

9-15-2025

## The Effect of Sample Size Differences on Calculating the Conditional Standard Error of Achievement Tests

Khalid Jamal Jasim

University of Baghdad/ University of Baghdad, Khalid.jamal@ircoedu.uobaghdad.edu.iq

Haris Sahib Naji

College of Education/ Ibn Rushed College of Education/ Ibn Rushed,

Hares.Naji2022m@ircoedu.uobaghdad.edu.iq

Follow this and additional works at: <https://alustath.uobaghdad.edu.iq/journal>

---

### Recommended Citation

Jasim, Khalid Jamal and Naji, Haris Sahib (2025) "The Effect of Sample Size Differences on Calculating the Conditional Standard Error of Achievement Tests," *Alustath Journal for Human and Social Sciences*: Vol. 64: Iss. 3, Article 11.

DOI: 10.36473/2518-9263.2431

Available at: <https://alustath.uobaghdad.edu.iq/journal/vol64/iss3/11>

This Article is brought to you for free and open access by Alustath Journal for Human and Social Sciences. It has been accepted for inclusion in Alustath Journal for Human and Social Sciences by an authorized editor of Alustath Journal for Human and Social Sciences.



Scan the QR to view  
the full-text article on  
the journal website

# The Effect of Sample Size Differences on Calculating the Conditional Standard Error of Achievement Tests

**Khalid Jamal Jasim, Haris Sahib Naji**

University of Baghdad/ University of Baghdad  
College of Education/ Ibn Rushed College of Education/ Ibn Rushed

## ABSTRACT:

The aim of the research was to study the effect of sample size differences (200, 400, 600) on calculating the conditional standard error using the Lord and Keats methods in achievement tests according to the classical measurement theory. The results showed that the difficulty coefficients ranged between 0.42 and 0.73 with an average of 0.60, and the discrimination coefficients ranged between 0.17 and 0.47 with an average of 0.36, with the exception of two items that did not meet the conditions. All the wrong alternatives were effective with negative values, and the Cronbach's alpha reliability coefficient was 0.879, The Lord and Keats methods were compared using arithmetic means and standard deviations, and the results showed statistically significant differences in favor of the Keats method due to the lower average standard errors compared to the Lord method. The effect of sample sizes was also studied using one-way analysis of variance, which showed statistically significant differences between the three sample sizes (200, 400, 600). The results showed that the conditional standard error decreased with increasing the sample size, and the larger sample size (600) was the better, the researcher presented a number of conclusions, recommendations and suggestions.

**Keywords:** Sample Size, Conditional Standard Error.

Received 4 December 2024; Revised 7 January 2025; Accepted 2 November 2024  
Available online 15 September 2025

Corresponding author: Khalid Jamal Jasim, Haris Sahib Naji  
E-mail address: [Hares.Naji2022m@ircoedu.uobaghdad.edu.iq](mailto:Hares.Naji2022m@ircoedu.uobaghdad.edu.iq), [Khalid.jamal@ircoedu.uobaghdad.edu.iq](mailto:Khalid.jamal@ircoedu.uobaghdad.edu.iq)

<https://doi.org/10.36473/2518-9263.2431>

2518-9263/© 2025 The Author(s). Alustath Journal for Human and Social Sciences. This is an open access article under the CC BY 4.0 Licence (<https://creativecommons.org/licenses/by/4.0/>).



# تأثير اختلاف حجم العينة على حساب الخطأ المعياري الشرطي للاختبارات التحصيلية

حارس صاحب ناجي، خالد جمال جاسم

قسم العلوم التربوية والنفسية/كلية التربية  
ابن رشد للعلوم الانسانية/جامعة بغداد

## المُلخَص

هدف البحث إلى دراسة تأثير اختلاف حجم العينة (200، 400، 600) على حساب الخطأ المعياري الشرطي باستخدام طريقتي لورد وكيّس في الاختبارات التحصيلية على وفق نظرية القياس الكلاسيكية، وأظهرت النتائج أن معاملات الصعوبة تراوحت بين 0.42 و 0.73 بمتوسط 0.60 ومعاملات التمييز تراوحت بين 0.17 و 0.47 بمتوسط 0.36 باستثناء فقرتين غير مستوفيتين للشروط، وكانت جميع البدائل الخاطئة فعّالة بقيم سالبة، وبلغ معامل الثبات ألفا كرونباخ 0.879، وتمت مقارنة طريقتي لورد وكيّس باستخدام المتوسطات الحسابية والانحرافات المعيارية وأظهرت النتائج فروقاً ذات دلالة إحصائية لصالح طريقة كيّس؛ بسبب انخفاض متوسط الأخطاء المعيارية مقارنة بطريقة لورد، كما تمت دراسة تأثير أحجام العينات باستخدام تحليل التباين الأحادي الذي أظهر وجود فروق ذات دلالة إحصائية بين أحجام العينات الثلاثة (200، 400، 600) وأظهرت النتائج أن الخطأ المعياري الشرطي يقل مع زيادة حجم العينة وكان حجم العينة الأكبر (600) هو الأفضل، وفي ضوء النتائج توصل الباحث إلى مجموعة من الاستنتاجات والتوصيات والمقترحات.

**الكلمات المفتاحية:** حجم العينة، الخطأ المعياري الشرطي.

تم الاستلام في 4 ديسمبر 2024 ؛ تم المراجعة في 7 يناير 2025 ؛ تم القبول في 2 نوفمبر 2024  
متاح على الانترنت 15 سبتمبر 2025

المؤلف المراسل: حارس صاحب ناجي، خالد جمال جاسم  
عنوان البريد الإلكتروني: Hares.Naji2022m@ircoedu.uobaghdad.edu.iq, Khalid.jamal@ircoedu.uobaghdad.edu.iq

<https://doi.org/10.36473/2518-9263.2431>

2518-9263/© 2025 The Author(s). Alustath Journal for Human and Social Sciences. This is an open access article under the CC BY 4.0 Licence  
(<https://creativecommons.org/licenses/by/4.0/>).

## مشكلة البحث

على الرغم من أهمية الخطأ المعياري في القياس، إلا أن الطريقة التقليدية المستخدمة لتقديره تعاني من بعض النقاط الضعف، ومن بينها: أن الخطأ المعياري في القياس يكون متساوياً لجميع الممتحنين بغض النظر عن قدراتهم، مما يقلل من قيمته التفسيرية، ولاسيما عندما يكون توزيع العلامات غير متوازن، مما يؤثر على دقة تقدير القدرة الحقيقية للممتحنين، فضلاً عن ذلك، فإن دقة فترات الثقة للقدرة الحقيقية غير محددة بشكل دقيق (عباس، 1993: 26) (Abbas, 1993: 26).

هذا من جهة ومن الجهة الأخرى فإنه من الصعوبة، بل إن لم يكن من المتعذر قياس السمات لكل أفراد المجتمع الإحصائي، ويمكن أن يكون صعباً ومكلفاً من حيث الجهد والوقت والمال، قد يكون من الصعب أيضاً التحكم في جميع أفراد المجتمع أو تحديدهم. ولذلك، يتعين الاعتماد على عينة في بعض الأحيان بدلاً من استهداف المجتمع بأكمله. فضلاً عن، فإن العينة يمكن أن تعطي مؤشرات مماثلة تقريباً لتلك التي يمنحها المجتمع الإحصائي (عبد السلام، 1987: 3) (Abdulsalam, 1983: 3).

إذ تعد أخطاء المعاينة من المصادر الأساس لإخطاء القياس التي لا يقل تأثيرها عن أخطاء المقياس، وقد تتأني أخطاء المعاينة من طريقة اختيار العينة أو من حجمها، بيد إنه يمكن السيطرة على أخطاء طريقة الاختيار إلى حد ما، لان هناك طرائق لاختيار العينة يمكن إن تحد منها (المياحي، 2005: 9) (Al-miyahhi, 2005: 9).

## أهمية البحث

يعد القياس أمراً على جانب كبير من الأهمية في أي علم من العلوم فالعلوم جميعها تسعى إلى تطوير اساليب موضوعية دقيقة لقياس الظواهر المتعلقة بها من فهم لهذه الظواهر وتفسيرها (علام، 2000: 13) (Allam, 2000: 13) في (جاسم وعيال، 2019: 275) (Jasim & Ayal, 2019: 275).

ويعتبر الخطأ المعياري في القياس احدى المؤشرات على دقة القياس؛ إذ يهدف إلى تقدير الدرجة الحقيقية للمفحوصين بشكل دقيق وصحيح، فضلاً عن الحصول على دقة في فترات الثقة لهذه الدرجة (الطاهات، 2016: 3) (Altahhat, 2016: 3).

إن دقة القياس هي أمر مهم ومتعلق بالاستخدام الفعال لأدوات القياس في مجالاتها المختلفة، ويمكن استخدام الخطأ المعياري في القياس كمؤشر لدقة القياس من بين مؤشرات أخرى، كما يُستخدم الخطأ المعياري لتقدير خاصية الثبات كواحدة من الخصائص السيكمترية لأدوات القياس (Gullikson, 1967: 142-130).

ويُعَدّ الخطأ المعياري في القياس وسيلة من وسائل التنبؤ بقيمة الثبات، إذ يرتبط الخطأ المعياري في القياس على نحو وثيق بالثبات، إذ يزداد الثبات عندما يكون الخطأ المعياري في القياس أقل، والعكس صحيح (عباس، 1993: 26) (Abbas, 1993: 26).

اضف إلى ذلك، لابد من الإشارة إلى وجود علاقة بين حجم العينة وبين أخطاء القياس، إذ كلما زاد حجم العينة قل خطأ الصدفة أو العشوائية؛ لان مثل هذا الخطأ لا يمكن ضبطه أو السيطرة عليه، ولكن يمكن إن يلغى بعضه من خلال حجم العينة الكبير (عبد الرحمن، 1998: 89) (Abdulrahman, 1998: 89).

ويشير «جي 1990» Gay إلى إن حجم العينة يعتمد على مدى تجانس المجتمع في السمة المقاسة، إذ كلما كان المجتمع متجانساً قل حجم العينة، وكلما قل تجانس المجتمع زاد حجم العينة، لذلك كلما كانت العينة كبيرة تقل أخطاء القياس فيها (جلال، 2001: 38-39) (Jalal, 2001: 38).

## ومما تقدم يمكن تلخيص أهمية البحث الحالي بما يأتي

1. يتناول البحث الحالي إحدى الموضوعات المهمة التي لم يتطرق لها أحداً (على حد علم الباحث) في مجال القياس والتقويم التربوي على المستوى المحلي، والتي تعد من الإضافات النوعية الجديدة للقياس والتقويم والتي لها اثر كبير على قيمة الخطأ المعياري الشرطي في القياس وذلك من خلال الكشف عن (تأثير اختلاف حجم العينة على حساب الخطأ المعياري الشرطي للاختبارات التحصيلية) كون ان الخطأ المعياري احد المؤشرات المهمة لدقة القياس.

2. إن دراسة طرائق حساب الخطأ المعياري الشرطي تزودنا ببيانات ومعلومات واضحة ومناسبة وموضوعية عن دقة القياس وكيفية تطبيق هذه الطرق.

## هدف البحث

### يهدف البحث الحالي إلى:

معرفة تأثير اختلاف حجم العينة على حساب الخطأ المعياري الشرطي للاختبارات التحصيلية عند استخدام طريقتين من طرائق حساب الخطأ المعياري الشرطي بغية الوصول إلى أفضل حجم للعينة بحيث يكون مناسب لحساب هذه الطرق.

### ولغرض تحقيق هذا الهدف وضع الباحث الفرضيتين الصفريتين التاليتين:

- لا توجد فروق ذات دلالة إحصائية بين طرائق حساب الخطأ المعياري الشرطي.
- لا توجد فروق ذات دلالة إحصائية بين أحجام العينة الثلاث (200,400,600) لطرائق حساب الخطأ المعياري الشرطي في القياس (لورد ، كيتس).

## حدود البحث

### يتحدد البحث الحالي بـ:

- طلبة الصف الرابع للأقسام غير الاختصاص التابعين لكليات التربية في جامعة ديالى لكلا الجنسين (ذكور - إناث) للدراسة الصباحية، للعام الدراسي (2023 - 2024).
- طريقة (لورد ، كيتس) في حساب الخطأ المعياري الشرطي في القياس.

## تحديد المصطلحات

- فيما يأتي تحديد المصطلحات الواردة في البحث الحالي والمتمثلة في (حجم العينة، العينة، الخطأ المعياري الشرطي للقياس، الاختبار التحصيلي):-

### اولا - حجم العينة Sample Size

لم يجد الباحث بعد الاطلاع على الأدبيات والدراسات المتخصصة في القياس النفسي والتربوي تعاريف تختلف في مضمونها عن ما اشار اليه ( زكري ، 2003) في حجم العينة على انه هو عدد وحدات المعاينة التي يتم اختيارها ويرمز له عادة بالرمز (N) ( زكري، 2003: 10) (Zakri,2003:10) في (العيناي، 2014: 13) (Al-Aithawi,2014:13).

### ثانيا - الخطأ المعياري الشرطي للقياس The Standard Error of Test:

كروكر و الجينا (1986)

«متوسط الانحراف المعياري لتوزيعات الخطأ لعدد كبير من المواقف الاختبارية المتكررة وهو مهم لإنتاج فترة ثقة حول درجات الاختبار الملاحظة التي لها احتمالية معروفة لتحتوي درجات المفحوص الملاحظة» (Crocker, & Algina (1986).

### ثالثاً: الاختبار التحصيلي Achievement Test:

جيزلي وآخرون (1981) Ghieselli& et al:

عينة من الفقرات أو الأسئلة أو المهام التي تقدم للطلب لتحديد مستوى تحصيله من المعلومات أو المهارات (Ghieselli& et al , 1981: 432).

## أطار نظري

### القياس النفسي والتربوي:

يعتبر القياس واحد من اهم الموضوعات المهمة لدى العلوم الطبيعية والانسانية وذلك لدور القياس في التقدير الكمي للخص والسمات، وبدون هذا التقدير يتحول العلم إلى مجرد انطباعات شخصية سطحية وملاحظات عابرة ووصف لفظي، وبالتالي سيمن ذلك من استيعابنا للظواهر والخصائص المدروسة فضلاً عن امكانية تدقيق الاستنتاجات المستخلصة من البيانات المتحصل عليها، وكل هذا يصب في مصلحة تطوير منهجيات البحث العلمي وطرق قياسه (بلال، 2023: 189) (Bilal,2023:189).

### نظرية القياس التقليدية ( الكلاسيكية ) Theory Test Classical (CTT):

تعود أصول نظرية القياس الكلاسيكية إلى العالم الإنجليزي سبيرمان (Spearman)، إذ تمكّن من طرح أسس نظرية الدرجة الحقيقية بين عامي 1904 و 1913. تُعرف هذه النظرية أيضًا باسم النظرية التقليدية للقياس، وتعتمد على فرضية أساسية تفيد بأن الدرجات التي يحصل عليها الأفراد المخضوعون للاختبارات أو المقاييس للسمة المقاسة تتبع توزيعًا اعتداليًا، والذي يتأثر بطبيعة فقرات الاختبار وخصائص العينة المخضوعة للقياس. (أبو ناهية، 1994: 51) (Abu Nahiya, 1994: 51).

تُركز نظرية القياس الكلاسيكية على موازنة أداء الفرد المفحوص مع أداء أقرانه وتفسير الدرجة التي يحصل عليها الفرد في ضوء المجموعة التي ينتمي إليها. فعندما يتم قياس الفرد، فإن القيمة الناتجة لا تكون ذات قيمة إلا إذا تمت مقارنتها بمقياس المجموعة المرجعية التي ينتمي إليها (الطيريري، 1997: 42) (AL-Tariri, 1997: 42).

### الخطأ المعياري للقياس Test of Error Standard The:

تهدف نظرية القياس الكلاسيكية إلى تقدير مقدار الخطأ بهدف تحقيق أقصى درجة من الدقة في القياس، من خلال تقدير النطاق الذي تقع ضمنه الدرجة الحقيقية.

يعرف الخطأ المعياري للقياس على أنه الانحراف المعياري لدرجات الخطأ، وبسبب صعوبة إعادة اختبار كل فرد في العينة عدة مرات لتحديد درجة الخطأ الفردية، فإننا غالبًا ما لا نستطيع تحديد قيمة هذا الخطأ بدقة مطلقة. ومع ذلك، يمكن تقدير قيمته من خلال حساب الانحراف المعياري للدرجات الملاحظة وقيمة الثبات، أن ارتفاع معامل الثبات يؤدي إلى تقليل الخطأ المعياري وتقليل التباين في درجات المفحوصين في عملية القياس، ويجب ملاحظة أن الخطأ المعياري للقياس يمثل الانحراف المعياري للخطأ، وهو قيمة مطلقة وليس نسبة بالنسبة لتوافق الدرجة الملاحظة مع الدرجة الحقيقية. لذلك يعتمد معامل الثبات على تباين درجات المجموعة، في حين الخطأ المعياري للقياس لا يتأثر بذلك التباين، وتقدير الخطأ المعياري للقياس لا يتطلب معرفة الدرجة الحقيقية (علام، 2006: 101) (Allam, 2006: 101).

### طرائق حساب الخطأ المعياري الشرطي في القياس

أولا/ طريقة لورد (Lord's Method):

يفترض لورد أن الاختبار عبارة عن اختيار عينات عشوائية للفقرات من مجتمع الفقرات. وأن الاختبارات المختارة بهذه الطريقة والتي تحتوي على العدد نفسه من الفقرات هي اختبارات عشوائية متوازنة. وأن الدرجة التي يحصل عليها فرد ما (عدد الإجابات الصحيحة) من خلال إجابته على عدد كبير من فقرات تتبع توزيع ذات الحدين (Binomial distribution).

المعادلة الرياضية:

$$\sigma_{Xa} = \sqrt{\frac{X_a \cdot (n - X_a)}{n - 1}}$$

حيث أن: n: عدد فقرات الاختبار.

Xa: الدرجة الناتجة للفرد (a).

Xa: تقدير الخطأ المعياري في القياس باستخدام طريقة لورد.

ومن الانتقادات التي وجهت إلى طريقة لورد فشلها في الأخذ بعين الاعتبار عدم التطابق الجيد بين نماذج الاختبارات المتوازنة ولاسيما فيما يتعلق بالمحتوى وصعوبة الفقرات وتساوي المتوسطات، وذلك لأن طريقة لورد تفترض أن الاختبار عبارة عن اختيار عينات عشوائية للفقرات من مجتمع الفقرات، وأن الاختبارات المختارة بهذه الطريقة والتي تحتوي على نفس العدد من الفقرات هي اختبارات متوازنة أي أن قيمة الخطأ المعياري في القياس المقدر بواسطة طريقة لورد ستكون أعلى من القيمة المتوقعة لها (عباس، 1993: 10) (Abbas, 1993: 10).

### ثانيا/ طريقة كيتس: (Keats Method)

للتخلص من الانتقادات التي وجهت إلى طريقة لورد في إيجاد الخطأ المعياري في القياس، وأهمها أن القيم التي نحصل عليها للخطأ المعياري في القياس تكون أعلى من القيم الموقعة لها، لذا اقترح كيتس تحسينا على طريقة لورد بحيث يصبح بالإمكان الحصول على قيم للخطأ المعياري في القياس ليست أعلى مما هو متوقع لها، ويكمن اقتراح كيتس بضرب النتائج التي نحصل عليها للخطأ المعياري في القياس بطريقة لورد بعدد ثابت، ومقدار هذا العدد الثابت يساوي:

حيث أن:  $R_x$ : معامل الثبات للاختبار.

KR21: معامل الثبات للاختبار باستخدام معادلة كودر-ريتشاردسون 21.

وبالتالي فإن معادلة كيتس لإيجاد تقدير الخطأ المعياري في القياس هي:

$$\sqrt{\frac{1 - R_x}{1 - KR21}}$$

$$s(X_a) = \sqrt{\frac{X_a \cdot (n - X_a)}{n - 1} \left[ \frac{1 - R_x}{1 - KR21} \right]} \quad (8) \dots\dots\dots$$

حيث أن:  $X_a$ : الدرجة الناتجة للفرد a.

n: عدد فقرات الاختبار.

$R_x$ : معامل الثبات للاختبار.

KR21: معامل الثبات للاختبار باستخدام معادلة كودر-ريتشاردسون 21.

واقترح كيتس أن نحصل على قيمة  $R_x$  من خلال استخدام علامات الاختبارات المتوازنة، أو من خلال معامل الثبات النصفى، أو باستخدام معامل كرونباخ الفا. ويمكن القول إنَّ قلة المتطلبات الافتراضية لطريقة كيتس بالإضافة لسهولة حساباتها يعتبر من حسنات هذه الطريقة (عباس، 1993:11) (Abbas, 1993:11).

### حجم العينة Sample Size:

يتداخل تمثيل العينة للمجتمع بشكل كبير مع حجم العينة، إذ يُفترض أن يكون حجم العينة مناسباً وفقاً لاعتبارات إحصائية محددة. وبالتالي، ينبغي على الباحث أن يأخذ بعين الاعتبار كلا الجانبين بشكل متزامن، مما يعني ضرورة التفكير في تمثيل الخصائص وحجم العينة في آن واحد (المياحي، 2005: 58-59) (Al-miyahhi, 2005:9).

### دراسات سابقة:

أولاً: دراسة فيلدت (Feldt, 1984) (المقارنة بين قيم الخطأ المعياري في القياس المقدرة بطريقة نموذج لورد والطريقة التقليدية). هدف الدراسة: المقارنة بين طريقتين من طرق حساب الخطأ المعياري الشرطي في القياس وهما ( لورد ، الطريقة التقليدية). الأداة: استخدم اختبار أيوا بمجالاته الثلاثة (القراءة ، التراكيب ، التفكير المنطقي) على عينة من الممتحنين لطلبة المرحلة النهائية في المدارس الأمريكية. النتائج: وأظهرت النتائج أن القيمة المقدرة للخطأ المعياري في القياس باستخدام نموذج لورد تتغير مع مستويات العلامات في حين أن القيم المقدرة للخطأ المعياري في القياس بالطريقة التقليدية لا تتغير ومستويات الدرجات، وأوصت الدراسة باستخدام نموذج لورد لحساب الخطأ المعياري في القياس بدلا من استخدام الطريقة التقليدية (Feldt, 1984:3). ثانياً: دراسة العيثاوي ، (2014) ( أثر اسلوب اختيار العينة وحجمها على الخصائص القياسية لاختبارات القدرات العقلية على وفق نظرية السمات الكامنة). هدف الدراسة: هدفت هذه الدراسة إلى تحديد تأثير طريقة اختيار العينة وحجمها على الخصائص القياسية لاختبارات القدرات العقلية التي تستند على نظرية السمات الكامنة.

العينة: تم تطبيق الاختبار على عينة عشوائية طبقية مكونة من (1100) طالبا وطالبة. الأداة: لتحقيق هدف الدراسة قام الباحث بأعداد اختبار (أونيس-لينون) للقدرة العقلية العامة للمستوى الابتدائي، وتم تكييفه مع البيئة العراقية وفقا لنظرية السمات الكامنة (العيثاوي، 2014: 6) (Al-Aithawi, 2014:6).

### منهج البحث:

المنهج هو الاسلوب العلمي الذي يسير على وفقه الباحث ليحقق هدفه المنشود في بحثه، ونظرا لتعدد المناهج في اجراء البحوث في العلوم الانسانية وتباينها فأُن طُبِيعَة الدراسة وموضوعها والهدف منها هو الذي سيحدد نوعية البيانات الاختبارية المراد جمعها وطبيعة المنهج المستعمل في اجرائها (عبدالرحمن وزكنة، 2008: 15) (Abdul Rahman and Zanganah, 2008: 15) (في (رحيم، 2024: 6) (Rahem, 2024:6)).



فقد استخدم الباحث منهج البحث الوصفي (الدراسات المقارنة للأسباب) بوصفه منهجاً ارتكزت عليه الدراسة من أجل تحقيق أهدافها والتحقق من فرضياتها، تم اختيار هذا المنهج لما يتميز به من مميزات تساعد في تحقيق أهداف البحث الحالي (عودة، 1987: 169) (Awda, 1987: 169).

### مجتمع البحث

يشتمل مجتمع البحث الحالي على طلبة الصف الرابع في كليات التربية للاقسام غير الاختصاص\* في جامعة دي إلى التابعة لوزارة التعليم العالي والبحث العلمي من العراقيين الدارسين في الدراسة الصباحية للعام الدراسي 2024/2023 ، ويتكون المجتمع الاحصائي من (1820) طالب وطالبة موزعين بحسب الكلية والجنس ، اما بالنسبة للجنس؛ فقد بلغ عدد الطلبة الذكور (573) طالب وتشكل نسبة (32%) من المجتمع، في حين بلغ عدد الطالبات الاناث (1247) وتشكل نسبة (68%) من المجتمع الاحصائي، اما بالنسبة للكلية فقد بلغ عدد الطلبة في كلية التربية للعلوم الإنسانية (1389) طالبا وطالبة ، وتشكل نسبة (76%) من المجتمع، في حين بلغ عدد الطلبة في كلية التربية للعلوم الصرفة (431) طالبا وطالبة وتشكل نسبة (24%).

### عينة البحث

- لأجل الوصول إلى نتائج دقيقة يمكن الوثوق بها، يتطلب ذلك عدداً من العينات والتي تمثلت فيما يأتي:
- 1- عينة وضوح التعليمات وفهم الفقرات البالغ قوامها (60) طالباً وطالبة بواقع (15) ذكور و (15) اناث من طلبة الصف الرابع لقسم التاريخ في كلية التربية للعلوم الانسانية، و(15) ذكور و(15) اناث من طلبة الصف الرابع لقسم علوم الحياة في كلية التربية للعلوم الصرفة.
  - 2- عينة التحليل الاحصائي البالغ قوامها (600) طالباً وطالبة بواقع (132) للصف الرابع في قسم اللغة العربية و (99) في قسم التاريخ و (124) في قسم الجغرافية في كلية التربية للعلوم الانسانية وكذلك (70) طالبا وطالبة في قسم علو الحياة و (135) طالبا وطالبة في قسم الكيمياء و (40) طالبا وطالبة في قسم علوم الحاسوب في كلية التربية للعلوم الصرفة في جامعة ديالى. والجدول (1) يوضح ذلك
  - 3- عينة الثبات البالغ قوامها (600) فرداً.

### أداة البحث

يعد الاختبار التحصيلي الذي تنبأه الباحث وسيلة لجمع البيانات المطلوبة بمثابة مرتكز اساسي للبحث الحالي، إذ ان فقرات الاختبار تم سحبها من بنك الاسئلة المعد من قبل (الدليمي، 2009) ، يتكون من (60) فقرة تم تبنيها من بنك الأسئلة المعد من قبل (الدليمي، 2009)، إذ تم اختيار الفقرات على اساس المجالات الثلاثة (معرفة ، فهم ، تطبيق) ، وحسب النسب المئوية لبنك الأسئلة ولكل مجال.

### التحليل المنطقي

التحليل المنطقي لفقرات الاختبار يجب ان يقيس الاختبار ظاهرياً ما وضع لأجل قياسه، أي إن يتوصل المختصون اثناء فحص الاختبار إلى ان الاختبار قد قام بالقياس بشكل صحيح (البطش وابو زينة، 2007: 821) (AL-Batch & Abu Zina) في (جاسم ومحمود، 2022: 373) (Jasim & Mahmood, 2022: 373).

جدول رقم (1): عينة البحث موزعة حسب الكلية والتخصص

| الكلية                   | التخصص        | العدد | المجموع الكلي |
|--------------------------|---------------|-------|---------------|
| التربية للعلوم الانسانية | اللغة العربية | 132   | 22%           |
|                          | الجغرافية     | 124   | 21%           |
|                          | التاريخ       | 99    | 16%           |
| التربية للعلوم الصرفة    | علوم الكيمياء | 135   | 23%           |
|                          | علوم الحياة   | 70    | 12%           |
|                          | علوم الحاسوب  | 40    | 6%            |
|                          | المجموع       | 600   | 100%          |



كخطوة أولية في التحقق من صدق الفقرات وصلاحياتها في قياس ما أُعدت من أجله، اعتمد الباحث على الصدق الظاهري كأساس للحكم المنطقي عليها. وذلك من خلال عرض الاختبار بصورته الأولية على مجموعة من الخبراء المتخصصين في العلوم التربوية والنفسية والذي بلغ عددهم (9) محكماً، وبإجماع جميع الخبراء بنسبة (100%)، إذ تم قبول الاختبار كما هو وتم اعتماده من قبل الباحث.

### عينة التحليل الإحصائي لفقرات الاختبار

إن التحليل الإحصائي للفقرات هو خطوة أساسية في إعداد المقاييس والاختبارات التربوية والنفسية. وذلك لأنه يساعد في الكشف عن مدى دقة فقرات أداة القياس وقدرتها على قياس الخاصية المراد قياسها (فرج، 1980:331 - 332).  
و قد اقترحت نانلي (Nunnally, 1978) أن يكون الحجم المناسب لعينة التحليل الإحصائي بما لا يقل عن خمسة أفراد يقابلون كل فقرة اختبارية. وعلى هذا الأساس تكونت عينة التحليل الإحصائي من (600) طالباً وطالبة من طلبة (كلية التربية للعلوم الإنسانية وكلية التربية للعلوم الصرفة) للأقسام غير الاختصاص، والذي تم اختيارهم بأسلوب العينة (الطبقية العشوائية). وكما هو موضح في الجدول رقم (1)

### التحليل الإحصائي للفقرات

وإن هدف تجربة التحليل الإحصائي لفقرات الاختبار التحصيلي هو حساب (معاملات الصعوبة، معاملات التمييز، معاملات فعالية البدائل الخاطئة، معاملات الصدق) لفقرات الاختبار، وقد قام الباحث بالإجراءات الآتية:

#### أ\_ صعوبة الفقرات

ويرى كلّا من «داوني، وبلوم وآخرون» (Dawni, 1967; Bloom et al, 1971)، أن أفضل مدى مسموح به لمعاملات الصعوبة يتراوح ما بين (0.20-0.80) (Bloom et al., 1971: 305)، (Downi, 1967: 215).  
وبالرجوع إلى الأدلة المتعلقة بمعاملات الصعوبة، فضلاً عن نتائج الدراسات السابقة التي استخدمت اختبارات مشابهة، فإن المدى المناسب والمقبول لمعامل صعوبة الفقرات يتطابق مع ما أشار إليه «داوني وبلوم»، وبعد أن قام الباحث بتصحيح إجابات أفراد العينة على الاختبار التحصيلي.

#### ب\_ تمييز الفقرات

بعد أن قام الباحث بحساب معاملات التمييز لفقرات، اتبع نفس الخطوات في حساب معاملات الصعوبة، الفرق يكمن في طريقة الحساب، إذ يتم حساب معامل التمييز من خلال الفرق بين عدد الإجابات الصحيحة في المجموعة العليا والمجموعة الدنيا، ثم تقسيم هذا الفرق على عدد أفراد إحدى المجموعتين. وقد أرفق الباحث معاملات الصعوبة والتمييز لحجم العينة الكلي كما يلي: كما موضح في الجدول رقم (2)

#### يتبين من الجدول (2) ما يأتي:

- 1- إن مدى معاملات الصعوبة لفقرات الاختبار التحصيلي يتراوح ما بين (-0.42 0.73) بمتوسط صعوبة مقداره (0.60) مما يدل على أن الفقرات متوسطة الصعوبة، وهذا يشير إلى أن معاملات الصعوبة هي ضمن المدى المقبول لمعامل الصعوبة الذي يتراوح بين (0.20 - 0.80).
- 2- أما مدى معاملات التمييز لفقرات الاختبار؛ فقد تراوح ما بين (0.17 - 0.47) بمتوسط تمييز مقداره (0.36) وهذا يدل على أن تمييز (58) فقرة ضمن المستوى المحدد من إيبيل، عدا فقرتين وهي (18، 52) لم تكن ضمن المستوى المحدد.

### الخصائص السيكومترية للاختبار التحصيلي

إنّ دقة البيانات والدرجات التي يحصل عليها الباحث من المقاييس والاختبارات تعتمد بشكل رئيس على خاصيتي الثبات والصدق. لذلك، أولى المختصون في مجال القياس التربوي والنفسي اهتماماً كبيراً بهاتين الخاصيتين، إذ سعوا إلى تطوير وتحسين دقة المقاييس والاختبارات من خلال دراسة وتحليل الخصائص السيكومترية لهذه الأدوات، بما يضمن قياس ما وضعت لقياسه بأقل قدر ممكن من الخطأ (عبد الرحمن، 1998: 227).

أولاً - الصدق (Validity):

الجدول ( 2 ): معاملات الصعوبة والتمييز لفقرات الاختبار التحصيلي

| معامل التمييز | معامل الصعوبة | عدد الاجابات الصحيحة |        | ت الفقره | معامل التمييز | معامل الصعوبة | عدد الاجابات الصحيحة |        | ت الفقره |
|---------------|---------------|----------------------|--------|----------|---------------|---------------|----------------------|--------|----------|
|               |               | الدنيا               | العليا |          |               |               | الدنيا               | العليا |          |
| 0.46          | 0.62          | 24                   | 99     | 31       | 0.41          | 0.53          | 42                   | 109    | 1        |
| 0.37          | 0.54          | 44                   | 104    | 32       | 0.31          | 0.58          | 42                   | 93     | 2        |
| 0.47          | 0.62          | 24                   | 100    | 33       | 0.36          | 0.47          | 57                   | 115    | 3        |
| 0.36          | 0.56          | 41                   | 100    | 34       | 0.37          | 0.45          | 59                   | 119    | 4        |
| 0.30          | 0.60          | 40                   | 89     | 35       | 0.37          | 0.50          | 51                   | 111    | 5        |
| 0.38          | 0.60          | 35                   | 96     | 36       | 0.39          | 0.42          | 62                   | 125    | 6        |
| 0.31          | 0.52          | 35                   | 103    | 37       | 0.39          | 0.63          | 29                   | 92     | 7        |
| 0.36          | 0.63          | 31                   | 90     | 38       | 0.45          | 0.44          | 55                   | 128    | 8        |
| 0.33          | 0.60          | 38                   | 92     | 39       | 0.43          | 0.57          | 35                   | 104    | 9        |
| 0.32          | 0.69          | 24                   | 76     | 40       | 0.38          | 0.65          | 27                   | 88     | 10       |
| 0.38          | 0.65          | 26                   | 88     | 41       | 0.38          | 0.59          | 36                   | 98     | 11       |
| 0.38          | 0.55          | 42                   | 103    | 42       | 0.38          | 0.51          | 48                   | 110    | 12       |
| 0.31          | 0.70          | 23                   | 73     | 43       | 0.44          | 0.69          | 15                   | 87     | 13       |
| 0.38          | 0.58          | 37                   | 98     | 44       | 0.31          | 0.69          | 25                   | 75     | 14       |
| 0.36          | 0.56          | 42                   | 101    | 45       | 0.30          | 0.60          | 40                   | 88     | 15       |
| 0.36          | 0.60          | 35                   | 94     | 46       | 0.36          | 0.56          | 41                   | 100    | 16       |
| 0.38          | 0.61          | 32                   | 93     | 47       | 0.38          | 0.48          | 53                   | 114    | 17       |
| 0.34          | 0.60          | 38                   | 93     | 48       | 0.17          | 0.69          | 37                   | 64     | 18       |
| 0.38          | 0.67          | 23                   | 85     | 49       | 0.37          | 0.64          | 29                   | 89     | 19       |
| 0.36          | 0.47          | 56                   | 115    | 50       | 0.33          | 0.65          | 30                   | 83     | 20       |
| 0.37          | 0.60          | 34                   | 94     | 51       | 0.30          | 0.70          | 24                   | 73     | 21       |
| 0.18          | 0.58          | 54                   | 83     | 52       | 0.38          | 0.63          | 29                   | 91     | 22       |
| 0.36          | 0.60          | 36                   | 95     | 53       | 0.31          | 0.64          | 34                   | 84     | 23       |
| 0.35          | 0.62          | 33                   | 89     | 54       | 0.34          | 0.61          | 36                   | 91     | 24       |
| 0.38          | 0.60          | 35                   | 96     | 55       | 0.30          | 0.64          | 34                   | 83     | 25       |
| 0.36          | 0.65          | 28                   | 86     | 56       | 0.32          | 0.71          | 21                   | 73     | 26       |
| 0.37          | 0.54          | 45                   | 105    | 57       | 0.38          | 0.60          | 34                   | 95     | 27       |
| 0.30          | 0.64          | 34                   | 83     | 58       | 0.36          | 0.66          | 25                   | 84     | 28       |
| 0.33          | 0.62          | 35                   | 89     | 59       | 0.31          | 0.60          | 39                   | 89     | 29       |
| 0.42          | 0.62          | 28                   | 96     | 60       | 0.39          | 0.66          | 23                   | 86     | 30       |

الصدق من أهم الخصائص السيكومترية التي يجب أن تتوافر في المقاييس والاختبارات النفسية. ذلك أن الصدق يُشير إلى قدرة الاختبار أو المقياس على قياس ما أُعد لقياسه بالفعل (Harrison,1983,11) وكذلك فإن الصدق يُعبّر بدقة عن مدى تمكن الاختبار من تحقيق الهدف الذي وُضع من أجله (عودة، 1998: 335). وقد حقق الباحث صدق البناء كما يأتي:

صدق البناء:

جدول (3): قيمة معامل الثبات حسب حجم العينة

| ت | حجم العينة | الثبات |
|---|------------|--------|
| 1 | 200        | 0,840  |
| 2 | 400        | 0,871  |
| 3 | 600        | 0.879  |

يُعد صدق التكوين الفرضي، أو ما يُسمى أيضًا بصدق البناء أو المفهوم، مؤشرًا على مدى قياس الاختبار لتكوين فرضي معين أو مفهوم نفسي محدد (ربيع، 1994:98)، وعندما تتطابق الدرجات التجريبية للاختبار مع الافتراضات النظرية التي استند عليها الباحث في بنائه، فإن ذلك يُعد دليلًا على أن الاختبار يقيس السمة أو المفهوم الذي أُعد لقياسه، وفقًا لتلك الافتراضات (Allen and Yen, 1976:108). وقد تحقق الباحث من صدق التكوين الفرضي للاختبار من خلال استنتاج بعض المؤشرات والدلائل، والتي تمثلت في:

1- صعوبة الفقرات

2- القوة التمييزية للفقرات

### ثانيا - الثبات (Reliability):

الثبات يُعد أحد المؤشرات الأساسية التي تستخدم للتأكد من دقة المقياس أو الاختبار وانسجام فقراته في قياس الخاصية المستهدفة. إذ يُشير الثبات إلى درجة الاستقرار والاتساق الداخلي التي يتمتع بها الاختبار. والغرض من حساب الثبات هو تقدير مدى أخطاء الاختبار بهدف التقليل منها (Murphy, 1988: 63).

ولحساب الثبات استخدم الباحث معادلة (ألفا كرونباخ) بالاعتماد على بيانات عينة التحليل الإحصائي للمفردا، إذ تم حساب معامل الثبات لأحجام العينة الثلاث كما موضح في الجدول رقم (3) ويعد معامل الثبات في الحالات الثلاث ولكلا المعاملين بمثابة معامل ثبات عال، كونهم قريبين من (1)، وهي أعلى قيمة ممكن الحصول عليها للثبات.

### الوسائل الإحصائية

الوسائل التي قام الباحث بحسابها عن طريق برنامج الاكسل (Excel).

1- معادلة صعوبة الفقرات.

2- معادلة تمييز الفقرات.

3- معامل الثبات كيوذر 21.

الوسائل التي قام الباحث بحسابها من خلال الحقيبة الإحصائية للعلوم الاجتماعية (spss).

- معامل الثبات ألفا كرونباخ
- تحليل التباين الاحادي
- اختبار شيفيه للمقارنات البعدية
- الاختبار التائي لعينتين مترابطتين

### مناقشة النتائج

هدف البحث: يتمثل بمعرفة تأثير اختلاف حجم العينة على حساب الخطأ المعياري الشرطي للاختبارات التحصيلية عند استخدام طريقتين من طرائق حساب الخطأ المعياري الشرطي بغية الوصول إلى أفضل حجم لعينة بحيث يكون مناسب لحساب هذه الطرق.

ولتحقيق هذا الهدف وضع الباحث الفرضيتين الصفريتين الاتيتين:

الأولى: لا توجد فروق ذات دلالة إحصائية بين طرائق حساب الخطأ المعياري الشرطي.

ولتحقيق هذا الهدف وأختبار الفرضية الصفرية قام الباحث باتباع الآتي:

تم حساب المتوسطات الحسابية والانحرافات المعيارية للأخطاء المعيارية في تقدير الخطأ المعياري الشرطي في القياس بأستخدام طريقتي ( لورد و كيتس)، وكذلك حساب الثبات من خلال معامل (ألفا كرونباخ و كيوذر 21) حسب حجم العينة، فضلاً عن ذلك فقد استخدم

**الجدول (4):** يبين نتائج الاختبار التائي للعينات المترابطة للمتوسطين الحسابيين للأخطاء المعيارية في تقدير الخطأ المعياري الشرطي في القياس بأستخدام طريقتي ( لورد و كيتس) تبعا لحجم العينة

| حجم العينة | الطريقة | العينة | المتوسط | الأنحراف المعياري | القيمة التائية |          | درجة الحرية | مستوى الدلالة عند مستوى 0.001 |
|------------|---------|--------|---------|-------------------|----------------|----------|-------------|-------------------------------|
|            |         |        |         |                   | المحسوبة       | الجدولية |             |                               |
| 200        | لورد    | 200    | 3.324   | 0.141             | 208.433        |          | 199         | دالة                          |
|            | كيتس    | 200    | 1.801   | 0.0381            |                |          |             |                               |
| 400        | لورد    | 400    | 3.236   | 0.157             | 255.142        |          | 399         | دالة                          |
|            | كيتس    | 400    | 1.780   | 0.043             |                |          |             |                               |
| 600        | لورد    | 600    | 3.227   | 0.225             | 220.948        | 3,291    | 599         | دالة                          |
|            | كيتس    | 600    | 1.764   | 0.063             |                |          |             |                               |

**جدول (5):** نتائج تحليل التباين الاحادي للتعرف على مصدر الفرق بين حجوم العينة الثلاث لطرائق حساب الخطأ المعياري الشرطي في القياس

| الطريقة | مصدر التباين   | مجموع المربعات S.S | درجة الحرية d.F | متوسط المربعات S.M | F      | مستوى الدلالة عند مستوى 0.001 |
|---------|----------------|--------------------|-----------------|--------------------|--------|-------------------------------|
| لورد    | بين المجموعات  | 1.474              | 2               | 0.737              | 19.921 | دالة                          |
|         | داخل المجموعات | 44.298             | 1197            | 0.037              |        |                               |
|         | الكل           | 45.573             | 1199            |                    |        |                               |
| كيتس    | بين المجموعات  | 0.216              | 2               | 0.108              | 37.728 | دالة                          |
|         | داخل المجموعات | 3.431              | 1197            | 0.003              |        |                               |
|         | الكل           | 3.647              | 1199            |                    |        |                               |

\*القيمة الفائتة الجدولية عند مستوى دلالة (0.001) ودرجة حرية (3 - 1197) تساوي (6.95).

الباحث الاختبار التائي لعينتين مترابطتين للكشف عن جوهرية الفرق الظاهري بين المتوسطين الحسابيين للأخطاء المعيارية ولكل حجم من احجام العينة الثلاث ، إذ كانت نتائج الاختبار التائي كما هي موضحة في الجدول (4).

ويتضح من الجدول اعلاه أن القيمة التائية المحسوبة أكبر من القيمة التائية الجدولية عند مستوى دلالة (0.001) وبدرجة حرية (599,399,199)، مما يشير ذلك إلى وجود فروق ذات دلالة إحصائية بين المتوسطين الحسابيين للأخطاء المعيارية بأستخدام طريقتي ( لورد و كيتس)، لصالح عملية عملية تقدير الخطأ المعياري الشرطي في القياس بأستخدام طريقة (كيتس) ، وذلك لأنخفاض قيمة المتوسط الحسابي لطريقة (كيتس) مقارنة بطريقة (لورد) ولكل حجم من احجام العينة.

الثانية: لا توجد فروق ذات دلالة احصائية بين احجام العينة الثلاث (200,400,600) لطرائق حساب الخطأ المعياري الشرطي في القياس ( لورد ، كيتس).

ولتحقيق هذا الهدف وأختبار الفرضية الصفرية قام الباحث باتباع الآتي:

تم استخدام تحليل التباين الاحادي للكشف عن جوهرية الفرق الظاهري بين المتوسطات الحسابية للأخطاء المعيارية في تقدير الخطأ المعياري الشرطي في القياس بأستخدام طريقتي (لورد و كيتس) تبعا لحجم العينة. والجدول (5) يوضح ذلك ويتضح من الجدول أعلاه أن القيمة الفائتة المحسوبة لطريقة لورد والبالغة (19.921)، والقيمة الفائتة المحسوبة لطريقة كيتس والبالغة (37,728) هي أكبر من القيمة الفائتة الجدولية، وهذا يعني أنه توجد فروق ذات دلالة احصائية بين طريقتي حساب الخطأ المعياري الشرطي في القياس (لورد وكيتس) ، وباستخدام اختبار شيفيه للمقارنات البعدية الذي يوضح الفروق بين متوسطات الطريقتين، وكما موضح في الجدول (8)، والذي يعتمد على متوسطات الخطأ المعياري الشرطي للطريقتين وحسب حجم العينة ، فضلاً عن حساب الثبات بواسطة معامل (الفا كرونباخ وكيودر 21) كونهم احد متطلبات طريقتي لورد وكيتس كما موضح في جدول (7) ، وتم حساب المتوسطات الحسابية والانحرافات المعيارية لطرائق حساب الخطأ المعياري الشرطي في القياس وحسب حجم العينة كما في الجدول رقم (6).

الجدول (6): المتوسطات الحسابية والانحرافات المعيارية لطريقتي (لورد ، كيتس) وحسب حجم العينة

| الطريقة | حجم العينة | المتوسط الحسابي | الانحراف المعياري |
|---------|------------|-----------------|-------------------|
| لورد    | 200        | 3.32422         | 0.14147           |
|         | 400        | 3.23602         | 0.15735           |
|         | 600        | 3.22724         | 0.22541           |
| كيتس    | 200        | 1.80090         | 0.03811           |
|         | 400        | 1.77992         | 0.04322           |
|         | 600        | 1.76399         | 0.06325           |

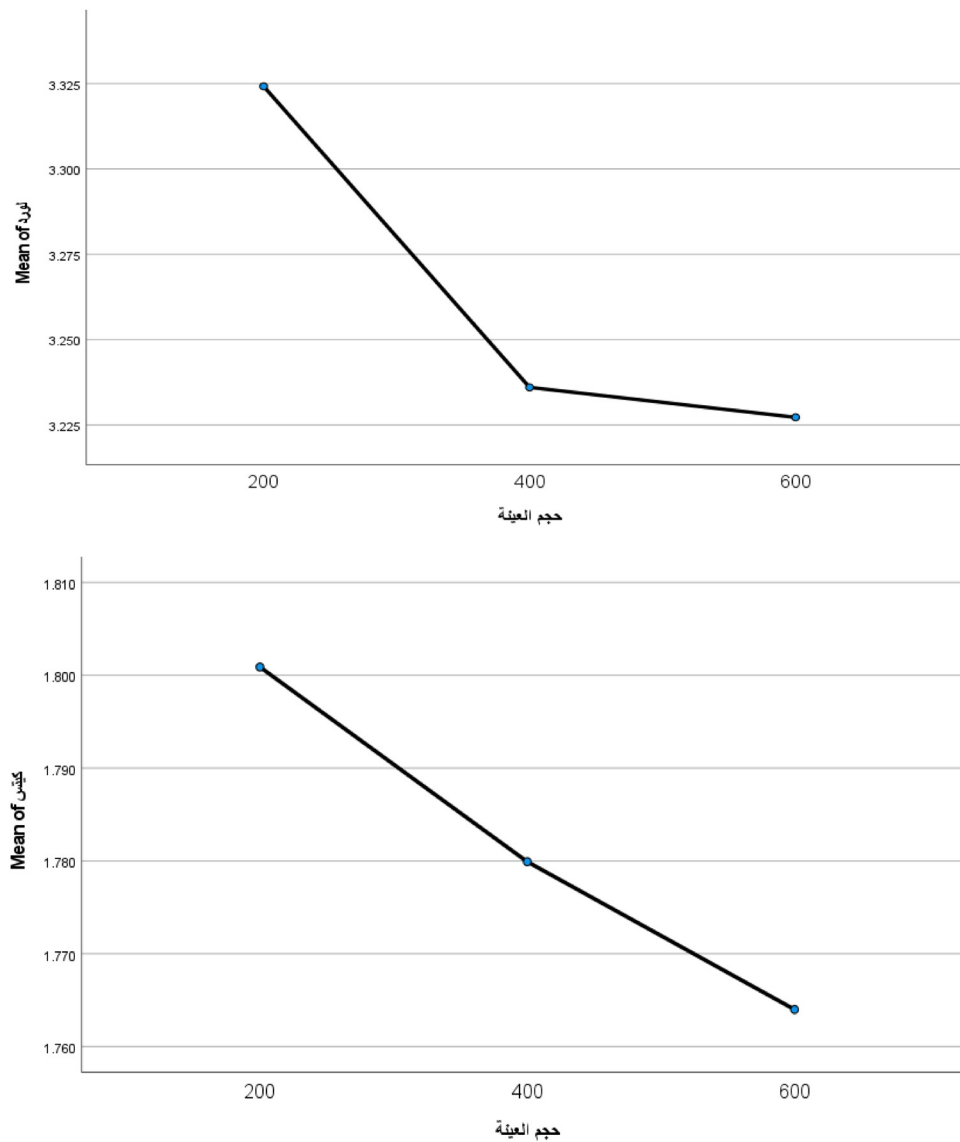
الجدول (7): حساب الثبات بواسطة معامل (الفا كرونباخ وكيودر 21)

| ت | حجم العينة | الثبات (ألفا كرونباخ) | الثبات (كيودر 21) |
|---|------------|-----------------------|-------------------|
| 1 | 200        | 0,840                 | 0.837             |
| 2 | 400        | 0,871                 | 0.868             |
| 3 | 600        | 0.879                 | 0.874             |

جدول (8): نتائج اختبار شيفيه للمقارنات البعدية للتعرف على مصدر الفرق بين أحجام العينة الثلاث لطرائق حساب الخطأ المعياري الشرطي في القياس

| الطريقة | حجم العينة | المتوسط الحسابي | الفرق بين المتوسطات | قيمة شيفيه الحرجة | الدلالة الأحصائية         |
|---------|------------|-----------------|---------------------|-------------------|---------------------------|
| لورد    | 200        | 3.32422         | 0.088201            | 0,0027            | دالة لصالح حجم العينة 400 |
|         | 400        | 3.23602         |                     |                   |                           |
|         | 200        | 3.32422         | 0.096974            | 0,0024            | دالة لصالح حجم العينة 600 |
|         | 600        | 3.22724         |                     |                   |                           |
|         | 400        | 3.23602         | 0.008773            | 0,0001            | دالة لصالح حجم العينة 600 |
|         | 600        | 3.22724         |                     |                   |                           |
| كيتس    | 200        | 1.80090         | 0.020980            | 0,0006            | دالة لصالح حجم العينة 400 |
|         | 400        | 1.77992         |                     |                   |                           |
|         | 200        | 1.80090         | 0.036906            | 0,0009            | دالة لصالح حجم العينة 600 |
|         | 600        | 1.76399         |                     |                   |                           |
|         | 400        | 1.77992         | 0.015926            | 0,0003            | دالة لصالح حجم العينة 600 |
|         | 600        | 1.76399         |                     |                   |                           |

يتضح من الجدول أعلاه أن قيمة شيفيه الحرجة عند مستوى دلالة (0.05) جاءت أصغر من المقارنات بين المتوسطات جميعها بالنسبة لطريقة (لورد ، كيتس)، وحسب حجم العينة (200,400,600)، إذ كانت الفروق لصالح المتوسط الحسابي الأقل للأخطاء المعيارية ولكل حجم عينة، ولصالح حجم العينة الأكبر، ويتضح من ذلك أنه كلما زاد حجم العينة قلت قيمة الخطأ المعياري الشرطي في القياس باستخدام طريقة (لورد ، كيتس). وكما موضح في الشكل رقم (1)



الشكل (1): يوضح تأثير اختلاف حجم العينة على طرائق حساب الخطأ المعياري الشرطي للاختبارات التحصيلية (لورد ، كيتس)

## الاستنتاجات

### 1- تأثير حجم العينة:

النتائج أظهرت أن زيادة حجم العينة من (200) إلى (400) ثم إلى (600) قد أدت إلى تقليل الخطأ المعياري الشرطي في القياس، وهذا يشير إلى أن البيانات الأكثر تمثيلاً تؤدي إلى تقديرات أكثر دقة وموثوقة.

### 2- الخطأ المعياري الشرطي:

أظهرت النتائج أن طريقة (كيتس) لحساب الخطأ المعياري الشرطي كانت الأفضل مقارنة بطريقة (لورد) المستخدمة في البحث، إذ كانت متوسطات الخطأ المعياري أقل بشكل ملحوظ، مما يشير إلى دقة أكبر في تقدير الأخطاء.

## التوصيات

في ضوء النتائج التي توصل إليها البحث الحالي ، يوصي الباحث بما يأتي:

- 1- دعوة للاستمرار في استخدام المتنوع لطرائق حساب الخطأ المعياري الشرطي في القياس في المقاييس النفسية والاختبارات التحصيلية.

- 2- حث الباحثين في التخصصات التربوية والنفسية على استخدام طريقة (كيتس) في حساب الخطأ المعياري الشرطي في القياس كونها تعطي خطأ معياري منخفض للوصول إلى أعلى مستويات الدقة في قياس الخواص النفسية والتربوية.
- 3- تطوير أدوات قياس يمكن ان تأخذ في الاعتبار حجم العينة عند حساب الخطأ المعياري الشرطي في القياس، لضمان دقة أكبر في نتائج الاختبارات.

## المقترحات

- من خلال نتائج البحث الحالي ، يقترح الباحث بعض الدراسات والبحوث الآتية:
- 1- توسيع نطاق الدراسة ليشمل أحجام عينات إضافية مثل (800، 1000) من أجل تحليل كيفية تأثيرها على الخطأ المعياري الشرطي في القياس المحسوب بطريقة (لورد وكيتس).
  - 2- إجراء دراسة مماثلة تأخذ بالحسبان تأثير (طريقة جمع البيانات ، نوع الاختبار، طبيعة العينة) على حساب الخطأ المعياري الشرطي في القياس.
  - 3- إجراء دراسة للمقارنة بين طريقة لورد وكيتس في حساب الخطأ المعياري الشرطي في القياس تبعاً لسلام التقدير (لفظي ، عددي ، لفظي عددي).

## المصادر

### المصادر العربية:

- أبو ناهية، صلاح الدين محمود (1994). القياس التربوي ، ط1، القاهرة ، مكتبة الانجلو للطباعة والنشر والتوزيع.
- البطش وابو زينة، (2007). منهجية البحث العلمي، تصميم البحث والتحليل الإحصائي، دار المسرة للنشر والتوزيع والطباعة، عمان.
- بلال، زيان يحيى (2023). الخصائص القياسية لاختبار التفكير التقني لطلبة التعليم المهني لاستخدام النموذج اللوجستي ثلاثي البارامتر، مجلة الاستاذ للعلوم الانسانية، مجلد(62)، العدد(1)، بغداد.
- جاسم، خالد جمال و محمود، محمد انور، (2022). استخدام نظرية المقياس النفسي في صياغة مقياس جيليام GARS-3 لتشخيص اضطراب طيف التوحد، مجلة العلوم الانسانية، مجلد(61)، العدد(4)، بغداد.
- جاسم، خالد جمال، عيال، ياسين حميد (2019). الخصائص القياسية لاختبار الاستدلال الاستقرائي لدى طلبة الجامعة على وفق نظرية الاستجابة للمفردة الاختبارية، مجلة الاستاذ للعلوم الانسانية والاجتماعية، المجلد (58)، العدد(4)، بغداد.
- جلال، سعد، (2001): القياس النفسي المقاييس والاختبارات، القاهرة، دار الفكر العربي.
- خليل، عبدالقادر صباح، (2021). دراسة مقارنة في الخصائص السيكمومترية لاختبار تحصيلي ذي اختيار من متعدد باختلاف طرق تقدير درجات المفردات وفقاً لنظرية القياس لاختبارات القدرات العقلية وفقاً لنظرية القياس الكلاسيكية، رسالة ماجستير، جامعة بغداد، كلية التربية ابن رشد، بغداد.
- ربيع، محمد شحاتة (1994): قياس الشخصية ، دار المعرفة، القاهرة.
- رحيم، هند صبيح (2024). بناء وتطبيق مقياس شخصية ذاتية الغرض لدى طلبة الجامعة، مجلة الاستاذ للعلوم الانسانية والاجتماعية، مجلد(63) ، العدد (1)، بغداد.
- الطاهات، خالد بسام (2016): المقارنة بين طرق تقدير الخطأ المعياري الشرطي في القياس تبعاً لطريقة التصحيح لاثـر التخمين وعدد البدائل، رسالة ماجستير، جامعة مؤتة، كلية الدراسات العليا، الاردن.
- الطريري، عبد الرحمن بن سليمان (1997): القياس النفسي والتربوي (نظرياته، اسسه، تطبيقاته) ، ط1 ، المملكة العربية السعودية، مكتب الرشيد للنشر والتوزيع.
- عباس، محمد خليل (1993): المقارنة بين خمس طرق لتقدير الخطأ المعياري الشرطي في القياس عند مستويات محددة لعلامات الاختبار، اطروحة دكتوراه، كلية الدراسات العليا، الجامعة الاردنية.
- عبد الرحمن (1998): القياس النفسي النظرية والتطبيق، ط3، دار الفكر العربي، القاهرة.
- عبد السلام، نادية محمد (1987): دراسة في حساب ثبات وصدق اختبار هـ في المرجع، مجلة علم النفس، الهيئة المصرية العامة للكتاب، العدد (2).
- علام (2006): الاختبارات والمقاييس التربوية والنفسية، عمان، دار الفكر العربي للنشر والتوزيع.
- علام، صلاح الدين محمود (2005). نماذج الاستجابة للمفردة الاختبارية احادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي، دار الفكر العربي، القاهرة.
- عودة (1998). القياس والتقويم في العملية التدريسية، ط2، اربد، دار الامل للنشر والتوزيع، عمان.
- عودة، احمد، فتحي، حسن مكاي، (1987) أساسيات البحث العلمي في التربية والعلوم الإنسانية، مكتبة المنار، عمان.



العيثاوي، محمد منعم، (2014). اثر اسلوب اختيار العينة وحجمها على الخصائص القياسية لاختبارات القدرات العقلية على وفق نظرية السمات الكامنة، اطروحة دكتوراه، جامعة بغداد، كلية التربية ابن رشد، بغداد.  
 فرج، صفوت (1980)، القياس النفسي، ط1، دار الفكر العربي، القاهرة.  
 المياحي، أمل اسماعيل، (2005): أثر اختلاف حجم العينة وطول المقياس في الخصائص السيكومترية لمقاييس الشخصية، (اطروحة دكتوراه)، بغداد.

### Arabic sources translated into English

- Abu Nahia, Salah Al-Din Mahmoud (1994): **Educational Measurement**, 1st Edition, Cairo, Anglo Library for Printing, Publishing, and Distribution.
- Al-Batch and Abu Zina (2007): **Methodology of Scientific Research, Research Design and Statistical Analysis**, Dar Al-Masara for Publishing, Distribution, and Printing, Amman.
- Abdul Rahman (1998): **Psychological Measurement: Theory and Application**, 3rd Edition, Cairo, Dar Al-Fikr Al-Arabi.
- Bilal, Ziyah Yahya (2023): **Measurement Properties of the Technical Thinking Test for Vocational Education Students Using the Three-Parameter Logistic Model**, Alustath Journal for Human Sciences, Volume (62), Issue (1), Baghdad.
- Jasim, Khalid Jamal and Mahmood, Mohammad Anwar (2022): **Using Psychometric Scale Theory to Formulate the Gilliam GARS-3 Scale for Diagnosing Autism Spectrum Disorder**, Journal of Human Sciences, Volume (61), Issue (4), Baghdad.
- Jasim, Khalid Jamal, Ayāl, Yaseen Hamid (2019): **The Measurement Properties of the Inductive Reasoning Test Among University Students According to Item Response Theory**, Alustath Journal for Human and Social Sciences, Volume (58), Issue (4), Baghdad.
- Jalal, Saad (2001): **Psychological Measurement: Scales and Tests**, Cairo, Dar Al-Fikr Al-Arabi.
- Khalil, Abdul-Qadir Sabah (2021): **A Comparative Study of the Psychometric Properties of a Multiple-Choice Achievement Test with Different Methods for Estimating Item Scores According to Classical Test Theory**, Master's Thesis, University of Baghdad, Ibn Rushd College of Education, Baghdad.
- Rabea, Mohammad Shehata (1994): **Personality Measurement**, Dar Al-Ma'rifa, Cairo.
- Rahim, Hind Sabeih (2024): **Construction and Application of a Self-Purpose Personality Scale for University Students**, Alustath Journal for Human and Social Sciences, Volume (63), Issue (1), Baghdad.
- Al-Tahat, Khaled Bassam (2016): **Comparison between Methods for Estimating Conditional Standard Error in Measurement According to Correction Method for Guessing Effect and Number of Alternatives**, Master's Thesis, Mu'tah University, Graduate Studies, Jordan.
- Al-Turaiir, Abdul Rahman bin Suleiman (1997): **Psychological and Educational Measurement (Theories, Foundations, Applications)**, 1st Edition, Kingdom of Saudi Arabia, Al-Rasheed Publishing and Distribution.
- Abbas, Mohammad Khalil (1993): **Comparison between Five Methods for Estimating Conditional Standard Error in Measurement at Specific Levels of Test Scores**, PhD Thesis, Graduate Studies, Jordan University.
- Abdul Rahman, (1998). **Psychological Measurement, Theory and Application**, 3rd Edition, Dar AL-Fikr AL-Arabi, Cairo.
- Abdel Salam, Nadia Mohammed (1987): **A Study in Calculating the Reliability and Validity of a Criterion-Referenced Test**, *Journal of Psychology*, Egyptian General Book Organization, Issue (2).
- Alam (2006): **Educational and Psychological Tests and Scales**, Amman, Dar Al-Fikr Al-Arabi for Publishing and Distribution.
- Alam, Salah Al-Din Mahmoud (2005): **Unidimensional and Multidimensional Item Response Models and Their Applications in Psychological and Educational Measurement**, Dar Al-Fikr Al-Arabi, Cairo.
- Awda (1998): **Measurement and Evaluation in the Teaching Process**, 2nd Edition, Irbid, Dar Al-Amal for Publishing and Distribution, Amman.
- Awda, Ahmed, Fathi, Hassan Makawi (1987): **Fundamentals of Scientific Research in Education and Humanities**, Al-Manar Library, Amman.
- Al-Ithaawi, Mohammad Moneim (2014): **The Effect of Sampling Method and Size on the Measurement Properties of Mental Ability Tests According to the Latent Trait Theory**, PhD Thesis, University of Baghdad, Ibn Rushd College of Education.
- Farag, Safwat (1980): **Psychological Measurement**, 1st Edition, Dar Al-Fikr Al-Arabi, Cairo.
- Al-Miyahi, Amal Ismail (2005): **The Effect of Sample Size and Scale Length on the Psychometric Properties of Personality Scales**, (PhD Thesis), Baghdad.

## المصادر الأجنبية

- Allen M.J & Yen,W.M (1976). **psychological testing**. 4th ed, New York, Macmillan.
- Bloom ,B.S .et al .(1971) .**Handbook on Formative and summative Evaluation of student Leaving**. Mc. Graw – Hill, New York.
- Crocker, l. & algian, j., (1986): **introduction to classical and modern test theory**, new york, cbs college publishing.
- Feldt, L.S. and Steffen, M.and Gubta, C.N. (1985) **A comparison of several methods for estimating the standard error of measurement at specific score levels**.
- Gulliksen, H.(1967). **The theory of mental tests**. New York Wiley. pp 130-142.
- Harrison. A. (1983). **A Language Testing Handbook**. London, The Macmillan Press.
- Murphy, K.R. & Davidshofer, C.O.(1988). **Psychological Testing Principles And Applications**. Prentice-Hall International, Inc.