تحديد حجم عينة فقرات الاختبار التقليدي باستخدام دالة معلومات الفقرة الاختبارية لنموذج "راش"

أ.د. محمد منصور محمد الشافعي

أستاذ القياس والإحصاء قسم علم النفس –كلية الدراسات الإنسانية والإدارية – كليات عنيزة بالقصيم

تحديد حجم عينة فقرات الاختبار التقليدي باستخدام دالة معلومات الفقرة الاختبارية لنموذج "راش"

أ.د. محمد منصور محمد الشافعي

أستاذ القياس والإحصاء قسم علم النفس -كلية الدراسات الإنسانية والإدارية - كليات عنيزة بالقصيم

تاريخ الإرسال: 1444/11/1هـ **تاريخ الإفادة:** 1445/4/1هـ

ملخص الدراسة:

استهدفت الدراسة الكشف عن إمكانية استخدام مؤشر أقصى قيم للمعلومات عند مستويات مختلفة من القدرة المستهدفة من القياس في خفض فقرات اختبار "أوتيس – لينون" ، اعتماداً على عينة من طلاب التعليم الابتدائي والإعدادي والثانوي، وقد استخدم مؤشر أقصى معلومات بتحليلات نموذج "راش" الأحادي البارامتر في تحليل بيانات الاختبار، كما استخدم مقياس التساعيات لتقسيم مدى القدرة المستهدف قياسها إلى تسعة مستويات مختلفة من حيث مستوى القدرة المستهدفة من القياس، ثم قام الباحث بسحب مفردات مختلفة الأعداد تبعاً لمؤشر أقصى قيمة للمعلومات و توصلت الدراسة إلى صور مصغرة من نفس الاختبار تتضمن كل منها عدد أقل من المفردات المحتواة بالاختبار الأصلي، وقد بلغ أقل عدد من المفردات بعد عملية الخفض (54) مفردةk وهي أقصر صورة من صور الاختبار، وكان لها القدرة ـــ أيضاً ـــ على إعطاء نفس تقديرات القدرة، وكذلك الخصائص السيكومترية التي تم الحصول عليها باستخدام الاختبار الأصلي.

الكلمات المفتاحية: أقصى معلومات للفقرة الاختبارية، نموذج راش، نظرية الاستجابة للفقرة.

Determining the sample size of the traditional test items using the test item information function of the "Rush" model

Dr. Muhammad Mansour Muhammad Al-Shafi'i
Professor of measurement and statistics.
Department of Psychology - College of Humanities and Administrative Studies - Unayzah Colleges in Qassim.

Abstract:

The study aimed to reveal the possibility of using the index of maximum values of information at different levels of measured ability in reducing the items of the "Otis-Lennon" test. The study used a sample of primary, preparatory and secondary education students, The maximum information index associated with the one-parameter "Rush" model was used in analyzing the test data. The study also used the ninth scale to divide the range of the targeted ability measured from the test into nine different ones that vary in terms of the level of measured ability. Then the researcher drew items of different lengths according to the indicator of the maximum value. For information, the study found miniature images from the same test, each of which contained a smaller number of items included in the original test. The lowest number of items after the reduction process was (54), which is the shortest form of the test. It also had the ability to give the same ability ratings as well. Psychometric characteristics obtained using the original test.

keywords: maximum information for the test item, Rasch model, item response theory.

المقدمة والخلفية النظرية

الاختبارات التقليدية المعروفة باسم اختبارات الورقة والقلم، من الوسائل الهامة التي يعتمد عليها في ميدان القياس والتقويم في جميع المجالات التعليمية، وفي مختلف المجالات والميادين الأخرى ، (Bond & Fox, 2020)، ولقد سعى المهتمون بمجال القياس السلوكي في بناء الاختبارات التي تُستخدم لقياس قدرات وخصائص الأفراد بمستويات مرتفعة من الدقة والموضوعية (Jeon & Kim, 2019)، ولقد وجدت بالآونة الأخيرة في ميدان القياس النفسي والسلوكي بعض والاتجاهات الحديثة التي تستهدف بناء الاختبارات والمقاييس في العديد من المجالات النفسية والمعرفية (Aimee,2019)، ومن بين هذه الاتجاهات استخدام النظرية الحديثة للقياس، والتي من بينها نظرية الاستجابة للفقرة (IRT) (Bond, & Fox, ,2020,p5) .

وقد تزامن ظهور هذه الاتجاهات الحديثة للقياس بظهور اختبارات الموائمة الكمبيوترية (CAT) (Computer Adaptive Testing) Kingsbury & Zara (,2020,) ، والاختبار الموائم هو الاختبار الذي يوازن بين صعوبة الفقرات الاختبارية وقدرات الأفراد، بحيث يحدد صعوبة الفقرات بما يناسب مستوى قدرات الأفراد (,lord,1980)،(علام ،2005)، وتختلف هذه النوعية من الاختبارات مع الاختبارات الورقية التقليدية حيث يتم عرض الفقرات الاختبارية على المفحوص تبعًا لقدرته بغض النظر عن ترتيب الفقرة في الاختبار (Linden &Pushley, 2015 ; Aimee,2019)

دقة تقديرات القدرة وعلاقتها بطول الاختبار

اهتمت بعض الدراسات بالعلاقة بين متغير طول الاختبار ودقة تقديراته للقدرة المستهدفة من القياس، وبعض شروطه السيكومترية، فقد توصلت الدراسة إلى أن دقة تقديرات القدرة ترتبط إيجابيا بطول الاختبار وكلما إزدادت فئات الاستجابة، و توصلت دراسة "Wells, Craig et.al.," (2021) إلى التأثر السلبي لتقديرات القدرة لعينة من طلبة المدارس الثانوية، وكذلك التأثر السلبي لمعاملات الثبات والصدق بطول الاختبارات التي تتضمن أعداد أقل من الفقرات، كما أثبتت دراسات أخرى تأثر تقديرات القدرة بطول الاختبار المستخدم، ومن بينها دراسة " Goodman " (2019) التي أثبتت أنه كلما قل طول الاختبار فإن ذلك يمكن أن يؤثر سلبيًا على تقدير القدرة بشكل عام.

كما تتأثر الخصائص السيكومترية للاختبار بطوله (فرج، 2007، 387) وقد أجريت في هذا المجال الكثير من الدراسات التي استهدفت تقدير الثبات للاختبارات مختلفة الطوال، ومن بين هذه الدراسات دراسة Kahraman & Thompson (2018) والتي استهدفت مقارنة تقديرات ثبات الاختبارات مختلفة الأطوال والأخرى القصيرة، وقد أظهرت النتائج أن الاختبارات الأطول تميزت بثبات مرتفع مقارنة بثبات الاختبارات الأقل طولا.

هذا وقد تبين للباحث من خلال عمله الأكاديمي في مجال القياس وبناء الاختبارات، ومروره على الأدبيات الخاصة بهذا الميدان، عدم وجود أسس دقيقة ومعايير علمية واضحة في تحديد العدد المناسب لفقرات الاختبار التقليدي، في حين أن الاتجاهات الحديثة للقياس أسفرت عن اختبارات حاسوبية موائمة

تستند إلى بعض المؤشرات الإحصائية في انتقاء فقراتها، ومن بين هذه المؤشرات مؤشر أقصى معلومات، الذي يمكن الحصول عليه من تحليلات النماذج اللوجيستية المرتبطة بنظرية الاستجابة للفقرة (IRT) tem Response Theory، وهذا ما حدا بالباحث إلى دراسة هذه الفكرة البحثية المطروحة على هذا البحث.

نظرية الاستجابة للفقرة (IRT) Item Response Theory

تمثل هذه النظرية أحد الاتجـــــاهات الحديثة للقياس، (De Ayala, 2022)، وبظهور هذه النظرية ظهرت العديد من النماذج اللوجيســتية المرتبطة بها، والتي حققت إلى حد بعيد الموضوعية التي تتصف بها أدوات قياس المجـــال الفيزيقي (Crall, 2018)، وقد اســـتخدمت هذه النماذج بهدف الحصـــول على المعالم الخاصة بفقرات تلك الاختبارات من بينها:

أ- **معلم الصعوبة difficulty parameter** : والذي يشير إلى "مواقع الفقرات على متصل تدريج الصفة المقاسة ويحدد امتلاكها لمستويات محددة من هذه الصفة" (Jason, Meyers, Edward. Miller , Walter & Way , 2009,42)

ب- **معلم التمييز**: **discrimination parameter**: وهو "مؤشر يفيد بقدرة الفقرة على التفريق بين المستويات المختلفة من القدرة المقاسة، ويقاس بظل الزاوية التي يميل بها المماس للمنحنى المميز للفقرة على محور القدرة" (1987, Hambleton, 26) .

ج- **معلم التخمين: uessing parameter**: وهو "يحدد إمكانية توصل المفحوص إلى الإجابة الصحيحة عن الفقرات التي لم تمكنه قدرته في الإجابة عنها عن طريق التخمين العشوائي" (Azevedo, 2009,93)

د- **مؤشر دالة المعلومات Information function**: وهي "مقدار ما تُعطيه الفقرة من معلومات عند مستويات القدرة المختلفة" Stocking ,2019 (, 4)

وقد اقترنت هذه النظرية بعدد من النماذج الرياضيةK، ومن بينها النموذج اللوجيستي أحادي المعلم الذي سوف يتم الاعتماد عليه بالدراسة الحالية، وبيانه فيما يلي (Hambleton ,1987, 6):

النموذج أحادي البارامتر: One-Parameter Logistic Model

توصل "راش" إلى الصورة الرياضية لهذا النموذج الموضحة فيما يلى **:(**Hambelton ,1985, 46**)**

$$p(x_{vi} = x/\theta_v b_i) = \frac{e^{(\theta_v - b_i)}}{1 + e^{(\theta_v - b_i)}}$$

حيث x = 0, 1 ،x هى استجابة الفرد على الفقرة، θv قدرة الفرد، bi صعوبة الفقرة.

e =هو أساس اللوغاريتم الطبيعى ويساوى (2.273).

P (X_{vi}) احتمال حدوث الاستجابة الصواب أوالخطأ.

يتميز نموذج "راش" بالخواص التالية (Hambelton ,1985):

والنموذج يتضمن عدد أقل من متغيرات الفقرة (Okewole, et.al.,2020)، كما أن كل الفقرات من المفترض أنها تمتلك قوة تمييز متساوية،

وكذلك مؤشر التخمين في هذا النموذج يكون أقل صوره. هذا ويمكن إيجاد دالة المعلومات للنموذج الأحادي البارومتر بالصيغة التالية:

$$I_j(\theta) = \frac{e^{\theta-\alpha_j}}{(1+e^{\theta-\alpha_j})^2}.$$

حيث $I_j(\theta)$ هي دالة المعلومات للفقرة (j) عند مستوى القدرة (θ)، ($e^{\theta-\alpha j}$) الأساس الطبيعي للفرق بين مستوى القدرة وصعوبة الفقرة.

هذا وتقتصر الدراسة في الإجراءات الخاصة بها على مؤشر أقصى معلومات للفقرة من خلال التحليلات الخاصة بالنموذج الأحادي البارمتر عند مختلف مستويات للقدرة المستهدفة من القياس، بهدف الحصول على أفضل الفقرات التي يمكن أن تعطي القدر الأكبر من المعلومات عن قدرة الفرد لكل مستوى من مستويات القدرة المقاسة، وسوف يتم تناول هذا المؤشر في الفقرة الآتية:

أقصى معلومات Maximum Information: وهو يشير إلى "أقصى ارتفاع لمنحنى دالة المعلومات للفقرة الاختبارية عند مستوى معين من السمة أوالقدرة" (Rose , Bjorner , Becker & Friesc ,2008,20).

إجراءات انتقاء الفقرة : عملية انتقاء الفقرة للاختبار هي إجراء يتبع لضوابط وقواعد تتحدد من خلال الأهداف والمستويات المعرفية المستهدف قياسها من الاختبار .(Penfield, 2016)، بينما يختلف الامر نسبيًا في حال إعداد الاختبارات التواؤمية، حيث يوجد هناك أسلوبان رئيسيان لانتقاء الفقرات هما (Azevedo, Bolfarine, ,Andrade, 2011):

أ- المعلومات القصوى Maximum Information Method

ويعتمد هذا المؤشر على بلوغ دالة فيشر للمعلومات الحد الأعلى لمستوى القدرة المقاسة للممتحن (Lord,1980). ويشير Kingsbury & Zara (2020,) إلى مميزات لاستخدام مؤشر أقصى معلومات وهي: أنه يوفر أعلى كفاءة للاختبار، نظرًا لأنه يتم انتقاء الفقرة التي تسفر عن أقصى معلومات عن قدرة الفرد.

ب- الأسلوب البييزي Bayesian Item Selection Method

وهذا الأسلوب يستخدم "التباين البعدي للقدرة كمحك لانتقاء الفقرة (Kingsbury & Zara, 2020, 363)، وقد أظهرت بعض الدراسات عدم حساسية الأسلوب البييزي في دقة التقدير ومنها دراسة Magis (2020) ، والتي استهدفت الكشف عن أثر استخدام متغير طول الاختبار في دقة تقدير المستويات القصوى وكذلك الدنيا للبراعة اعتماداً على نموذج "راش"

وقد اعتمد الباحث على مؤشر أقصى معلومات، في انتقاء الفقرات في الاختبارات الخطية الفرعية المختلفة من حيث الأطوال، نظرا لمزايا هذا المؤشر وتفوقه على بعض الأساليب الأخرى، وهذا ما أظهرنه دراسة Altaf, Saim, Aslam & Muhammad " (2011)، وتشابهه مع البعض الآخر من نفس الأساليب وهذا ما أوضحته دراسةMarinagi & Kaburlasos (2021) وقد أُستخدم كذلك في إجراءات بناء واختيار الفقرات في الاختبارات الموائمة المحوسبة، فلقد توصلت الكثير والعديد من الأدبيات أن الاختبارات الموائمة تتمتع بالمزايا التي من اهمها إمكانية تقدير قدرة الممتحن من خلال عدد أقل

من الفقرات التي لها أقصى قيم للمعلومات. (Siang.& Fritz,2006)، وقد أشارت دراسة Aimee (2019) أن الاختبار الموائم يكون أقصر في الغالب من الاختبار الخطي التقليدي، كما يسفر عن قياس أعلى من حيث الدقة.

لذا يحاول الباحث من خلال دراسته الحالية على إمكانية استخدام مؤشر أقصى معلومات للفقرات الاختبارية في انتقاء فقرات الاختبار التقليدي الورقي المستخدم بدراسته من اجل التحقق من دقة هذا المؤشر في تقليص عدد هذه الفقرات إلى الحد الأقل المناسب الذي يمكن أن يكون له نفس مميزات الاختبار الموائم في تقدير قدرات الممتحنين بعدد أقل من الفقرات الاختبارية.

مشكلة الدراسة:

يمكن تلخيص مشكلة الدراسة في الإجابة على السؤال العام الآتي:

هل تختلف تقديرات قدرات العينة البحثية المشتقة من الاختبار الأصلي للدراسة مع التقديرات المناظرة لها المشتقة من الاختبارات الفرعية المتضمنة لأعداد أقل من الفقرات، والتي لها أقصى قيم لمؤشر المعلومات على جميع مستويات القدرة؟ ومن هذا السؤال العام تتفرع الأسئلة الآتية:

١. هل تختلف تقديرات القدرة المتناظرة للعينة البحثية الناتجة عن تحليلات الاختبار الأصلي والأخرى الناتجة من الاختبار الفرعي باختلاف مستوى أقصى معلومات للفقرات المتضمنة بالاختبار الفرعي (منخفضة / مرتفعة)؟

٢. هل تختلف تقديرات القدرة المتناظرة الناتجة عن تحليلات الاختبار الأصلي، والأخرى الناتجة عن الاختبارات الفرعية المكونة من الاختبار الأصلي،

والمتضمنة لأعداد مختلفة من الفقرات التي لها أقصى قيم للمعلومات على جميع مستويات القدرة (تبعاً لمقياس التساعيات) ؟

٣. ما أقل عدد مناسب من الفقرات الاختبارية التي لها أقصى معلومات والتي يمكن أن تُعطي نفس تقديرات القدرة المشتقة من الاختبار الأصلي؟

هل تختلف الشروط السيكومترية للاختبارات الفرعية ذات الأطوال المختلفة، والتي تحتوي على فقرات لها أقصى معلومات على فترات مختلفة من القدرة تبعاً للمقياس التساعي عن الشروط السيكومترية للاختبار الأصلي؟

مصطلحات الدراسة:

الاختبار التقليدي Classic test : هو الاختبار التقليدي المعروف باسم (الورقة والقلم) . وإجرائياً سوف يُستخدم اختبار " أوتيس لينون" في الدراسة الحالية كاختبار تقليدي.

أقصى معلومات Maximum Information : وهو يشير إلى أقصى ارتفاع لمنحنى دالة المعلومات للفقرة الاختبارية عند مستوى معين من السمة أوالقدرة، (Rose , Bjorner , Becker ,&. Friesc, ,2008, 20) ، وإجرائياً سوف يستخدم برنامج بايلوج MG المتضمن لنموذج راش الأحادي البارمتر في استخراج قيم أقصى معلومات للفقرات الاختبارية المتضمنة بأداة بالدراسة الحالية.

نموذج راش Rasch Model : ويهتم بتحديد موقع الفقرة الاختبارية على تدريج صعوبة جميع الفقرات التى تشكل الاختبار" (صلاح الدين علام،

2005، 693). وإجرائياً: يتم استخدام هذا النموذج في تحليل البيانات الخاصة بفقرات اختبار أوتيس لينون المستخدم بالدراسة الحالية.

أهمية الدراسة:

أولاً : الأهمية النظرية: الدراسة تستخدم مؤشر إحصائي متطور، وهو مؤشر المعلومات القصوى، في إمكانية خفض عدد الفقرات الاختبارية، كما تهتم أيضا بتحليل فقرات اختبار أوتيس – لينون للمستوى المتوسط باستخدام أحد النماذج الرياضية اللوغارتمية المتصلة بنظرية الاستجابة للفقرة، وهو: النموذج الأحادي المعلم ، وهذا له أهميته النظرية في ميدان القياس بوجه عام، وبصفة خاصة في مجال إجراءات بناء أدوات القياس وإعدادها، التي تهتم بخفض وتقليص الاختبارات المشهورة عالميًا إلى صور مختصرة.

ثانياً : الأهمية التطبيقية: تسعى الدراسة إلى الكشف عن الحجم المناسب للفقرات الاختبارية التي يمكن اشتقاقها من العينة الكلية للفقرات المحتواة بالاختبار الأصلي لمستخدم بالدراسة الحالية، والذي قد يعطي نفس تقديرات القدرة للأفراد المشتقة من العينة الكلية لفقرات الاختبار الأصلي، وهذا له أهميته التطبيقية في ميدان الاستخدامات الشائعة والكثيرة لهذا الاختبار ، وكذلك له أهميته التطبيقية في الحصول على صور مختصرة من لمختلف الاختبارات وأدوات القياس المشهورة عالمياً، التي يمكن أن توفر لكل من المفحوصين والممتحنين الكثير من الوقت والجهد.

أهداف الدراسة:

١. الكشف عن تأثر تقديرات القدرة للأفراد باستخدام اختبارات مختلفة الأطوال تتضمن فقرات تمتلك أقصى قيم للمعلومات.

٢. الكشف عن أقل عدد من الفقرات الاختبارية – التي يمكن سحبها من اختبار الدراسة، والتي يمكن أن يسفر استخدامها عن نفس تقديرات القدرة المشتقة من الاختبار الأصلي.

٣. الكشف عن حقيقة تأثر الخصائص السيكومترية للاختبارات المختلفة، من حيث الطول، والتي لها أقصى معلومات على المستويات المختلفة من القدرة المستهدفة من القياس.

حدود الدراسة:

١. تقتصر الدراسة على استخدام النموذج الأحادي المعلم في تحليلات استجابات العينة البحثية على اختبار الدراسة والاختبارات الفرعية المشتقة من هذا الاختبار.

٢. تعتمد الدراسة فقط على اختبار " اوتيس لينون" في تحقيق الأهداف الخاصة بها.

٣. تقتصر الدراسة الحالية على عينة من طلبة التعليم الابتدائي والإعدادي والثانوي بجمهورية مصر العربية.

٤. تم تطبيق الأداة الخاصة بها على العينة البحثية خلال الفصل الدراسي الأول من العام الدراسي 2022/2021.

إجراءات الدراسة:

العينة : تكونت العينة المتاحة الدراسة من طلاب التعليم العام الابتدائي والإعدادي والأول الثانوي، بمحافظة جنوب القاهرة، وقد توزعت على المرحلة الدراسية تبعاً للأعداد الموضحة بالجدول التالي:

جدول (2) : توزيع أفراد عينة الدراسة تبعاً للمراحل التعليمية والعمر

مسلسل	المرحلة الدراسية	الصف	حجم العينة	متوسط العمر	الانحراف المعياري
1	الابتدائية	السادس	360	11.2	1.8
2	المتوسط	الأول	461	12.4	2.1
3		الثاني	290	13.3	1.3
4		الثالث	132	14.5	1.1
5	الثانوية	الأول	287	15.6	1.5
المجموع		العينة الكلية	1530	13.4	-

الأدوات: اعتمد الباحث على اختبار أوتيس لينون المستوى المتوسط الصورة (Level J) ، وهو من تعريب وتقنين (مراد،صلاح وعبدالغفار،محمد، 1985) على عينة من المجتمع المصري، كما أجرت دراسة (زكري، 2008) تقنين نفس الاختبار على عينة سعودية، وقد توصلت الدراسة إلى وجود معاملات ثبات وصدق جيدة للاختبار، وقد تم إعداد هذا الاختبار بهدف قياس القدرة العقلية العامة للأفراد من سن (11-16) سنة، ويحتوي الاختبار على (80) فقرة من نوع الاختيار من متعدد، وتتكون كل فقرة اختبارية من خمسة بدائل للإجابة، وتتوزع فقرات الاختبار تبعاً للمجالات الآتية (مراد،صلاح وعبدالغفار،محمد،1985):

الاستيعاب اللفظي : ويشمل (20) فقرة وهي الفقرات أرقام:

8 10 14 19 27 32 36 44 46 49 51 53 56 59
61 64 67 73 75 77

الاستدلال اللفظي.:ويشمل (32) فقرة، وهي أرقام:

33	30	28	25	23	21	17	15	12	11	6	5	1	
80	79	76	71	69	65	62	60	58	54	52	47	42	
							41	40	39	38	37	34	

الاستدلال الشكلي : ويشمل (15) فقرة وهي أرقام:

78	74	68	55	48	43	35	26	22	20	18	16	9	4	2

الاستدلال الكمي :ويشمل (13) فقرة وهي أرقام:

72	70	66	63	57	50	45	31	29	24	13	7	3

وقد قام الباحث الحالي بحساب ثبات الاختبار على عينة استطلاعية حجمها (50) طالب بطريقة معامل (الفا كرونباخ)، حيث بلغت قيمته (0.88)، كما تم التحقق من الصدق بإيجاد معامل الارتباط بين درجات العينة الاستطلاعية على الاختبار ومجموع درجاتهم الحاصلين عليها في نهاية الفصل الدراسي، باعتبار أن المستوى التحصيلي للطالب هو المحك الخارجي على صدق الاختبار، وقد بلغ معامل الارتباط (0.91)؛ مما يعني تمتع الاختبار بشروط سيكومترية جيدة، وقد وقع اختيار الباحث على هذا الاختبار نظرًا لاتساع مدى معلم الصعوبة لفقراته (مراد، وعبدالغفار، 1985 وشمول الاختبار على عدد كبير من الفقرات يتيح إجراء المعالجات البحثية، وبما يوفر مدى متسع لمؤشر أقصى معلومات .(Lihua Yao.,2019).

الأساليب الإحصائية:

١. مقياس التساعيات كأحد إجراءات الإجابة عن السؤال الثاني.

٢. التحليل العاملي الاستكشافي بهدف التحقق من تمتع الاختبار افتراض أحادية البعد.

٣. تحليل تباين القياس المتعدد، الخطأ المعياري لقياس القدرة للإجابة عن السؤال الأول والثاني.

٤. النسبة الحرجة للكشف عن دلالة الفرق بين نسبتين المصححة باختبار بنفروني (Bonferroni method) adjested p-values لمعالجة تضخم الخطأ من النوع ألفًا نتيجة المقارنات المتعددة باستخدام تحليلات (SPSS) إصدار (26). وذلك للإجابة عن السؤال الأول والثاني.

إجراءات الإجابة عن تساؤلات الدراسة:

١. تحليل استجابات العينة البحثية على فقرات اختبار الدراسة باستخدام نموذج "راش" لتحديد الأفراد، وكذلك الفقرات غير الصالحين تبعاً لمؤشر اختبار حسن المطابقة (chi^2). (Zimowski, et.al., 2007) .

٢. التحقق من تحقق افتراضات أحادية البعد، استقلالية القياس،السرعة (Chon, Lee, & Dunbar, ,2010).

٣. إجراء التحليلات النهائية لفقرات الاختبار، واستخراج مؤشر أقصى معلومات لها، ثم تقسيم الفقرات باستخدام مقياس الوسيط إلى اختبارين فرعيين أحدهما مرتفع، والآخر منخفض مع مراعاة تمثيل الفقرات لمجالات الاختبار الأربعة.

٤. استخدام المقياس التساعي في تقسيم مدى الصعوبة الكلي فاقرات الاختبار الأصلي إلى تسعة مستويات متباينة من حيث الصعوبة، والتي تناظر نفس مستويات القدرة المستهدفة من القياس بالاختبار.

٥. ترتيب فقرات الاختبار الكلي، وذلك وفقًا لمتغير أقصى معلومات.

٦. اختيار (8) فقرات صاحبة أقصى معلومات من كل من المجالات الأربع المحتواة بالاختبار من بين الفقرات المناظرة لكل فترة من الفترات التسع للقدرة المقاسة، فيصبح عدد فقرات الاختبار الفرعي الناتج عن هذا الإجراء (72) فقرة، ثم تكرار نفس الإجراء، ولكن باختيار فقرات مختلفة في عددها [(7)، (6)، (5)، (4)، (3)، (2) ، (1)]، وبالتالي فإن الاختبارات الفرعية الناتجة تحتوي فقرات عددها على الترتيب (63، 54، 45، 36، 27، 18، 9).

٧. تحليل الاختبارات الفرعية الناتجة عن الخطوات السابقة، للحصول على التدريج الجديد لتقديرات قدرات أفراد العينة البحثية لكل اختبار من الاختبارات الفرعية.

٨. تعيين فروق تقديرات القدرة المتناظرة لكل فرد على حدة الناتجة عن تحليل الاختبار الأصلي وكل اختبار فرعي.

٩. تعيين مجموع الخطأين المعيارين الناتجين عن تحليل الاختبار الأصلي وأي اختبار فرعي لجميع أفراد العينة البحثية، ومن ثم إيجاد دلالة الفروق بين تقديرات القدرة الناتجة من الاختبار الأصلي وكل اختبار فرعي اعتماداً على القاعدة الآتية: عندما يكون الفرق بين تقديري القدرة الناتجين من الاختبار الأصلي والاختبار الفرعي (< أو =) مجموع الخطأين المعيارين لهما كان ذلك

فرقاً جوهريًا، وعندما يكون أقل كان ذلك فرقًا غير جوهريًا. (Chon, Lee & Dunbar, 2010) .

١٠.تعيين النسبة المئوية للفروق (الدالة) في تقديرات القدرة الناتجة عن الاختبار الأصلي، وكلا الاختبارين الفرعيين (مرتفع ومنخفض) من حيث أقصى المعلومات.

١١.اختبار دلالة الفرق بين النسب المئوية المتناظرة اعتماداً على اختبار (z) للتحقق من دلالة الفرق بين النسب.

١٢.تعيين النسبة المئوية للفروق (الدالة) بين تقديرات القدرة الناتجة من الاختبار الأصلي وكل اختبار فرعي، وعندما تتجاوز النسبة المئوية القيمة (0.05).

١٣.إجراء تحليلات تباين القياس المتعدد لتقديرات القدرة الناتجة عن الاختبار الأصلي، وجميع الاختبارات الثماني الفرعية.

١٤.اختبار تكافؤ الشروط السيكومترية (الصدق والثبات) لجميع لاختبارات الفرعية مع نظائرها من الشروط السيكومترية للاختبار الأصلي، من خلال استخدام معامل ارتباط بيرسون لجميع تقديرات القدرة الناتجة من الاختبار الأصلي مع تقديرات القدرة الناتجة عن كل اختبار فرعي.

نتائج الدراسة

1-نتائج تحليل مفردات اختبار الدراسة باستخدام النموذج الأحادي البارومتر

أسفرت تحليلات البيانات الخاصة باختبار الدراسة باستخدام نموذج "راش" عن النتائج المبينة بالجدول التالي:

جدول (4) :تقديرات الصعوبة واقصى معلومات المشتقة من تحليل فقرات اختبار الدراسة باستخدام النموذج الأحادي البارومتر

رقم الفقرة	مستوى الصعوبة	الخطاء المعياري	مؤشر أقصى معلومات	الخطاء المعياري	مربع كاي	رقم الفقرة	مستوى الصعوبة	الخطاء المعياري	مؤشر أقصى معلومات	الخطاء المعياري	مربع كاي
1.	1.039	0.147	0.2215	0.0711	5	41.	1.969	0.224	0.4027	0.0907	3.8
2.	1.134	0.070	0.7786	0.0418	9.3	42.	1.094	0.055	1.2133	0.0421	14.5
3.	0.46	0.155	0.2918	0.0617	35	43.	2.59	0.495	0.3226	0.0012	3.6
4.	0.0119	0.243	0.2084	0.0353	26.5	44.	2.06	0.238	0.9374	0.0363	1.7
5.	-0.266	0.172	0.4048	0.0614	57.2	45.	2.824	0.629	0.4699	0.0401	5.2
6.	0.982	0.049	1.8057	0.0899	12.9	46.	3.443	1.093	0.438	0.0433	11.7
7.	0.98	0.106	0.3597	0.1122	15.4	47.	2.862	0.643	0.2876	0.0904	27.1
8.	0.603	0.072	0.8427	0.0811	9.3	48.	3.071	0.826	0.5212	0.0148	6.5
9.	1.167	0.070	0.797	0.0411	5	49.	1.951	0.208	2.0637	0.0919	7.6
10.	0.563	0.118	0.3772	0.0839	18.9	50.	1.683	0.142	0.719	0.0131	2.3
11.	1.221	0.062	1.243	0.0955	8.3	51.	1.274	0.059	1.495	0.0148	4.5
12.	1.392	0.073	1.4156	0.0177	4.6	52.	1.436	0.072	1.7413	0.0695	16.5
13.	1.469	0.138	0.3503	0.1012	9.8	53.	2.624	0.568	0.6224	0.0656	7.6
14.	0.262	0.253	0.1629	0.0329	22.3	54.	1.988	0.190	1.2835	0.0854	8.9
15.	0.53	0.191	0.2236	0.0544	8.8	55.	2.092	0.269	0.4028	0.0079	7.3
16.	1.474	0.076	1.8031	0.0421	8.1	56.	2.255	0.332	0.2596	0.0371	10.9
17.	5.747	2.978	0.2138	0.0602	6.3	57.	3.265	0.796	0.1589	0.0968	12.7
18.	0.435	0.158	0.1583	0.0305	55.4	58.	2.027	0.246	0.2331	0.0074	6.1
19.	1.497	0.118	0.5081	0.0935	9.2	59.	1.416	0.083	0.9709	0.0426	1.3
20.	1.025	0.129	0.2295	0.0610	12	60.	1.15	0.084	0.5284	0.0664	5.3
21.	1.346	0.045	4.1644	0.1816	6.6	61.	1.633	0.106	1.1302	0.0528	7.8
22.	1.319	0.162	0.1821	0.0648	8.5	62.	2.829	0.685	0.6766	0.0479	5.9
23.	1.236	0.101	0.3858	0.0259	6.1	63.	1.702	0.222	0.1267	0.0491	22.2
24.	1.146	0.066	0.8365	0.0361	4	64.	1.675	0.106	2.626	0.0805	11.5
25.	0.902	0.053	1.1824	0.093	7	65.	2.286	0.351	0.3312	0.0828	7.5

رقم الفقرة	مستوى الصعوبة	الخطاء المعياري	مؤشر أقصى معلومات	الخطاء المعياري	مربع كاي	رقم الفقرة	مستوى الصعوبة	الخطاء المعياري	مؤشر أقصى معلومات	الخطاء المعياري	مربع كاي
26.	1.959	0.212	0.6768	0.0471	2.1	66.	3.412	1.121	0.4227	0.0322	1.9
27.	2.125	0.291	0.2103	0.0028	14.3	67.	2.052	0.216	1.4175	0.0552	19
28.	1.52	0.101	0.8663	0.0345	11.6	68.	2.034	0.240	0.8858	0.0062	5.3
29.	0.505	0.099	0.5187	0.0072	16	69.	3.553	1.218	0.4626	0.0728	9.9
30.	1.604	0.094	1.635	0.0901	10.9	70.	1.928	0.237	0.184	0.0817	12.6
31.	0.913	0.092	0.4312	0.0195	9.8	71.	1.618	0.085	2.4539	0.0609	7.7
32.	0.797	0.073	0.7182	0.0782	6.5	72.	1.564	0.135	0.4016	0.0562	5.5
33.	1.235	0.065	1.0346	0.0218	8.3	73.	1.552	0.093	1.2848	0.0011	8.3
34.	1.088	0.080	0.5136	0.0452	7.3	74.	1.132	0.144	0.1982	0.0575	17.9
35.	0.843	0.188	0.1484	0.0353	24.8	75.	1.211	0.067	1.0795	0.0606	6.5
36.	1.205	0.109	0.328	0. 084	5.4	76.	2.023	0.236	1.3996	0.0315	4.4
37.	1.146	0.085	0.4923	0.0485	2.7	77.	2.839	0.663	0.5402	0.0057	12.7
38.	1.328	0.096	0.5152	0.0789	6.5	78.	2.5	0.473	0.6221	0.0398	6.8
39.	1.465	0.095	0.784	0.0813	7.6	79.	1.968	0.230	0.2013	0.0871	6.4
40.	1.507	0.091	1.1395	0.0311	8.6	80.	2.493	0.437	1.0634	0.0873	6.5

** دال عند مستوى (**0.01**) * دال عند مستوى (**0.05**)

يتضح من الجدول السابق أن قيم اختبار مربع كاي غير دالة بما يفيد بصلاحية جمع فقرات الاختبار لاستكمال التحليلات.

2- نتائج فحص افتراضات نظرية الاستجابة للفقرة:

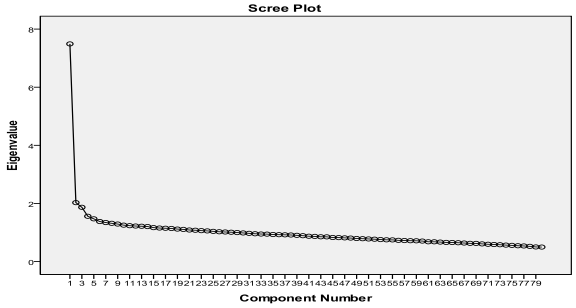
تم استخدام التحليلات الخاصة بالتحقق من توافر افتراضات نظرية (IRT) بالاختبار المستخدم وجاءت النتائج على النحو التالي:

أولاً: التحقق من افتراض أحادية البعد: بعد استخدام تحليلات التحليل العاملي الاستكشافي لفقرات الاختبار المستخدم تم الحصول على النتائج المبينة بالجدول التالي (وقد اكتفت الدراسة باستعراض العوامل الثلاثة الأولى الناتجة عن التحليلات) :

جدول (3):الجذور الكامنة للعوامل الناتجة عن التحليل العاملي الاستكشافي لاختبار الدراسة

العوامل	الجذور الكامنة		
	الكلي	النسبة المئوية للتباين	التباين المفسر
1	7.49	22.362	22.362
2	2.028	7.535	29.897
3	1.867	3.334	33.231

يبدو من الجدول رقم (3) أن نسبة التباين المفسر للعامل الأول تجاوزت نسبة (20%) وفقًا لمحك ريكاس (Recase,cited in Hattie,1985)؛ مما يشير إلى وجود عامل واحد مهيمن يكفي لإثبات أحادية البعد للاختبار حتى مع وجود عوامل أخرى، كما يبدو ايضاً من مقياس Scree plot، وجود انحدار شديد بين قيم الجذر الكامن بين العامل الأول والثاني، مما يشير أيضاً إلى تمتع الاختبار بأحادية البعد (Gessaroli & De Champlain , 2005).. كما يبدو من شكل (1).



شكل (1): رسم بياني لمقياس (Scree plot) للجذور الكامنة لفقرات الاختبار

- **ثانياً التحقق من توافر افتراض استقلالية المحل:** حيث يتم التحقق من عدم اعتماد فقرات الاختبار بعضها على بعض، وقد تبين عدم تجاوز أي فقرة من فقرات الاختبار مؤشر (مربع كاي) (Zimowski, et.al., 2007)،

حيث جاءت كل مستويات الدلالة لهذا المؤشر تزيد عن (0.05)، بما يشير إلى توافر افتراض الاستقلالية لفقرات الاختبار,

- **التحقق من افتراض السرعة:** حيث تم التحقق من أن الفقرات المتروكة بنهاية الاختبار لم تترك بسبب عدم قدرة الممتحن في الإجابة عنها، وليس لعدم توافر الوقت الكافي لمحاولة الممتحن أن يجيب عنها (كاظم، 1988، 92.) ، وقد كان الباحث حريصاً عند تطبيق اختبار الدراسة في إعطاء وقت غير محدد للإجابة عن فقرات الاختبار؛ لتجنب تأثير عامل السرعة في الإجابة.

٢- نتائج إيجاد الاختبارين الفرعيين (المرتفع / المنخفض) من حيث مؤشر أقصى معلومات

انتهت نتائج استخدام مقياس الوسيط بهدف الحصول على فقرات الاختبار الأصلي (المرتفعة / المنخفضة) لقيم مؤشر أقصى معلومات بعد إجراء الترتيب التنازلي لها، عن البيانات الموضحة بالجدول التالي:

جدول (5): الاختباران الفرعيان (المرتفع / المنخفض) الناتجان من فقرات الاختبار الأصلي وفقًا لمؤشر أقصى معلومات

الاختباران الفرعيان	مدى أقصى معلومات لفقرات الاختبار	أرقام الفقرات	الفقرات موزعة على أبعاد الاختبار				
			استيعاب لفظي	استدلال لفظي	استدلال شكلي	استدلال كمي	العدد الكلي
مرتفع	أقص قيم للمعلومات **0.51 <**	2,6,8,9,11,12,16, 21,24,25,26,28,30,32,33,39,40,42 ,44,48,49,50,51, 52,53,54,59,60,61,62,64,67,68,71 ,73,75,76,77,78, 80	13	18	7	2	40

الاختباران الفرعيان	مدى أقصى معلومات لفقرات الاختبار	أرقام الفقرات	الفقرات موزعة على أبعاد الاختبار				
			استيعاب لفظي	استدلال لفظي	استدلال شكلي	استدلال كمي	العدد الكلي
منخفض	أقص قيم للمعلومات 0.51 >	1,3,4,5,7,10,13,14,15,17,18,19,20,22,23,27,29,31,34,35,36,37,38,41,43,45,46,47,55,56,57,58,63,65,66,69,70,72,74,79	7	15	8	11	40

٣- نتائج الإجابة عن السؤال الأول:

اعتمدت الدراسة على إجراءين مختلفين للإجابة عن السؤال الأول والذي ينص على " هل تختلف تقديرات القدرة المتناظرة للعينة البحثية الناتجة عن تحليلات الاختبار الأصلي، والأخرى الناتجة من الاختبار الفرعي باختلاف مستوى أقصى معلومات للفقرات المتضمنة بالاختبار الفرعي (منخفضة / مرتفعة)؟

أولا: نتائج الإجراء الأول:

استخدمت الدراسة اختبار (Z) للكشف عن دلالة الفرق بين نسبتين، وانتهت التحليلات لهذا الإجراء عن البيانات المتضمنة بالجدول الآتي:

جدول (6): تكرار الفروق الجوهرية بين تقديرات القدرة الناتجة عن الاختبار الأصلي والاختبارين الفرعيين (مرتفع /منخفض) المعلومات، ودلالة اختبار (z)

الفروق بين تقديرات القدرة	العدد لفروق الكلي	تكرارات الفروق الجوهرية	نسبة تكرارات الفروق	النسبة المئوية	قيمة (z)	مستوى الدلالة
للاختبار الكلي	1530	16	0.011	1.1	5.3	0.000
للاختبار الكلي	1530	64	0.042	4.2		

يتبين من الجدول السابق رقم (6) أن تكرار الفروق الجوهرية بين تقديرات القدرة الناتجة عن الاختبار الأصلي والاختبار مرتفع المعلومات كانت لعدد (16) فرد من جملة (1530) فرد بنسبة قدرها (1.1%)، أما الفروق الدالة في تقديرات القدرة المشتقة من الاختبار الكلي والاختبار الفرعي المنخفض كانت لعدد (64) فرد بنسبة مئوية قدرها (4.2%)، وقد بلغت قيمة اختبار (Z) لفرق هاتين النسبتين القيمة (5.3) عند (P < 0.01)، وذلك في صالح النسبة المئوية الأكبر، وهي (4.2%)، وبالتالي فإن تقديرات القدرة الناتجة عن تحليل الاختبار مرتفع المعلومات تتكافأ مع تقديرات القدرة الناتجة عن تحليل الاختبار الأصلي، بشكل أعلى من تكافؤ تقديرات القدرة الناتجة عن الاختبار المنخفض مع الاختبار الأصلي بما يدعم الفكرة البحثية بأن المفردات التي تمتلك أقصى معلومات تسفر عن تقديرات للقدرة تتكافأ مع تقديرات القدرة المشتقة من الاختبار الكلي المكون من عدد أكبر من المفردات.

ثانيا: النتائج الإجراء الثاني:

انتهت التحليلات الخاصة باستخدام تحليل تباين القياس المتعدد Repeated Measure Anova عن البيانات المتضمنة بالجدول الاتي:

جدول (7) : الإحصاء الوصفي لتقديرات القدرة الناتجة من تحليلات الاختبار الأصلي والاختبارين الفرعيين (المرتفع المنخفض) من حيث مؤشر أقصى معلومات

الاختبار المستخدم في التحليلات	حجم العينة	متوسط تقديرات القدرة مقدرة بالوات	الانحراف المعياري
الكلي (الأصلي)	1530	43.024517	7.0121310
فرعي مرتفع		42.980241	6.9224571
فرعي منخفض		40.001210	6.0013204

يتبين من الجدول السابق رقم (7) أن متوسط تقديرات القدرة الناتجة من الاختبار مرتفع المعلومات بلغت (42.98) تقريبًا، بينما بلغت قيمة متوسط تقديرات القدرة للاختبار منخفض المعلومات (40) تقريبًا، ينما بلغت من الاختبار الأصلي (43.02.) تقريباً.

جدول (8): دلالة تباين تقديرات قدرات الأفراد المشتقة من تحليلات الاختبار الأصلي والاختبارين الفرعيين (المرتفع / المنخفض) تبعاً لمؤشر أقصى معلومات

الاختبار	قيمة الاختبار	قيمة التباين	درجات الحرية	مستوى الدلالة
Wilks' Lambda	0.554	433.653	2	0.00

يتضح من الجدول رقم (8) أن قيمة اختبار ولكس لامدا 'Wilks Lambda بلغت(0.554) بتباين قدره (433.653)، وهي دالة عند مستوى دلالة قدره (0.00) ؛ مما يعني وجود دلالة لتباين تقديرات القدرة الناتجة عن تحليلات الاختبار الأصلي والاختبارين الفرعيين (المرتفع /المنخفض) تبعاً لمؤشر أقصى معلومات تعود إلى الاختبار الجاري استخدامه.

جدول (9): قيمة ودلالة اختبار "ماكولي" لشرط تقديرات القدرة الثلاثة المقاسة الناتجة من الاختبار الأصلي والاختبارين الفرعيين (مرتفع /منخفض اقصى معلومات)

قيمة "ماكولي"	القيمة التقريبية لمربع كاي	درجات الحرية	مستوى الدلالة
0.265	205.56	2	0.001

يتضح من الجدول السابق رقم (9) أن قيمة اختبار ماكولي بلغت (265.)، وهي قيمة دالة ند مستوى (0.01) ، بما يفيد عدم تجانس القياسات الثلاثة للاختبار الأصلي والاختبارين الفرعيين، وعدم تحقق شرط الدورية مما يقتضي استخدام اختبار " جرين هاوس – جيسر" Greenhouse-Geisser

عند فحص دلالة التباين بين القياسات الثلاثة للقدرات كما بالجدول الآتي:

جدول (10):قيمة ودلالة تباين القياسات الثلاثة الناتجة عن الاختبار الأصلي والاختباران الفرعيان (المرتفع /المنخفض) تبعاً لأقصى معلومات

مصدر التباين	الاختبار	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة (ف)	مستوى الدلالة
بين القياسات	Greenhouse-Geisser	5364.143	1.140	4705.389	473.856	.000
داخل القياسات	Greenhouse-Geisser	17141.421	1726.252	9.930		

يتبين من الجدول السابق رقم (10) أن اختبار "جرين هاوس – جيسر " بلغت قيمته (474) تقريبًا، وهي دالة بما يفيد بوجود فروق دالة إحصائياً بين متوسطات قياسات تقديرات القدرة الناتجة من الاختبارات الثلاثة ولتحديد اتجاه الفروق استخدمت المقارنات الثنائية Pairwise Comparisons بطريقةLSD، كما مبين بالجدول الاتي:

جدول (11): المقارنات بين متوسطات تقديرات القدرة الناتجة عن الاختبار الأصلي والاختبارين الفرعيين (مرتفع /منخفض أقصى معلومات) باستخدام اختبار (LSD)

طرف المقارنة الأول	طرف المقارنة الثاني	الفرق بين متوسطي الطرف (الأول – الثاني)
الاختبار الأصلي	الاختبار مرتفع المعلومات	-.355
	الاختبار منخفض المعلومات	1.373*

* دال عند مستوى (0.01)

يتضح من الجدول السابق رقم (11) أن هناك فروقًا دالة إحصائياً بين متوسطي تقديرات القدرة الناتجة من الاختبار الأصلي والتقديرات الناتجة عن تحليلات الاختبار منخفض المعلومات في صالح الاختبار الأصلي، كما تبين عدم وجود فروق دالة بين متوسطات تقديرات القدرة الناتجة عن الاختبار الأصلي والاختبار مرتفع المعلومات، بما يفيد أن تحليلات الاختبار مرتفع المعلومات تسفر عن نفس التقديرات التي تم الحصول عليها من الاختبار الأصلي، وهذا ما أشار إليه "Kingsbury & Zara" (2020) من أن استخدام أسلوب أقصى معلومات في انتقاء المفردات الاختبارية يوفر أعلى كفاءة للاختبار؛ مما يؤدي إلى تقديرات للقدرة أكثر دقة، ويبدو أن هذه النتيجة تتفق مع ما توصل إليه "Wang & Vispoel" (2017) بأن اختيار الفقرة تبعاً لمؤشر أقصى معلومات يعطي دقة لتقديرات القدرة، كما اتفقت هذه النتيجة مع نتائج دراسة (Aimee, 2019) التي أشارت إلى أن انخفاض عدد مفردات الاختبار الموائم، والتي تمتلك أقصى معلومات لا تتأثر به دقة تقديرات القدرة.

٤- نتائج الإجابة على السؤال الثاني:

اعتمدت الدراسة في إجراءات الإجابة عن السؤال الثاني والذي ينص على " هل تختلف تقديرات القدرة المتناظرة الناتجة عن تحليلات الاختبار الأصلي والأخرى الناتجة عن الاختبارات الفرعية المكونة من الاختبار الأصلي والمتضمنة لأعداد مختلفة من الفقرات التي لها أقصى قيم للمعلومات على جميع مستويات القدرة (تبعاً لمقياس التساعيات) على عدد من الإجراءات بيانها فيما يلي:

أ- اشتقاق الاختبارات الفرعية المختلفة من حيث الطول والتي لفقراتها أعلى قيم لمؤشر أقصى معلومات على تسعة مستويات مختلفة على متصل القدرة المستهدفة من القياس تبعاً للمقياس التساعي.

واعتماداً على هذا الإجراء تم الحصول على النتائج الموضحة بالجدول رقم (12)، والتي تضمنت نقاط مقياس التساعيات، والتي تم من خلالها تحديد تسعة مستويات مختلفة على متصل القدرة المستهدفة من القياس بالاختبار الأصلي.

جدول (12):رتب وقيم نقاط مالقياس التساعي للقدرة المقاسة بالاختبار الأصلي

رتبة التساعي	قيمة التساعي	رتبة التساعي	قيمة التساعي
الأول	0.707	الخامس	1.532
الثاني	1.080	السادس	1.851
الثالث	1.202	السابع	2.074
الرابع	1.400	الثامن	2.755

وقد تم اختيار الفقرات التي لها أعلى قيم لمؤشر أقصى معلومات في المستويات التسعة، ومن خلال هذا الإجراء تم اختيار فقرات الاختبارات الفرعية الموزعة على هذه المستويات التسعة، كما هو موضح بالجدول الاتي:

جدول (13): فقرات الاختبارات الفرعية (الثماني) موزعة على المستويات التسعة المحددة بنقاط المقياس التساعي

مستويات القدرة	فترة القدرة المقاسة	فقرات الاختبار	الفقرات التي لها أكبر مؤشر لأقصى معلومات	عدد الفقرات
الأول	.707 > القدرة	الفرعي الأول	29	1
		الفرعي الثاني	29 ، 5	2
		الفرعي الثالث	29 ، 15 ،5	3
		الفرعي الرابع	29 ،15 ،10 ، 5	4
		الفرعي الخامس	29 ، 15، 10 ، 5 ، 3	5
		الفرعي	29 ، 15، 10 ، 5 ،4، 3	6
		الفرعي السابع	29 ، 15 ، 14 ،10، 5 ،4، 3	7
		الفرعي الثامن	، 18 ، 14 ،10، 5 ،4، 3	8

الثاني	> القدرة ≤ 707. 1.080	الفرعي الأول	6
		الفرعي الثاني	25 ، 6
		الفرعي الثالث	25 ، 8 ، 6
		الفرعي الرابع	32، 25 ، 8 ، 6
		الفرعي الخامس	32 ، 31، 25 ، 8 ، 6
		الفرعي	32 ، 31 ، 25، 8 ، 7 ، 6
		الفرعي السابع	32 ، 31 ، 25، 20، 8 ، 7 ، 6
		الفرعي الثامن	، 25، 20 ، 8 ، 7 ، 6 ،1
الثالث	القدرة ≤ 1.088 < 1.202	الفرعي الأول	42
		الفرعي الثاني	42 ، 24
		الفرعي الثالث	42 ،24 ، 2
		الفرعي الرابع	60 ، 42 ،24 ، 2
		الفرعي الخامس	60 ، 42،34 ،24 ، 2
		الفرعي	60 ، 42،34 ، 24، 9 ، 2
		الفرعي السابع	، 42 ، 37، 34 ، 24، 9 ، 2
		الفرعي الثامن	، 37 ، 36، 34 ، 24، 9 ، 2
الرابع	القدرة ≤ 1.202 < 1.400	الفرعي الأول	51
		الفرعي الثاني	12، 51
		الفرعي الثالث	51 ،21 ، 12
		الفرعي الرابع	51 ، 21 ،12، 11
		الفرعي الخامس	75، 51 ، ،21، 12 ، 11
		الفرعي	75، 51 ،33 ، 21، 12 ، 11
		الفرعي السابع	، 51 ،38، 33 ، 21، 12 ، 11
		الفرعي الثامن	،38، 33،23 ، 21، 12 ، 11
الخامس	القدرة ≤ 1.400 < 1.532	الفرعي الأول	16
		الفرعي الثاني	52 ، 16
		الفرعي الثالث	73، 52 ، 16
		الفرعي الرابع	73 ،52، 40، 16
		الفرعي الخامس	73 ،52، 59 ، 40، 16
		الفرعي	73 ،52، 59 ،40، 28 ، 16
		الفرعي السابع	، 59 ، 40 ، 39 ، 28 ، 16
		الفرعي الثامن	، 40 ، 39 ، 28 ،19، 16
السادس	القدرة ≤ 1.532 < 1.851	الفرعي الأول	64
		الفرعي الثاني	71 ، 64
		الفرعي الثالث	71 ، 64 ،49
		الفرعي الرابع	71 ، 64 ،49 ،30
		الفرعي الخامس	71 ، 64، 61 ، 50 ،49 ، 30
		الفرعي	، 71 ، 64، 61 ، 50 ،49 ، 30
		الفرعي السابع	، 71 ، 64، 61 ، 50 ،49 ، 30
		الفرعي الثامن	70 ، 64، 61 ، 50 ،49 ، 30
السابع	القدرة ≤1.581 < 2.074	الفرعي الأول	67
		الفرعي الثاني	76، 67
		الفرعي الثالث	76 ،67 ، 54
		الفرعي الرابع	76 ، 67 ، 54، 44
		الفرعي الخامس	76 ، 68، 67 ، 54، 44
		الفرعي	76 ،68، 67 ، 54 ،44، 26

		الفرعي السابع	،68، 67 ، 54 ،44 ، 41، 26	7
		الفرعي الثامن	،67 ، 58 ، 54 ،44 ، 41، 26	8
الثامن	القدرة ≤ 2.074 < 2.755	الفرعي الأول	80	1
		الفرعي الثاني	80 ، 78	2
		الفرعي الثالث	80 ،78 ، 45	3
		الفرعي الرابع	80 ، 78 ،53 ، 45	4
		الفرعي الخامس	80 ، 78 ، 55 ،53 ، 45	5
		الفرعي	80 ، 78، 65 ، 55 ،53 ، 45	6
		الفرعي السابع	78، 65 ، 55، 53 ، 45 ،43	7
		الفرعي الثامن	65، 56 ، 55، 53 ، 45 ،43	8
التاسع	القدرة ≥ 2.755	الفرعي الأول	62	1
		الفرعي الثاني	77 ، 62	2
		الفرعي الثالث	77 ،62، 48	3
		الفرعي الرابع	77 ،69 ،62، 48	4
		الفرعي الخامس	77، 69، 66 ،62، 48	5
		الفرعي	77، 69 ، 66 ، 62 ، 48، 46	6
		الفرعي السابع	69 ، 66 ، 62 ، 48 ، 47 ،46	7
		الفرعي الثامن	66 ، 62 ، 48 ، 47 ،46 ،17	8

اتضح من الجدول السابق رقم (13) أن فقرات الاختبارات الفرعية الثمانية التي لها أقصى معلومات، والموزعة على المستويات التسعة، انتمت إلى الأبعاد الأربعة للاختبار الأصلي وهي: الاستيعاب اللفظي، الاستدلال اللفظي، الاستدلال الشكلي، الاستدلال الكمي، وبالتالي فإن الاختبارات الفرعية مثلت الاختبار الأصلي بصور جيدة، كما موضح بالجدول الآتي:

جدول (14): توزيع فقرات الاختبارات الفرعية التي لها أقصى معلومات على أبعاد الاختبار الأصلي

الاختبار الفرعي	عدد الفقرات				
	الكلي	أبعاد الاختبار الأربعة			
		الاستيعاب اللفظي	الاستدلال اللفظي	الاستدلال الشكلي	الاستدلال الكمي
الأول	72	20	29	12	11
الثاني	63	16	26	11	10
الثالث	54	14	22	9	9
الرابع	45	14	19	5	7
الخامس	36	11	18	4	3
السادس	27	7	13	4	3
السابع	18	4	9	3	2
الثامن	9	3	4	1	1

ب– استخدام اختبار النسبة الحرجة للكشف عن دلالة الفرق بين نسبتين.

وأسفر هذا الإجراء عن الجدول رقم (15) المتضمن التكرارات والنسب المئوية للفروق الجوهرية بين تقديرات القدرة الناتجة عن الاختبار الأصلي والأخرى الناتجة عن الاختبارات الفرعية ، والجدول رقم (16) المتضمن لقيم اختبار (Z) لدلالة نسب الفروق الدالة إحصائياً بين تقديرات القدرة الناتجة من تحليلات من الاختبار الأصلي والأخرى الفرعية:

جدول (15): التكرارات والنسب المئوية للفروق الجوهرية بين تقديرات القدرة الناتجة من تحليل الاختبار الأصلي، والأخرى الناتجة عن الاختبارات الثماني الفرعية

الفروق في تقديرات القدرة الناتجة من طرفي المقارنة	تكرارات الفروق الجوهرية	نسبة الفروق الجوهرية إلى الفروق الكلية	النسبة المئوية
الاختبار الكلي والاختبار الفرعي الأول (72)	0	0	0
الاختبار الكلي والاختبار الفرعي الثاني (63) مفردة	0	0	0
الاختبار الكلي والاختبار الفرعي الثالث (54) مفردة	4	0.003	0.3
الاختبار الكلي والاختبار الفرعي الرابع (45) مفردة	70	0.046	4.6
الاختبار الكلي والاختبار الفرعي الخامس (36) مفردة	71	0.067	6.7
الاختبار الكلي والاختبار الفرعي السادس (27) مفردة	79	0.052	5.2
الاختبار الكلي والاختبار الفرعي السابع (18) مفردة	172	0.113	11.3
الاختبار الكلي والاختبار الفرعي الثامن (9) مفردة	223		14.7

يبدو من الجدول (15) السابق أن النسب المئوية للتكرارات الخاصة بالفروق الجوهرية بين تقديرات القدرة الناتجة عن تحليل الاختبار الأصلي والتقديرات الناتجة من تحليلات الاختبارات الفرعية مختلفة الأطوال التي تتضمن فقط من (54) إلى (72) فقرة، لم تتجاوز النسبة المئوية (5%) تقريباً مما يشير إلى تكافؤ التقديرات الناتجة عن هذه الاختبارات مع نظائرها الناتجة عن الاختبار الأصلي، والجزء التالي يعرض لنتائج اختبار (z) دلالة الفروق الجوهرية

في تقديرات القدرة الناتجة عن الاختبار الأصلي والاختبارات الأخرى مختلفة الأطوال.

جدول (16): قيمة ودلالة اختبار (Z) بين نسب الفروق الجوهرية في تقديرات القدرة الناتجة عن الاختبار الأصلي، والأخرى الناتجة عن الاختبارات الفرعية مختلفة الأطوال

		النسبة المئوية للفروق الجوهرية بين تقديرات القدرة المشتقة من الاختبار الكلي المرجعي (خ م) وكل اختبار فرعي على حدة				
		اختباري المقارنة	الأول/ خ م	الثاني/ خ م	الثالث/ خ م	الرابع/ خ م
	اختباري المقارنة	النسبة المئوية للفروق الجوهرية	0	0	0.2	1.1
النسبة المئوية للفروق الجوهرية بين تقديرات القدرة المشتقة من الاختبار الكلي (المرجعي) وكل اختبار فرعي	الأول/ خ م	**0**	-	-	-	-
	الثاني / خ م	**0**	**0**	**0**	-	-
	الثالث / خ م	**0.2**	**1.9***	**1.9***	**0**	-
	الرابع / خ م	**1.1**	**4.1****	**4.1****	**2.8****	**0**
	الخامس / خ م	**2.8**	**6.8****	**6.8****	**5.8****	**3.3****
	السادس / خ م	**50**	**8.9****	**8.9****	**8.1****	**6.0****
	السابع / خ م	**109**	**13.9***	**13.9****	**13.3****	**11.6****
	الثامن / خ م	**13.8**	**12.2****	**12.2****	**15.6****	**14****

		اختباري المقارنة	الخامس / خ م	السادس / خ م	السابع / خ م	الثامن / خ م
النسبة المئوية للفروق الجوهرية بين تقديرات القدرة المشتقة من الاختبار الكلي (المرجعي) وكل اختبار فرعي	اختباري المقارنة	النسبة المئوية للفروق الجوهرية	2.8	5.0	10.9	13.8
	الأول /خ م	**0**	-	-	-	-
	الثاني /خ م	**0**	-	-	-	-
	الثالث /خ م	**0.2**	-	-	-	-
	الرابع / خ م	**1.1**	-	-	-	-
	الخامس / خ م	**2.8**	**0**	-	-	-
	السادس / خ م	**50**	**2.9****	**0**	-	-
	السابع / خ م	**109**	**8.3****	**5.9****	**0**	-
	الثامن / خ م	**13.8**	**11.3****	**8.3****	**2.6****	**0**

يبدو من الجدول رقم (16) أن هناك فروقاً جوهرية بين نسب الفروق الجوهرية بين تقديرات القدرة الناتجة من الاختبار الأصلي والاختبارات الفرعية مختلفة الأطوال، والتي تشتمل على عدد من المفردات ينحصر بين (9 إلى 54) ، عند مستوى (0.01) عدا الاختبار الذي يحتوي (54) فقرة، فقد جاءت الفروق

عند مستوى أقل من (0.05)،كما تبين عدم وجود فروق جوهرية بين فروق نسب الفروق الجوهرية الناتجة عن الاختبار الأصلي وكل من الاختبارين الفرعيين المتضمنين لعدد من الفقرات (63 ، 72) فقرة، وتشير هذه النتائج ان كل من الاختبار الفرعي الأول والمحتوى على(72) فقرة، والفرعي الثاني المحتوي على (63) مفردة لم ينتج عنهما فروق جوهرية بين تقديرات القدرة الناتجة من تحليل كل منهما، وتقدير القدرة الناتج عن الاختبار الأصلي، وبالرغم من هذه النتائج فإن الباحث يرى أن الاختبارات الفرعية المسحوب فقراتها من الاختبار الأصلي تبعاً لمؤشر أقصى معلومات، والتي تتضمن (54) فقرة لم تتخط نسب الفروق الجوهرية بين تقديرات القدرة الناتجة عن تحليلها والأخرى التي الناتجة من الاختبار الأصلي (0.02) تقريباً، وهي لم تتخط النسبة (0.05) من جملة الفروق بين التقديرات المتناظرة للقدرة؛ مما يثبت قبول تقديرات القدرة الناتجة عن تحليلات هذه الاختبارات الفرعية مختلفة الأطوال بقدر عالي من الثقة كما لو كانت ناتجة من الاختبار الأصلي.

ويستطيع الباحث بناءً على ما سبق من نتائج أن يقبل بقدر عالٍ من الثقة إمكانية خفض فقرات الاختبار باستخدام متغير أقصى معلومات للفقرة.

ج- **الإجراء الثالث: والذي يستهدف استخدام اختبار تباين القياس المتعدد بين قياسات القدرة الناتجة من تحليل الاختبار الأصلي من جهة، وقياسات القدرة الناتجة عن تحليلات الاختبارات الفرعية مختلفة الأطول.**

وبناءً على هذا الإجراء تم الحصول على النتائج الموضحة بالجداول الآتية:

جدول (17): المتوسط والانحراف المعياري لتقديرات القدرة للعينة الناتجة من الاختبار الأصلي والاختبارات الفرعية مختلفة الأطوال

الاختبار المستخدم في تقدير القدرة	متوسط القدرات المقدرة بوحدة	الانحراف المعياري بوحدة الوات	الاختبار المستخدم في تقدير القدرة	متوسط القدرات المقدرة بوحدة	الانحراف المعياري بوحدة الوات
الاختبار الكلي (80) مفردة	41.31	5.40	الاختبار الفرعي الرابع) مفردة	41.19	5.70
الاختبار الفرعي الأول	40.96	5.45	الاختبار الفرعي	41.00	6.22
الاختبار الفرعي الثاني مفردة	41.38	5.40	الاختبار الفرعي السادس (27)	41.70	6.30
الاختبار الفرعي الثالث مفردة	41.12	5.73	الاختبار الفرعي السابع (18) مفردة	42.87	6.77
الاختبارالفرعي الثامن	43.94	8.82			

يتضح من الجدول رقم (17) أن أكبر قيمة لمتوسط تقديرات القدرة القيمة (43.94) تقريبًا، وهي ناتجة عن الاختبار الفرعي الثامن، بينما بلغت أقل قيمة (40.96) تقريبًا، واشتقت من الاختبار الفرعي الأول.

جدول (18): دلالة تباين قياسات القدرة الناتجة عن تحليل الاختبار الأصلي والاختبارات الفرعية مختلفة الأطول

الاختبار	قيمة الاختبار	قيمة التباين	درجات الحرية	مستوى الدلالة
Wilks' Lambda	0.822	91. 303	8	0.040

يبدو من الجدول رقم (18) الذي يتضمن نتائج دلالة تباين قياسات القدرة الناتجة عن التحليلات الخاصة بالاختبار الأصلي والاختبارات الفرعية مختلفة الأطوال، أن اختبار "ولكس لامدا " Wilks' Lambda قد بلغت قيمته (.822) بتباين قدره (91.303) وهي دالة عند مستوى (0.05).

مما يعني وجود دلالة لتباين قياسات القدرة الناتجة عن تحليل الاختبار الأصلي والاختبارات الأخرى الفرعية مختلفة الأطوال تُعزى إلى الاختبار المستخدم.

جدول (19): اختبار ماكولي للتحقق من شرط الدورية لقياسات القدرة المشتقة من الاختبار الكلي المرجعي والاختبارات الفرعية

قيمة اختبار ماكولي	مربع كاي التقريبية	درجات الحرية	مستوى الدلالة
0.751	751.328	8	0.00

يبدو من جدول (19) السابق أن قيمة اختبار الدورية لماكولي للدورية بلغت (0.751) تقريبًا وهي دالة؛ مما يعني عدم وجود تجانس بين القياسات الناتجة عن الاختبارات التسعة، وبالتالي عدم توافر شرط الدورية؛ لذا يجب استخدام اختبار " جرين هاوس – جيسر" Greenhouse-Geisser عند اختبار دلالة التباين الخاص بين تلك القياسات، وهو موضح بالجدول الآتي:

جدول (20): دلالة تباين قياسات تقديرات القدرة الناتجة عن تحليل الاختبار الأصلي والاختبارات الفرعية مختلفة الأطوال

مصدر التباين	الاختبار	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة (ف)	مستوى الدلالة
بين القياسات	**Greenhouse-Geisser**	272.721	1.992	136.91	3.4615	0.030
داخل القياسات	**Greenhouse-Geisser**	4224.922	111.571	37.867		

يبدو من الجدول (20) السابق أن قيمة اختبار "جرين هاوس – جيسر " بلغت (3.46) تقريباً دالة عند ($p<0.05$)؛ مما يعني وجود فروق جوهرية بين متوسطات القياسات التسعة للاختبارات المختلفة من حيث الطول،

وللتعرف على اتجاه الفروق في صالح أي قياس سوف يجرى استخدام المقارنات الثنائية Pairwise Comparisons بطريقة LSD، وهي مبينة بالجدول التالي:

جدول (21): المقارنات الثنائية بين متوسطات القدرة الناتجة من الاختبار الأصلي والاختبارات مختلفة الأطوال باستخدام اختبار (LSD)

الطرف الأول للمقارنة	الطرف الثاني للمقارنة	فرق متوسطي الطرفين (الأول – الثاني) للمقارنة
الاختبار الأصلي	الفرعي الأول	.185
	الفرعي الثاني	.175-
	الفرعي الثالث	.120-
	الفرعي الرابع	.982*-
	الفرعي الخامس	1.289*-
	الفرعي السادس	1.790-*
	الفرعي السابع	2.777-*
	الفرعي الثامن	3.110-*

* دال عند مستوى (**0.05**)

يبدو من جدول (21) السابق وجود فروق جوهرية بين متوسطي تقديرات القدرة الناتجة من تحليل الاختبار الأصلي وتقديرات القدرة الناتجة عن تحليلات الاختبارات الفرعية مختلفة الأطوال: الفرعي (الرابع) المكون من (45) فقرة ، والخامس المكون من (36) فقرة إلى الثامن المكون من (9) فقرات، في صالح الاختبارات الفرعية (الرابع ، والخامس... ، حتى الثامن)، وهي الاختبارات الفرعية التي تمتلك المتوسط الأكبر، وقد تبين عدم وجود فروق جوهرية بين متوسطات تقديرات القدرة الاختبار الأصلي المشتمل على (80) فقرة ، والاختبارات الفرعية: الأول المكون من (72) فقرة والثاني المكون من (63) فقرة، والثالث المكون من (54) فقرة؛ مما يفيد بإمكانية استخدام متغير

أقصى معلومات للفقرة الاختبارية في تخفيض عدد اختزال فقرات الاختبار من (80) فقرة إلى عدد (54) فقرة.

وهذه نتائج مقبولة حيث إن فقرات الاختبار الأصلي بعضها تمتلك أقصى معلومات، وبالتالي فإن تقديرات القدرات الناتجة من هذا الاختبار تتقارب مع التقديرات الناتجة من تحليل الاختبار الفرعي الناتج من الاختبار الأصلي، والذي يتضمن فقط فقرات تمتلك أقصى معلومات، وهذا ما أشار إليه "Kingsbury & Zara" (2020) وهو أن الاعتماد على مؤشر أقصى معلومات يعطي أعلى كفاءة للاختبار، والذي يسفر عن تقديرات أكثر دقة، و هذه النتائج أيضاً تتفق مع ما أشار إليه "Wang& Vispoel" (2017) بأن انتقاء الفقرة تبعاً لمتغير أقصى معلومات يسفر عن تقديرات دقيقة للقدرة المقاسة، كما تتفق هذه النتيجة كذلك مع نتائج تقديرات القدرة التي يمكن تحقيقها باستخدام اختبارات المواءمة الحاسوبية، التي تؤدي إلى تقديرات أدق للقدرة مع العدد الأقل من الفقرات " Schmitt" (2010) ، كما أسهمت هذه النتيجة في دعم النتائج التي توصلت إليها دراسة "Aimee" (2019) ، وهي أن النقص في عدد فقرات الاختبار الموائم التي لها أقصى معلومات لا يؤثر بالسلب في دقة تقديرات الاختبار للقدرة، ولا يؤثر في خصائصه السيكومترية، في حين تعارضت هذه النتائج مع ما أفادت به دراسة "Goodman" (2019) والتي توصلت نتائجها إلى أنه كلما قل طول الاختبار كلما أثر ذلك سلبيًا على تقدير القدرة.

٥- نتائج الإجابة عن السؤال الثالث:

أسفرت إجراءات الدراسة للإجابة عن السؤال الثالث، والذي ينص على "ما أقل عدد مناسب من الفقرات الاختبارية التي لها أقصى معلومات، والتي يمكن أن تُعطي نفس تقديرات القدرة المشتقة من الاختبار الأصلي؟

لقد تبين من خلال الإجراءات التي استُخدمت للإجابة عن السؤال الثاني من أن أقل عدد من الفقرات التي أدت إلى نفس تقديرات القدرة التي اشتقت من الاختبار الأصلي هي فقرات بالاختبار الفرعي الثالث، والتي بلغ عددها (54) فقرة، وهي تقريبًا تساوي ثُلثي عدد الفقرات المتضمنة بالاختبار الأصلي، والبالغ عددها (80) فقرة.

وهذه النتيجة جاءت بشكل منهجي وإجراءات مدروسة، تم الاعتماد عليها في الإجابة عن هذا السؤال، وهذه النتيجة تعتبر مقبولة حيث أن فقرات الاختبار الذي يحتوي (54) فقرة هي فقرات تمتلك أعلى قيم لمؤشر أقصى لمعلومات من بين فقرات الاختبار الأصلي، وهذا ما أشار إليه " Kingsbury & Zara "(2020) من أن الاعتماد على أسلوب أقصى معلومات في اختيار الفقرات الاختبارية يعطي تقديرات أكثر دقة للقدرة.

٦- نتائج الإجابة عن السؤال الرابع:

1- لقد أسفرت إجراءات الإجابة عن هذا السؤال، والذي ينص على: "هل تختلف الشروط السيكومترية للاختبارات الفرعية ذات الأطوال المختلفة والتي تحتوي على فقرات لها أقصى معلومات على فترات مختلفة من القدرة تبعاً

للمقياس التساعي عن الشروط السيكومترية للاختبار الأصلي؟" عن النتائج التالية:

انتهت التحليلات الخاصة بالتحقق من الخصائص السيكومترية (الثبات والصدق) للاختبار الأصلي والاختبارات الفرعية (الثمانية) مختلفة الأطوال والمسحوبة من نفس الاختبار عن النتائج التالية:

(أ) نتائج حساب الثبات:

تم الاعتماد على معامل "ألفا كرونباخ" لتقدير ثبات الاتساق الداخلي لفقرات الاختبار الأصلي، والاختبارات الفرعية مختلفة الأطوال، وقد أسفر هذا الإجراء عن النتائج المبينة بالجدول الآتي:

جدول (22): معامل ألفا كرونباخ لحساب ثبات الاتساق الداخلي

الاختبار	عدد الفقرات	معامل ألفا كرونباخ لثبات للاتساق الداخلي
الأصلي	80	0.84
الفرعي الأول	72	0.83
الفرعي الثاني	63	0.82
الفرعي الثالث	54	0.82
الفرعي الرابع	45	0.60
الفرعي الخامس	36	0.61
الفرعي السادس	27	0.58
الفرعي السابع	18	0.55
الفرعي الثامن	9	0.43

يبدو من النتائج المتضمنة بجدول (22) السابق أن معامل الثبات للاختبار الأصلي بلغت قيمته (0.84)، كما تبين أن قيمة معامل ألفا قد انخفض بشكل واضح للاختبار الفرعي الرابع، والخامس السادس والسابع والثامن، والتي اشتملت على فقرات أحجامها (45، 36 ، 27، 18، 9) حيث

انخفض بشدة معامل ألفا للاختبارات السابقة بين القيمتين (0.60 إلى 0.43) على الترتيب، مما يفيد بوجود فروق جوهرية بين قيمة معامل ثبات ألفا للاختبار الأصلي وقيم نفس المعامل للاختبارات الخمسة السابقة، في حين تبين أن الاختبارات الفرعية من الأول وحتى الثالث والتي تشتمل على (72 ، 63 ، 54) فقرة على الترتيب تراوحت قيم ثبات الاتساق لها بين القيمتين (0.82 إلى 0.84) ، وهي قيم لم تتخط الفروق بينها وبين قيمة معامل ألفا المحسوب من الاختبار الأصلي الحد (0.05) بما يفيد بعدم وجود فروق جوهرية بين قيم معاملات الثبات لهذه الاختبارات الفرعية الثلاثة، وقيمة معامل الثبات المحسوب من الاختبار الأصلي، وبالتالي فإن الاختبارات الفرعية مختلفة الأطوال، والتي اشتملت على فقرات يتراوح عددها بين (0.54 إلى 72) فقرة معاملات الثبات الخاصة بها تتساوى تقريباً به مع معامل الثبات للاختبار الأصلي، وهذا عكس نتائج بعض الدراسات التي تفيد بأن ثبات الاختبار يتأثر سلباً بطول الاختبار (فرج ، 2007)..

(ب) نتائج حساب الصدق:

اعتمد إجراءات حساب الصدق على استخدام معامل ارتباط "بيرسون" بين تقديرات القدرة المتناظرة الناتجة عن تحليل الاختبار الأصلي، وتحليلات الاختبارات الفرعية مختلفة الأطوال وقد أسفر هذا الإجراء عن النتائج المتضمنة بالجدول التالي:

جدول (23): قيم معاملات الارتباط بين تقديرات القدرة الناتجة من الاختبارات الفرعية وتقديرات القدرة الناتجة من الاختبار الأصلي

الاختبارات الفرعية	عدد الفقرات	معامل ارتباط القدرة الناتجة من الاختبار الفرعي مع تقدير القدرة الناتجة من الاختبار الأصلي
الأول	72	.981**
الثاني	63	.980**
الثالث	54	.979**
الرابع	45	.831**
الخامس	36	.818**
السادس	27	.752**
السابع	18	.747**
الثامن	9	.621**

** دال عند مستوى (0.01)

يبدو من جدول (23) السابق أن معاملات الارتباط بين تقديرات القدرة الناتجة عن الاختبار والأخرى الناتجة من الاختبارات الفرعية (الأول والثاني والثالث) كان لها قيم تراوحت بين القيمتين (0.97 إلى 0.99) تقريباً، وهي دالة عند($p < 0.01$) مما يعني وجود علاقة طردية قوية بين التقديرات الناتجة من الاختبار الأصلي، والاختبارات الفرعية (الأول والثاني والثالث)، بينما انخفض معامل الارتباط عن القيمة (0.90) بين التقديرات الناتجة من الاختبار الأصلي، والأخرى الناتجة عن الاختبارات الفرعية (من الرابع حتى الثامن) فقد تراوحت قيم معاملات ارتباطها بين القيمتين (0.62 إلى 0.83) تقريباً وهي دالة عند ($p < 0.01$) كذلك، مما يفيد بأن الاختبارات الفرعية مختلفة الأطوال، والتي تشمل على فقرات حجمها (72 ، 63 ، 54،) تقيس نفس المتغير المقاس بالاختبار الأصلي كاختبار محكي. مما يعني أن هذه الاختبارات الفرعية أعلى صدقاً من الاختبارات الفرعية، والتي تشمل على

فقرات حجمها (45، 36 ،27 ، 18 ،9)، وهذه النتائج تشير إلى أن نتائج التحليلات التي انتهت إليها التحليلات الخاصة بالاختبارات الفرعية الثلاثة، والتي تتضمن فقرات عددها (54، 45 ، 36) يمكن قبولها بدرجة مرتفعة من الثقة حال مقارنتها بالاختبارات الفرعية الأخرى. وهذه النتيجة تتفق من نتائج دراسة (Aimee ,2019) التي توصلت إلى أن انخفاض عدد فقرات الاختبار الموائم المحوسب، والتي لها أقصى معلومات لا يؤثر بالسلب في خصائص الاختبار السيكومترية.

قائمة المراجع

أولاً: المراجع العربية:

ذكري، علي(2008). الخصائص السيكومترية لاختبار (أوتيس – لينون) للقدرة العقلية مقدرة وفق القياس الكلاسيكي ونموذج راش لدى طلبة المرحلة المتوسطة بمحافظة صبيا التعليمية ، (رسالة ماجيستير،غير منشورة)، جامعة أم القرى،كلية التربية.

علام ، صلاح الدين (1986): تطورات معاصرة في القياس النفسي والتربوي، الكويت، مطابع القبس التجارية.

علام ، صلاح الدين (2005) : نماذج الاستجابة للفقرة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي ، ط1 ، القاهرة،دار الفكر العربي.

فرج ،صفوت (2007).القياس النفسي، القاهرة، مكتبة الأنجلو المصرية.

كاظم، أمينة (1988). دراسة نظرية نقدية حول التفسير الموضوعي للسلوك (نموذج راش) ،ط1، الكويت ،مؤسسة الكويت للتقويم العلمي.

مراد ،صلاح و الشافعي،محمد (1998). أثر حجم العينة فى دقة وكفاءة ضم اختبارين فى تدريج مشترك (بنك الأسئلة)". مجلة البحوث النفسية والتربوية، جامعة المنوفية، كلية التربية ، 12(3) ، 55-85.

مراد،صلاح وعبدالغفار،محمد.(1985). اختبار أوتيس – لينون للقدرة العقلية العامة المستوى المتوسط (11- 16) سنة، كراسة الأسئلة ، القاهر ة : دار النهضة العربية.

ثانياً : المراجع الإنجليزية

Aimee,M,(2019). A comparison of (CAT) scores based on based on different Measurement Models for testlets, *Educational and psychological measurement*,31(3),371-398.

Altaf, Saim, Aslam, Muhammad.(2011)**.** Bayesian Analysis Of The Davidson Model For Paired Comparison With Order Effect Using Noninformative Priors, *Pakistan Journal of Statistics*, 27 (2): ,171-185.

Azevedo, Caio L.N , Bolfarine, Heleno: Andrade, Dalton F .(2011).Bayesian inference for a skew-normal IRT model under the centred parameterization. *Computational Statistics & Data Analysis*, 55 (1),353-365.

Azevedo,C.L.N. (2009). Some Observations on the Identification and Interpretation of the 3PL IRT Model, *Applied Psychological Measurement* , 33(7),.89–114.

Bond, T. G., & Fox, C. M. (2020). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

Chon, K.; Lee, W.; Dunbar, S.,(2010) . A Comparison of Item Fit Statistics for Mixed **IRT** Models, *Journal of Educational Measurement*, 45 (2),320-344.

Crall, J. J. (2018). Classical test theory and item response theory analyses of multi-item scales assessing parents' perceptionsof their children's dental care. *Medical Care*, 44(2),560–568.

De Ayala,R.J.(2022).*The Theory and Practice of Item Response Theory*,2Ed.NY.the Guilford press.

de la Torre, Jimmy.(2017). Multidimensional Scoring of Abilities: The Ordered Polytomous Response Case, *Applied Psychological Measurement*, 32 (5),355-370.

Dodd, B. (2017). The effect of item selection procedure and stepsize on omputerized adaptive attitude measurement using the rating scale model. *Applied Psychological Measurement,31(5) 255-269.*

Edward, H. (2010) . Interpretation of the Three-Parameter Testlet Response Model and Information Function. *Applied Psychological Measurement*, 34 (7),467-482.

Gessaroli ME & De Champlain A. (2005). *Test dimensionality*: Assessment of. In: Everitt BS, Howell DC, Eds.Encyclopedia of Statistics in Behavioral Science. Chichester:John Wiley & Sons.

Goodman, Joshua T.(2019). Identification of Differential Item Functioning in Assessment Booklet Designs With Structurally Missing Data**,** *Educational & Psychological Measurement***,** 71 (1),80-94.

Hambleton,R.(1987).*Item Response Theory Principles And Applications* :N.Y.,Kluwer.Nijhoff Publishing.

Hattie,J.:Methodology Review(1985). Assessing Unidimensionality of Tests and Items, *Applied Psychological Measurement*, 9 (2),139-164.

Jeon, K. S., & Kim, H.-Y. (2019). Development of relativization in Koreanas a foreign language: The noun phrase accessibility hierarchy in headinternaland head-external relative clauses. *Studies in Second LanguageAcquisition*, 29(2),253–276.

Kahraman, Nilufer, Thompson, Tony.(2018). Relating Unidimensional IRT Parameters to a Multidimensional Response Space: A Review of Two Alternative Projection IRT Models for Scoring Subscales**,***Journal of Educational Measurement,* 48 (2),146-164.

Kingsbury & Zara .(2020). Item Selection in Computerized Adaptive Testing: Should More Discriminating Items be Used First? *Journal of Educational Measurement*, 33 (2),221-255.

Liden,W., & Pushley ,P.(2015). *item selction and ability estimation in adaptive testing. Computerized adaptive testing : theory and practice* ,kluwer academic publishers .

Lihua Yao.(2019). Multidimensional Linking for Domain Scores and Overall Scores for Nonequivalent Groups. *Applied Psychological Measurement*, 33 (2), 40-68.

Lord, F. M. (1980*). Applications of item response theory to practical testing problems.* Hillsdale, NJ: Erlbaum.

Magis, David**.(**2020). A Test-Length Correction to the Estimation of Extreme Proficiency Levels**,** *Applied Psychological Measurement*, 35 (2):p p91-109.

Marinagi, Catherine C, Kaburlasos, Vassilis G.(2021). Bayesian Decision Theory for Multi-Category Adaptive Testing, *Educational & Psychological Measurement* , 20(1),37-59.

. Okewole, D. M., Olubusoye, O. E, Shangodoyin, D. K. (2020). A Comparative Study of the Classical and **Bayesian** Methods of Estimating a Just-Identified Simultaneous Equations Model. *Asian Social Science*,7 (5),222-235

Penfield, Randall D.(2016). Applying Bayesian Item Selection Approaches to Adaptive Tests Using Polytomous Items. *Applied Measurement in Education*, 17 (1),1-20.

Rose ,B. Bjorner , J. Becker , F. Friesc,.E. (2008) . Evaluation Of A Preliminary Physical Function Item Bank Supported The Expected Advantages Of The Patient-Reported Outcomes Measurement Information System (Promis) , *Journal of Clinical Epidemiology* ,61(4),17- 33.

Schmitt, T. A.(2010). A Monte Carlo Simulation Investigating the Validity and Reliability of Ability Estimation in Item Response Theory with Speeded Computer Adaptive Tests. *International Journal of Testing*, 10 (3),230-261.

Siang C. & Fritz D.(2006). How Big Is Big Enough? Sample Size Requirements for CAST Item Parameter Estimation**,** *Applied Psychological Measurement*, 19(3),241–255.

Stocking, M. (2019). *Two Simulated Feasibility Study in Computerized Adaptive Testing*. (Research Rep. 84-15). Princeton NJ: Educational Testing Service.

Uttl, Bob .(2018). Measurement of Individual Differences. *Psychological Science* (Wiley-Blackwell), 16 (6),460-467.

Wang, T. & Vispoel, W. (2017). Properties of ability estimation methods in computerized adaptive testing. *Journal of Educational Measurement*, 35(2),109-135.

Wang, W., Liu, C.(2018). Computerized Classification Testing Under the Generalized Graded Unfolding Model**,** *Educational & Psychological Measurement*, 71 (1),114-128**.**

Wells, Craig S. , Subcoviak, Michael J. ,Serlin, Ronald C.(2021). The Effect of Item Parameter on Examinee Ability Estimates. *Educational & Psychological Measurement*, (1),33-52.

Zimowski, M. ,Muraki,E.,. Mislevy,R., Bock ,R.(2007).*BILOG -MG 3 for Windows*, Chicago:Scientific Software Inc.

al-Marājiʻ al-ʻArabīyyah:

Dhikrī, ʻAlī (2008). al-Khaṣāʼiṣ alsykwmtryh lākhtbār (awtys – lynwn) llqdrh al-ʻaqlīyah mqdrh wafqa al-qiyās alklāsyky wa-namūdhaj rāsh ladá ṭalabat al-marḥalah al-mutawassiṭah bi-Muḥāfaẓat ṣbyā al-taʻlīmīyah, (Risālat mājīstīr, ghayr manshūrah), Jāmiʻat Umm al-Qurá, Kullīyat al-Tarbiyah.

Allām, Ṣalāḥ al-Dīn (1986) : taṭawwurāt muʻāṣirah fī al-qiyās al-nafsī wa-al-tarbawī, al-Kuwayt, Maṭābiʻ al-Qabas al-Tijārīyah.

Allām, Ṣalāḥ al-Dīn (2005) : namādhij al-istijābah llfqrh alākhtbāryh aḥādīyat al-Buʻd wmtʻddh al-abʻād wa-taṭbīqātuhā fī al-qiyās al-nafsī wa-al-tarbawī, Ṭ1, al-Qāhirah, Dār al-Fikr al-ʻArabī.

Faraj, Ṣafwat (2007). al-qiyās al-nafsī, al-Qāhirah, Maktabat al-Anjlū al-Miṣrīyah.

Kāẓim, Amīnah (1988). dirāsah Naẓarīyat naqdīyah ḥawla al-tafsīr al-mawḍūʻī lil-sulūk (namūdhaj rāsh), Ṭ1, al-Kuwayt, Muʼassasat al-Kuwayt lil-taqwīm al-ʻIlmī.

Murād, Ṣalāḥ wa al-Shāfiʻī, Muḥammad (1998). Athar ḥajm al-ʻayyinah fī diqqat wa-kafāʼat ḍamm akhtbāryn fī tdryj mushtarak (Bank al-asʼilah) ". Majallat al-Buḥūth al-nafsīyah wa-al-tarbawīyah, Jāmiʻat al-Minūfīyah, Kullīyat al-Tarbiyah, 12 (3) Ṣ Ṣ, 55-85.

Murād, Ṣalāḥ wʻbdālghfār, Muḥammad. (1985). ikhtibār awtys – lynwn llqdrh al-ʻaqlīyah al-ʻĀmmah al-mustawá al-Mutawassiṭ (11-16) sanat, Kurrāsah al-asʼilah, al-Qāhir ah : Dār al-Nahḍah al-ʻArabīyah.

Arabic References:

Dhikrī, ʻAlī (2008): al-Khaṣāʼiṣ alsykwmtryh lākhtbār (awtys – lynwn) llqdrh al-ʻaqlīyah mqdrh wafqa al-qiyās alklāsyky wa-namūdhaj rāsh ladá ṭalabat al-marḥalah al-mutawassiṭah bi-Muḥāfaẓat ṣbyā al-taʻlīmīyah, (Risālat mājīstīr, ghayr manshūrah), Jāmiʻat Umm al-Qurá, Kullīyat al-Tarbiyah.

ʻAllām, Ṣalāḥ al-Dīn (1986): taṭawwurāt muʻāṣirah fī al-qiyās al-nafsī wa-al-tarbawī, al-Kuwayt, Maṭābiʻ al-Qabas al-Tijārīyah.

ʻAllām, Ṣalāḥ al-Dīn (2005): namādhij al-istijābah llfqrh alākhtbāryh aḥādīyat al-Buʻd wmtʻddh al-abʻād wa-taṭbīqātuhā fī al-qiyās al-nafsī wa-al-tarbawī, Ṭ1, al-Qāhirah, Dār al-Fikr al-ʻArabī.

Faraj, Ṣafwat (2007). al-qiyās al-nafsī, al-Qāhirah, Maktabat al-Anjlū al-Miṣrīyah.

Kāẓim, Amīnah (1988). dirāsah Naẓarīyat naqdīyah ḥawla al-tafsīr al-mawḍūʻī lil-sulūk (namūdhaj rāsh), Ṭ1, al-Kuwayt, Muʼassasat al-Kuwayt lil-taqwīm al-ʻIlmī.

Murād, Ṣalāḥ wa al-Shāfiʻī, Muḥammad (1998). Athar ḥajm al-ʻayyinah fī diqqat wa-kafāʼat ḍamm akhtbāryn fī tdryj mushtarak (Bank al-asʼilah) ". Majallat

al-Buḥūth al-nafsīyah wa-al-tarbawīyah, Jāmi‘at al-Minūfīyah, Kullīyat al-Tarbiyah, 12 (3), 55-85.
Murād, Ṣalāḥ w‘bdālghfār, Muḥammad. (1985). ikhtibār awtys – lynwn llqdrh al-‘aqlīyah al-‘Āmmah al-mustawá al-Mutawassiṭ (11-16) sanat, Kurrāsah al-as’ilah, al-Qāhir ah : Dār al-Nahḍah al-‘Arabīyah.