Sana'a University Journal of Human Sciences
مجلة جامعة صنعاء للعلوم الإنسانية
Vol. 5 | No. 7 | Page 73 – 107 |2026|
https://journals.su.edu.ye/index
ISSN : 2958-8677

Psychometric Study of Raven's Standard Progressive Matrices Using Item Response Theory among University Students in Hajjah Governorate, Yemen

Alia a Ahmed Mohamed Alawamy [1,*], Najah Abdel Rahim [2], Amt Al-Razzak Al-Washli[1]

[1] Faculty of Educational - Sana'a University, Sana'a, Yemen.

[2] Faculty of Educational–Tazia University, Tazia,.

***Corresponding author:** alia.alawamy@su.edu.ye*

Keywords

1. Item Response Theory
2. Four-Parameter
3. Psychometric
4. Raven Matrices Test
5. Standardized Progressive Matrices Test (SPM)

Abstract:

The aim of this study was to estimate the psychometric properties of Raven's Standard Progressive Matrices using the four-parameter item response theory (4PL IRT) model, focusing on item characteristics, ability estimation, and information function. The sample consisted of 1,035 university and college students from Hajjah Governorate, Yemen. Data were analyzed using R (mirt package) and SPSS after verifying the assumptions of the model.

Item fit was assessed using Chi-square tests and the G^2 statistic, and the overall fit indices indicated acceptable results. The items demonstrated a wide range of difficulty and discrimination, reflecting their suitability for different ability levels. The average guessing parameter was 0.073, and all values were within the acceptable limit, indicating a low level of random responding. The average carelessness parameter was 0.961 with a low standard deviation, suggesting that most participants responded attentively—except for items (35) and (48), which recorded the lowest values (0.51 and 0.61, respectively) despite being difficult and highly discriminating. These findings confirm the strong psychometric quality of the test and its effectiveness in estimating mental ability.

Sana'a University Journal of Human Sciences
مجلة جامعة صنعاء للعلوم الإنسانية
Vol. 5 | No. 7 | Page 74 – 107 | 2026|
https://journals.su.edu.ye/index
ISSN : 2958-8677

دراسة سيكومترية لاختبار المصفوفات المتتابعة المعياري لرافن باستخدام نظرية الاستجابة للمفردة لدى طلبة الجامعات في محافظة حجة

علياء أحمد محمد العوامي [1]*، نجاح عبد الرحيم محمد عثمان [2]، أمّه الرزاق محمد أحمد الوشلي[1]

[1] كلية التربية - جامعة صنعاء ، صنعاء ، اليمن.
[2] كلية التربية - جامعة تعز ، تعز ، اليمن.
المؤلف: *alia.alawamy@su.edu.ye*

الكلمات المفتاحية

1. نظرية الاستجابة للمفردة
2. نموذج رباعي المعلمة
3. معالم الاختبار
4. اختبار رافن للمصفوفات المتتابعة
5. اختبار المصفوفات المتتابعة المعياري

الملخص:

تهدف هذه الدراسة إلى تقدير معالم اختبار مصفوفات رافن المعيارية باستخدام نموذج نظرية الاستجابة للمفردة رباعي المعلمة (4 PL IRT) مع التركيز على معالم المفردات، وتقدير القدرة، ودالة المعلومات. شملت العينة 1035 طالبًا وطالبة من الجامعات والكليات في محافظة حجة باليمن. تم تحليل البيانات باستخدام برنامجي (R حزمة mirt و SPSS) بعد التحقق من افتراضات النظرية.

تم تقييم مطابقة الفقرات للنموذج باستخدام اختبارات مربع كاي ومربع كاي المعدل وإحصاء G^2، وأظهرت مؤشرات المطابقة الكلية نتائج مقبولة، أظهرت الفقرات مدى متنوعًا من الصعوبة والتمييز، بما يعكس ملاءمة الفقرات لمستويات القدرة المختلفة. تشير إلى انخفاض الاستجابات العشوائية، كما بلغ متوسط معلمة الإهمال أو معلمة المقارب العلوي (0.961) بانحراف معياري منخفض، مما يدل على التزام المشاركين بالإجابة بدقة باستثناء الفقرتين (35) و(48) اللتين سجلتا أدنى قيم (0.51 و0.61 على التوالي) رغم كونهما مفردتين صعبتين وذات تمييز عالٍ. وتؤكد النتائج جودة معالم الاختبار وفاعلية نموذج الاستجابة للمفردة رباعي المعلمة في تقدير القدرة العقلية.

المقدمة:

يهتم العلماء في مختلف مجالات العلوم السلوكية بابتكار وتطوير طرائق وأساليب جديدة للقياس وذلك بهدف البحث والتقصي عن الطرق والوسائل والأدوات التي تحقق دقة القياس وموضوعيته (كاظم، 1988: 13) حيث يشهد ميدان التقويم التربوي تقدما ملموسًا وذلك مع التطور التكنولوجي الذي يمثل السمة الرئيسية للعصر الحالي فمنذ منتصف القرن الماضي، ونتيجة لهذا التقدم المستمر في مجال القياس النفسي والتربوي، وتزايد اهتمام العلماء في الدول المتقدمة بضبط إجراءات وأساليب بناء المقاييس والاختبارات وتحليل مفرداتها والتأكد من صلاحيتها باستخدام النظريات الحديثة في القياس؛ دعت الضرورة إلى استخدام هذه النظريات الحديثة وما ينبثق عنها من نماذج في محاولة جاهدة للتغلب على عيوب النظرية الكلاسيكية والتي سيطرت على منهجيات القياس طوال القرن الماضي ، وما تزال أغلب البحوث والقياسات في العالم العربي على مبادئها وقوانينها في عملية بناء وتقنين المقاييس(عبدالوهاب،2010 :2).

وتعد نظرية الاستجابة للمفردة من أهم وسائل التقويم الحديث التي تعمل على تقدير قدرات الفرد وتهدف الى تحقيق متطلبات الموضوعية في القياس السلوكي (عبد الوهاب، 2010: 14). كما تمثل الاستجابة للمفردة نموذجا احتمالياً، نظراً لأن طبيعة البيانات تسمح بالتنبؤ باحتمال النجاح على أي مفردة من المفردات المدرجة كما توفر تقديرات لقدرات الافراد بوحدات النموذج المستخدم (اسماعيل،2007: 13)

هذا وقد انبثقت عن هذه النظرية مجموعة من نماذج الاستجابة للمفردة لتناسب السمات المختلفة المراد قياسها وأنواع وخصائص المفردات المستخدمة في هذا القياس، فتنقسم هذه النماذج على أساس المتغير المراد قياسه إلى نماذج أحادية البعد – تقيس قدره أو سمة واحدة فقط – أو نماذج متعددة الأبعاد – تقيس عدة سمات مركبة– وذلك تبعاً لعدد السمات الكامنة التي يقيسها هذا المتغير، وهذه النماذج أو تلك تختلف باختلاف نوع المفردات سواء كانت ثنائية الاستجابة (أي تحصل على الدرجتين: صفر، ١) أو متعددة الاستجابات (أي تحصل على الدرجات: ١، ٢، ٣ ،٤ ... .).

كما تختلف هذه النماذج فيما بينها أيضا باختلاف عدد معالم المفردات(Harvey,1999:358-364). فالنموذج احادي المعلم (راش)يحدد معلمة الصعوبة فقط بينما النموذج ثنائي المعلم فيحدد معلمة الصعوبة ومعلمة التمييز بينما يتميز النموذج ثلاثي المعلم حيث انه يحدد معلمة الصعوبة ومعلمة التمييز ومعلمة التخمين لكل مفردة وبخاصة مفردات الاختيار من متعدد مما يؤدي الى نتائج أكثر واقعية في العملية الاختبارية وبناء المقاييس(علام، 2005 : 697) والغرض من اضافة هذا المعلم في النموذج هو محاولة مراعاة عدم مطابقة المنحنيات المميزة للمفردة عند الطرف الادنى لمتصل السمة حيث يكون التخمين العشوائي احد المتغيرات التي تؤثر في أداء الاختبار (علام، 2005: 73؛ محاسنة، 2013: 19) توثيق

وفي إطار السعي إلى تحسين دقة التقدير وكفاءة القياس، جاء النموذج رباعي المعلمات (4PL) ليضيف معلمة رابعة يُشار إليها غالبًا في الأدبيات العربية بـ "معلمة الإهمال/اللامبالاة"، وتُعرف في الأدبيات الإحصائية باسم "المقارب العلوي (Upper Asymptote)"؛ إذ تعكس احتمال حدوث خطأ غير مقصود لدى ذوي القدرة العالية نتيجة التسرع أو ضعف الانتباه، بما يُعرف أيضًا بـ عدم الإتقان أو الانزلاق. وتكمن أهمية هذه المعلمة في قدرتها على التقليل من التحيز في تقدير قدرات الممتحنين الناتج عن الإهمال أو قلة التركيز، لا سيّما في المواقف الاختبارية العملية، مما يسهم في تحسين الجودة (Liao et al., 2012:1680–1683)

كما يتمتع النموذج الرباعي المعلمة بقدرة محسّنة على تمثيل سلوك الاستجابة بدقة، خاصةً في أطراف طيف القدرة، مما يجعله مفيدًا بشكل خاص في الاختبارات عالية الأهمية التي تتطلب تقديرات دقيقة (Loken & Rulison, 2010:525).

كما ان هذه النماذج تصف التفاعل بين قدرة الفرد الممتحن ومفردات الاختبار من خلال دالة احتمالية تربط بين معلمين احداهما يتعلق بمعالم المفردة الاختبارية والاخر تتعلق بالفرد بغرض الوصول الى قيم تقديرية عددية واحدة تتعلق بالفرد وتسمى بمعلم القدرة او السمة المقاسة لدى الفرد وقيمه تقديرية عددية واحدة او أكثر تتعلق بالمفردة الاختبارية وتسمى بمعلم او معلمات المفردة (Baker, 2002:85).

ولما يُسهم ذلك في رفع دقّة تقدير المعالم وتحسين دالة المعلومات، فإنه يؤدي دورًا مهمًا في بناء الاختبارات وتكوينها وتقييم فقراتها لدى المفحوصين؛ إذ تعتمد دالة معلومات الاختبار على تجميع إسهامات كل مفردة من مفرداته من المعلومات عند مستوى قدرة معين، بافتراض استقلالية هذه الإسهامات. (علام،2005: 117).

وقد اوضح كثير من خبراء القياس المعاصر انه لا يوجد اسلوب وحيد يمكن استخدامه في ملاءمة البيانات للنموذج بحيث يكون كافيا وأكدوا ضرورة الحصول على ادلة متعددة للمراجعة الشاملة لمدى ملاءمة النماذج والحكم على ملاءمة النموذج الذي يتم اختياره واستخدامه في تطبيق معين (علام،2005 :107–106).

ويعتبر الذكاء واحدا من العوامل المهمة المحددة للسلوك الانساني والمسؤولة عن الفروق بين الافراد وهو ما جعله محط انظار المهتمين من الباحثين بدراسة السلوك الانساني وأسبابه وكذلك لكونه احد اهم المؤشرات المستخدمة في تحديد الموهوبين او فرز الاطفال ذوي صعوبات التعلم ويتضح من الادبيات المتخصصة ان اختبارات الذكاء بنوعيها الفردية والجمعية يمكن الاعتماد عليها في التشخيص والكشف عن الفروق الفردية في القدرات ويعتبر اختبار المصفوفات المتتابعة من اختبارات الذكاء الجمعية وغير اللفظية حيث لقي اهتماما غير مسبوق بين جميع ادوات الفرز المبدئي ولهذا يستخدم بشكل فعال في الدراسات التي تهتم بفرز الطلبة الموهوبين. وقد تم استخدام اختبار المصفوفات المتتابعة المعياري لرافن بصورة كبيرة في كثير من الاقطار والبلدان وقنن على مختلف الثقافات مثل بريطانيا والولايات المتحدة الامريكية والارجنتين والعراق والمملكة العربية

السعودية ولم يطرأ عليه الا تعديلات طفيفة (زمزمي،1998: 16)

وفي هذا الإطار، برزت أهمية استخدام النموذج رباعي المعلمات (4PL) في تقدير معالم اختبارات الذكاء غير اللفظي، مثل اختبار المصفوفات المتتابعة لرافن، نظرًا لما يوفره من دقة عالية في تقدير معلمات الفقرات وسلوك الاستجابة، خاصة في حالات التخمين أو الإهمال.

فقد طبّقت دراسة علوان (2023) النموذج رباعي المعلمات على اختبار أوتيس–ليون للقدرة المدرسية، وأظهرت النتائج ملاءمة عالية للنموذج ودقة في تقدير معلمات الفقرات لدى طلبة المرحلة الإعدادية.

أما دراسة مرشود (2020)، ورغم أنها لم تستخدم اختبار رافن، فإنها تُعد من الدراسات العربية البارزة التي استعانت بالنموذج رباعي المعلمة في تحليل اختبار تحصيلي في مادة علم النفس والحياة، وقد كشفت النتائج عن فروق دالة إحصائيًا في معلمة "الاهمال"، مما يعكس أهمية هذا البُعد في التقدير التربوي.

وفي السياق ذاته، أجرى النفيعي (2023) تدريجًا لاختبار المصفوفات المتتابعة المعياري المطوّر باستخدام النموذج اللوغاريتمي المتداخل ثنائي المعلمات، وأظهرت نتائجه توافقًا جيدًا مع خصائص النموذج، ما يؤكد أهمية استخدام النماذج الحديثة في تحليل اختبارات الذكاء غير اللفظي.

كما قدّم (Georgiev, 2008) تحليلًا تفصيليًا لسلاسل C و D و E من مصفوفات رافن القياسية (SPM) لدى طلبة المرحلة الثانوية في بلغاريا، باستخدام النموذج ثنائي المعلمات(2PL) ، مع تفسير شامل لخصائص الفقرات بناءً على منحنيات الخصائص ودوال المعلومات.

وتشير مراجعة الأدبيات والدراسات السابقة إلى أن الدراسات العربية التي تناولت نموذج الاستجابة للمفردة رباعي المعلمة لا تزال محدودة نسبيًا، مقارنة بتلك التي ركزت على نموذج راش أو النماذج الأخرى الأقل تعقيدًا. وعلى الرغم من وجود بعض الجهود البحثية في هذا المجال، إلا أن استخدام النموذج رباعي المعلمة في تقدير معالم اختبارات الذكاء ما زال نادرًا، مما يبرز الحاجة إلى المزيد من الدراسات التي تستكشف إمكاناته في السياقات التعليمية العربية، ومنها البيئة الجامعية اليمنية.

مشكلة الدراسة وأسئلتها

تعددت نماذج نظرية الاستجابة للمفردة المستخدمة في الدراسات والابحاث بتطور حركة القياس فكل دراسة تعتمد في تحليل نتائجها على نموذج معين وهذه الدراسات هدفها تحقيق الموضوعية والدقة في القياس. حيث لم تتفق الدراسات السابقة على أي نماذج الاستجابة للمفردة هو النموذج الاكثر ملاءمة لتقدير معالم الاختبارات والمقاييس النفسية. بالإضافة إلى ندرة الدراسات العربية واليمنية – في حدود ما تم الاطلاع عليه – التي تناولت معالم الاختبارات النفسية باستخدام نموذج الاستجابة للمفردة رباعي المعلمة كونه أحد النماذج المهمة والحديثة. فضلاً عن انه يغلب في بناء وتقنين الاختبارات والمقاييس النفسية المستخدمة في مجال العلوم الانسانية في البيئة العربية عامة والبيئة اليمنية خاصة الاعتماد على نظرية

القياس التقليدية.

وفي ضوء ما سبق تتلخص مشكلة الدراسة في الاجابة على التساؤل الآتي:

ما هي تقديرات معالم اختبار المصفوفات المتتابعة المعياري وفق نموذج الاستجابة للمفردة رباعي المعلمة؟

أهداف الدراسة

تهدف الدراسة إلى التعرف على تقدير معالم اختبار المصفوفات المتتابعة المعياري لطلبة الكليات والجامعات في محافظة حجة اليمنية وفق نموذج الاستجابة للمفردة رباعي المعلمة. ويتفرع من هذا الهدف الرئيس الأهداف الفرعية الآتية:

1- التحقق من افتراضات نظرية الاستجابة للمفردة في بيانات اختبار مصفوفات رافن العادية.

2- دراسة مدى ملاءمة بيانات اختبار المصفوفات المتتابعة المعياري لنموذج الاستجابة للمفردة الرباعي المعلمة (4PL).

- تحديد مدى ملاءمة الاختبار ككل لنموذج الاستجابة للمفردة الرباعي المعلمة (4 PL).
- دراسة مدى ملائمة مفردات اختبار مصفوفات رافن لنموذج الاستجابة للمفردة الرباعي المعلمة (4 PL).
- تحديد مدى ملاءمة بيانات الأفراد المستجيبين لنموذج الاستجابة للمفردة الرباعي المعلمة (4PL).

3- تقدير معالم مفردات اختبار المصفوفات المتتابعة المعياري وفق نموذج الاستجابة للمفردة رباعي المعلمة.

4- تقدير قدرات المستجيبين في اختبار المصفوفات المتتابعة المعياري وفق نموذج الاستجابة للمفردة رباعي المعلمة.

5- تقدير كمية المعلومات التي يقدمها اختبار المصفوفات المتتابعة المعياري لرافن عند مستويات القدرة المختلفة وفق النموذج رباعي المعلمة.

أهمية الدراسة

تبرز أهمية الدراسة الحالية من خلال الآتي:

6- تُسهم الدراسة في دعم الأدبيات العربية واليمنية حول نموذج الاستجابة للمفردة رباعي المعلمة (4PL)بوصفه من النماذج الحديثة التي تعزز دقّة القياس وجودة تقدير معالم البنود، مع ندرة استخدامه محليًا.

7- وفر مؤشرات سيكومترية حديثة لاختبار رافن للمصفوفات المتتابعة المعياري على عينة محلية من طلبة الجامعات مما يُسهل استخدامه في الدراسات المستقبلية في القياس والتقويم. وكذلك كونها الدراسة الأولى في البيئة اليمنية – في حدود ما تم الاطلاع عليه – التي درست الخصائص السيكومترية لاختبار المصفوفات المتتابعة المعياري باستخدام نموذج الاستجابة للمفردة رباعي المعلمة.

8- تُبرز إمكانية استخدام برنامج (R) في تحليل بيانات الاختبارات وفق نماذج نظرية الاستجابة للمفردة، بما يسهل على الباحثين تطبيق النموذج وتكرار التحليل.

9- تناولها اختبار المصفوفات المتتابعة المعياري لرافن كونه من الاختبارات المهمة المستخدمة في

الميادين التربوية والنفسية للتوجيه التربوي والارشاد النفسي والاختيار المهني والكشف عن الفئات الخاصة كالموهوبين والمتخلفين.

حدود الدراسة ومحدداتها

تقتصر الدراسة الحالية على استخدام نموذج الاستجابة للمفردة رباعي المعلمة (4PL) في تقدير معالم مفردات اختبار رافن للمصفوفات المتتابعة المعياري لدى طلبة الكليات والجامعات في محافظة حجة بالجمهورية اليمنية، وذلك خلال العام الدراسي 2022–2023.

مصطلحات الدراسة

1- نظرية الاستجابة للمفردة (Item Response Theory – IRT)

نظرية حديثة في القياس النفسي تفترض وجود سمة كامنة واحدة أو أكثر لدى الفرد تقف وراء استجاباته على مفردات الاختبار. وتقوم على أن احتمال الإجابة عن المفردة يتحدد بوظيفة تجمع بين قدرة المفحوص (θ)ومعالم المفردة (التقي، 2009: 64-63).

2- النموذج رباعي المعلمة (Four-Parameter Logistic Model – 4PL)

أحد نماذج IRT يتضمن أربع معالم: الصعوبة، التمييز، التخمين، الإهمال (de Ayala, 2013:227).

3- القدرة (Ability – θ)

تمثل السمة الكامنة التي تعكس مستوى أداء الفرد المتوقع على الاختبار، وتُقدَّر إحصائيًا اعتمادًا على نمط استجاباته للمفردات (Beker, 2001: 85)

تعرف الدراسة الحالية القدرة إجرائيًا بانها: تقدير القدرة (θ) الذي ينتجه نموذج 4P لكل طالب/طالبة بناءً على استجاباته/استجابتها على اختبار المصفوفات المتتابعة المعياري المعتمد في الدراسة الحالية.

4- معلمة الصعوبة (Difficulty – b)

تمثل مقدار القدرة اللازمة لتحقيق احتمال معيّن للإجابة الصحيحة) وغالبًا يُشار إلى نقطة 50% في النماذج الأبسط، مع ملاحظة أن وجود التخمين واللامبالاة قد يغيّر المعنى الدقيق لهذه النقطة (Hambleton et al., 1991:13)

تعرف الدراسة الحالية معلمه الصعوبة اجرائيا بانها: قيمة (b) المقدّرة لكل مفردة في اختبار رافن وفق نموذج 4PL، وتمثل مستوى القدرة الذي ترتفع عنده احتمالية الإجابة الصحيحة على المفردة.

5. معلمة التمييز (Discrimination – a)

تعكس مدى حساسية المفردة في التفريق بين الأفراد ذوي القدرات المختلفة، أي مدى انحدار منحنى خصائص المفردة حول مستوى الصعوبة (Hambleton et al., 1991: 13).

تعرف الدراسة الحالية معلمة التمييز إجرائيًا بأنها: قيمة (a) المقدّرة لكل مفردة وفق نموذج 4PL، وتمثل فاعلية مفردات رافن في التمييز بين الطلبة ذوي القدرة الأعلى والأدنى.

5- معلمة التخمين (Guessing – c)

تمثل الحد الأدنى لاحتمال الإجابة الصحيحة الذي قد يحدث نتيجة التخمين لدى ذوي القدرة المنخفضة (Hambleton et al., 1991:17).

تعرف الدراسة الحالية معلمة التخمين إجرائيًا بانها: قيمة (c) المقدّرة لكل مفردة وفق نموذج4PL، وتمثل احتمال الإجابة الصحيحة بالصدفة لدى المفحوص منخفض القدرة.

6 . معلمة الإهمال (d – Carelessness)

تشير إلى الحد الأعلى لاحتمال الإجابة الصحيحة على المفردة، وقد تُفسَّر تربويًا على أنها احتمال وقوع خطأ لدى ذوي القدرة العالية بسبب عدم الإتقان/التسرع/تشتت الانتباه (de Ayala, 2013 :227).

a. **تعرف الدراسة الحالية معلمة الإهمال إجرائيًا بأنها:** قيمة (d) المقدّرة لكل مفردة وفق نموذج 4PL، وتمثل الحد الأعلى لاحتمال الإجابة الصحيحة على المفردة (وقد تعكس أخطاء غير مقصودة لدى ذوي القدرة المرتفعة).

6–دالة معلومات المفردة (Item Information Function)

تعبر عن مقدار المعلومات (الدقة) التي توفرها المفردة عند مستويات مختلفة من القدرة على متصل السمة الكامنة. (علام،2005:113)

تعرف الدراسة الحالية دالة معلومات المفردة إجرائيًا بانها: دالة معلومات كل مفردة من مفردات اختبار رافن المحسوبة في إطار نموذج 4PL، والتي تبين مقدار مساهمة المفردة في دقة تقدير القدرة (θ)

7–داله معلومات الاختبار (Test Information Function): هي مجموع دوال معلومات مفردات الاختبار عند كل مستوى من مستويات القدرة، إذ تُعد مساهمة كل مفردة مستقلة في المعلومات (علام، 2005: 117).

وتُعرَّف إجرائيًا بأنها: داله المعلومات الكلية لاختبار رافن وفق نموذج 4PL، الناتجة عن تجميع معلومات جميع المفردات عبر مستويات القدرة (θ) .

8–اختبار المصفوفات المتتابعة المعياري لرافن (Raven's Standard Progressive Matrices – SPM)

اختبار غير لفظي يهدف إلى قياس الذكاء السائل من خلال إدراك الأنماط والعلاقات البصرية. طوّره جون رافن عام 1938 لقياس العامل العام للذكاء وفق تصور سبيرمان . (Raven, 2000: 1–2) .

منهجية الدراسة وإجراءاتها

منهج الدراسة

اتبعت الدراسة الحالية المنهج الوصفي اذ تسعى لوصف معالم اختبار رافن للمصفوفات المتتابعة المعياري باستخدام نموذج الاستجابة للمفردة رباعي المعلمة وذلك لمناسبته لجمع وتحليل وتفسير البيانات في مثل هذا النوع من الدراسات بالإضافة إلى انه يتميز بقدرته على تحليل البيانات بشكل كمي وتقديم صورة واضحة عن مدى استجابة المفحوصين على مفردات الاختبار.

مجتمع الدراسة

تكون مجتمع الدراسة من جميع الكليات والجامعات وتخصصات مختلفة في محافظة حجة اليمنية للعام الجامعي (2023/2022م) بلغ عددهم

(4135) طالبًا وطالبة، موزعين على ست مؤسسات تعليمية تشمل جامعتين (جامعة حجة وجامعة الرازي) وأربع كليات مستقلة، وينتمون إلى (18)تخصصًا متاحًا في هذه المؤسسات جميعًا، وذلك وفقًا للإحصاءات الرسمية الموثقة لكل مؤسسة وتخصص.

عينة الدراسة

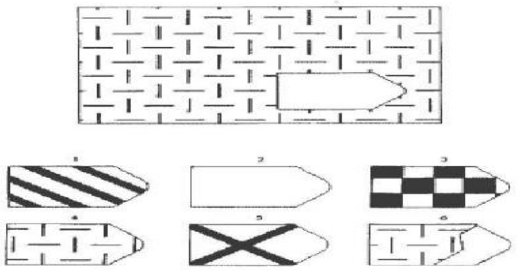
تم اختيار عينة الدراسة الأساسية بالطريقة العشوائية البسيطة من طلبة الجامعات والكليات في محافظه حجة للعام الجامعي (2022-2023)، وبلغ عددها (1035) طالب وطالبة (509) ذكور (526) اناث بنسبة (25%) من المجتمع الكلي، والموضحة بالجدول (1) وفقا للتخصص.

جدول (1) توزيع أفراد عينة الدراسة وفقًا لمتغير التخصص

التخصص	العلوم التربوية	العلوم التطبيقية	العلوم الطبية	الاجمالي
عدد المستجيبين	297	302	436	1035

أداة الدراسة

استخدمت الدراسة الحالية اختبار المصفوفات المتتابعة المعياري، وهو اختبار غير لفظي، أعده العالم جون رافن عام (1938) كأداة تشخيص اكلينيكية لقياس العامل العام للذكاء بمفهوم سبيرمان، واجريت عليه تعديلات طفيفة عامي (1947) و(1958م). إجمالي عدد مفرداته (60) مفردة مرتبة في خمس مجموعات هي (أ، ب، ج، د، هـ) تحتوي كل مجموعة منها على (12) مفردة مصممة لتقييم القدرة على التفكير المجرد والاستنتاج البصري-المكاني، وتتضمن مشكلات بدرجات متفاوتة من الصعوبة تبدأ بمشكلات أساسية تزداد تعقيدًا تدريجيًا. وتظهر كل مفردة بالحبر الأسود على خلفية بيضاء. وتتألف كل مفردة من شكل أساسي يحتوي رسم أو تصميم هندسي أو نمط شكلي حذف جزء منه؛ اسفله ستة أو ثمانية بدائل تتضمن بديل واحد فقط يمثل الإجابة الصحيحة شكل (1). والبيانات التي يمكن الحصول عليها من تطبيق الاختبار هي بيانات ثنائية التقسيم حيث تقدر الاستجابة الصحيحة بالدرجة (واحد) والخاطئة بالدرجة (صفر) لكل مفردة من فقرات الاختبار (Raven, 2003: 223-225) ويستخدم الاختبار على الأفراد من عمر احدى عشر سنه إلى ستين سنة.



شكل (1) أول مفردة في الاختبار

الخصائص السيكومترية للاختبار في البيئة العربية:

أشارت الدراسات التي أجريت في البيئة العربية الى تمتع اختبار المصفوفات المتتابعة بدرجات عالية من الصدق والثبات، حيث تراوح معامل الثبات بين (0.78) و(0.96) في البيئة السعودية (أبو حطب ،1979: 209) وبلغ (0.92) في البيئة القطرية (مبارك،2001:143) وتراوح بين (0.83) و(0.93) في البيئة الكويتية (Abdel-Khalek & Raven, 2006:171) وبلغ (0.88) في البيئة الفلسطينية (ابو غالي وابو مصطفى،2014

: 101). كما ظهر ارتباط قوي بين اختبار رافن واختبار تيرمان–بينيه بلغ (0.86) في السياق العربي، مما يعكس فعاليته في قياس الذكاء عبر الثقافات المختلفة (بترجي واخرون، 2013).

الخصائص السيكومترية للاختبار في البيئة اليمنية:

1- **صدق الاختبار**: تم التحقق من صدق اختبار المصفوفات المتتابعة المعياري في البحث الحالي بعدة طرق هي:

- الصدق الظاهري حيث عرض الاختبار على (8) متخصصًا في مجال علم النفس والقياس النفسي ملحق (1) للتحقق من صدقه الظاهري وملائمته للبيئة اليمنية، وقد اتفق المحكمين على ملاءمة المفردات لقياس القدرة العقلية العامة في البيئة اليمنية حيث حصلت المفردات على نسبة اتفاق (100٪).
- مؤشر صدق البناء: حيث تم اجراء التحليل العاملي لبيانات الاختبار والذي اشارت نتائجه الى ان الاختبار يقيس قدرة واحدة هي القدرة العقلية العامة كما هو مبين في نتائج الدراسة.

2- **ثبات الاختبار**: تم التحقق من ثبات الفاكرومباخ بتطبيق الاختبار على عينه البحث الاستطلاعية الثانية وبلغ معامله (.92) وهي قيمة مرتفعة تؤكد تمتع الاختبار بدرجة عالية من الثبات.

3- **تعليمات الاختبار وتصحيحه:** اعتمدت تعليمات الاختبار كما وردت في دليله، كما يتم التصحيح وفقًا لمفتاح تصحيح الاختبار؛ حيث تعطى الإجابة الصحيحة الدرجة (1) والإجابة الخاطئة الدرجة (صفر).

4- **تجهيز الاختبار**: تم إعداد مفردات الاختبار الـ (60) في كراسة خاصة بحسب تسلسلها في الاختبار الأصلي، كما أعدت ورقة مستقلة لتعليمات الاختبار، وورقة مستقلة للإجابة.

5- **التجربة الاستطلاعية للاختبار**: طبق الاختبار على عينة استطلاعية عددها (36) طالباً وطالبة من طلبة الكليات والجامعة بمحافظة حجة في المستوى الدراسي الأول، للتأكد من مدى وضوح تعليمات الاختبار ومفرداته، ولمعرفة الوقت الذي تستغرقه الاجابة على جميع مفردات الاختبار، إذ طلب منهم الاستفسار عن أي غموض في الاختبار وتعليماته، والتعرف على أهم الصعوبات والمعوقات التي قد تحدث أثناء تطبيق الاختبار لتلافيها. وقد تبين من هذه التجربة أن تعليمات الاختبار ومفرداته واضحة، اما الوقت المستغرق فقد كان بين (70) و(90) دقيقة للانتهاء من الاجابة على جميع مفردات الاختبار وبمتوسط (80) دقيقة، علمًا بأنه لا يوجد زمن محدد للإجابة عن هذا الاختبار.

6- **إجراءات التطبيق الميداني**

اتبعت الاجراءات التالية لاستخراج الخصائص السيكومترية للاختبار:

1- الحصول على البيانات والاحصائيات الخاصة بأعداد الطلبة في الكليات والجامعات بمحافظة حجة في الفصل الدراسي الثاني من العام الجامعي (2022/ 2023م)، وعلى الموافقات اللازمة لتطبيق أداة البحث.

2-اختيار عينة الدراسة الاساسية بالطريقة العشوائية البسيطة من الكليات والتخصصات والمستويات المختلفة، ووضع خطة زمنية للتطبيق.

3- تطبيق اختبار المصفوفات المتتابعة المعياري في الفترة من (2/ نوفمبر/ 2022م) إلى (5/ نوفمبر/2022م) على عينة ضمت (113) طالباً وطالبة، ثم تصحيح الاستجابات يدويًا وادخالها إلى برنامج (SPSS) لحساب ثبات ألفا-كرونباخ وفقًا لنظرية القياس التقليدية كما في النتائج.

4- تطبيق اختبار المصفوفات المتتابعة المعياري في الفترة من (3/ ديسمبر/ 2022 م) إلى (24/ يناير/ 2023م) على عينة الدراسة الاساسية، ثم تصحيح الاستجابات يدويًا ومعالجة البيانات وتحليلها باستخدام برنامج (SPSS) وبرنامج (R) من خلال حزمتي (mirt) و(ltm) لتحقيق أهداف الدراسة كما في النتائج.

المعالجة الإحصائية

تم التحليل الاحصائي للبيانات باستخدام البرنامجين الإحصائيين التاليين:

1- برنامج (SPSS) باستخدام الاساليب الاحصائية التالية:

2- المتوسطات والانحرافات المعيارية للتأكد من خصائص التوزيع الاحصائي للدرجات المتحصل عليها من اختبار المصفوفات المتتابعة لرافن.

3- التحليل العاملي لمعرفة الجذور الكامنة لفحص افتراض احادية البعد.

4- برنامج R الاحصائي حزمتي (ltm،mirt) (Chalmers, 2012)باستخدام الاساليب الاحصائية التالية : شكرًا للملاحظة. تم الإبقاء على أسماء حزم R بصيغتها الرسمية كما تظهر في CRAN وكما تُستدعى برمجيًا، لأن R حساس لحالة الأحرف (case-sensitive)، لذا تُكتب الحزم: mirt و ltm (Chalmers, 2012)، بينما كتابة MIRT قد تُفهم على أنها اسم مختلف وغير قابل للاستدعاء في R. إحصائيات المطابقة لاستخراج مؤشرات حسن المطابقة.

[1] **مؤشرات مربع كاي:** تُفسَّر دلالة اختبار كاي تربيع(χ^2) على النحو الآتي:

1. إذا كانت القيمة الاحتمالية ($p > 0.05$) فإن الفروق بين الاستجابات الملاحظة والاستجابات المتوقعة وفق النموذج لا تُعد دالة إحصائيًا، مما يشير إلى ملاءمة جيدة للمفردة.

2. إذا كانت ($p \leq 0.05$) فيُحتمل وجود سوء ملاءمة. ومع ذلك، يُراعى أن هذا الاختبار قد يكون حساسًا لحجم العينة.

[2] مؤشر RMSEA (Root Mean Square Error of Approximation)

يقيس RMSEA مدى ملاءمة النموذج مع البيانات لكل درجة حرية، مع مراعاة تعقيد النموذج. تشير القيم المنخفضة إلى توافق أفضل بين البيانات الفعلية والمتوقعة:

- **توافق جيد** :أقل من 0.05
- **توافق مقبول** :بين 0.05 و0.08

- **سوء توافق** :أكبر من 0.10

5- حساب معالم الفقرات والافراد.

نتائج الدراسة ومناقشتها

نتائج الهدف الفرعي الأول: ينص الهدف الفرعي الأول على: " **التحقق من مطابقة بيانات اختبار المصفوفات المتتابعة المعياري لافتراضات نظرية الاستجابة للمفردة**"،

وقد تم اتباع الاجراءات الاتية للتحقق من الافتراضات:

1. **التحقق من افتراض أحادية البعد:** تم اتباع أكثر من إجراء؛ إذ إن الاعتماد على مؤشر واحد لا يُعد كافيًا للحكم على أحادية البعد، لذلك جرى استخدام مؤشرات تمهيدية ومؤشرات تأكيدية على النحو الآتي:

أولًا: تحليل تمهيدي باستخدام طريقة المكونات الأساسية(PCA)

1. التحقق من كفاية حجم العينة ومدى الاعتمادية على العوامل المستخلصة من التحليل العاملي في تفسير الاداء باستخدام اختبار كايزر- ماير- أولكن (KMO) والنتائج مبينة في الجدول (2).

جدول (2) نتائج اختبار (KMO) وبارتليت للتكور

الاختبار	القيمة		مستوى الدلالة
(KMO)	0.92		دالة
بارتليت للتكور	مربع كاي	16112.817	دالة

يتضح من الجدول (2) أن قيمة اختبار (KMO) تدل على أن العينة مناسبة وكافية لإجراء التحليل العاملي، كما أن قيمة اختبار بارتليت دالة إحصائيًا مما يعني أن بنود الاختبار مترابطة بما يكفي لإجراء التحليل العاملي في الدراسة، وتؤكد هذه النتائج تمتع بيانات اختبار المصفوفات المتتابعة بالشروط اللازمة للتحليل العاملي.

تم إجراء التحليل العاملي للمصفوفة الارتباط باستخدام طريقة المكونات الأساسية. وقد استخدمت الباحثة محك جليفورد لتحديد الدلالة الإحصائية للتشبعات الناتجة، حيث اعتُبرت التشبعات التي تساوي 0.30 أو أكثر بأنها دالة إحصائيًا. بناءً على نتائج التحليل العاملي، تم استخراج 15 عاملًا رئيسيًا، حيث كان الجذر الكامن لكل منها أكبر من الواحد الصحيح. الجدول التالي يوضح الجذور الكامنة والتباين المُفسَّر لكل من العوامل المستخرجة من التحليل العاملي.

جدول (3) الجذور الكامنة ونسبة التباين المفسر

العامل	الجذر الكامن	النسبة المئوية للتباين المفسر	النسبة المئوية للتباين التراكمي
1	11.08	18.47	18.47
2	3.04	5.06	23.53
3	2.07	3.45	26.99
4	1.87	3.11	30.10
5	1.61	2.67	32.78
6	1.58	2.63	35.40
7	1.42	2.37	37.77
8	1.30	2.16	39.93
9	1.26	2.09	42.02
10	1.18	1.96	43.98
11	1.15	1.91	45.89

12	1.13	1.88	47.78
13	1.07	1.78	49.56
14	1.06	1.77	51.33
15	1.02	1.71	53.03

في الجدول (3) افرزت طريقة المكونات الأساسية (15) عامل جذورها الكامنة أكبر من الواحد الصحيح، فسرت مجتمعة (53.03%) من التباين الكلي، وهي قيمة أكبر من الحد الادنى المناسب والفعال للاختزال الاحصائي للبيانات (50٪ – 75٪) من التباين الكلي، وهذا يدعم إمكانية تفسير هذه العوامل للأداء. يتضح من الجدول أن نسبة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني تبلغ 3.65، وهو ما يزيد عن 2 وفقًا لمعيار لورد، الذي يحدد أنه يجب أن تكون هذه النسبة أكبر من 2 لتأكيد أحادية البُعد. هذا يعني أن العامل الأول يفسر الجزء الأكبر من التباين في استجابات الأفراد، مما يدعم فكرة أن الاختبار يعتمد بشكل رئيسي على عامل واحد مهيمن.

على الرغم من أن التباين المفسر للعامل الأول(18%) ليس مرتفعاً جداً، إلا أن النسبة بين الجذر الكامن للعامل الأول والثاني ومعيار لورد تؤكد بوضوح أحادية البعد.

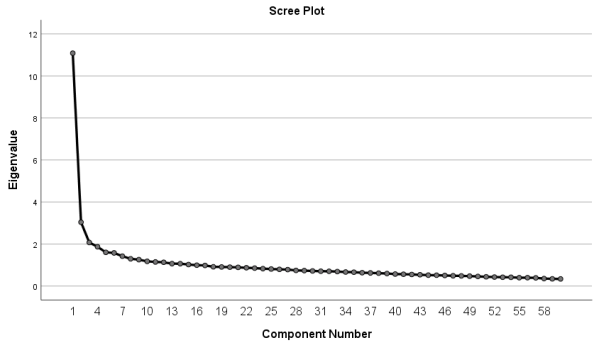
أما بالنسبة لمؤشر دلالة التشبع، الذي يُستخدم لتقييم مدى مساهمة كل مفردة في العامل المستخلص، فقد أظهرت النتائج أن جميع مفردات الاختبار تتشبع بشكل جيد بالعامل العام. حيث تراوحت التشبعات المعيارية بين 0.404 و0.635، مما يشير إلى أن جميع المفردات تُساهم بشكل كافٍ في تفسير العامل، وبالتالي يدعم افتراض أحادية البُعد للاختبار بدرجة كافية في اختبار رافن كماهو موضح في جدول (4).

جدول (4) درجة تشبع المفردات على العامل الأول (العام)

الفقرة	التشبع	الفقرة	التشبع	الفقرة	التشبع
q1	0.56	q21	0.56	q41	0.55
q2	0.63	q22	0.64	q42	0.53
q3	0.60	q23	0.54	q43	0.45
q4	0.63	q24	0.54	q44	0.51
q5	0.60	q25	0.46	q45	0.48
q6	0.53	q26	0.54	q46	0.47
q7	0.49	q27	0.53	q47	0.44
q8	0.57	q28	0.52	q48	0.55
q9	0.50	q29	0.51	q49	0.53
q10	0.48	q30	0.42	q50	0.40
q11	0.58	q31	0.52	q51	0.46

q12	0.61	q32	0.53	q52	0.61
q13	0.56	q33	0.45	q53	0.63
q14	0.46	q34	0.52	q54	0.51
q15	0.49	q35	0.46	q55	0.63
q16	0.59	q36	0.55	q56	0.48
q17	0.61	q37	0.54	q57	0.52
q18	0.54	q38	0.52	q58	0.49
q19	0.51	q39	0.58	q59	0.56
q20	0.44	q40	0.52	q60	0.62

التحليل البياني للجذور الكامنة (Scree Plot) كما هو موضح بالشكل البياني (2).



شكل (2) يوضح التحليل البياني للجذور الكامنة

يبين الشكل (2) ان الخط المنحنى بدأ يغير ميله بشكل مفاجئ بين النقطتين المناظرتين للعاملين الاول والثاني كما يلاحظ ان قيم الجذور الكامنة للعوامل الباقية متقاربة مع بعضها، والتغير في الميل طفيف، مما يعد مؤشر لوجود عامل سائد واحد يستدل منه على أحادية البعد، أي أن هناك سمة كامنة واحدة يقيسها الاختبار، هي المسؤولة عن تفسير ما يحدث من تباين في درجات الطلبة.

ثانياً: التحليل العاملي التوكيدي (CFA)

تم إجراء التحليل العاملي التوكيدي (CFA) لاختبار نموذج عامل واحد بوصفه نموذجًا أحادي البعد، وتم تقييم ملاءمة هذا النموذج باستخدام مؤشرات المطابقة مؤشر المطابقة المقارن (CFI) ومؤشر لويس تاكر (TLI). وقد جاءت قيم المطابقة ضمن الحدود المقبولة (CFI= 0.957، TLI =0.952)

بما يدعم أن بنية الاختبار تتسم بـ أحادية بعد أساسية (Essential Unidimensionality) وأن وجود عامل عام مهيمن يفسر جانبًا مهمًا من تباين استجابات الطلبة. كما تتسق هذه النتيجة مع المؤشرات التمهيدية المستخلصة من التحليل بطريقة المكونات الأساسية(PCA) ، والتي أشارت بدورها إلى هيمنة العامل الأول.

بناءً على المؤشرات السابقة يمكن القول بثقة إن اختبار المصفوفات المتتابعة المعياري يتسم بأحادية البعد، حيث تقيس جميع الفقرات بُعدًا كامنًا واحدًا بدقة، مما يؤكد الصلاحية البنائية لأداة الدراسة. وتتفق هذه النتيجة مع الدراسات السابقة كدراسة شكري (2022) ودراسة عبد الرازق (2022) اللتان اكدتا احاديه البعد وعززتا الأساس المنطقي لتصميم الاختبار الذي يركز على قياس الذكاء العام.

2- التحقق من افتراض الاستقلال المحلي

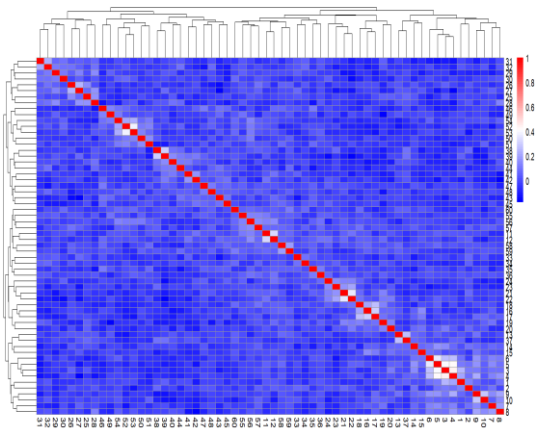
تم التحقق من افتراض الاستقلال المحلي بالطرق الاتية:

1. اختبار مربع كاي (Chi-Square Test) بين المفردات:

تم تطبيق اختبار مربع كاي (χ^2)على 1770 زوجًا من المفردات من أجل فحص الاستقلال المحلي بين مفردات الاختبار، وجميع النتائج كانت غير دالة إحصائياً عند مستوى دلالة ($P>0.05$) الذي يعد مؤشر لتحقق الاستقلال الموضعي. يعني هذا أن المفردة الأولى تم اختبارها ضد جميع المفردات الأخرى الـ 59 في الاختبار. بعد ذلك، يتم اختبار المفردة الثانية ضد الـ 58 مفردة المتبقية، وهكذا حتى يتم اختبار جميع المفردات ضد بعضها البعض.

كما تم فحص الاستقلال المحلي بعد ملاءمة نموذج (4PL)باستخدام معامل ارتباط البواقي (Yen's Q3) بوصفه مؤشرًا شائعًا للكشف عن الاعتمادية المحلية بين المفردات. ويُنظر عادةً إلى القيم التي تبلغ $|Q3| \geq 0.20$على أنها مؤشر مبدئي لاحتمال وجود اعتماد محلي، بينما تشير القيم الأعلى (مثل 0.30 فأكثر) إلى اعتمادٍ أقوى. وبتحليل **(1770)** زوجًا من المفردات (لعدد 60 مفردة)، أظهرت الخريطة الحرارية شكل (3) أن معظم معاملات Q3 كانت قريبة من الصفر وخارج القطر الرئيسي، بما يدعم تحقق الاستقلال المحلي بدرجة كبيرة. كما أظهرت النتائج أن 24 زوجًا فقط تجاوزت $|Q3| \geq$ 0.20بنسبة تقارب 1.36%من إجمالي الأزواج، في حين انخفض العدد إلى 8 أزواج عند $|Q3| \geq$ 0.30، وإلى3 أزواج فقط عند $|Q3| \geq 0.40$، وهو ما يشير إلى أن الاعتمادية المحلية—إن وُجدت—فهي محدودة ومحصورة في عدد قليل من الأزواج.

بمعنى آخر، فإن الاستجابات للمفردات تعتمد أساسًا على السمة الكامنة المقاسة من الاختبار ولا يوجد تأثير متبادل قوي بين المفردات خارج هذه السمة.



شكل (3) يوضح الخريطة الحرارية

كما لوحظ في الخريطة شكل (3) وجود بعض المربعات ذات اللون الوردي، ما يُشير إلى احتمال وجود درجة من الاعتمادية بين المفردات ذات ذلك اللون. وبتحليل المفردات (3,4,5,6,11) وجد انها مفردات سهلة جدًا وهناك ارتباط بينها إذا تمكن المستجيبون من التعرف على نمط أو قاعدة في الاختبار تُمكنهم من إعطاء الإجابة الصحيحة دون الحاجة الفعلية لفهم المحتوى أو استخدام السمة المقصود قياسها. أما بالنسبة للمفردة (52) فهي مفردة صعبة جدًا قد يلجأ المستجيبون عليها إلى التخمين، والذي قد يؤدي إلى ارتباط عالٍ بين المفردات الصعبة جدًا إذا كان المستجيبون يميلون إلى التخمين فيها،

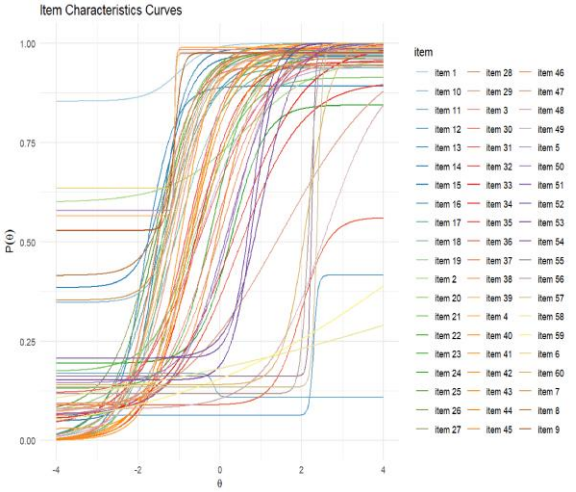
وقد ينتج عن ذلك ارتفاع في الارتباطات المحلية نظرًا لقلة التباين في الاستجابات.

3) التحقق من افتراض التحرر من عامل السرعة:

للتحقق من هذا الافتراض تم منح الطلبة وقتًا كافيًا للإجابة على مفردات الاختبار، وترك الوقت مفتوحًا أمامهم للإجابة عن جميع مفردات الاختبار، مما يضمن عدم تأثير الوقت على الأداء، ولذلك يمكننا القول بأن عامل السرعة لم يكن له أي دور في الإجابة على مفردات الاختبار، وأن عدم قدرة الطالب على الإجابة على أية مفردة من مفردات الاختبار يرجع فقط لقدرته.

4) المنحنى المميز للمفردة

يعتبر تحليل الاستجابة للمفردة (Item Response Theory – IRT) أحد الأدوات الأساسية في تقييم فاعلية الأدوات القياسية المستخدمة في القياس النفسي والتعليمي. يتسم هذا النوع من التحليل بقدرته على تقييم العلاقة بين قدرة المستجيب والمفردات المعروضة عليه، مما يسمح بتحديد مدى ملاءمة الأسئلة وفعاليتها في قياس القدرات المستهدفة. من بين الأدوات الأساسية المستخدمة في تحليل IRT، يبرز المنحنى المميز للمفردة (Item Characteristic Curve – ICC) كأداة رئيسية لفهم كيفية تفاعل المفردات مع طيف واسع من قدرات المستجيبين.

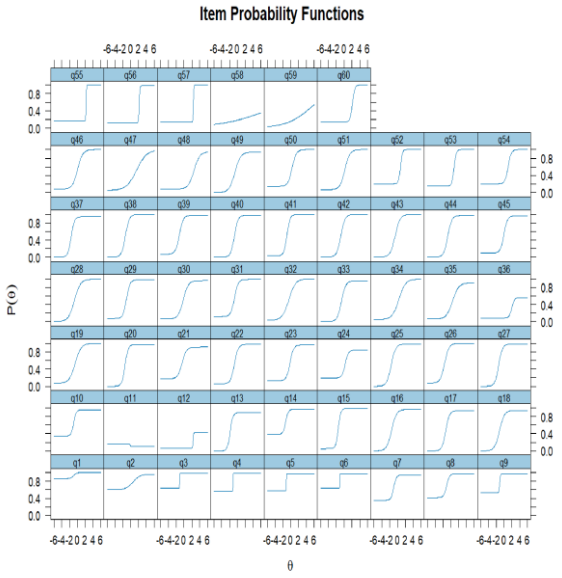


شكل (4) المنحنى المميز للفقرات

يسهم رسم المنحنيات المميِّزة للمفردات في تقديم تقييمٍ متكامل لفعالية المفردات ودقتها في التمييز عبر مستويات القدرة المختلفة. صياغات. كما يتيح تقييم كيفية تفاعل المفردات مع قدرات المستجيبين المختلفة، وبالتالي تحديد مدى قدرتها على التمييز بين الأفراد ذوي القدرات المختلفة.

في الشكل (5)، يتم عرض دالة الاحتمال للمفردة، التي تمثل علاقة القدرة الكامنة للمستجيب (θ)مع احتمالية الإجابة الصحيحة على مفردة معينة. على الرغم من أن هذا الشكل يشبه الـICC ، إلا أن الفارق الرئيسي هو أنه يعكس الاحتمال الفردي لمفردة واحدة في اختبار معين، وليس توزيع مفردات متعددة.

تُظهر هذه الدالة كيفية تغير الاحتمالية لكل مفردة بناءً على قدرة المستجيب. حيث يُظهر منحنى بشكل مباشر كيف تزداد احتمالية النجاح في الإجابة مع زيادة القدرة للمستجيب.



شكل (5) دالة الاحتمال للمفردة (Item Probability Function)

- في المنحنى المميز للمفرد(ICC) ، يتم رسم عدة مفردات معًا، ما يتيح مقارنة تفاعل مختلف الأسئلة مع طيف واسع من القدرات في وقت واحد. أما في دالة الاحتمال، فإنك تركز على مفردة واحدة فقط وتدرس احتمال الإجابة الصحيحة بناءً على قدرة المستجيب.
- بينما يُظهر ICCبشكل شامل أداء المفردات في الاختبار ككل، فإن دالة الاحتمال يعزز الفهم الدقيق للتمييز الخاص بكل مفردة على حدة، ويساعد في تحسين تصميم الأسئلة لمستوى أفضل من التمييز.
- من خلال مقارنة هذه الأشكال، يمكننا الحصول على رؤى هامة حول فعالية كل مفردة داخل الاختبار وكيفية تعاملها مع المستجيبين المختلفين، المنحنى المميز للمفرد يقدم تقييمًا عامًا لجميع المفردات، في حين أن دالة الاحتمال يتيح تقييمًا دقيقًا لكل مفردة على حدة.

نتائج الهدف الفرعي الثاني

ينص الهدف الفرعي الثاني على "التحقق من حسن مطابقة وملاءمة بيانات اختبار المصفوفات المتتابعة المعياري لنموذج الاستجابة للمفردة الرباعي المعلمة (4PL)" من خلال المستويات التالية:

1. ملاءمة اختبار المصفوفات المتتابعة المعياري ككل لنموذج الاستجابة للمفردة الرباعي المعلمة.
2. ملاءمة مفردات اختبار مصفوفات رافن لنموذج الاستجابة للمفردة الرباعي المعلمة (4PL) على مستوى المفردة.
3. ملاءمة بيانات الأفراد المستجيبين لنموذج الاستجابة للمفردة الرباعي المعلمة.

- **تحديد مدى ملاءمة اختبار المصفوفات المتتابعة المعياري ككل لنموذج الاستجابة للمفردة الرباعي المعلمة (4PL)."**

من خلال حساب قيم المؤشرات الاتية:

– تم حساب قيم مؤشرات والرصان (CFI) وتاكر– لويس (TLI) وجذر متوسط مربعات الخطأ التقريبية (RMSEA) كما موضح في الجدول (5).

جدول (5) قيم مؤشرات ملاءمة الاختبار ككل لنموذج الاستجابة للمفردة رباعي المعلمة

المؤشر	القيمة	معيار الملائمة الجيدة للنموذج
(RMSEA)	0.0405	القيمة أقل من 0.06 تعتبر مثالية

الرصان (CFI)	0.9571	القيمة تقترب من الواحد
تاكر – لويس (TLI)	0.9522	القيمة أعلى من 0.90

من الجدول (5) يتضح أن قيم جميع المؤشرات تقع في إطار المعيار مما يدل على ملائمة جيدة للاختبار ككل لنموذج الاستجابة للمفردة رباعي المعلمة.

- **مدى ملاءمة مفردات اختبار مصفوفات رافن لنموذج الاستجابة للمفردة الرباعي المعلمة (4PL) على مستوى المفردة.**

من خلال حساب قيم مؤشرات إحصائيات شملت مربع كاي مع درجات الحرية المناسبة (df) كاي تربيع (χ^2) حسب طريقة بوك (1972)، واس كاي تربيع (χ^2 –S) وفق طريقة أورلاندو وثيسن (2000، 2003)، وإحصاء G^2 وفق طريقة ماكينلي وميلز (1985)، ومؤشرات RMSEA لكل اختبار (Liu & Alberto, 2014)، كما هي موضحة في الجدول (6).

جدول (6) مؤشرات مربع كاي(χ^2) واس كاي تربيع(S_χ^2) وإحصاء G^2 و RMSEAلتقييم مطابقة المفردات للنموذج رباعي المعلمة

م	كاي تربيع طريقة بوك				اس كاي تربيع طريقة أورلاندو وثيسن				G^2طريقة ماكينلي وميلز			
	X^2	Df	$p.X^2$	RMSEA	S_X^2	df	p S_X^2	RMSEA	G^2	df	p. G2	RMSEA
q1	14.96	6	0.02	0.04	19.24	15	0.20	0.01	15.08	1	0.00	0.12
J, q2	4.52	6	0.61	0.00	43.50	41	0.37	0.00	4.55	6	0.60	0.00
q3	7.70	6	0.26	0.02	29.36	27	0.34	0.02	7.49	5	0.19	0.02
q4	5.06	6	0.54	0.00	25.22	26	0.51	0.02	5.76	5	0.33	0.01
q5	3.25	6	0.78	0.00	42.77	30	0.06	0.02	3.34	6	0.77	0.00
q6	7.05	6	0.32	0.01	18.69	30	0.95	0.01	7.72	4	0.10	0.03
q7	8.34	6	0.21	0.02	46.60	36	0.11	0.03	8.20	6	0.22	0.02
q8	16.41	6	0.01	0.04	42.90	34	0.14	0.01	16.94	5	0.00	0.05
q9	14.47	6	0.02	0.04	42.79	30	0.06	0.03	14.76	5	0.01	0.04
q10	8.58	6	0.20	0.02	30.94	33	0.57	0.01	9.13	6	0.17	0.02
q11	12.16	6	0.06	0.03	57.66	32	0.00	0.02	11.27	6	0.08	0.03
q12	7.96	6	0.24	0.02	35.98	26	0.09	0.02	7.31	6	0.29	0.01
q13	16.70	6	0.01	0.04	47.07	36	0.10	0.00	17.17	6	0.01	0.04
q14	8.57	6	0.20	0.02	38.19	36	0.37	0.01	9.01	6	0.17	0.02

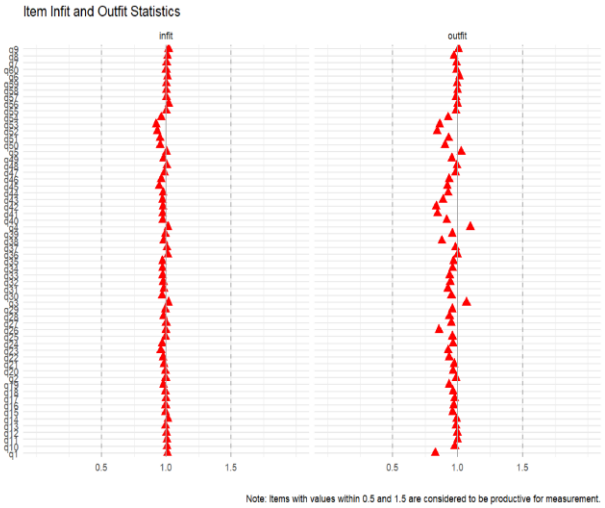
0.01	0.29	5	6.13	0.02	0.44	28	28.48	0.00	0.64	6	4.25	q15
0.00	0.66	6	4.16	0.01	0.31	36	39.65	0.00	0.67	6	4.05	q16
0.03	0.05	6	12.87	0.00	0.78	35	28.44	0.03	0.04	6	13.13	q17
0.00	0.54	6	5.05	0.00	0.41	35	36.31	0.00	0.54	6	5.03	q18
0.02	0.20	6	8.52	0.02	0.07	35	48.05	0.02	0.24	6	8.02	q19
0.02	0.24	6	7.97	0.03	0.08	33	45.16	0.01	0.30	6	7.23	q20
0.00	0.76	6	3.36	0.01	0.19	38	45.39	0.00	0.78	6	3.26	q21
0.01	0.40	6	6.22	0.03	0.25	30	34.87	0.00	0.46	6	5.72	q22
0.02	0.16	6	9.31	0.01	0.32	32	35.07	0.02	0.17	6	9.04	q23
0.02	0.28	6	7.45	0.01	0.42	38	39.05	0.01	0.30	6	7.26	q24
0.00	0.74	6	3.52	0.00	0.14	35	44.18	0.00	0.76	6	3.40	q25
0.00	0.42	5	4.94	0.02	0.02	31	48.97	0.00	0.62	6	4.39	q26
0.03	0.07	5	10.25	0.02	0.23	32	37.64	0.02	0.22	6	8.24	q27
0.00	0.80	6	3.08	0.01	0.73	34	28.63	0.00	0.80	6	3.04	q28
0.01	0.30	6	7.25	0.00	0.32	33	36.19	0.01	0.32	6	6.99	q29
0.04	0.02	6	15.36	0.01	0.63	32	28.70	0.04	0.02	6	14.71	q30
0.02	0.22	6	8.29	0.00	0.58	33	30.68	0.02	0.25	6	7.83	q31
0.01	0.32	6	7.03	0.01	0.44	35	35.53	0.01	0.34	6	6.78	q32
0.00	0.58	6	4.70	0.02	0.05	30	44.17	0.00	0.61	6	4.49	q33
0.02	0.22	6	8.24	0.01	0.38	32	33.73	0.02	0.25	6	7.89	q34
0.04	0.01	6	18.32	0.02	0.45	33	33.41	0.04	0.01	6	17.69	q35
0.06	0.00	6	27.85	0.02	0.03	32	48.29	0.06	0.00	6	28.01	q36
0.05	0.00	4	15.59	0.02	0.04	28	42.32	0.03	0.05	6	12.48	q37
0.02	0.23	5	6.87	0.02	0.27	30	34.37	0.01	0.36	6	6.57	q38
0.00	0.45	6	5.80	0.01	0.26	33	37.67	0.00	0.50	6	5.31	q39
0.00	0.43	6	5.93	0.01	0.47	30	29.85	0.00	0.48	6	5.50	q40
0.03	0.07	4	8.67	0.00	0.51	26	25.19	0.01	0.31	6	7.14	q41
0.02	0.15	5	8.14	0.01	0.13	29	37.66	0.01	0.32	6	7.01	q42
0.03	0.07	6	11.47	0.00	0.63	29	25.89	0.03	0.09	6	10.80	q43
0.01	0.39	6	6.32	0.01	0.58	30	27.90	0.00	0.45	6	5.75	q44
0.01	0.29	6	7.31	0.00	0.40	30	31.41	0.01	0.37	6	6.54	q45
0.01	0.33	6	6.92	0.02	0.05	32	45.94	0.00	0.45	6	5.80	q46

0.06	0.00	6	26.62	0.00	0.84	34	25.82	0.06	0.00	6	26.99	q47
0.04	0.02	6	15.26	0.02	0.05	32	46.19	0.04	0.01	6	15.82	q48
0.03	0.06	6	12.09	0.01	0.27	34	38.45	0.03	0.07	6	11.78	q49
0.04	0.01	6	15.89	0.00	0.57	32	29.97	0.03	0.04	6	13.45	q50
0.00	0.84	6	2.78	0.02	0.31	31	34.31	0.00	0.86	6	2.62	q51
0.04	0.04	5	11.69	0.02	0.01	31	51.15	0.02	0.17	6	9.09	q52
0.04	0.02	6	15.18	0.02	0.17	28	35.10	0.03	0.04	6	13.30	q53
0.06	0.00	6	25.70	0.00	0.23	36	41.88	0.05	0.00	6	24.27	q54
0.06	0.00	6	26.51	0.01	0.03	34	51.12	0.06	0.00	6	26.22	q55
0.07	0.00	6	34.44	0.02	0.93	31	20.21	0.08	0.00	6	46.00	q56
0.03	0.12	6	10.18	0.00	0.34	31	33.68	0.03	0.08	6	11.41	q57
0.05	0.00	6	20.30	0.03	0.00	34	68.09	0.05	0.00	6	21.48	q58
0.05	0.00	6	20.92	0.00	0.22	34	39.93	0.05	0.00	6	21.14	q59
0.03	0.06	6	12.00	0.00	0.98	34	18.75	0.03	0.08	6	11.37	q60

يتضح من الجدول (6) أن بيانات الاختبار تتلاءم بشكل عام مع النموذج الرباعي المعلمة. الا انه لوحظ أن بعض المفردات ذات الصعوبة العالية أظهرت مؤشرات ملاءمة غير مرضية في بعض الفقرات. أُخذ في الاعتبار الاحتفاظ بجميع المفردات، إذ لم تتفق مؤشرات الملاءمة المختلفة على رفض المفردات نفسها بصورة موحّدة، كما يُعتبر التدرج في مستوى الصعوبة من السمات المميزة لاختبار رافن، حيث يجب أن يمتد من الأسئلة السهلة جدًا إلى الأسئلة الصعبة جدًا لقياس القدرة العقلية بشكل شامل. هذا الجانب يُبرر الحاجة لشمول المفردات بمختلف درجات الصعوبة لضمان تكامل الاختبار وشموليته.

[1] كما تم استخدام مؤشرات الملاءمة الداخلية (Infit) والخارجية (Outfit) بالإضافة إلى القيم القياسية المتصلة Z.outfit و Z.infit. كما هو موضح في ملحق (2).

ومن الشكل البياني (6) يتضح أن جميع مفردات الاختبار مطابقة لنموذج الاستجابة للمفردة رباعي المعلمة وفقًا لمؤشرات Infit وOutfit، فقيمة احصائي المطابقة الخارجية والداخلية المقابلة لتقديرات معالم المفردات تقع داخل حدود المطابقة المناسبة (0.5- 1.5).



شكل (6) مؤشرات الملاءمة الداخلية (infit) والخارجية (Outfit) الحرف الأول كبير

تظهر الفقرات التوافق مع معايير الملاءمة للبيانات وفقًا لمؤشرات Infit وOutfit، مما يعزز من صلاحيتها ضمن الاختبار.

تحديد مدى ملائمة بيانات الأفراد المستجيبين لنموذج الاستجابة للمفردة الرباعي المعلمة (4PL).

لتحديد مدى ملائمة بيانات الأفراد المستجيبين لنموذج الاستجابة للمفردة الرباعي المعلمة (4PL) من خلال استخدام عدة مؤشرات وتحليلات إحصائية. تم فحص الاستجابات باستخدام إحصائيات(Infit, Outfit , Z.Outfit, Z.Infit) بالإضافة إلى إحصائية Zh لتقييم مدى توافق الأفراد مع النموذج.

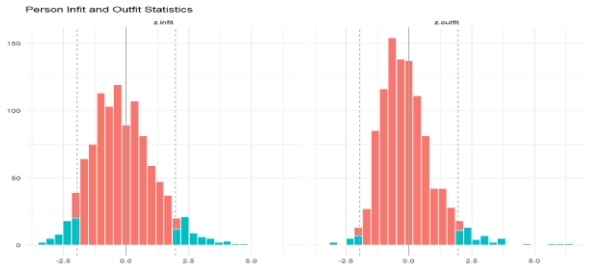
فراغ بين الكلمات الانجليزية والعربية

إحصائية Zh تُعد من المؤشرات الهامة في تحديد مدى التوافق العام للمشاركين مع النموذج، حيث أنها تُظهر الفرق بين القيم الفعلية للنموذج (المقاسة) والتوقعات النظرية. في حال كانت القيمة ضمن النطاق المقبول (−2 إلى 2)، يعتبر النموذج مناسبًا للمستجيبين.

جدول (7) نتائج ملاءمة الأفراد لنموذج الاستجابة للمفردة:

المؤشر	عدد الأفراد المتجاوزين للنطاق المقبول	النسبة المئوية
Z.Outfit	72	7%
Z.Infit	133	12.9%
Outfit	43	4.2%
Infit	3	0.3%
Zh	91	9%

وفي هذه الدراسة، أظهرت إحصائية Zh توافقًا جيدًا للنموذج، على الرغم من وجود بعض الاستجابات التي تحتاج إلى تدقيق إضافي. كما تشير النتائج إلى أن غالبية الأفراد المستجيبين أظهروا توافقًا جيدًا مع نموذج الاستجابة للمفردة الرباعي المعلمة(4PL) ، حيث كانت نسبة الأفراد الذين لم يتوافقوا مع النموذج منخفضة وفقًا لهذه الإحصائيات. هذه النتائج تدعم فعالية النموذج في قياس الاستجابات بشكل دقيق وموثوق، بهذا يمكن التأكيد على أن البيانات تظهر ملاءمة جيدة لنموذج الاستجابة للمفردة الرباعي المعلمة (4PL) في قياس الاستجابات وفقًا للمعايير المقبولة.



شكل (7) يوضح المستجيبين الذين لا يتوافقون مع النموذج الرباعي المعلمة باللون الأخضر مقابل أولئك المتوافقين مع النموذج الذين تم تمييزهم باللون الأحمر.

نتائج الهدف الفرعي الثالث

ينص الهدف الفرعي الثالث على " **تقدير معالم مفردات اختبار المصفوفات المتتابعة المعياري وفق نموذج الاستجابة للمفردة رباعي المعلمة.**"

لتحقيق هذا الهدف تم استخدام برنامج (R) في تقدير كما هو موضح في جدول (8)

جدول (8) تقدير معالم مفردات اختبار المصفوفات المتتابعة المعياري وفق نموذج الاستجابة للمفردة رباعي المعلمة

المفردة	التمييز	الصعوبة	التخمين	الإهمال
q1	1.50	−2.95	0.00	0.999
q2	0.45	−2.30	0.00	0.999
q3	1.84	−2.25	0.00	0.996
q4	3.59	−1.58	0.30	0.99
q5	2.18	−2.03	0.00	0.984
q6	1.72	−2.21	0.00	0.997
q7	1.74	−1.35	0.12	0.966
q8	1.60	−1.74	0.00	0.989
q9	2.58	−1.86	0.00	0.983
q10	2.00	−1.35	0.08	0.968
q11	8.68	2.70	0.10	0.999
q12	11.87	2.87	0.05	0.999
q13	3.36	−1.99	0.00	0.929
q14	1.93	−1.91	0.00	0.974
q15	2.90	−1.53	0.08	0.988
q16	1.61	−1.39	0.10	0.968
q17	1.96	−1.34	0.00	0.939
q18	1.49	−1.21	0.00	0.954

0.999	0.09	−0.54	1.32	q19
0.971	0.00	−1.49	2.50	q20
0.914	0.20	−0.45	1.91	q21
0.981	0.06	−0.78	2.02	q22
0.961	0.13	−0.20	2.23	q23
0.844	0.19	0.27	2.50	q24
0.983	0.00	−1.65	1.72	q25
0.999	0.07	−1.63	1.74	q26
0.982	0.08	−1.34	2.18	q27
0.994	0.00	−0.77	1.26	q28
0.975	0.08	−1.11	2.27	q29
0.951	0.06	−0.23	1.58	q30
0.982	0.15	−0.71	2.30	q31
0.999	0.00	−0.56	1.13	q32
0.948	0.00	−0.69	1.84	q33
0.994	0.06	0.38	1.21	q34
0.885	0.06	0.39	1.32	q35
0.511	0.09	1.84	3.44	q36
0.982	0.00	−1.68	2.95	q37

0.996	0.03	−1.33	2.09	q38
0.972	0.11	−1.14	2.17	q39
0.974	0.05	−0.72	2.31	q40
0.999	0.00	−1.03	2.38	q41
0.999	0.05	−0.75	1.98	q42
0.999	0.00	−0.58	1.61	q43
0.972	0.00	−0.78	1.95	q44
0.955	0.11	−0.06	2.37	q45
0.999	0.11	0.11	1.66	q46
0.851	0.10	1.17	1.02	q47
0.605	0.09	1.83	1.99	q48
0.934	0.00	−0.67	1.42	q49
0.999	0.15	0.33	1.91	q50
0.968	0.09	0.16	1.84	q51
0.999	0.18	0.73	3.69	q52
0.999	0.13	0.79	3.38	q53
0.999	0.20	1.03	2.61	q54
0.999	0.15	2.35	21.66	q55
0.999	0.09	2.16	13.21	q56
0.999	0.11	2.37	22.56	q57

q58	22.13	2.44	0.17	0.999
q59	27.11	2.49	0.15	0.999
q60	2.99	2.15	0.14	0.999

ويلاحظ من الجدول (8) ان قيم معلمة التمييز تراوحت بين (0.45) للفقرة (2) وهي قيمه الأصغر. و (27.11) للمفردة (59) وهي الأكبر، بمتوسط بلغ قيمة (3.94) وانحراف معياري (5.69).

وقد توزعت قيم معلمة تمييز فقرات الاختبار وفقا لمعيار بيكر (Baker, 2001.35) **كما في جدول (9)**

جدول (9) توزيع قيم معلمة تمييز فقرات الاختبار وفقا لمعيار بيكر

مدى قيم معلمة التمييز	مستوى التمييز	عدد الفقرات	الفقرات
اقل من 0.34	منخفض جدا	–	–
0.64 -0.35	منخفض	1	2
1.34-0.65	متوسط	6	19,28,32,34,35,47
1.69 -1.35	مرتفع	8	1,8,16,18,30,44,47,49
أكثر من 1.70	مرتفع جدا	45	3–7، 159–، 17، 2521–، 29، 31، 33، 4036–، 41، 42، 44، 45، 5048–، و51 حتى 60.

ومن خلال الجدول يتضح أن (8) مفردات جاءت ذات تمييز مرتفع، بينما بلغت المفردات ذات التمييز المرتفع جدًا (45) مفردة، وهو ما يعكس قدرةً عالية للاختبار على التفريق بين ذوي القدرة المرتفعة والمنخفضة، ومع ذلك، لوحِظ ارتفاعٌ غير معتاد في قيم بعض معاملات التمييز للمفردات (55–59)، وهي ظاهرة قد ترتبط بخصائص تقدير نموذج الاستجابة للمفردة رباعي المعلمة (4PL) عند وجود مفردات شديدة الصعوبة أو عند انحراف توزيع الاستجابات، حيث قد تميل طرق التقدير إلى إعطاء ميلٍ حاد لمنحنى خصائص المفردة بما يرفع قيمة معامل التمييز بصورة تتجاوز المدى الشائع، وهو ما أشار الية (Loken & Rulison, 2010: 523) ويُلاحظ أن هذا التفسير يتسق أيضًا مع طبيعة بناء اختبار رافن؛ إذ تُرتَّب مفرداته تدريجيًا من السهل إلى الصعب، ما يجعل المفردات الأخيرة أكثر صعوبة

بحكم التصميم، وقد يؤدي ذلك إلى قلة الاستجابات الصحيحة في نهايات الاختبار ومن ثمّ ظهور تقديرات تمييز مرتفعة لبعض تلك المفردات. وبالنظر الى قيم معلمة صعوبة المفردات في الجدول (8) يلاحظ انها تروحت ما بين (2.87) للمفردة (12) وهي أكبر قيمة (صعبة جدا) و (−2.95) للمفردة (1) وهي أصغر قيمة (سهلة جدا) وبمتوسط (−0.355) وانحراف معياري (1.486).

وقد توزعت قيم معلمة الصعوبة للمفردات الاختبار وفقا لمعيار بيكر كما في جدول (10)

جدول (10) توزيع قيم معلمة الصعوبة الاختبار وفقا لمعايير بيكر

مدى قيم معلمة الصعوبة	مستوى الصعوبة	عدد الفقرات	الفقرات
من (−2) إ6لى (−3)	سهلة جدا	5	1,2,3,5,6
من (−1) إلى (−2)	سهلة	20	4,7,8,9,10,13,14,15 16, 17, 18,20, 25 27, 29, 37, 38, 39, 41
من (1) إلى (−1)	متوسطة	23	19, 21,22, 23, 24, 28, 30, 31, 32, 33,34,35,40, 42, 43, 44, 45, 46, 49, 50, 51, 52, 53
من (2) الى (1)	صعبة	4	36, 47, 48, 54
من (3) الى (2)	صعبة جدا	8	11, 12, 55, 56, 57, 58, 59, 60

نلاحظ من الجدول (10) أن 5 مفردات كانت سهلة جدا و8 مفردات صعبة جدا بينما كانت اغلب المفردات سهله ومتوسطة وتأكد أن مفردات اختبار المصفوفات المتتابعة المعياري يغطى مدى واسع من الصعوبة.

بالنسبة إلى معامل التخمين في الجدول (8)، يُلاحظ أنه تراوح بين (0.00–0.03) وجميع القيم تقع ضمن النطاق المقبول للتخمين وفقًا لمحك بيكر، الذي ينص على أن معامل التخمين لا يجب أن يتجاوز 0.35. يوضح هذا أن المفردات في الاختبار لا تعتمد بشكل كبير على التخمين، حيث أن الاستجابات تكون معتمدة بشكل أساسي على قدرات الأفراد ومعرفتهم الفعلية بدلاً من التخمين العشوائي.

وقد بلغ متوسط معاملات التخمين 0.073، مما يعكس أن التخمين في الاختبار بشكل عام منخفض جدًا ولا يؤثر بشكل كبير على النتائج. أما الانحراف المعياري

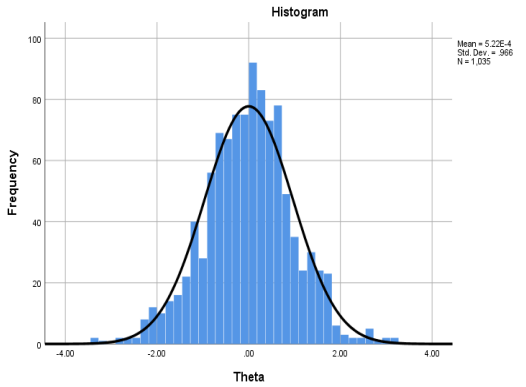
البالغ 0.069فيشير إلى أن التفاوت بين معاملات التخمين للفقرات يعتبر ضئيلًا جدًا، مما يعزز فكرة أن جميع الفقرات تحتوي على تخمين منخفض للغاية، ويعكس بذلك دقة عالية في قياس قدرات الأفراد.

بالنسبة لمعاملات الإهمال بلغ متوسط معامل الإهمال0.961 بانحراف معياري 0.083، ما يشير إلى أن أغلب مفردات الاختبار سجلت قيماً قريبة من الواحد الصحيح، وهو ما يعكس انخفاض الإهمال وارتفاع التزام المفحوصين ودقتهم في الاستجابة، مع تشتت محدود يدل على اتساق السلوك عبر المفردات. ويُعزَّز هذا التفسير بكون العينة من **طلبة الجامعات** الذين يُتوقع عادةً أن يكون سلوكهم أكثر **جدية** وانضباطًا أثناء الإجابة.

سجلت مفردتان أدنى قيم لمعامل الإهمال، وهما المفردة (35) والمفردة(48) وكان قيمة معلمة الاهمال على التوالي (0.51و 0.61) وهما فقرتين صعبة وذات تمييز عالٍ، ولكن تحتوي على مستوى من الاهمال، فقد يكون ذلك نتيجة لصعوبة المفردة التي قد تؤدي إلى استجابة غير متقنة من بعض المستجيبين، خصوصًا إذا كانت المفردة تتطلب تركيزًا أعلى أو مستوى معرفي معقد. من الممكن أن بعض المستجيبين، عند مواجهتهم لهذه الفقرات الصعبة، قد لا يولون اهتمامًا كافيًا للتمكن من الإجابة بدقة، وهو ما ينعكس في معامل الاهمال المرتفع نسبيًا.

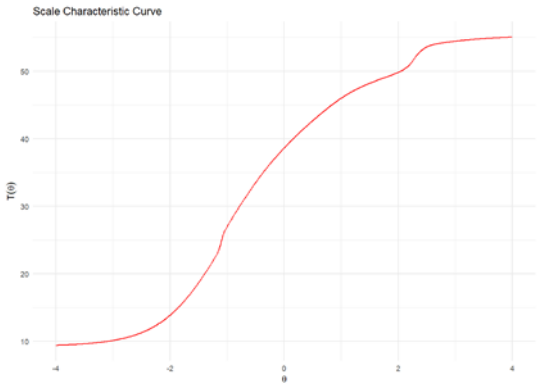
نتائج الهدف الفرعي الرابع:

10- ينص الهدف الفرعي الرابع على " **تقدير قدرات المستجيبين اختبار المصفوفات المتتابعة المعياري وفق نموذج الاستجابة للمفردة رباعي المعلمة**." لتحقيق هذا الهدف، تم استخدام برنامج (R)لتقدير قدرات المستجيبين، حيث تراوحت قيم القدرات بين (3.28-و3.17)، مع اعتبار 0.00 هو متوسط القدرة. تشير هذه النتائج إلى وجود نطاق واسع من القدرات بين أفراد العينة.



الشكل (8) يوضح الرسم البياني لمستويات القدرة المقدرة ويمثل المحور x السمة الكامنة(θ)

يوفر منحنى خصائص المقياس (Scale Characteristic Curve) الكلمات تبأ بحروف كبير نظرة تفصيلية حول كيفية استجابة درجات الاختبار من مصفوفات رافن للتغيرات في القدرة الكامنة للمتقدم للاختبار (θ)، والتي يتم تحليلها هنا باستخدام النموذج الرباعي المعلمة (4PL) في (IRT). يوضح الشكل (9) منحنى خصائص اختبار رافن.



شكل(9) منحنى خصائص المقياس

ويظهر ان منحنى على شكل حرف S وهو نموذجي للعديد من نماذجIRT ، وهو ما يعكس الطبيعة الاحتمالية للإجابة على الفقرات بشكل صحيح كدالة للقدرة. اما المحور الأفقي القدرة (θ) ويمثل المحور الأفقي السمة الكامنة أو مستوى القدرة للأفراد الذين يجرون الاختبار. تتمحور هذه السمة عادة حول الصفر، حيث تشير القيم الإيجابية إلى قدرة أعلى من المتوسط والقيم السلبية تشير إلى قدرة أقل من المتوسط. اما المحور الرأسي T(θ) – درجة الاختبار فيُظهر المحور الرأسي درجة الاختبار المتوقعة المقابلة لكل مستوى قدرة. وتزداد النتيجة مع زيادة القدرة، ولتوفير تفسير واضح لمنحنى خصائص المقياس رافن باستخدام النموذج الرباعي المعلمة في نظرية الاستجابة للمفردة، لنفترض أن هذه النقاط موضوعة بشكل استراتيجي عبر المنحنى لتسليط الضوء على قدرات محددة ودرجات الاختبار المقابلة.

وفيما يلي تفسير مفصل بناء على هذا الافتراض:

عند النقطة θ = −4: تقع هذه النقطة على الطرف الأيسر من المنحنى، وتمثل الأفراد ذوي القدرة الأقل بكثير من المتوسط. درجة الاختبار هنا منخفضة وقريبة من الخط المقارب الأدنى، مما يشير إلى أن الاختبار قد لا يميز بشكل فعال بين المستويات المختلفة للقدرة المنخفضة.

عند النقطة θ = −2: لا تزال هذه النقطة في النصف السفلي من المنحنى ولكنها أقرب إلى حد ما من القدرة المتوسطة. يتمتع الأفراد في هذه المرحلة بقدرة أقل من المتوسط ولكنها أعلى من مجموعة القدرة المنخفضة جدًا. درجات الاختبار المقابلة أعلى قليلاً، مما يدل على بعض القدرة على التمييز ضمن هذا النطاق، ولكنها لا تزال ليست في ذروة الحساسية.

عند θ = 0: تمثل عادةً متوسط مستوى القدرة، وهذه النقطة هي حيث يبدأ المنحنى في الانحدار بشكل ملحوظ. درجة الاختبار هنا تقع حول النطاق المتوسط، وهذه المنطقة غالبًا ما يصبح فيها الاختبار أكثر حساسية للتغيرات في القدرة، إنه أمر بالغ الأهمية لتحديد فعالية الاختبار في التمييز بين أصحاب الأداء المتوسط.

عند θ = 1: أعلى بقليل من المتوسط في القدرة، وتقع هذه النقطة في القسم الأوسط شديد الانحدار من المنحنى. تكون درجات الاختبار المقابلة أعلى وتزداد بسرعة مع زيادات طفيفة في القدرة. وهذا يشير إلى قوة تمييزية عالية للاختبار عند هذا المستوى.

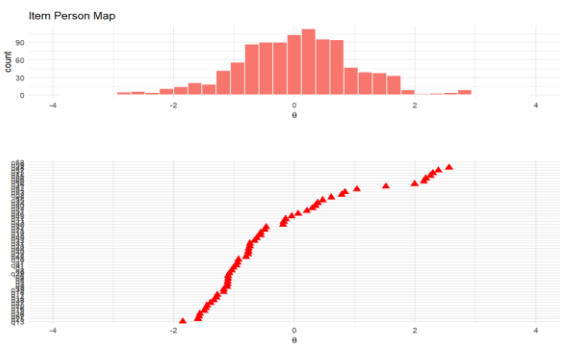
عند النقطة θ = 2: تقع هذه النقطة في النصف العلوي من قسم المنحنى الحاد، مما يشير إلى قدرة أعلى بكثير من المتوسط. تستمر درجات الاختبار في الارتفاع، ويميز الاختبار بشكل فعال بين الأفراد في نطاق القدرة العالية هذا، على الرغم من أن معدل زيادة الدرجات قد يبدأ في التباطؤ.

النقطة عند θ = 4: على الطرف الأيمن من المنحنى، تمثل هذه النقطة الأفراد ذوي القدرة العالية جدًا، بالقرب من الخط المقارب العلوي. الدرجات هنا عالية ولكنها تظهر زيادة طفيفة مع زيادات أخرى في القدرة، مما يشير إلى أن الاختبار قد يصل إلى مستوى ثابت ويوفر تمايزًا محدودًا عند مستويات القدرة العالية للغاية.

توضح هذه النقاط مجتمعة كيف تختلف درجات الاختبار باختلاف مستويات القدرة وأين يكون الاختبار أكثر وأقل فعالية في التمييز بين قدرات الأفراد.

في الشكل التالي (10)، خريطة الافراد والمفردات(Item Person Map) ومنحنى خصائص المقياس الخاص بنموذج الاستجابة للمفردة الرباعي المعلم . (4PL) توفر هذه الخريطة تمثيلاً بصريًا واضحًا لكيفية توزيع معلم الفقرات الاختبارية مقارنةً بمستويات قدرات المشاركين في العينة، حيث يتم عرض توزيع قدرات الأشخاص وصعوبة الفقرات على نفس المقياس، مما يسمح برؤية مدى توافق الفقرات مع مجموعة القدرات المتنوعة للمستجيبين.

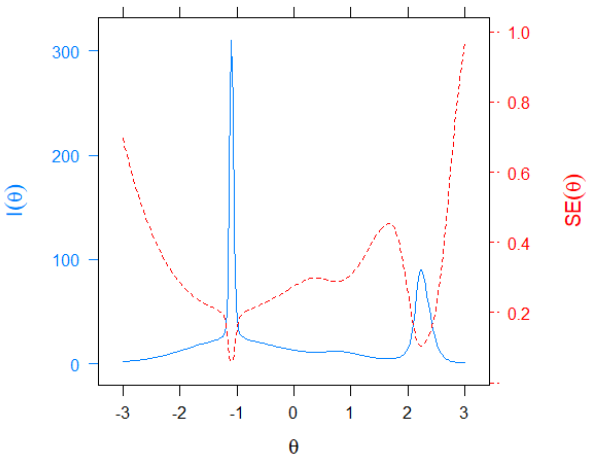
من ناحية أخرى، يوضح منحنى خصائص المقياس احتمالية الإجابة الصحيحة عبر متوالية مستويات القدرة، وهو ما يساعد في فهم خصائص كل مفردة من حيث الصعوبة والتمييز والتخمين والاهمال يمكن استخدام هذا الشكل لتقييم الفجوات في صعوبة المفردات أو التجمعات في قدرات الأشخاص، وكذلك لتعديل الاختبار وضمان أنه يوفر قياسًا موثوقًا وصالحًا للمفاهيم التي نحاول قياسها.



شكل (10) خريطة توزيع المفردات والمفحوصين في اختبار رافن(Item–Person Map)

نتائج الهدف الفرعي الخامس

ينص الهدف الفرعي الخامس على " **تقدير كمية المعلومات التي يقدمها اختبار المصفوفات المتتابعة المعياري لرافن عند مستويات القدرة المختلفة وفق نموذج الاستجابة للمفردة رباعي المعلمة** "، لتحقيق



هذا الهدف تم استخدام برنامج (R) في تقدير كمية المعلومات التي يقدمها اختبار المصفوفات المتتابعة المعياري لرافن عند مستويات القدرة المختلفة وفق نموذج الاستجابة للمفردة رباعي المعلمة وشكل (12) يُظهر الرسم البياني لمنحنيات دالة معلومات الفقرات (θ) والخطأ القياسي في منحنيات التقدير SE(θ) لمفردات الاختبار، التي تم تحليلها وفق نموذج 4PL. حيث يمثل الخط الأزرق الصلب المعلومات المقدمة من الفقرات عبر مستويات مختلفة من السمة (θ)، بينما يوضح الخط الأحمر المتقطع الخطأ القياسي المرتبط بتقدير θ."

شكل(12) يوضح دالة معلومات الاختبار (باللون الأزرق) و الخطأ المعياري للتقدير (باللون الأحمر المتقطع) عبر متصل السمة.

يُظهر منحنى دالة معلومات المفردات لنموذج 4PL قمتين او منوالين، تشير هذه القمم إلى وجود مجموعتين مختلفتين من العناصر داخل بيانات اختبار المصفوفات المتتابعة المعياري والتي توفر الحد الأقصى من المعلومات عند مستويات مختلفة من القدرة (θ). تشير الذروة العالية عند θ = −1 إلى مجموعة من المفردات المفيدة جدًا لمستويات القدرة المنخفضة يشير هذا إلى أن عناصر محددة توفر الحد الأقصى من المعلومات في هذه النقاط، مما يشير إلى أنها أكثر حساسية وتمييزًا للمستجيبين بالقرب من مستويات القدرة المنخفضة، بينما تشير الذروة الأصغر حول θ = 2 إلى مجموعة أخرى من المفردات الفعالة لمستويات القدرة الأعلى. كما يشير هذا التوزيع إلى أن الاختبار يميز بشكل فعال بين مستويات مختلفة من القدرة، خاصة بالنسبة لأولئك الذين يقعون في الأطراف الدنيا والعليا من الاختبار.

وتُظهر الخطوط المتقطعة الحمراء الخطأ المعياري، الذي يرتبط عكسيًا بكمية المعلومات المتاحة. عندما تصل المعلومات إلى ذروتها، ويكون الخطأ المعياري في أدنى مستوياته، مما يشير إلى دقة أعلى في تقدير قدرة الممتحنين في تلك النطاقات.

المقترحات

من خلال نتائج الدراسة الحالية ومن خلال النظر في عدد من الدراسات السابقة يمكن اقتراح ما يلي:

- إجراء المزيد من الدراسات على اختبار المصفوفات المتتابعة المعياري باستخدام نماذج الاستجابة للمفردة مع فئات اخرى من المجتمع اليمني.
- تدريج اختبار المصفوفات المتتابعة المعياري باستخدام نماذج نظرية الاستجابة للمفردة.
- المقارنة بين نموذج الاستجابة للمفردة رباعي المعلمة والنموذج ثلاثي المعلمة في تقدير معالم اختبار المصفوفات المتتابعة المعياري.

التوصيات

- توصي الدراسة باعتماد نموذج الاستجابة للمفردة رباعي المعلمة في تحليل بيانات الاختبارات المعيارية، لما يتميز به من دقة في تقدير معلمات الفقرات وقدرات المفحوصين مقارنة بالنماذج الأبسط.
- توصي الدراسة بتطوير بنوك فقرات معيارية مستندة إلى معالم دقيقة تم تقديرها وفق نماذج نظرية الاستجابة للمفردة، لتكون أساسًا لبناء اختبارات ذكاء وقدرات أكثر موضوعية في البيئة اليمنية.
- يوصى بتطبيق النموذج رباعي المعلمة على عينات أوسع من طلبة الجامعات في محافظات مختلفة، للتحقق من ثبات معلمات الفقرات وملاءمة النموذج لمختلف البيئات التعليمية.

قائمة المصادر والمراجع

أولاً: المراجع باللغة العربية

[1] أبو غالي، عطاف محمود، أبو مصطفى، نظمي عودة. (2014). تقنين اختبار المصفوفات المتتابعة العادي لرافن للفئة العمرية من (8–18) سنة على طلبة التعليم العام في محافظات غزة .مجلة جامعة طيبة للعلوم التربوية،(1)9،108–90.

[2] آل ثاني، العنود بنت مبارك بن أحمد. (2020). *تكافؤ نموذجي اختبار رافن (المحوسب – التقليدي) لدى عينة من طلاب وطالبات جامعة قطر في ضوء قلق الاختبار وقلق الحاسب والاتجاه نحو الحاسب* (رسالة دكتوراه غير منشورة). جامعة أم القرى، المملكة العربية السعودية.

[3] إسماعيل، ميمي السيد أحمد. (2007). الخصائص السيكومترية لاختبار القدرة العقلية باستخدام نموذج راش لدى طلبة المرحلة الثانوية العامة (رسالة ماجستير غير منشورة). كلية التربية، جامعة الزقازيق.

[4] التقي، احمد محمد. (2009). النظرية الحديثة في القياس، دار المسيرة عمان الاردن

[5] بترجي، عادل عبد الجليل، الخليفة، عمر هارون، خليل، يوسف أحمد. (2013). ارتفاع معدل الذكاء القومي في السعودية بين 1977–2010م: دراسة تحليلية لعينة من منطقة مكة المكرمة. المجلة التربوية, 28.(109)

[6] زمزمي، عبد الرحمن. (1998). تقنيين اختبار المصفوفات المتتابعة الملون لجون رافن على الطلاب الصم في معاهد الامل للمرحلة الابتدائية بالمملكة العربية السعودية.

[7] عبد الرازق، وفاء محمود نصار (2022). الخصائص الإحصائية لفقرات مقياس رافن للقدرات العقلية العامة في ضوء النموذج الثنائي والثلاثي المعلم–دراسة مقارنة. مجلة الطفولة والتربية (جامعة الإسكندرية)، 52(3)، 590–626.

[8] عبد الوهاب، محمد محمود .(2010). *استخدام نماذج الاستجابة للمفردة الاختبارية في تدريج مفردات بعض الاختبارات المعرفية* (أطروحة دكتوراه غير منشورة). جامعة المنيا – كلية التربية، مصر.

[9] علام، صلاح الدين .(2005). *نماذج الاستجابة للمفردة الاختبارية: أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي* .القاهرة: دار الفكر العربي.

[10]علوان، بلال طارق .(2023) .الخصائص القياسية لاختبار (OLSAT)للقدرة المدرسية المتقدم (G) وفقًا للأنموذج رباعي المعلمة لدى طلبة المرحلة الإعدادية .*مجلة القادسية في الآداب والعلوم التربوية*.(2)*1* ,

[11]كاظم، أمينة. (1987). دراسة نظرية نقدية حول القياس الموضوعي للسلوك نموذج "راش". سلسلة الكتب المتخصصة. الكويت: مؤسسة الكويت للتقدم العلمي .*المجلة التربوية والنفسية – البحرين*،(4)7

[12]محاسنة، إبراهيم .(2013) .*القياس النفسي في ظل النظرية التقليدية والحديثة* .عمّان: دار جرير للنشر والتوزيع.

[13]مرشود، محمد فايق .(2020) .ثر عدد البدائل والفقرات في تقدير معالم الأفراد والفقرات بإستخدام النموذج رباعي المعلم البارامتري .*مجلة جامعة القدس المفتوحة للأبحاث والدراسات التربوية* , *11*(32) , 83–101.

[14]النفيعي، عبد الرحمن عبد الله أحمد .(2023) . تدريج اختبار المصفوفات المتتابعة المعياري المطور (SPM Plus)في ضوء نموذج الاختيار من متعدد اللوغاريتمي المتداخل .*مجلة كلية التربية جامعة بني سويف*،(119)*20*، 828–876.

ثانياً: المراجع باللغة الانجليزية

[1] Abdel-Khalek, A. M. (2005). Reliability and factorial validity of the Standard Progressive Matrices among Kuwaiti children ages 8 to 15 years. *Perceptual and Motor Skills, 101*, 409–412

[2] Abdel-Khalek, A. M., & Raven, J. (2006). Normative Data From The Standardization Of Raven's Standard Progressive Matrices In Kuwait In An International Context. *Social Behavior and Personality: An International Journal*, *34*(2), 169–180. https://doi.org/10.2224/sbp.2006.34.2.169

[3] Antoniou, F., Alkhadim, G., Mouzaki, A., & Simos, P. (2022). A Psychometric Analysis of Raven's Colored Progressive Matrices: Evaluating Guessing and Carelessness Using the 4PL Item Response Theory Model. *Journal of Intelligence*, *10*(1), 6. https://doi.org/10.3390/jintelligence10010006

[4] Baker, F. B. (2001). *The basics of item response theory*. ERIC.

[5] Chalmers, R. P. (2012). mirt: A Multidimensional Item Response Theory Package for the R Environment. *Journal of Statistical Software*, *48*(6).

[6] de Ayala, R. J. (2013). *The Theory and Practice of Item Response Theory*. Guilford Publications. https://books.google.com/books?id=NchXAQAAQBAJ

[7] Embretson, S. E., & Reise, S. P. (2013). *Item Response Theory*. Psychology Press. https://doi.org/10.4324/9781410605269

[8] Georgiev, N. (2008). Item Analysis of C, D and E Series from Raven's Standard Progressive Matrices with Item Response Theory Two-Parameter Logistic Model. *Europe's Journal of Psychology*, *4*(3). https://doi.org/10.5964/ejop.v4i3.431

[9] Harvey, R. J., & Hammer, A. L. (1999). Item response theory. *The Counseling Psychologist*, *27*(3), 353-383.

[10] Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory* (Vol. 2). Sage.

[11] Liao, W. W., Ho, R. G., Yen, Y. C., & Cheng, H. C. (2012). The four-parameter logistic item response theory model as a robust method of estimating ability despite aberrant responses. *Social Behavior and Personality*, *40*(10), 1679–1694. https://doi.org/10.2224/sbp.2012.40.10.1679

[12] Liu, Y., & Alberto, M.-O. (2014). Identifying the Source of Misfit in Item Response Theory Models. *Multivariate Behavioral Research*, *49*(4), 354–371. https://doi.org/10.1080/00273171.2014.910744

[13] Loken, E., & Rulison, K. L. (2010). Estimation of a four-parameter item response theory model. *British Journal of Mathematical and Statistical Psychology*, *63*(3), 509–525. https://doi.org/10.1348/000711009X474502

[14] Raven, J. (1989). The Raven Progressive Matrices: A review of national norming studies and ethnic and socioeconomic variation within the United States. *Journal of Educational Measurement*, *26*(1), 1–16.

[15] Raven John, Raven J.C., & Court J.H. (2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales Section 3: Progressive*

الملاحق

ملحق (1) أسماء السادة المحكمين

	الاسم	اللقب العلمي	التخصص
1	محمد الصوفي	برفسور	علم النفس التربوي
2	محمد الخضر	أستاذ مشارك	قياس وتقويم
3	ياسر المسهلي	أستاذ مساعد	علم النفس التربوي
4	سامي كميت	أستاذ مساعد	علم النفس التربوي
5	نجاة مرشد	أستاذ مساعد	علم النفس
6	ايمان العوامي	أستاذ مساعد	احصاء
7	ميمونة الغيلي	مدرس	قياس والتقويم
8	عبد العزيز زهرة	اخصائي نفسي	علم نفس

ملحق (2) مؤشرات الملائمة للفقرات اختبار رافن

م	Outfit	Z.Outfit	Infit	Z.Infit	م	Outfit	Z.Outfit	Infit	Z.Infit
q1	0.73	−0.87	1.01	0.11	q31	0.93	−0.67	0.98	−0.49
q2	0.99	−0.26	1.00	−0.07	q32	0.94	−1.21	0.98	−0.78
q3	0.95	−0.12	1.01	0.09	q33	0.94	−1.01	0.97	−0.73
q4	0.98	0.01	1.01	0.13	q34	0.96	−1.08	0.97	−1.19
q5	0.98	−0.04	1.01	0.12	q35	0.97	−0.86	0.97	−1.10
q6	0.94	−0.17	0.99	−0.08	q36	1.00	0.00	1.01	0.13
q7	0.98	−0.24	1.00	−0.06	q37	0.97	−0.11	0.99	−0.07
q8	0.96	−0.26	1.00	−0.02	q38	0.91	−0.45	0.97	−0.38
q9	0.98	−0.07	1.00	0.00	q39	0.96	−0.32	0.99	−0.20
q10	0.97	−0.24	1.00	0.04	q40	0.93	−0.80	0.96	−0.74
q11	1.00	−0.02	1.00	0.00	q41	0.80	−1.01	0.96	−0.65
q12	0.99	0.00	0.99	−0.05	q42	0.85	−1.19	0.96	−0.77
q13	0.99	−0.07	0.99	−0.06	q43	0.89	−1.30	0.97	−0.88
q14	0.98	−0.10	1.00	0.07	q44	0.93	−0.87	0.97	−0.66
q15	0.96	−0.15	1.00	−0.01	q45	0.92	−1.41	0.95	−1.53
q16	0.98	−0.25	0.99	−0.08	q46	0.91	−1.59	0.96	−1.36
q17	0.98	−0.25	0.99	−0.08	q47	0.99	−0.37	0.99	−0.43
q18	0.97	−0.47	0.99	−0.25	q48	1.00	−0.02	1.00	0.06
q19	0.94	−1.03	0.98	−0.72	q49	0.96	−0.98	0.98	−0.71
q20	0.96	−0.25	0.99	−0.15	q50	0.90	−1.72	0.95	−1.56
q21	0.97	−0.61	0.98	−0.58	q51	0.94	−1.39	0.95	−1.55

−1.96	0.93	−2.17	0.85	q52	−0.61	0.97	−0.76	0.93	q22
−1.94	0.91	−1.95	0.86	q53	−1.18	0.96	−1.25	0.93	q23
−1.26	0.96	−1.37	0.92	q54	−1.20	0.97	−1.12	0.96	q24
0.02	1.00	−0.12	0.99	q55	−0.08	0.99	−0.32	0.96	q25
0.32	1.03	0.06	1.00	q56	−0.02	1.00	−0.62	0.86	q26
0.03	1.00	−0.07	0.99	q57	−0.05	1.00	−0.25	0.96	q27
0.06	1.00	−0.09	0.99	q58	−0.67	0.98	−0.88	0.95	q28
−0.02	1.00	−0.14	0.99	q59	−0.20	0.99	−0.36	0.96	q29
−0.01	1.00	−0.13	0.99	q60	−1.10	0.97	−1.09	0.95	q30