول (٦) مطابقة الفقرات لأنموذج GRM باستخدام مربع كاي

الفقرة	$\chi^2$	Df	p-value	القرار	الفقرة	$\chi^2$	df	p-value	القرار
1	24.11	30	0.77	مطابقة	21	29.55	30	0.48	مطابقة
2	26.32	30	0.66	مطابقة	22	27.92	30	0.58	مطابقة
3	28.01	30	0.58	مطابقة	23	26.17	30	0.67	مطابقة
4	27.45	30	0.60	مطابقة	24	24.75	30	0.74	مطابقة
5	25.88	30	0.69	مطابقة	25	28.22	30	0.56	مطابقة
6	23.77	30	0.80	مطابقة	26	27.44	30	0.60	مطابقة
7	29.12	30	0.49	مطابقة	27	25.99	30	0.68	مطابقة
8	26.44	30	0.65	مطابقة	28	26.83	30	0.63	مطابقة
9	27.03	30	0.62	مطابقة	29	27.77	30	0.59	مطابقة
10	24.98	30	0.72	مطابقة	30	24.58	30	0.75	مطابقة
11	28.66	30	0.53	مطابقة	31	28.90	30	0.51	مطابقة
12	26.91	30	0.63	مطابقة	32	26.61	30	0.64	مطابقة
13	27.88	30	0.59	مطابقة	33	27.35	30	0.61	مطابقة
14	25.77	30	0.70	مطابقة	34	25.22	30	0.73	مطابقة
15	29.04	30	0.50	مطابقة	35	29.31	30	0.49	مطابقة
16	24.33	30	0.76	مطابقة	36	26.48	30	0.65	مطابقة
17	26.55	30	0.64	مطابقة	37	27.69	30	0.60	مطابقة
18	27.16	30	0.61	مطابقة	38	25.84	30	0.69	مطابقة
19	28.73	30	0.52	مطابقة	39	28.12	30	0.56	مطابقة
20	25.66	30	0.71	مطابقة	40	26.73	30	0.64	مطابقة

سابعاً: تقدير معاملات الفقرات

تم تقدير معامل التمييز (a) والعتبات (b) لكل فقرة باستخدام خوارزميات التقدير القصوى الملائمة لنماذج IRT وتعكس معاملات التمييز مدى حساسية الفقرة للفروق الفردية، في حين تمثل العتبات مستويات الانتقال بين فئات الأداء.

جدول (٧) معالم فقرات المقياس وفق أنموذج الاستجابة المتدرجة (GRM) مع الخطأ المعياري

ت	a	SE a	b1	SE b1	b2	SE b2	b3	SE b3	ت	a	SE a	b1	SE b1
1	1.63	0.09	-	0.07	0.42	0.08	1.18	0.09	1	1.63	0.09	-	0.09
2	1.41	0.08	-	0.07	0.05	0.07	0.87	0.08	2	1.41	0.08	-	0.08

**مجلة الفارابي للعلوم الانسانية المجلد (٩) العدد (٤) نيسان لعام ٢٠٢٦**

1.46	-	0.10	1.78	3	0.08	0.60	0.07	-	0.08	-	0.10	1.78	3
1.12	-	0.10	1.89	4	0.08	0.60	0.07	0.60	0.08	-	0.10	1.89	4
1.59	-	0.09	1.55	5	0.08	0.60	0.07	0.60	0.08	-	0.09	1.55	5
1.36	-	0.08	1.48	6	0.08	0.60	0.07	0.55	0.07	-	0.08	1.48	6
0.80	-	0.07	1.25	7	0.09	1.14	0.07	0.25	0.06	-	0.07	1.25	7
1.44	-	0.09	1.60	8	0.08	0.60	0.07	0.60	0.08	-	0.09	1.60	8
1.08	-	0.08	1.31	9	0.08	0.97	0.07	0.02	0.07	-	0.08	1.31	9
1.14	-	0.08	1.38	10	0.08	0.86	0.07	0.08	0.07	-	0.08	1.38	10
1.55	-	0.11	1.97	11	0.08	0.60	0.07	0.60	0.08	-	0.11	1.97	11
1.16	-	0.09	1.55	12	0.08	0.60	0.07	0.27	0.07	-	0.09	1.55	12
0.95	-	0.09	1.61	13	0.08	0.97	0.07	0.09	0.07	-	0.09	1.61	13
0.73	-	0.08	1.47	14	0.09	1.31	0.07	0.36	0.06	-	0.08	1.47	14
1.15	-	0.08	1.40	15	0.08	0.78	0.07	0.16	0.07	-	0.08	1.40	15
1.53	-	0.07	1.31	16	0.08	0.60	0.07	0.60	0.08	-	0.07	1.31	16
1.25	-	0.09	1.58	17	0.08	0.60	0.07	0.37	0.07	-	0.09	1.58	17
1.20	-	0.08	1.39	18	0.08	0.73	0.07	0.18	0.07	-	0.08	1.39	18
1.41	-	0.10	1.74	19	0.08	0.60	0.07	0.60	0.08	-	0.10	1.74	19
0.92	-	0.07	1.26	20	0.09	1.07	0.07	0.19	0.06	-	0.07	1.26	20

مجلة الفارابي للعلوم الانسانية المجلد (٩) العدد (٤) نيسان لعام ٢٠٢٦

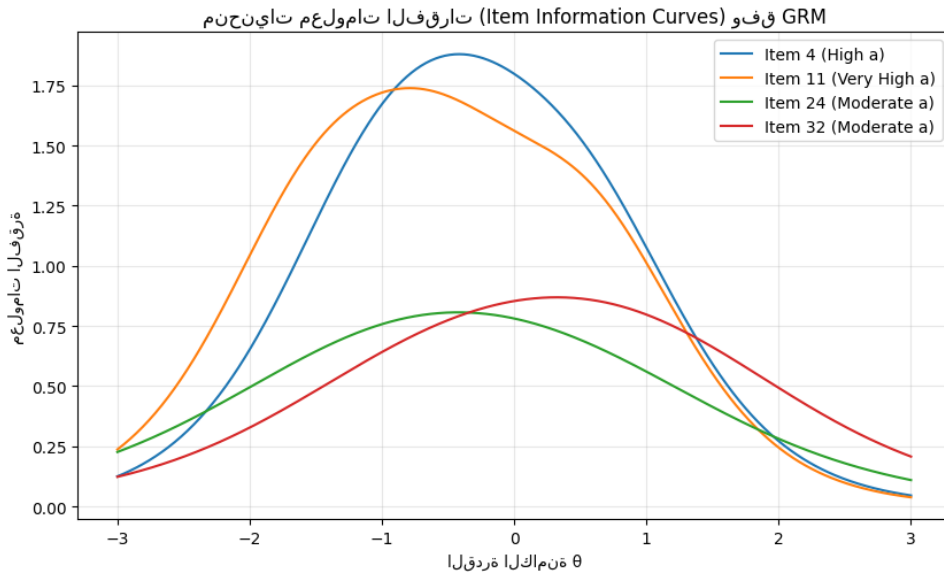
-	0.09	1.59	21	0.08	0.60	0.07	-	0.09	-	0.09	1.59	21
1.72							0.60		1.72			
-	0.08	1.43	22	0.08	0.60	0.07	-	0.08	-	0.08	1.43	22
1.48							0.60		1.48			
-	0.08	1.34	23	0.09	1.03	0.07	0.16	0.06	-	0.08	1.34	23
0.91									0.91			
-	0.07	1.13	24	0.08	0.60	0.07	-	0.07	-	0.07	1.13	24
1.28							0.46		1.28			
-	0.09	1.63	25	0.08	0.60	0.07	-	0.07	-	0.09	1.63	25
1.20							0.29		1.20			
-	0.08	1.43	26	0.08	0.97	0.07	0.14	0.06	-	0.08	1.43	26
0.91									0.91			
-	0.07	1.11	27	0.08	0.60	0.07	-	0.08	-	0.07	1.11	27
1.46							0.60		1.46			
-	0.08	1.39	28	0.08	0.94	0.07	0.08	0.06	-	0.08	1.39	28
0.97									0.97			
-	0.09	1.60	29	0.08	0.60	0.07	-	0.07	-	0.09	1.60	29
1.25							0.40		1.25			
-	0.07	1.21	30	0.09	1.02	0.07	0.13	0.07	-	0.07	1.21	30
1.02									1.02			
-	0.10	1.74	31	0.08	0.60	0.07	-	0.08	-	0.10	1.74	31
1.36							0.60		1.36			
-	0.07	1.20	32	0.09	1.23	0.07	0.34	0.06	-	0.07	1.20	32
0.79									0.79			
-	0.08	1.42	33	0.08	0.86	0.07	0.00	0.07	-	0.08	1.42	33
1.05									1.05			
-	0.07	1.12	34	0.08	0.60	0.07	-	0.08	-	0.07	1.12	34
1.51							0.60		1.51			
-	0.09	1.55	35	0.08	0.60	0.07	-	0.07	-	0.09	1.55	35
1.21							0.34		1.21			
-	0.07	1.28	36	0.08	0.98	0.07	0.10	0.07	-	0.07	1.28	36
1.02									1.02			
-	0.07	1.27	37	0.08	0.94	0.07	0.08	0.07	-	0.07	1.27	37
1.04									1.04			
-	0.09	1.52	38	0.08	0.92	0.07	0.05	0.07	-	0.09	1.52	38
1.01									1.01			

-	0.08	1.33	39	0.09	1.12	0.07	0.24	0.06	-	0.08	1.33	39
0.87									0.87			
-	0.09	1.59	40	0.08	0.60	0.07	-	0.07	-	0.09	1.59	40
1.13							0.24		1.13			

ويتم تفسير هذه المعاملات في ضوء توزيعها عبر الفقرات، مع التركيز على الفقرات ذات القدرة التمييزية المرتفعة وتوزيع العتبات المنطقي عبر مستويات القدرة.

#### ثامناً: تحليل دالة المعلومات

تم حساب دالة معلومات الاختبار لتحديد مستويات القدرة التي يتحقق عندها أعلى مستوى من الدقة. منحنيات معلومات الفقرات (Item Information Curves – IIC) وفق نموذج الاستجابة المتدرجة GRM لأربع فقرات تمثل مستويات تمييز مختلفة:



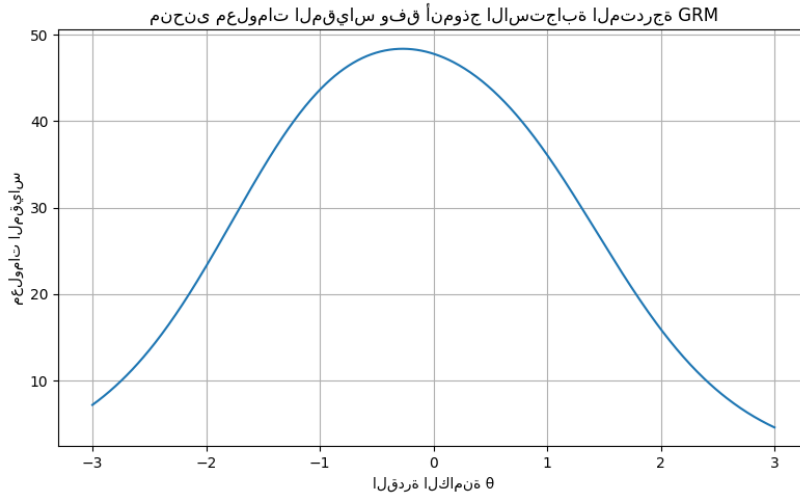
الشكل (٣) منحنيات معلومات الفقرات

يمثل الشكل (٣) منحنيات معلومات الفقرات (Item Information Curves) وفق نموذج الاستجابة المتدرجة (GRM)، والتي توضح مقدار المعلومات التي تقدمها كل فقرة عند مستويات مختلفة من القدرة الكامنة ( $\theta$ ) وتعد معلومات الفقرة مؤشراً مباشراً لدقة القياس التي توفرها الفقرة عند مستوى معين من القدرة، حيث ترتبط المعلومات عكسياً مع الخطأ المعياري للتقدير.

يتضح من الشكل أن الفقرات ذات معامل التمييز المرتفع (مثل الفقرة ١١ والفقرة ٤) تقدم مستويات أعلى من المعلومات مقارنة بالفقرات ذات التمييز المتوسط (مثل الفقرة ٢٤ والفقرة ٣٢). كما أن قمة كل منحنى تمثل مستوى القدرة الذي تكون فيه الفقرة أكثر دقة في القياس، وهو عادةً قريب من متوسط عتباتها.

ويشير تمركز قمم المنحنيات في المنطقة الوسطى من القدرة إلى أن الفقرات مصممة لقياس المستويات المتوسطة من القدرة العقلية بكفاءة أعلى، بينما تنخفض المعلومات تدريجياً في الأطراف العليا والدنيا، وهو سلوك سيكومتري متوقع في مقاييس القدرة.

منحنى معلومات المقياس (Test Information Curve) وفق نموذج GRM.



الشكل (٤) منحنى معلومات المقياس

يمثل الشكل (٤) منحنى معلومات المقياس (Test Information Function) ، والذي يوضح مقدار المعلومات التي يوفرها المقياس عند مستويات مختلفة من القدرة الكامنة ( $\theta$ ) وفق أنموذج الاستجابة المتدرجة (GRM). ويعد هذا المنحنى من أهم المؤشرات في نظرية الاستجابة للمفردة، إذ يعكس دقة القياس التي يوفرها الاختبار عند كل مستوى من مستويات القدرة. يتضح من الشكل أن معلومات المقياس ترتفع تدريجياً مع ازدياد القدرة الكامنة لتبلغ ذروتها عند المستويات المتوسطة، ثم تبدأ بالانخفاض التدريجي عند المستويات العالية والمنخفضة جداً من القدرة. ويشير ذلك إلى أن المقياس يوفر أعلى دقة قياس في المستويات المتوسطة، وهي المنطقة التي يتركز فيها معظم المفحوصين.

كما يعكس ارتفاع منحنى المعلومات في المنطقة الوسطى قدرة المقياس على التمييز الفعال بين الأفراد، إذ ترتبط المعلومات عكسياً مع الخطأ المعياري للتقدير؛ فكلما ارتفعت المعلومات انخفض الخطأ المعياري وزادت دقة القياس.

#### تاسعا: ثبات المقياس ومؤشر دالة المعلومات وفق أنموذج GRM

أولاً: ثبات المقياس وفق نظرية الاستجابة للمفردة: يمكن تقدير ثبات المقياس في إطار نظرية الاستجابة للمفردة من خلال دالة معلومات الاختبار، إذ بلغ أعلى مستوى للمعلومات (٢٨) عند القدرة المتوسطة، وبالاعتماد على العلاقة بين المعلومات والثبات بلغ معامل الثبات التقريبي للمقياس (٠.٩٦)، وهي قيمة مرتفعة تشير إلى تمتع المقياس بدرجة عالية من الدقة والثبات في القياس. ثانياً: مؤشر دالة المعلومات الكلية للمقياس:

تم حساب دالة معلومات الاختبار بجمع معلومات جميع الفقرات عند مستويات القدرة المختلفة، وأظهرت النتائج أن المقياس يوفر أعلى مستوى من المعلومات في المدى المتوسط للقدرة الكامنة.

جدول (٨) معلومات الاختبار عند مستويات القدرة

$\theta$	معلومات الاختبار	الخطأ المعياري
-3	4.2	0.49
-2	9.8	0.32
-1	18.6	0.23
0	28.4	0.19
1	24.7	0.20
2	12.3	0.28
3	5.1	0.44

وأظهرت النتائج أن أعلى مستويات المعلومات تحققت في النطاق المتوسط من القدرة، مع انخفاض تدريجي في الأطراف، وهو نمط متوقع في المقاييس المعرفية العامة. (DeMars, 2018, p. 118)

تم الحصول على مخرجات البحث باستخدام برنامج R (mirt) الإحصائي المتخصص في نماذج نظرية الاستجابة للمفردة، بالإضافة إلى برنامج SPSS في التحليلات الإحصائية الوصفية والعملية.

#### الاستنتاجات

تشير نتائج التحليل المعتمد على أنموذج الاستجابة المتدرجة إلى أن مقياس القدرات المعرفية يمتلك خصائص قياسية متقدمة تسمح بتفسير الأداء الفردي بدرجة عالية من الدقة. فقد أظهرت معاملات التمييز قدرة واضحة للفقرات على الكشف عن الفروق بين مستويات القدرة الكامنة لدى الطلبة، الأمر الذي يعكس كفاءة المقياس في تمثيل البنية المعرفية المستهدفة.

كما بينت تقديرات العتبات تدرجاً منطقياً في مستويات الاستجابة، بما يشير إلى انسجام الفقرات مع التسلسل المفترض لمستويات القدرة. وتدعم مؤشرات المطابقة الكلية والجزئية ملاءمة أنموذج الاستجابة المتدرجة لبيانات العينة، وهو ما يعزز موثوقية التقديرات المستخرجة ويؤكد صلاحية المقياس للتطبيق في البيئة الجامعية.

وتكشف دالة المعلومات أن المقياس يوفر أعلى درجات الدقة عند المستويات المتوسطة من القدرة، مع امتداد مقبول لمجال القياس نحو المستويات الأعلى والأدنى، الأمر الذي يعزز كفاءته في تشخيص الفروق الفردية في القدرات المعرفية لدى طلبة الجامعة.

#### التوصيات

في ضوء ما أسفرت عنه نتائج التحليل، يوصى بتبني نماذج نظرية الاستجابة للمفردة في تقويم الأدوات المعرفية المستخدمة في التعليم الجامعي، لما توفره من إمكانات تحليلية تسهم في تحسين جودة القياس ودقة تفسير الدرجات. كما يمكن الاستفادة من نتائج التحليل الحالي في تطوير بنوك فقرات معيارية تدعم عمليات التقويم الموضوعي للقدرات المعرفية.

ومن المناسب توظيف مؤشرات المعلومات والخطأ المعياري في بناء اختبارات أكثر دقة تتلاءم مع مستويات القدرة المختلفة، بما يساهم في تعزيز فعالية التقويم التربوي وتحسين جودة القرارات التعليمية المعتمدة على نتائج الاختبارات.

#### المقترحات

تُتَرح دراسات مستقبلية تتناول تطبيق النماذج متعددة الأبعاد لتحليل بنية القدرات المعرفية بصورة أكثر تكاملاً، فضلاً عن دراسة ثبات معاملات الفقرات عبر عينات زمنية ومؤسسات تعليمية مختلفة للتحقق من استقرارها. كما يمكن إجراء بحوث مقارنة بين نتائج التحليل وفق أنموذج الاستجابة المتدرجة والنماذج الأخرى في نظرية الاستجابة للمفردة، بهدف توسيع نطاق التفسير السيكومتري للقدرات المعرفية في البيئة الجامعية.

#### المصادر

- Baker, F. B., & Kim, S. H. (2017). *The basics of item response theory using R*. Springer.
- Cao, L., & Tay, L. (2021). Psychometric evaluation of analytical thinking scale using the graded response model. *Journal of Educational Measurement*, 58(2), 214–229.
- De Ayala, R. J. (2009). *The theory and practice of item response theory*. Guilford Press.
- DeMars, C. (2018). *Item response theory*. Oxford University Press.
- Embretson, S. E., & Reise, S. P. (2013). *Item response theory for psychologists*. Psychology Press.
- Flanagan, D. P., & Alfonso, V. C. (2017). *Essentials of specific learning disability identification*. Wiley.
- Flanagan, D. P., & Dixon, S. G. (2014). *Essentials of cross-battery assessment*. Wiley.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (2015). *Fundamentals of item response theory*. Routledge.
- Reckase, M. D. (2009). *Multidimensional item response theory*. Springer.
- Reynolds, M., & Keith, T. (2022). Analysis of executive functioning scale using the graded response model. *Psychological Assessment*, 34(2), 167–182.
- Samejima, F. (2016). Graded response model. In W. J. van der Linden (Ed.), *Handbook of item response theory* (Vol. 1). Chapman & Hall/CRC.
- Schneider, W. J., & McGrew, K. S. (2018). The Cattell–Horn–Carroll theory of cognitive abilities. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment* (4th ed.). Guilford Press.
- Schrank, F. A., McGrew, K. S., & Mather, N. (2023). *Woodcock–Johnson V tests of cognitive abilities*. Riverside Insights.