تأثير حجم العينة على دقة تقديرات كل من : معلم الصعوبة للمفردات الاختبارية، ومعلم القدرة لأفراد عينة التحليل باستخدام النموذج اللوجيستي ثلاثي المعلم

تاريخ استلام البحث: ٢٠٢٤\٦\٢٧ تاريخ قبول البحث للنشر: ٢٠٢٤\٧\٨

أ.د. محمد منصور محمد الشافعي*

المستخلص

استهدفت الدراسة الكشف عن تأثير حجم العينة على تقديرات صعوبات المفردات وقدرات الأفراد باستخدام تحليلات النموذج اللوجيستي ثلاثي المعلم ، وقد بلغ حجم عينة الدراسة من (١٨٠٠) فرد، حيث قام الباحث بتطبيق اختبار لمادة الجبر للصف الثاني الثانوي (مكون من ٤٠ مفردة اختبارية من موع الاختيار من متعدد)، على أربعة عينات ذات أحجام مختلفة (٣٠٠،٦٠٠،٩٠٠، ١٨٠٠) فرد، واستخدم الباحث في التحليلات الخاصة بالدراسة برنامجي التحليلات الإحصائية SPSS v.27 , -Bilog_MG3 لإجراء التحليلات الخاصة بمفردات الاختبار، وقد استخدمت النتائج أخطاء القياس كمؤشر لدقة القياس والكشف عن الفروق الدالة بين تقديرات الصعوبة المتناظرة كما استخدمت الدراسة أسلوب تحليل تباين القياس المتعدد للكشف عن الفروق الدالة بين تقديرات القدرة المتناظرة الناتجة عن تحليل استجابات الحجام المختلفة من العينة ، أسفرت نتائج الدراسة الى وجود فروق ذات دلالة إحصائية عند تقدير صعوبات الفقرات الاختبارية لاختلاف حجم العينة، كما تبين أي ان دقة تقدير صعوبات الفقرات الاختبارية وكذلك دقة تقديرات قدرات الأفراد تزداد بزيادة حجم العينة.

الكلمات المفتاحية : معلم الصعوبة- معلم القدرة - حجم عينة التحليل - نظرية الاستجابة للمفردة - النموذج اللوجيستي الثلاثي المعلم.

The effect of sample size on the accuracy of estimates of: the difficulty parameter of the test items, and the ability parameter for individuals in the analysis sample using the three-parameter logistic model

Prof. Dr. Muhammad Mansour Muhammad Al-Shafi'i

Abstract

The study aimed to reveal the effect of sample size on estimates of vocabulary difficulties and individuals' abilities using analyzes of the three-teacher logistic model. The size of the study sample was (1800) individuals, where the researcher applied an algebra test for the second year of secondary school (consisting of 40 test items from the selection type). Multiple), on four samples of different sizes (900, 600, 300, 1800) individuals. In the analyzes of the study, the researcher used the statistical analysis programs SPSS v.27 and -Bilog_MG3 to conduct analyzes of the test items. The results used measurement errors as an indicator of accuracy. Measurement and detection of significant differences between corresponding difficulty estimates. The study also used the multiple measurement analysis of variance method to detect significant differences between corresponding ability estimates resulting from analyzing the responses of different sizes of the sample. The results of the study resulted in the presence of statistically significant differences when estimating the difficulties of the test items due to differences in size. The sample also shows that the accuracy of estimating the difficulties of the test items, as well as the accuracy of estimating individuals' abilities, increases with increasing sample size

Keywords: difficulty parameter - ability parameter - analysis sample size - item response theory - three-parameter logistic model.

(١)أستاذ القياس والإحصاء- قسم علم النفس – كلية الدراسات الإنسانية والإدارية- كليات عنيزة بالقصيم

مقدمة

تهتم إجراءات القياس النفسي بتحقيق الموضوعية المنشودة للأدوات المستخدمة في تقدير درجات الأفراد على السمات والخصائص السلوكية المختلفة، ومن اجل تحقيق هذا الغرض .(De Ayala,R.J.,2022) ، يجب ان تشير هذه التقديرات إلى دقة مرتفعة في وضع الافراد على متصل السمة المستهدفة من إجراء القياس، وذلك من اجل الوصول الى أدق التفسيرات للأداء الحقيقي للمفحوص على أداة القياس دون أخطاء تقود الى قرارات غير دقيقة، ونظرا لكثرة استخدام الاختبارات الموضوعية في الكثير من اختبارات القدرات والتحصيل، فإن استجابات متلقين هذه النوعية من الاختبارات قد لا تشير إلى قدرة المفحوص الحقيقية بسبب عدة عوامل من بينها الغش، والتخمين (Armstrong & Belov, 2010, 36) .

هذا وتعتمد إجراءات القياس النفسي في بناء وإعداد أدوات القياس على مدخلين مختلفين وهما : المدخل الكلاسيكي المعتمد على نظرية القياس التقليدية (الكلاسيكية) والمدخل الحديث الذي يعتمد على نظرية الاستجابة للمفردة (IRT) . وهيمنت نظرية القياس الكلاسيكية على إجراءات بناء وإعداد أدوات القياس لفترة طويلة من الزمان ، وقد اعتمدت هذه النظرية على نموذج للقياس ينص على ان درجة الفرد الملاحظة بالاختبار تساوي مجموع الدرجة الحقيقية له ودرجة الخطأ بالاختبار . وبالتالي فإن إجراء القياس في النظرية الكلاسيكية. يتأثر بخصائص عينة المتلقين للاختبار كما يتأثر كذلك بصعوبة المفردات المتضمنة بالاختبار (Stone&Yumoto,2004,41). أما وعلى نحو آخر يتضح أن نظرية الاستجابة للمفردة تتجه إلى نمذجة البيانات من خلال نمذجة العلاقة بين المتغير الكامن وهو قدرة الفرد، واحتمال حدوث الاستجابة الصواب على المفردة وهو المتغير الذي يخضع للملاحظة .

وتوفر النظرية الحديثة للقياس إمكانية التنبؤ بأداء الفرد وتقدير قدرته الكامنة خلف هذا الاداء بصرف النظر عن الاداة المستخدمة فبي إجراء القياس كما يسفر استخدامها عن معالم إحصائية ثابتة ومستقرة لأداة القياس ومفرداته لا تختلف أو تتغير باختلاف وتغير عينة الافراد، كما تتميز هذه النظرية بقياس اكثر دقة بأخطاء اقل للقياس(Nitko,2001) ,وتبعا لهذه النظرية يمكن تفسير درجة الفرد على الاختبار من خلال محكات بشكل مسبق ولا تفسر بمتوسط العينة التي ينتمي اليها الفرد،(Wells, S. , Subcoviak, J. &Serlin, C. 2021)ولذلك اعتبرت نظرية الاستجابة للمفردة حلا مناسبا للكثير من المشكلات المتعلقة باستخدام النظرية التقليدية في القياس ، وتتسم ادوات للقياس التي يتم إعدادها في ضوء هذه النظرية بالموضوعية المعهودة من الادوات المستخدمة في القياس الفيزيقي . (Embretson & Reise, 2000, 85)

وتعد نظرية الاستجابة للمفردة (IRT) أتجاه حديث في القياس النفسي وتستهدف تحقيق الموضوعية وتعتبر من النظريات السيكو مترية المعاصرة التي اتضحت فائدتها في التغلب على كثير من مشكلات القياس التي فشلت النظرية التقليدية للقياس في مواجهتها (القيسي ،٢٠١٣)، فهذه النظرية تقدم نموذج للعلاقة القائمة بين مستوى الصفة المقاسة من اختبار محدد، لد ى الفرد واستجابته لمفردة من مفردات الاختبار . وحيث أن السمة كامنة بطبيعتها وتؤثر في المتغيرات المشاهدة، فإن تقديرها يكون من خلال هذه المتغيرات الملاحظة . وهذه النظرية ترتبط نماذج رياضية تعتمد على افتراضات قوية لها علاقة بما يحدث عند استجابة الفرد لمفردة اختبارية بيانها فيما يلي (صلاح الدين علام، ٢٠٠٥، ٥٣)

١- احادية البعد Uni dimensionality

ويعني هذا الافتراض أن هناك عا مل واحد فقط يكمن خلف الأداء على المقياس، وهو القدرة او السمة المقاسة، وأحادية البعد تتحقق في مقياس (راشد الدوسري ،٢٠٠٤، ٤٥)

٢- الاستقلال المحلي Local Independence

ويقصد بالاستقلال المحلي أن احتمال حدوث الإجابة الصواب للفرد على مفردة اختباريه يكون مستقلا عن استجاباته عن أي مفردة أخرى بنفس الاختبار، بمعنى عدم تأثر الإجابة عن أي

مفـــردة اختباريـــة بالإجابـــة عـــن أي مفـــردة أخـــرى بالاختبـــار بالســـلب او بالإيجـــاب (Reeve,2004,12) .

٣- المنحنى المميز للمفردة Item Characteristic Curve

ويسمى أيضاً بمنحنى خصائص المفردة، ويمثل هذا المنحنى بقدرة المفحوص على المحور الأفقي واحتمـال حـدوث الإجابـة الصـواب علـى المحـور الرأسـي، ويعتمـد احتمـال الاجابـة الصـواب للمفحوص على مفردة ما على شكل ذلك المنحني، وهذه الإجابة الصواب تعتمد على الصعوبة فقط (Lord١٩٨٠,٣٥)

٤- افتراض السرعة assumption of speediness

ويشير هذا الافتراض إلى ان إجابة المفحوص على أي مفردة اختبارية يجب ان تعتمد على مقدار القدرة التي المقاسة بالاختبار، وليس على أن وقت الاختبار لم يسمح له بالوصول إلى هذه المفردة (Hambelton & Swamina, 1985, 30).

نماذج الاستجابة للمفردة

١ - النماذج احادية البعد Unidimensional Models :

وهذه النماذج تفيد بأن احادية البعد تعني ان هنالك سمة واحدة متصلة تكمن خلف استجابات الافراد على مفردات المقياس، وان الفروق بين الافراد في هذه السمة يعود إليها كل التباينات بين استجابات الافراد على المفردات الاختبارية.

٢ -النماذج متعددة الابعاد Multidimensional Models

وتفترض هذه النماذج ان بوجود أكثر من بعد خلف استجابات الافراد على المفردات الاختبارية، وتشكل هذه النماذج امتدادا لنماذج الاستجابة للمفردة الاحادية البعد، ,تتوافق مع الاختبارات التي تشمل مفردات تقيس سمات مركبة او سمة تتكون من عدة لأبعاد.

ونظرا لأن النموذج المستخدم في الدراسة هو النموذج اللوجيستي الثلاثي المعلم هو أحد النماذج التي تنتمي إلى طائفة النماذج أحادية البعد فسوف يتم الاكتفاء بتناول هذا النموذج بالفقرة القادمة

النموذج اللوجستي ثلاثي المعلم Three Parameter Logistic Model

يسمى هذا النموذج بنموذج لوردLord Model وتستهدف التحليلات الخاصة به تدريج المفردات الاختبارية التي تمتلك مستويات مختلفة من الصعوبة والتمييز، وهذه المفردات يكون احتمال الإجابة الصواب عنها من الأفراد ذوي القدرات المتدنية جدا يكون عن طريق التخمين ، وبالتالي يمكن التعامل مع مشكلة التخمين في الإجابة عن الفقرات الموضوعية مثل الصواب والخطأ والاختيار من متعدد. حيث يتضمن هذا النموذج متغيرا ثالثا ويسمى بالخط المقارب السفلي او معلم التخمين . ويقوم هذا النموذج على افتراض أن المفردات تختلف من حيث مستويات الصعوبة والتمييز، كما أن مستويات التخمين تختلف من مفردة لأخرى داخل الاختبار، ولهذا النموذج الصيغة الرياضية التالية

(Emberson, & Reise, ,2015) (Magno, 2019) :

$$P(\theta) = c + \frac{1-c}{1+e^{-1.7a(\theta-b)}}$$

معادلة حساب احتمالية الاستجابة في النموذج اللوجيستي ثلاثي البارامتر

حيث ترمز a وb وc للبارامترات المميزة للمفردة كما يلي : يشير البارامتر (a) إلى معلم التمييز للمفردة discriminating power أما (b) فهي تشير إلى معلم صعوبة المفردة item difficulty ويشير البارامتر c إلى الخط المقارب السفلي lower

asymptote للمنحنى المميز للمفردة وهو معلم التخمين guessing parameter وقد أطلق عليه (Lord ,1980: 12-13) مسمى مستوى شبه الصدفة Pseudo-chance level parameter وهذا البارامتر يشير إلى احتمالية أن الفرد ذي القدرة المنخفضة أن يجيب صواباً عن المفردة، وعندما لا تحدث الإجابة الصواب عن المفردة فإن قيمة معلم التخمين تكون c = صفر. (Shih , Kuo & Hsin ,2010) .

هذا وقد أجريت بعض الدراسات التي تناولت التحقق من دقة التقدير باستخدام النموذج اللوجيستي الثلاثي المعلم ومن بين هذه الدراسات دراسة كلير (٢٠٢٢) Keller, التي استهدفت الكشف عن دقة تقدير معالم المفردة والقدرة باستخدام اسلوب المحاكاة، وقد تم توليد البيانات عن طريق ثلاث اختبارات افتراضية تتضمن ١٥، ٢٥ ، ٤٠. مفردة اختبارية وعينات ذات أحجام مختلفة تتألف من ١٠٠، ٢٠٠، ٥٠٠، ١٠٠٠ من البيانات المحاكاة وقد استخدمت الدراسة برنامج MG-BILOG في التحليلات الخاصة بالنموذج المستخدم ، وأسفرت النتائج عن تأثر دقة تقدير معلم الصعوبة والتمييز مع اختلاف أحجام العينات فقد ازدادت دقة التقدير بازدياد حجم العينة، وعلى جانب آخر لم تتأثر دقة تقدير التخمين بحجم العينة .كما استهدفت دراسة دي لا توري ويوان (Yuan, & Torre La De(2020 لتعرف على اثر حجم العينة على تقديرات معالم القدرة و المفردة في اختبارات تم اعدادها في ضوء نماذج نظرية الاستجابة للمفردة .(IRT) واستخدمت الدراسة أسلوب مونت كارلو للمحاكاة في توليد البيانات، من اجل التعرف على أثر حجم العينة في دقة تقدير معلم القدرة المفردة، وقد أسفرت النتائج بأن حجم العينة يؤثر على قدرة الاختبار في تقدير معالم القدرة و المفردة، .وفي دراسة أخرى اجراها بني عطا (٢٠١٤) والتي استهدفت الكشف عن دقة تقدير النموذج اللوجستي الثلاثي المعلم لمعالم المفردة والقدرة، من خلال تغير طول الاختبار وحجم العينة .وقد اعتمدت الدراسة على توليد بيانات محاكاة ثنائية الاستجابة لستة من أطوال الاختبار،١٠ ،٢٥ ،٥٠ ،٧٥ ،١٠٠ مفردة وستة أحجام مختلفة من العينة ،١٠٠ ،٢٥٠ ،٥٠٠ ،١٠٠٠ من خلال استخدام برنامج توليد البيانات WINGEN وباستخدام برنامج Mg -Bilog3 . وأظهرت نتائج الدراسة عن تأثير دال إحصائياً لكل من متغير طول الاختبار ومتغير حجم العينة والتفاعل بينهما في دقة تقديرات معالم المفردات والأفراد.

وفي دراسة الزبون (٢٠١٣) التي استهدفت الكشف عن أثر حجم العينة على تقدير دالة المعلومات وكذلك الخطأ المعياري للتقدير. ولتحقيق أهداف الدراسة تم الاعتماد على استجابات ٧٥٠٠ من طلبة الصف الثامن الأساسي موزعين عشوائيا على خمسة مجموعات وهي: ٢٥٠٠،٢٠٠٠،١٥٠٠،١٠٠٠،٥٠٠ وقد استجابوا على الاختبار الوطني لضبط نوعية التعلم مقرر الرياضيات المتضمن ل ٤٠ مفردة موضوعية من نوع الاختيار من متعدد، وقد استخدم برنامج Mg-Bilog3 لتقدير معالم المفردات والخطأ المعياري لدالة المعلومات وفقا النموذج اللوجيستي الثلاثي المعلم، وأشارت النتائج إلى ان تقديرات دالة المعلومات تتغير بتغير حجم العينة، حيث تزداد تبعا لزيادة حجم العينة، كما أظهرت النتائج أيضا ان الخطأ المعياري في تقدير دالة المعلومات تتغير كذلك بتغير حجم العينة، اذ يتناقص هذا الخطأ مع ازدياد حجم العينة

وفي دراسة شما (٢٠١٣) التي استهدفت التعرف على تأثير حجم العينة على دقة تقدير صعوبة المفردات وقدرة الافراد باستخدام نموذج راش ولتحقيق الهدف من الدراسة فقد تم تطبيق اختبار أوتيس- لينون للقدرة العقلية على ثلاث عينات مختلفة من حيث الحجم تتكون من (٢٠٠ ،٨٠٠ ، ١٥٠٠) من طلبة المرحلة الاساسية وقد تم استخدام برنامج mg-Bilog3 لتقدير صعوبات مفردات الاختبار وكذلك تقدير قدرات أفراد العينات مختلفة الحجم، وأظهرت نتائج الدراسة إلى أن دقة تقديرات صعوبة المفردات تزداد مع زيادة حجم العينة كما أن اختلاف حجم العينة ليس له تأثير على دقة تقدير قدرات الافراد .

كما استهدفت دراسة الشريفين (٢٠١٢) كشف تأثير تغير حجم العينة على تقدير معالم المفردات وقدرت الافراد وكذلك والخصائص السيكومترية للاختبار، وتم بناء اختبار تحصيلي في الفيزياء مكون من ٣٣ مفردة من نوع الاختيار من متعدد، وأجرت الدراسة تطبيق

الاختبار على عينة مكونة من ١٠٠٠ طالب وطالبة من طلبة الصف الثاني ثانوي علمي ، وتم استخدام التحليلات الخاصة بنتائج التطبيق اعتمادا على النموذج اللوجيستي الثلاثي المعلم باستخدام برنامج .MG Bilog3 وأظهرت النتائج عدم وجود فروق ذات دلالة إحصائية تعزى لمتغير حجم العينة بين متوسطات الأخطاء المعيارية لتقديرات القدرة للأفراد وكذلك تقديرات معالم المفردات ، كما أظهرت النتائج عدم وجود فروق دالة بين معاملات الثبات المقدرة تبعا لنظرية الاستجابة للمفردة مع أحجام العينة المختلفة (١٠٠، ٥٠٠ ، ١٠٠٠) فرد، وأظهرت النتائج الى ان دقة تقديرات معالم القدرة تزداد مع عينة الافراد من ذوي القدرة المنخفضة عند استخدام طريقة بيز، وتزداد الدقة مع الافراد من ذوي القدرة المتوسطة باستخدام طريقة الارجحية العظمى بغض النظر عن اختلاف حجم العينة.

وقد تبين للباحث من خلال تفقد الدراسات السابقة أن استخدام النموذج الثلاثي المعلم في التحقق من تأثير حجم العينة على معلمي الصعوبة والقدرة كان من خلال دراسات المحاكاة في توليد أحجام مختلفة من العينات ، عدا دراسة الشريفين (٢٠١٢) التي استمدت البيانات الخاصة بها من استخدام اختبار تحصيلي في الفيزياء ، على عينات مختلفة لم تزيد حجمها عن (١٠٠٠) فرد، واختلفت نتائجها مع نتائج جميع الدراسات السابقة التي كان لها نفس الأهداف وهذا مما حدا بالباحث بإجراء هذه الدراسة للتحقق من حقيقة هذا الاختلاف باستخدام بيانات حقيقة ناتجة عن تطبيق اختبار في مادة الجبر للصف الثاني الثانوي ، وعينات تختلف من حيث الحجم حيث يبلغ حجم العينة الكبر (١٨٠٠) فرد مما يعطي نتائج أوضح عن تأثير حجم العينة على كل من تقديرات الصعوبة لفقرات الاختبار وتقديرات القدرة للعينات موضع المعالجة.

مشكلة الدراسة

حيث ان النماذج اللوجيستية المرتبطة بالنظرية الحديثة للقياس تعتمد في تقديرها الموضوعي للمتغيرات على مطابقة بيانات الاختبار للنموذج المستخدم في التحليلات، لذا يحاول الباحث من خلال دراسته الحالية الكشف عن أثر حجم العينة بوصفه عامل مؤثر على المطابقة، وبناء على ذلك قد يؤثر بدوره على دقة تقدير صعوبات مفردات الاختبار وتقدير قدرات الافراد، لذلك اهتمت هذه الدراسة بالكشف عن تأثير اختلاف حجم العينة على دقة تقدير صعوبات المفردات الاختبارية الخاصة بالأداة المستخدمة بالدراسة وكذلك تقديرات القدرات أفراد العينة البحثية المتلقين لهذه الأداة باستخدام النموذج اللوجيستي الثلاثي المعلم. وتتلخص مشكلة الدراسة في الإجابة عن الأسئلة التالية:

١. ما تأثير حجم العينة على دقة تقدير صعوبات المفردات الاختبارية لأداة الدراسة باستخدام التحليلات الخاصة بالنموذج اللوجيستي الثلاثي المعلم
٢. ما تأثير حجم العينة على دقة تقدير قدرات أفراد العينة البحثية باستخدام تحليلات النموذج اللوجيستي الثلاثي المعلم .

أهداف الدراسة

استهدفت هذه الدراسة الكشف عن تأثير حجم العينة البحثية على دقة تقدير كل من صعوبات المفردات الاختبارية المتضمنة بالأداة المستخدمة بالدراسة وكذلك تأثيرها على دقة تقدير كل من قدرات أفراد العينة البحثية المتلقين لهذه الأداة باستخدام تحليلات النموذج اللوجيستي ثلاثي المعلمة

أهمية الدراسة

تبدو اهمية الدراسة في النقاط التالية :

١- محاولة الاسهام في القاء الضوء على مدى أثر حجم العينة على تقديرات صعوبة المفردات وتقديرات القدرة لأفراد العينة .بما يثري المكتبة العربية في هذا المجال ويكون له انعكاساته على الأدوات المستخدمة بإجراءات القياس النفسي والسلوكي.

٢- قد تسفر الدراسة عن معلومات تساعد المهتمين في مجال القيا عن تطوير الاختبارات باستخدام الحجم المناسب للعينة .

٣- قد تسفر نتائج الدراسة عن تزويد الباحثين المستخدمين لنظرية الاستجابة للمفردة بالمعلومات التي قد تفيد باختيار الحجم المناسب للعينة للحصول على معالم للمفردات والأفراد أكثر صدقا.

حدود الدراسة

١- **الحدود الزمنية:** اقتصرت الدراسة في تطبيق أداتها على الفصل الأول من العام الدراسي ٢٠٢٣ / ٢٠٢٤م.

٢- **الحدود المكانية:** أجريت الدراسة على المدارس الثانوية بمحافظة جنوب القاهرة التعليمية.

٣- **الحدود البشرية:** اقتصرت الدراسة على طلبة الصف الثاني الثانوي من التعليم العام.

٤- **الحدود الموضوعية:** اقتصرت الدراسة في معالجة المتغيرات الخاصة بها على متغيري :تقديرات الصعوبة للمفردات الاختبارية المتضمنة بأداة الدراسة وتقديرات قدرات الأفراد المتلقين لنفس الأداة باستخدام التحليلات الخاصة بالنموذج اللوجيستي الثلاثي المعلم.

مصطلحات الدراسة

حجم العينة : يقدر حجم العينة بعدد الأفراد الذين سيجرى عليهم تطبيق أداة الدراسة، ويحدد إجرائياً باستخدام ثلاث عينات بأحجام مختلفة (٣٠٠ ، ٦٠٠ ، ٩٠٠) فرد

(أ) بارامتر التمييز: Discrimination parameters

وهو مؤشر للمفردة الاختبارية الذي يشير إلى قدرتها على التمييز أو التفريق بين المستويات المتفاوتة من القدرة المستهدفة من القياس , (Azevedo,C.L.N., 2009) (Urry ,1977) (Wang &Vispoel ,1998: 116). ويقدر إجرائياً من خلال تحليلات النموذج اللوجيستي الثلاثي المعلم المستخدم بالدراسة الحالية

(ب) معلم الصعوبة: Difficulty parameters

وهو مؤشر للمفردة الاختبارية يشير إلى مواقع المفردة على متصل الصفة المقاسة من الاختبار ويفيد بامتلاكها لمستويات محددة من الصعوبة (De Ayala ,1992:516)، (Ponsoda& et al. 1997, 213)، (Jason, Edward& Walter , 2009) (McBride &Martine, 1983)، (De Ayala & et al. ,1990) ومن المفترض أن تتوزع قيم بارامترات صعوبة المفردات بصورة واسعة على تدريج الصعوبة وتقدر إجرائياً من خلال تحليلات النموذج اللوجيستي الثلاثي المعلم المستخدم بالدراسة الحالية

(ج) معلم التخمين: Guessing parameters

وهو مؤشر المفردة الاختبارية الذي يشير إلى توصل الممتحن إلى الإجابة الصواب عن المفردة التي لم تسمح قدرته بالإجابة عنها عن طريق التخمين (Azevedo,C.L.N., 2009) (Urry ,1977) (Ponsoda& et al. 1997:213).

إجراءات الدراسة

مجتمع وعينة الدراسة :

تكون مجتمع الدراسة من جميع طلبة الصف الثاني الإعدادي بالمدارس الحكومية في محافظة جنوب القاهرة للفصل الأول من لعام الدراسي ٢٠٢٣ / ٢٠٢٤م. وتكونت عينة الدراسة من (١٨٠٠) طالبا من مدارس المحافظة .

أداة الدراسة (اختبار الجبر للصف الثاني الإعدادي للفصل الدراسي الأول)

تم اعداد هذا الاختبار من جانب الباحث، ويتكون الاختبار في صورته النهائية من عدد (٤٨) مفردة من نوع الاختيار من متعدد وتكون كل مفردة اختبارية من أربعة بدائل و تستهدف قياس تحصيل الطلبة في مقرر الرياضيات للصف الثاني الثانوي ، بهدف الكشف عن تأثير حجم العينة على تقديرات صعوبات المفردات وقدرات الفرد، وقد قام الباحث بتحليل محتوى مقرر الرياضيات الفصل الأول للصف الثاني الثانوي ، وصياغة الأهداف المعرفية للوحدات الدراسية التي يتكون منها الكتاب المقرر وقد قام الباحث بعد هذه المرحلة ببناء جدول المواصفات لمحتوى الاختبار : في ضوء تحليل المنهج اعتمادا على المحتوى والأهداف المعرفية، وتصنيف هذه الأهداف المعرفية وفقا لتصنيف بلوم للأهداف المعرفية إلى مستويات الدنيا للأهداف(المعرفة، الاستيعاب، التطبيق) والمستويات العليا (التحليل، التركيب / الابتكار، والتقويم)، ثم تم تحديد الأوزان النسبية للوحدات الدراسية ومستويات الأهداف التي تضمنها لاختبار بناءً على حجم المادة الدراسية، وعدد الأهد اف، والزمن الذي يستغرقه تدريس الوحدة الدراسية، وفي النهاية تم عرض نتائج التحليل على مشرفي الرياضيات للتحقق من مدى ملاءمة الاختبار للمنهج الدراسي

التحقق من الخصائص السيكومترية للاختبار

صدق الاختبار: تم استخدام نوعين من الصدق وهما :

صدق المحتوى : تم عرض الاختبار على المحكمين، للكشف عن مدى ملائمة نتائج تحليل المحتوى والاهداف لجدول مواصفات الاختبار وكذلك مدى سلامة اللغة وسلامة صياغة المفردات، وتم تنفيذ توجيهاتهم بحذف ٥ مفردات.

.الصدق التلازمي للاختبار : تم تطبيق الاختبار على عينة استطلاعية تكونت من (٧٠) طالب، وحساب معامل الارتباط بين درجة الطالب الكلية في الاختبار ودرجته الحاصل عليها في مقرر الرياضيات بالصف الأول الثانوي. وبلغت قيمة معامل الصدق.٠,٩٢٢.

ثبات الاختبار : تم حساب معامل ثبات الاختبار باستخدام نوعين من الثبات:

تم حساب ثبات الاختبار بطريقة ألفا كرونباخ وقد بلغت قيمته (٠.٩٠) بعد حذف مفردة واحدة من المفردات الاختبارية ، ثم قام الباحث بحساب معامل الثبات بطريقة التجزئة النصفية وبلغت قيمته (٠.٨٩) وبالتالي بلغ عدد مفردات الاختبار بعد إجراءات التحقق من الصدق والثبات (٤٠) مفردة.

هذا ومن خلال إجراءات التحقق من الخصائص والشروط السيكومترية السابقة، فقد تم الحصول على معاملات للصدق والثبات تظهر أن الاختبار يتمتع خصائص سيكومترية مرتفعة يمكن من خلالها قبول النتائج التي تسفر عن تطبيقه بالواقع الميداني للدراسة

إجراءات الإجابة عن أسئلة الدراسة:

١- تطبيق الاختبار على العينة البحثية الكلية.

٢- سحب عينات عشوائية فرعية باستخدام برنامج SPSS v.27 من استجابات العينة الكلية المكونة من (١٨٠٠) فرد حجمها (٩٠٠ ، ٦٠٠ ، ٣٠٠) من العينة الكلية ، بهدف استخدامها في التحليلات الخاصة بإجراءات الدراسة.

٣- تحليل استجابات العينة الكلية من الطلاب لمفردات الاختبار باستخدام النموذج الثلاثي المعلم، وذلك باستخدام برنامج BILOG-MG3 ، ثم تكرار نفس التحليلات على العينات الفرعية الثلاثة التي حجمها (٩٠٠ ، ٦٠٠ ، ٣٠٠) طالب ، يليها تكرار نفس الخطوة على عينة حجمها (٣٠٠) طالب .

٤- إيجاد تقديرات الصعوبة الخاصة بمفردات الاختبار والمشتقة من التحليلات الخاصة بالعينات : العينة الكلية المكونة من (١٨٠٠) فرد الثلاثة مختلفة الحجم (٩٠٠ ، ٦٠٠ ، ٣٠٠) على حدة.

٥- إيجاد تقديرات القدرة الخاصة بقدرات الأفراد والمشتقة من التحليلات الخاصة بالعينات العينة الكلية والعينات الفرعية مختلفة الحجم كل على حدة .

أولاً : لإجراءات الإجابة عن السؤال الأول:

١- إيجاد القيم المطلقة للفروق بين تقديرات الصعوبة المشتقة من التدريج المرجعي الناتج من تحليل استجابات العينة الكلية المكونة من (١٨٠٠) فرد وكل من التدريج المشتق من تحليل استجابات العينات الفرعية والتي حجمها (٩٠٠) ، (٦٠٠) ، (٣٠٠) فرد

٢- إيجاد مجموع الأخطاء المعيارية لصعوبات مفردات التدريج المرجعي المشتقة من التدريجين المرجعي المكون من العينة الكلية والتدريج الناتج عن تحليل عينة حجمها (٩٠٠)فرد.

٣- تكرار الخطوة السابقة بالخطوة (رقم ٢) ولكن مع التدريجين المشتقين من العينتين (٦٠٠ ، ٣٠٠) فرد

٤- حساب دلالة الفروق بين تقديرات صعوبة التدريج المرجعي والتدريج الناتج عن العينات الفرعية (٩٠٠ ، ٦٠٠ ، ٣٠٠) فرد، وذلك بمقارنة القيم المطلقة للفروق (الناتجة عن الخطوة ١) بمجموع الأخطاء المعيارية الناتجة من الخطوة (رقم ٣) وفقا للقواعد التالية:

- إذا كانت الفروق في تقديرات الصعوبة للمفردات أكبر من مجموع الأخطاء المعيارية لها اعتبرت هذه الفروق دالة من الوجهة الإحصائية، أما إذا كانت الفروق في تقديرات الصعوبة للمفردات أقل من أو تساوي مجموع الأخطاء المعيارية لها اعتبرت هذه الفروق غير دالة من الوجهة الإحصائية، وإذا بلغ عدد الفروق الدالة إحصائيا بين تقديرات الصعوبة المتناظرة للمفردات بين التدريجين المرجعي والمشترك المعدل (وفقا للقاعدة السابقة) نسبة (٠.٠٥) أو أقل من عدد الفروق دل ذلك على دقة التدريج المشترك المعدل (بنك الأسئلة)، أما إذا بلغ عدد الفروق الدالة إحصائيا بين تقديرات الصعوبة المتناظرة للمفردات بين التدريجين المرجعي والمشترك المعدل أكبر من النسبة (٠.٠٥) من عدد الفروق دل ذلك على عدم دقة التدريج المشترك المعدل (بنك الأسئلة) (Lidwine B.,2006)(Aimee ,M, 2008.)

إجراءات الإجابة عن السؤال الثاني

١- نظراً لضخامة حجم العينة البحثية الكلية والتي تكونت من (١٨٠٠) فرد وصعوبة استخدام نفس الإجراءات المتبعة في الاجابة عن السؤال الأول مما يسفر عن استخراج نتائج ضخمة من التحليلات التي لا يتسع لها البحث الحالي فقد تم استخدام التحليلات الخاصة بأسلوب يتسم بإيجاز النتائج وهو اسلوب تحليل تباين القياس المتكرر (repeated measure) للقياسات الأربعة الخاصة بتقديرات القدرة الناتجة عن تحليل استجابات العينات الأربعة المختلفة من حيث الحجم.

٢- استخدام اختبار "بنفروني" لإجراء المقارنات الثنائية بين متوسطات تقديرات الصعوبة وكذلك متوسطات تقديرات القدرة المشتقة من التحليلات الخاصة بالعينات الثلاثة موضع المقارنة.

منهج الدراسة و الأساليب الإحصائية:

استخدمت الدراسة المنهج الوصفي (التحليلي- المقارن) في الإجراءات الخاصة بها مع استخدام الأساليب الإحصائية التالية:

- التحليل العاملي الاستكشافي للتحقق من مدى توافر افتراض أحادية البعد في البيانات الخاصة بالاختبار المستخدم بالدراسة.

- مؤشر الملاءمة الإحصائي مربع كاي المتضمن بتحليلات برنامج BILOG-MG3 للتحقق من مدى توافر افتراض استقلالية المحل لمفردات اختبار الدراسة
- أسلوب **تحليل تباين القياس المتكرر** Repeated Measures بتحليلات برنامج SPSS v.27 للتحقق من وجود فروق في تباين تقديرات الصعوبة للمفردات وكذلك تقديرات القدرة لأفراد العينات البحثية موضع المقارنة
- أسلوب فريدمان كبديل في حال عدم تحقق شروط استخدام تباين القياس المتكرر
- استخدام اختبار (LSD) لعقد المقارنات الثنائية Pairwise Comparisons نظرا لأن التحليلات تُجرى على مجموعة واحدة فقط من الأفراد .
- اختبار (**كلمجروف – سميرنوف**) المتضمن بتحليلات برنامج SPSS للتحقق من اعتدالية توزيع تقديرات متغيري (الصعوبة والقدرة) موضع المقارنة.

عرض ومناقشة نتائج الدراسة:

(أ) التحقق من افتراض احادية البعد لبيانات اختبار الدراسة

استخدمت الدراسة أسلوب التحليل العاملي الاستكشافي للبيانات الخاصة باستجابات العينة الكلية للتحقق من توافر هذا الافتراض وأسفرت التحليلات عن النتائج الموضحة بالجدول التالي:

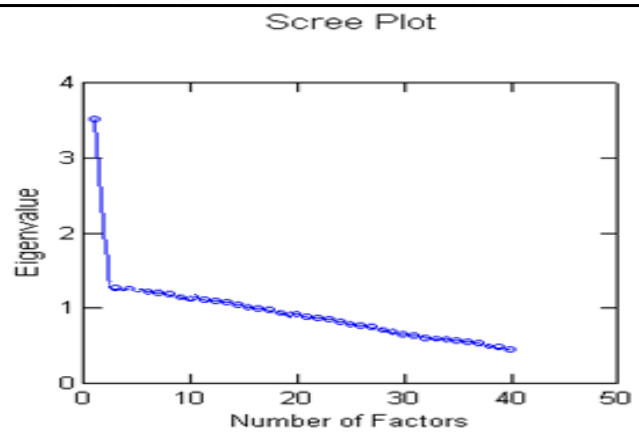
جدول (١)

الجذور الكامنة للثلاثة عوامل الأولى الناتجة عن التحليل العاملي لمفردات اختبار الدراسة

العوامل	الجذور الكامنة	
	الكلي	النسبة المئوية للتباين
1	3.722	25.254
2	1.712	7.301
3	1.631	5.122

يتضح من جدول التباين الكلي المفسر(السابق) أن قيمة الجذر الكامن للعامل الأول بلغت (11.301) كما بلغت قيمة الجذر الكامن للعامل الثاني (4.012) ، وقد لوحظ أن نسبة التباين المفسر بالعامل الأول من النسبة الكلية للتباين المفسر بجملة العوامل التي استخرجت من التحليل بلغت (25) تقريبا كما نسبة قيمة الجذر الكامن للعامل الأول إلى قيمة الجذر الكامن للعامل الثاني بلغت (2.8) تقريبا، أي أن النسبة تزيد عن القيمة (2)، وهذه النتائج تفيد بتوافر افتراض أحادية البعد، حيث أشار "ريكاس" أن نسبة تباين العامل الأول عندما تزيد عن (20٪) وهو أكبر تباين (Recase,cited in Hattie,1985)، فإن ذلك يعتبر دليلاً على أحادية البعد، كما تحقق هذه النتائج مؤشر لورد(lord , 1980) الذي يشير إلى أن المفردات التي تحقق خاصية أحادية البعد إذا كانت نسبة الجذر الكامن للعامل الأول إلى نسبة الجذر الكامن للعامل الثاني يجب ان تكون أكبر من (2) مما يعني أن الاختبار متشبعا بعامل واحد مهيمن. (Gessaroli & De Champlain , 2005)

كما اتضح من التمثيل البياني لاختبار الفرز scree plot بالشكل التالي رقم (١) وجود انحدار شديد بين الجذر الكامن للعامل الأول والجذور الكامنة للعوامل المتبقية مما يشير إلى تمتع الاختبار بأحادية البعد



شكل (١) تمثيل بياني للجذور الكامنة الناتجة عن تحليل فقرات اختبار الدراسة

(ب) التحقق من افتراض استقلالية المحل لبيانات اختبار الدراسة

ويفيد هذا الافتراض بعدم اعتماد المفردات الاختبارية بعضها على البعض الآخر (بمعنى ان الإجابة على أي مفردة اختبارية لا ينبغي ان تتأثر بالإجابة على المفردات الأخرى)، وقد أفادت نتائج التحليلات الخاصة باختبار هذا الافتراض والمتضمنة ببرنامج BILOG-MG3 بعدم تجاوز أي مفردة اختبارية لحدود مؤشر حسن المطابقة الإحصائية (Chi-square) كما أن تحقق افتراض أحادية البعد يعني تحقق افتراض استقلالية المحل(Hambleton,Swaminathan,1985)

(ج) التحقق من توافر افتراض السرعة

ويعني هذا الافتراض أن الأسئلة التي تركت بنهاية الاختبار ينبغي ان تكون بسبب أن قدرة المفحوص لم تمكنه من الإجابة الصحيحة عنها وليس بسبب عدم وجود وقت كاف لمحاولة المتحن أن يجيب عنها، وقد روعي ذلك في تطبيق اختبار الدراسة حيث تم السماح بإعطاء وقتا كافي للإجابة عن جميع فقرات الاختبار ، ولم يتضح للباحث شكاوى من الطلاب عن عدم كفاية الوقت المخصص للإجابة عن فقرات الاختبار ،وبالتالي فلم يكن لعامل السرعة أي أثر في الاستجابة على فقرات الاختبار بما يفيد بتحقق افتراض السرعة.

(د) التحقق من مطابقة بيانات الاختبار للنموذج (Goodness-of-Fit-Test)

لقد أسفرت نتائج المطابقة للأفراد والمفردات عن النتائج التالية:

- مطابقة المفردات Item Fit لتوقعات النموذج الثلاثي المعلم ، استناداً إلى دليل برنامج BILOG_MG3 ، (Zimowski, et.al.,2007) وذلك بعد حذف المفردات غير المطابقة ثم إعادة التحليل ، فالمفردة غير المطابقة لنموذج التحليل يكون قيمة مؤشر المطابقة الداخلي لها (Chi_Squre) دال عند مستوى (0.05)، وقد تبين من نتائج التحليل عدم وجود أية مفردة تخطت حدود المطابقة للنموذج المستخدم ، وبالتالي فقد أسفرت التحليلات عن نتائج نهائية قد تم الاعتماد عليها في المعالجات الإحصائية الخاصة بمتغيرات الدراسة.
- مطابقة الأفراد Person Fit لتوقعات لنموذج التحليل، وقد اتضح أن إجمالي عدد الطلاب وعددهم (900) طالب جاء مؤشر المطابقة لقدراتهم داخل حدود المطابقة المسموح بها .

ثانيا : نتائج اختبار اعتدالية توزيعات تقديرات جميع متغيرات الدراسة:

لقد أسفرت التحليلات الخاصة باختبار" **كلموجروف - سميرنوف**" للتحقق من مدى اعتدالية توزيع التقديرات الخاصة بكل من : تقديرات صعوبة المفردات الاختبارية ، وتقديرات قدرات أفراد العينة البحثية ، والتي أجري التوصل إليها من التحليلات الخاصة بالنموذج الثلاثي المعلم - عن النتائج التالية:

١- نتائج التحقق من فحص اعتدالية توزيع صعوبات المفردات الاختبارية:

جدول (٢): قيمة ودلالة اختبار كلمجروف - سميرنوف لتوزيع تقديرات الصعوبة لمفردات اختبار الدراسة

قيمة (z) لاختبار كلمجروف-سميرنوف	0.823
مستوى الدلالة	0.23

يتضح من الجدول السابق أن قيمة اختبار (كلموجروف - سميرنوف) لتوزيع تقديرات صعوبة مفردات اختبار الدراسة ، قد بلغت (0.823) وهي قيمة غير دالة عند مستوى (0.05) مما يفيد باعتدالية توزيع تقديرات الصعوبة للاختبار.

٢- نتائج التحقق من فحص اعتدالية توزيع تقديرات القدرة لأفراد العينة البحثية

جدول (٣) : قيمة ودلالة اختبار كلمجروف - سميرنوف لتوزيع تقديرات القدرة لأفراد العينة البحثية

قيمة (z) لاختبار كلمجروف-سميرنوف	0.970
مستوى الدلالة	0.522

يتضح من الجدول السابق أن قيم اختبار (كلموجروف - سميرنوف) لتقديرات القدرة لأفراد العينة البحثية بلغت (0.970) وهي غير دالة عند مستوى (0.05) مما يعني أن توزيع تقديرات القدرة جاء بشكل اعتدالي.

ويتضح للباحث من النتائج السابقة أن جميع التوزيعات الخاصة بتقديرات الصعوبة لمفردات اختبار الدراسة وكذلك تقديرات القدرة لأفراد العينة البحثية جاءت على نحو اعتدالي، مما يتيح عقد المقارنات في ضوء جميع المتغيرات موضوع الدراسة باستخدام أسلوب تحليل تباين القياس المتكرر.

يستعرض الباحث في هذا الجزء نتائج الدراسة وتفسيرها من خلال عرض نتائج الإجابة عن الأسئلة البحثية في ضوء الاطلاع على نتائج الدراسات السابقة.

عرض ومناقشة نتائج الدراسة :

أولاً: عرض نتائج تحليل مفردات اختبار الدراسة باستخدام النموذج الثلاثي المعلم

لقد أسفرت نتائج تحليل البيانات الاستجابات الخاصة بمفردات اختبار الدراسة للحصول على متغيري الدراسة (تدريج الصعوبة للمفردات الاختبارية و تقديرات القدرة للعينات مختلفة الحجم) باستخدام نموذج التحليل ثلاثي المعلم عن النتائج الموضحة بالجدول التالي:

جدول (٤): تقديرات الصعوبة لمفردات الاختبار والأخطاء المعيارية لها الناتجة عن تحليلات استجابات العينة الكلية والعينات الفرعية ذات الأحجام المختلفة

المفردة	التحليلات الناتجة عن عينة حجمها (١٨٠٠) فرد		التحليلات الناتجة عن عينة حجمها (٩٠٠) فرد		التحليلات الناتجة عن عينة حجمها (٦٠٠) فرد		التحليلات الناتجة عن عينة حجمها (٣٠٠) فرد	
	مستوى الصعوبة باللوجيت	الخطأ المعياري	مستوى الصعوبة باللوجيت	الخطأ المعياري	مستوى الصعوبة باللوجيت	الخطأ المعياري	مستوى الصعوبة باللوجيت	الخطأ المعياري
ITEM1	1.059	0.147	1.041	0.147	1.052	0.147	1.925	0.147
ITEM2	1.224	0.07	1.136	0.07	1.147	0.07	1.258	0.07
ITEM3	0.76	0.155	0.462	0.155	0.473	0.155	0.39	0.155
ITEM4	-0.14	0.243	-0.117	0.243	-0.106	0.243	-0.258	0.243
ITEM5	-0.266	0.172	-0.264	0.172	-0.253	0.172	-0.39	0.172
ITEM6	0.982	0.049	0.984	0.049	0.995	0.049	1.988	0.049
ITEM7	0.98	0.106	0.982	0.106	0.993	0.106	1.92	0.106
ITEM8	0.603	0.072	0.605	0.072	0.616	0.072	0.577	0.072
ITEM9	1.157	0.07	1.169	0.07	1.18	0.07	2.137	0.07
ITEM10	0.552	0.118	0.565	0.118	0.576	0.118	0.493	0.118
ITEM11	1.241	0.062	1.223	0.062	1.234	0.062	1.18	0.062
ITEM12	1.41	0.073	1.394	0.073	1.405	0.073	1.322	0.073
ITEM13	1.369	0.138	1.471	0.138	1.482	0.138	1.499	0.138
ITEM14	0.255	0.253	0.264	0.253	0.275	0.253	0.192	0.253
ITEM15	0.531	0.191	0.532	0.191	0.543	0.191	0.46	0.191
ITEM16	1.454	0.076	1.476	0.076	1.487	0.076	1.123	0.076
ITEM17	6.887	2.978	5.749	2.978	5.76	2.978	3.377	2.978
ITEM18	0.42	0.158	0.437	0.158	0.448	0.158	0.365	0.158
ITEM19	1.5976	0.118	1.499	0.118	1.51	0.118	1.427	0.118
ITEM20	1.125	0.129	1.027	0.129	1.038	0.129	1.388	0.129
ITEM21	1.059	0.147	1.041	0.150	1.052	0.176	1.925	0.201
ITEM22	1.224	0.07	1.136	0.092	1.147	0.261	1.258	0.115
ITEM23	0.76	0.155	0.462	0.177	0.473	0.159	0.39	0.159
ITEM24	-0.14	0.243	-0.117	0.248	-0.106	0.249	-0.258	0.305
ITEM25	-0.266	0.172	-0.264	0.18	-0.253	0.177	-0.39	0.199
ITEM26	0.982	0.049	0.984	0.065	0.995	0.081	1.988	0.109
ITEM27	0.98	0.106	0.982	0.115	0.993	0.117	1.92	0.126
ITEM28	0.603	0.072	0.605	0.099	0.616	0.088	0.577	0.192

المفردة	التحليلات الناتجة عن عينة حجمها (١٨٠) فرد		التحليلات الناتجة عن عينة حجمها (٩٠) فرد		التحليلات الناتجة عن عينة حجمها (٦٠) فرد		التحليلات الناتجة عن عينة حجمها (٣٠) فرد	
	مستوى الصعوبة باللوجيت	الخطأ المعياري	مستوى الصعوبة باللوجيت	الخطأ المعياري	مستوى الصعوبة باللوجيت	الخطأ المعياري	مستوى الصعوبة باللوجيت	الخطأ المعياري
ITEM29	1.157	0.07	1.169	0.12	1.18	0.115	2.137	0.136
ITEM30	0.552	0.118	0.565	0.125	0.576	0.121	0.493	0.195
ITEM31	1.241	0.062	1.223	0.170	1.234	0.091	1.18	0.109
ITEM32	1.41	0.073	1.394	0.089	1.405	0.091	1.322	0.102
ITEM33	1.369	0.138	1.471	0.153	1.482	0.144	1.499	0.251
ITEM34	0.255	0.253	0.264	0.291	0.275	0.277	0.192	0.344
ITEM35	0.531	0.191	0.532	0.201	0.543	0.198	0.46	0.211
ITEM36	1.454	0.076	1.476	0.088	1.487	0.085	1.123	0.099
ITEM37	6.887	2.978	5.749	2.701	5.76	2.991	3.377	3.554
ITEM38	0.42	0.158	0.437	0.177	0.448	0.174	0.365	0.198
ITEM39	1.5976	0.118	1.499	0.125	1.51	0.129	1.427	0.121
ITEM40	1.125	0.129	1.027	0.171	1.038	0.145	1.388	0.128
المتوسط	-	**0.189**	-	**0.1988**	-	**0.2133**	-	**0.2395**

ثانياً: عرض ومناقشة نتائج الإجابة عن السؤال الأول:

لقد أسفرت الإجراءات الخاصة بالإجابة عن السؤال الأول الذي ينص على: "ما تأثير حجم العينة على دقة تقدير صعوبات المفردات الاختبارية لأداة الدراسة باستخدام التحليلات الخاصة بالنموذج اللوجيستي الثلاثي المعلم" عن النتائج التالية:

جدول (٥): الفروق في تقديرات الصعوبة المتناظرة المشتقة من التدريج المرجعي والأخرى المشتقة تدريج العينات الفرعية ومجموع الأخطاء المعيارية لها المشتقة من التدريج المرجعي والأخرى المشتقة من العينات الفرعية

الفروق في تقديرات الصعوبة المتناظرة بين التدريج المرجعي والتدريج الناتج عن تحليل	مجموع الخطأين المعياريين الناتجين عن التدريج المرجعي والتدريج الناتج عن	الفروق في تقديرات الصعوبة المتناظرة بين التدريج المرجعي والتدريج الناتج عن تحليل العينة (٦٠٠) فرد	مجموع الخطأين المعياريين الناتجين عن التدريج المرجعي والتدريج الناتج عن	الفروق في تقديرات الصعوبة المتناظرة بين التدريج المرجعي والتدريج الناتج عن تحليل العينة (٣٠٠) فرد	مجموع الخطأين المعياريين الناتجين عن التدريج المرجعي والتدريج الناتج عن العينة (٣٠٠) فرد
0.018	0.297	0.007	0.323	0.866*	0.348
0.088	0.162	0.077	0.331	0.034	0.185
0.298	0.332	0.287	0.314	0.37	0.314
0.023	0.491	0.034	0.492	0.118	0.548
0.002	0.352	0.013	0.349	0.124	0.371
0.002	0.114	0.013	0.13	1.006*	0.158
0.002	0.221	0.013	0.223	0.94*	0.232
0.002	0.171	0.013	0.16	0.026	0.264
0.012	0.19	0.023	0.185	0.98*	0.206
0.013	0.243	0.024	0.239	0.059	0.313
0.018	0.232	0.007	0.153	0.061	0.171

الفروق في تقديرات الصعوبة المتناظرة بين التدريج المرجعي والتدريج الناتج عن تحليل	مجموع الخطأين المعياريين الناتجين عن التدريج المرجعي والتدريج الناتج عن	الفروق في تقديرات الصعوبة المتناظرة بين التدريج المرجعي والتدريج الناتج عن تحليل العينة (٦٠٠) فرد	مجموع الخطأين المعياريين الناتجين عن التدريج المرجعي والتدريج الناتج عن	الفروق في تقديرات الصعوبة المتناظرة بين التدريج المرجعي والتدريج الناتج عن تحليل العينة (٣٠٠) فرد	مجموع الخطأين المعياريين الناتجين عن التدريج المرجعي والتدريج الناتج عن تحليل العينة (٣٠٠) فرد
0.016	0.162	0.005	0.164	0.088	0.175
0.102	0.291	0.113	0.282	0.13	0.389
0.009	0.544	0.02	0.53	0.063	0.597
0.001	0.392	0.012	0.389	0.071	0.402
0.022	0.164	0.033	0.161	0.331*	0.175
1.138	5.679	1.127	5.969	3.51	6.532
0.017	0.335	0.028	0.332	0.055	0.356
0.0986	0.243	0.0876	0.247	0.1706	0.239
0.098	0.3	0.087	0.274	0.263*	0.257
0.002	0.189	0.009	0.132	0.074	0.146
0.012	0.321	0.001	0.344	0.084	0.344
0.017	0.208	0.002	0.234	0.081	0.223
0.011	0.137	0.007	0.164	0.79*	0.167
0.026	0.106	0.015	0.152	0.098	0.13
0.339	0.426	0.328	0.429	0.411	0.431
0.009	0.585	0.002	0.592	0.081	0.593
0.033	0.93	0.022	0.243	0.105	0.208
0.035	0.17	0.024	0.208	0.107	0.197
0.027	0.175	0.016	0.193	0.099	0.296
0.03	0.099	0.019	0.193	0.102	0.187
0.089	0.106	0.1	0.16	0.068	0.182
0.043	0.07	0.032	0.162	0.115	0.142
0.022	0.082	0.011	0.197	0.094	0.252
0.009	0.296	0.002	0.383	0.081	0.389
0.006	0.128	0.017	0.266	0.066	0.227
0.034	0.14	0.023	0.19	0.106	0.189
0.019	0.172	0.03	0.203	0.053	0.204
0.045	0.1	0.056	0.204	0.027	0.202
0.018	0.173	0.007	0.212	0.13	0.215

جدول (٦) : متوسطات الأخطاء المعيارية لتقديرات مفردات اختبار الدراسة تبعا لحجم عينة التحليل

الإحصاءات	أخطاء تقدير الصعوبة الناتجة من تحليل العينة (١٨٠٠)	أخطاء تقدير الصعوبة الناتجة من تحليل العينة (٩٠٠)	أخطاء تقدير الصعوبة الناتجة من تحليل العينة (٦٠٠)	أخطاء تقدير الصعوبة الناتجة من تحليل العينة (٣٠٠)
عدد المفردات	40	40	40	40
المتوسط الحسابي	0.1894	0.1988	0.2133	0.2395

يتبين من الجدول السابق عدم وجود فروق دالة بين تقديرات الصعوبة المتناظرة الناتجة عن تحليل العينة الكلية البلاغ حجمها (١٨٠٠) فرد (التدريج المرجعي) وكل من التدريجين : التدريج الناتج من تحليل استجابات العينة التي بلغ حجمها (٩٠٠) فرد ، وكذلك التدريج الناتج من تحليل استجابات العينة التي بلغ حجمها (٦٠٠) فرد ، كما تبين وجود فروق جوهرية بين تقديرات الصعوبة المتناظرة بين التدريج المرجعي وتدريج الصعوبة الناتج عن تحليل استجابات العينة البالغ حجمها (٣٠٠) فرد حيث بلغ عدد الفروق الدالة بين تقديرات الصعوبة المتناظرة (٧) فروق (وهي التي تجاوزت فروق تقديرات الصعوبة المتناظرة لها مجموع أخطائها المعيارية) وذلك بنسبة قدرها (% 17.5) من جملة الفروق وهي النسب التي تخطت النسب المقبولة وهي (٠.٠٥)، ممار يفيد بتأثير حجم العينة على تقديرات الصعوبة الخاصة بالمفردات الاختبارية ، وهذه النتيجة تتسق مع نتائج دراسات (٢٠٢٢) Keller, والتي أظهرت تأثر دقة تقدير معلم الصعوبة مع اختلاف أحجام العينات فقد ازدادت دقة التقدير بازدياد حجم العينة وكذلك دراسة دي لا توري ويوان (2020)Yuan & Torre La De التي أظهرت نتائجها وجود تأثير دال إحصائياً لمتغير حجم العينة على دقة تقديرات معالم المفردات. وكذلك دراسة الزبون (٢٠١٣) التي أظهرت أثر حجم العينة على تقدير الخطأ المعياري للتقدير.، وكذلك دراسة شما (٢٠١٣) التي أظهرت أن دقة تقديرات صعوبة المفردات تزداد مع زيادة حجم العينة بينما اختلفت هذه النتائج مع دراسة الشريفين (٢٠١٢) التي أظهرت عدم وجود فروق ذات دلالة إحصائية تعزى لمتغير حجم العينة بين متوسطات الأخطاء المعيارية لتقديرات معالم المفردات، وقد تفسر نتائج دراسة الشريفين والتي جاءت مناقضة لنتائج جميع الدراسات التي كان لها نفس الأهداف بوجود فروق صغيرة إلى حد ما بين حجم العينات موضع المقارنة مما قد يفسر وجود عدم اختلاف في تقديرات صعوبة المفردات الناتجة عن تحليلات الاستجابات الخاصة بهذه العينات.

هذا ويرى الباحث بأن تفسير هذه النتائج له علاقة بالأخطاء المعيارية الخاصة بتقدير صعوبات الفقرات فمن المعروف من الوجهة الإحصائية أن ازدياد حجم العينة يؤدي إلى انخفاض الأخطاء المعيارية الخاصة بتقدير صعوبة المفردات الاختبارية نظرا لن الخطأ المعياري للتقدير يرتبط بحجم عينة التحليل مما قد يؤثر على تقديرات الصعوبة التي يمكن أن تختلف باختلاف حجم العينات موضع التحليل، وهذا ما تأكد من النتائج المتضمنة بجدول (٦) حيث تبين أن متوسط الأخطاء المعيارية لتقديرات صعوبة الفقرات تنخفض بازدياد حجم العينة مما يفيد أن دقة تقديرات الصعوبة تزداد بازدياد حجم عينة التحليل..

ثالثاً: عرض ومناقشة نتائج الإجابة عن السؤال الثاني:

لقد أسفرت الإجراءات الخاصة بالإجابة عن السؤال الثاني الذي ينص على: ما تأثير حجم العينة على تقديرات قدرات أفراد العينة البحثية باستخدام تحليلات النموذج اللوجيستي الثلاثي المعلم .

لقد أسفرت نتائج التحليلات الخاصة بالإجابة عن هذا السؤال عن النتائج المبينة بالجداول التالية:

جدول(٧): اختبار ماكولي للدورية Sphericity Mauchly's Test of

التأثير داخل المجموعات	اختبار ماكولي Mauchly's W	قيمة مربع كاي	درجات الحرية	اختبار مستوى الدلالة
factor1	0.681	530.12	3	0.000

يتبين من بالجدول السابق أن قيمة اختبار مربع كاي بلغت (٥٣٠.١٢) وهي دالة عند مستوى (٠.٠١) مما يدل على أن قياسات القدرات الأربعة لا تحقق شرط الدورية مما يستلزم تصحيح قيمة (ف) لتحليل تباين القياس المتعدد باستخدام اختبار جرين هاوس- جيسر Greenhouse-Geisser

جدول(٨) : تحليل تباين القياس المتعدد

مصدر التباين	الاختبار	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة (ف)	مستوى الدلالة
بين القياسات	Greenhouse-Geisser	٩٣٧٢.٩	١.٣١٨	٧١١١.٥	٥٥٥.٦	٠.٠٠٠
الخطأ	Greenhouse-	٩٢٧٨.٧	٧٢٣.٢	١٢.٨		

تبين من الجدول السابق وجود تباين دال إحصائيا بين القياسات الأربعة مما يشير إلى وجود فروقا جوهرية بين تقديرات القدرة (بوحدة الوات) المقابلة لكل درجة خام متوقعة من الاختبار والتي تم الحصول عليها من خلال تحليلات استجابات العينات الأربعة مختلفة الحجم ، والجدول التالي يوضح اتجاه هذه الفروق باستخدام اختبار LSD للمقارنات الثنائية.

جدول(٩) : المقارنات الثنائية بين متوسطات تقديرات القدرة لكل درجة متوقعة للعينات الفرعية مختلفة الحجم على اختبار الدراسة

القياس (أ)	القياس (ب)	متوسط الفرق (أ- ب)
العينة الكلية (١٨٠٠)	العينة الفرعية (٩٠٠)	٧.٢١٥
	العينة الفرعية (٦٠٠)	١٠.٤٣٢
	العينة الفرعية (٣٠٠)	**٨٥.٠١١
العينة الفرعية (٩٠٠)	العينة الكلية (١٨٠٠)	-٧.٢١٥
	العينة الفرعية (٦٠٠)	٩.١١٠
	العينة الفرعية (٣٠٠)	*٥٠.١٧٢
العينة الفرعية (٦٠٠)	العينة الفرعية (٣٠٠)	*٤٧.٢٣١
	العينة الفرعية (٩٠٠)	-٩.١١٠
	العينة الكلية (١٨٠٠)	-١٠.٤٣٢

◆ دال عند مستوى دلالة(٠.٠١) ◆◆ دال عند مستوى دلالة(٠.٠٥)

جدول (١٠) : متوسطات الأخطاء المعيارية لتقديرات مفردات اختبار الدراسة تبعاً لحجم عينة التحليل

الإحصاءات	أخطاء تقدير القدرة الناتجة من تحليل العينة (١٨٠٠)	أخطاء تقدير القدرة الناتجة من تحليل العينة (٩٠٠)	أخطاء تقدير القدرة الناتجة من تحليل العينة (٦٠٠)	أخطاء تقدير القدرة الناتجة من تحليل العينة (٣٠٠)
المتوسط الحسابي	0.1078	0.1991	0.2342	0.2694

يتضح من النتائج المبينة بالجدول رقم (٩) أن هناك فروق جوهرية عند مستوى (٠.٠١) بين متوسط تقديرات القدرة لكل درجة متوقعة على الاختبار المرجعي إذا ما قورنت بمتوسط لتقديرات القدرة لكل درجة متوقعة على نفس الاختبار الناتج من تحليل استجابات العينة الفرعية (٣٠٠) فرد في صالح العينة (١٨٠٠) فرد ونظرا لأن متوسط أخطاء تقدير القدرات الناتجة عن تحليل استجابات العينة البالغ حجمها(١٨٠٠) بلغ قيمته (0.1078) وهي أقل من متوسط الأخطاء المعيارية التي بلغت قيمتها (0.2694) الناتجة عن تحليل استجابات العينة التي بلغ حجمها (٣٠٠)مما يشير إلى الاختلاف الجوهري في تقديرات القدرة لكل درجة متوقعة الناتجة عن تحليلات استجابات العينات الفرعية مما يمكن عزوه لاختلاف حجم عينتي الأفراد كما يفيد بدقة تقديرات القدرة الناتجة عن تحليلات العينة البالغة (١٨٠٠) عندما تقارن بالعينة (٣٠٠) ، كما تبين من نفس الجدول السابق وجود فروق جوهرية عند مستوى (٠.٠٥) بين متوسط تقديرات القدرة لكل درجة متوقعة على الاختبار الذي تم تحليله وفقا لاستجابات العينة الفرعية (٩٠٠) فرد إذا ما قورنت بمتوسط لتقديرات القدرة لكل درجة متوقعة على نفس الاختبار الناتج من تحليل استجابات العينة الفرعية (٣٠٠) فرد في صالح العينة (٩٠٠) فرد، مما

يشير إلى الاختلاف الجوهري في تقديرات القدرة لكل درجة متوقعة الناتجة عن تحليلات استجابات العينات الفرعية مما يمكن عزوه لاختلاف حجم عينتي الأفراد ، كما تبين من نفس الجدول السابق وجود فروق جوهرية عند مستوى (٠.٠٥) بين متوسط تقديرات القدرة لكل درجة متوقعة على الاختبار الذي تم تحليله وفقا لاستجابات العينة الفرعية (٦٠٠) فرد إذا ما قورنت بمتوسط لتقديرات القدرة لكل درجة متوقعة على نفس الاختبار الناتج من تحليل استجابات العينة الفرعية (٣٠٠) فرد في صالح العينة (٦٠٠) فرد ووفقا للنتائج الخاصة بمتوسط الأخطاء المعيارية الواردة بجدول(١٠) تبين أن تقديرات القدرة الناتجة عن تحليل العينة التي بلغ حجمها (٦٠٠) أكثر دقة من العينة التي بلغ حجمها (٣٠٠) ، مما يشير إلى الاختلاف الجوهري في تقديرات القدرة لكل درجة متوقعة الناتجة عن تحليلات استجابات العينات الفرعية مما يمكن عزوه لاختلاف حجم عينتي الأفراد وبما يفيد بدقة تقديرات القدرة للعينة الأكبر حجما ، وتتفق هذه النتائج مع ما توصلت إليه نتائج دراسة (٢٠٢٢) Keller, التي أظهرت تأثر دقة تقدير معلم الصعوبة والتمييز مع اختلاف أحجام العينات فقد ازدادت دقة التقدير بازدياد حجم العينة، وكذلك دراسة دي لا توري ويوان (2020) Yuan ,& Torre La De والتي أظهرت نتائجها أن حجم العينة يؤثر على قدرة الاختبار في تقدير معالم القدرة و المفردة، وكذلك دراسة بني عطا (٢٠١٤) والتي أظهرت نتائجها عن تأثير دال إحصائيا لمتغير حجم العينة والتفاعل بينهما في دقة تقديرات معالم المفردات والأفراد. كما اتسقت مع الزبون (٢٠١٣) التي استهدفت الكشف عن أثر حجم العينة على تقدير دالة المعلومات وكذلك الخطأ المعياري للتقدير، والتي أشارت نتائجها إلى ان تقديرات دالة المعلومات تتغير بتغير حجم العينة، حيث تزداد تبعا لزيادة حجم العينة، كما أظهرت النتائج أيضا ان الخطأ المعياري في تقدير دالة المعلومات تتغير كذلك بتغير حجم العينة، اذ يتناقص هذا الخطأ مع ازدياد حجم العينة مما يفيد بأن حجم العينة يؤثر على تقديرات معالم الاختبار بصورة عامة، في حين اختلفت هذه النتائج مع نتائج دراسة الشريفين (٢٠١٢) التي أظهرت نتائجها عدم وجود فروق ذات دلالة إحصائية تعزى لمتغير حجم العينة بين متوسطات الأخطاء المعيارية لتقديرات القدرة للأفراد وكذلك تقديرات معالم المفردات، ومن خلال هذه النتائج يتضح تأثر تقديرات قدرات الأفراد بحجم عينة التحليل، مما يستوجب ضرورة إجراء مزيد من الدراسات المستقبلية باستخدام عينات أخرى مختلفة للتوصل إلى الحجم المناسب من العينة التي يمكن أن تسفر عن دقة أكبر من تقدير القدرة.

التوصيات

في ضوء النتائج التي انتهت إليها الدراسة لذا يوصي الباحث بما يلي:

- يجب النظر بعين الاعتبار لحجم العينة التحليل عند تقدير تقديرات الصعوبة للمفردات الاختبارية، خلال إجراءات بناء الاختبارات باستخدام نموذج الاستجابة للمفردة، وذلك بزيادة حجم عينة الأفراد المتلقين للاختبار، بهدف خفض أخطاء القياس وبالتالي رفع مستوى دقة هذه تقديرات.
- للحصول على أعلى دقة لتقديرات القدرة للأفراد المتلقين للاختباريجب النظر بعين الاعتبار لحجم العينة التحليل بهدف خفض أخطاء القياس وبالتالي رفع مستوى دقة هذه التقديرات أيضا.

دراسات مقترحة

- إجراء دراسة تستهدف الكشف عن تأثير أحجام مختلفة من عينات التحليل على تقدير قدرة الفرد وكذلك دالة المعلومات باستخدام النموذج الثنائي المعلم .
- إجراء دراسة للمقارنة بين نماذج الاستجابة للمفردة للكشف عن مدى تأثرها باختلاف أحجام عينات التحليل.

المراجع

- حابس الزبون (٢٠١٣). أثر حجم العينة على تقدير دالة المعلومات للاختبار والخطأ المعياري في تقديرها باستخدام النظرية الحديثة في القياس . *مجلة جامعة النجاح للأبحاث (العلوم الانسانية)* ،٢٧(٦)، ٢٢-٤٩.
- حسين عبد النبي القيسي (٢٠١٣): دقة تقدير معالم الفقرة والقدرة باستخدام نماذج نظرية استجابة الفقرة البارامترية واللابارامترية باختلاف حجم العينة وطول الاختبار : دراسة مقارنة . (رسالة دكتوراه غير منشورة)، كلية التربية ، جامعة اليرموك.
- صلاح الدين محمود علام (٢٠٠٥). *نماذج الاستجابة للمفردات الاختبارية احادية البعد ومتعددة الابعاد وتطبيقاتها في القياس النفسي والتربوي*، الطبعة الاولى، القاهرة : دار الفكر العربي.
- .ماجد محمد الخياط (٢٠١٢). درجة مطابقة اختبار تحصيلي وفق نموذج راش احادي المعلمة في الكشف عن مستوى المعرفة العلمية في المهارات الرياضية لدى طلبة الصف الثامن الاساسي. *مجلة جامعة الاقصى (سلسلة العلوم الإنسانية*،١٦(١)، ٤٤-٨٧.
- نضال كمال الشريفين (٢٠١٢). أثر طريقة تقدير معالم الفقرة وقدرات الافراد على قيم معالم الفقرة، والخصائص السيكومترية للاختبار، في ضوء تغير حجم العينة . *المجلة التربوية*، الاردن، ٢٦، ١٣٨-١٧٧.
- يمان نزار شما (٢٠١٣). اثر حجم العينة على دقة تقدير صعوبة المفردات وقدرة الافراد باستخدام نموذج راش. **مجلة الآداب**، جامعة بغداد، ١٠٥ ،٦٩٨-٦٧٣.

-Belov, D. I. & Armstrong, R. D. (2010). Automatic detection of answer copying via kullback-leibler divergence and K- index. Applied psychological Measuement, 34(6), 379-392.

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

- De la Torre, J.; Yuan H. (2020). Parameter Estimation with Small Sample Size A Higher-Order IRT Model Approach. *Applied Psychological Measurement*. 28 (4), 255-272.

- Emberson, S. & Reise, S (2015). *Item Response Theory for Psychologists* ,New Jersey: Lawrence Earlbaum Associates, Publishers Mahwa.

- Hambleton, R., Swaminathan, H. &Rogers, H. (1991). Fundamentals of item response theory, *International Educational and Professional*. Publisher Newbury Park.

- 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.

. -Keller, L., (2022). Small –sample item parameter estimation in the three parameter logistic model: Using collateral information. *Doctoral dissertations* Available from proquest. AA13068572.

- Lord, F. (1980). *Applications of Item Response Theory to Practical testing problems*. New Jersey, Lawrence Erlbaum Associates,

- Magno, C. (2019). Demonstrating the Difference between. Classical Test Theory and item Response Theory Using Derived Test Data. *The International Journal of Educational and Psychological Assessment*,23(7),124-166
- Nitko, A. (2001). Educational Assessment of Students. Amazon, UAS: Merrill. - Reeve, B. (2004). *An introduction to modern measurement theory*. Division of fencer control and population science, National Cancer Institute.
-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.