معامل ألفا للتحقق من ثبات درجات أدوات القياس بين الحقائق والمعتقدات الخاطئة لدى الباحثين

أ.د. السيد محمد أبو هاشم حسن*

المستخلص

تهدف الدراسة الحالية إلى رصد وتحليل أهم المعتقدات الخاطئة الشائعة حول معامل ألفا لدى الباحثين، وأهم هذه المعتقدات: (١) أن كرونباخ هو أول من قدم معامل ألفا، (٢) استخدام ألفا دون التحقق من مسلماته، (٣) معامل ألفا يساوي ثبات الدرجات ،(٤) القيمة المرتفعة لمعامل ألفا مؤشر للاتساق الداخلي والتجانس، (٥) يمكن تحسين ثبات درجات المقياس عن طريق حذف بعض البنود، (٦) لابد أن تساوي أو تزيد قيمة معامل ألفا عن (٠.٧ أو ٠.٨) ، (٧) معامل ألفا هو الأفضل بين كل معاملات الثبات المنشورة، (٨) ألفا خاصية ثابتة لوصف المقياس، (٩) يقيس معامل ألفا الاتساق الداخلي للمقياس، (١٠) معامل ألفا الأكبر هو الأفضل دائماً، (١١) قيمة معامل ألفا تمتد بين الصفر والواحد ، (١٢) معامل ألفا دال إحصائياً ،(١٣) زيادة عدد البنود يؤدى إلى قيم مرتفعة لمعامل ألفا، (١٤) زيادة عدد البدائل يؤدى إلى قيم مرتفعة لمعامل ألفا، (١٥) اختلاف اتجاه البند يؤدى إلى قيم مرتفعة لمعامل ألفا، (١٦) زيادة حجم العينة يؤدى إلى قيم مرتفعة لمعامل ألفا. وفي النهاية تم تقديم مجموعة من التوصيات للعمل على تصحيح هذه المعتقدات الخاطئة لدى الباحثين.

الكلمات المفتاحية: الثبات، معامل ألفا، الاتساق الداخلي، المعتقدات الخاطئة.

مدخل الدراسة :

يتم تقدير ثبات درجات أداة القياس بناء على أربع طرق: طريقة الإعادة التي تستهدف تقدير استقرار الدرجات، وطريقة الصور المتكافئة التي تستهدف تقدير التكافؤ، وطريقة التجزئة النصفية التي تستهدف تقدير الاتساق، وأخيراً طريقة الاتساق الداخلي التي تستهدف تقدير التجانس والاتساق. ولقد كانت طريقة التجزئة النصفية باستخدام تصحيح سبيرمان – بروان، وإلى وقت قريب، أكثر طرق تقدير الثبات استخداماً لأسباب عملية في الغالب أنها تقوم على تطبيق الأداة مرة واحدة. ومع تطور البرامج الإحصائية أصبح معامل ألفا Coefficient Alpha هو الأكثر استخداماً بدون منازع في جميع مجالات البحوث النفسية والتربوية والاجتماعية.

ويلخص Dorst (2011:109) طرق التحقق من الثبات الأكثر شيوعاً في بحوث العلوم الاجتماعية في ثلاثة طرق رئيسة هي: الاتساق الداخلي Internal Consistencyويشمل التجزئة النصفية-Split Half، وثبات المقيمين Inter-rater، وألفا لكرونباخ Cronbach Alpha. والتكافؤ Equivalence ويشمل الصور البديلة Alternative Forms، والاستقرار عبر الزمن Stability over timeويشمل إعادة التطبيق Test-Retest.

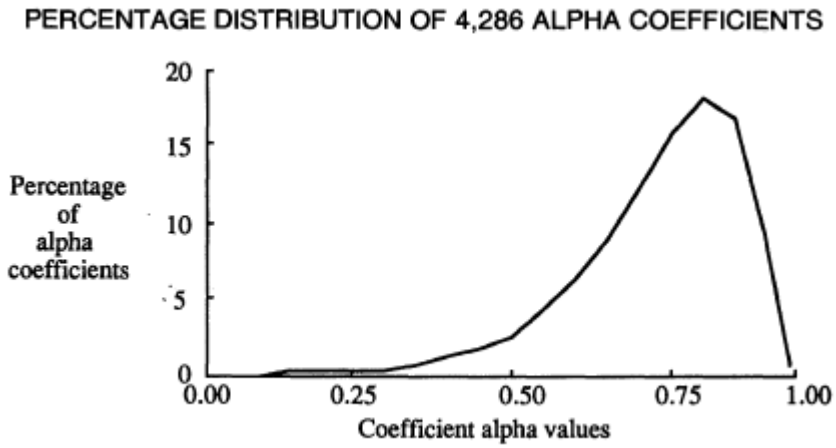
* أستاذ علم النفس – كلية التربية – جامعتي الملك سعود و الزقازيق
البريد الالكتروني:shashem@ksu.edu.sa

ويقدم كل من Souza. , Alexandre. & Guirardello. (2017:653) وصفاً لثلاثة أنواع من الثبات، وهي: (أ) الاستقرار Stability ويشمل إعادة التطبيق Test-Retest، ويعنى استقرار القياس بمرور الوقت، ويمكن التحقق من ذلك بحساب معامل الارتباط بين التطبيقين وقيمته هي مؤشر لمعامل الثبات، وكذلك معامل الارتباط بين الطبقات أو الداخلي Intraclass correlation coefficient (ICC)، (ب) الاتساق الداخلي Internal Consistency ويعنى ارتباط بنود الاختبار بعضها ببعض. وأكثر الطرق استخداماً للتحقق من ذلك ألفا لكرونباخ (المتغيرات المستمرة)،و كيودر ريتشاردسون (المتغيرات الثنائية)، (ج) التكافؤ Equivalence ويعنى درجة الاتفاق أو الانسجام بين أثنين أو أكثر من المقيمين بشأن الدرجة على الأداة، ويمكن التحقق من ذلك باستخدام ثبات الملاحظين (كابا) Inter-observer reliability(Kapp). ويقارن Finch & French (2019: 59) بين طرق التحقق من ثبات درجات أدوات القياس موضحاً المعادلات والتوصيات للاستخدام، كما في الجدول التالي:

جدول (١) مقارنة طرق تقدير الثبات

تقدير الثبات	المعادلة	التوصيات للاستخدام
الصور البديلة Alternate Forms	معامل الارتباط بين الأداء في الصورتين	يستخدم عند توفر أكثر من نسخة للأداة
إعادة الاختبار Test-retest	معامل الارتباط بين التطبيقين	يستخدم عندما يكون الاستقرار الزمنى للأداة مهم
ألفا لكرونباخ Cronbach's α	$\alpha = \left(\frac{K}{K-1}\right)\left[1-\frac{\sum_{k=1}^{P} S_k^2}{S_{Total}^2}\right]$	يستخدم عند الحاجة إلى تقدير الحد الأدنى للثبات
التجزئة النصفية Split halves	معامل الارتباط بين النصفين مع التصحيح باستخدام معادلة سبيرمان بروان $SB = \frac{2r_{H1,H2}}{1+r_{H1,H2}}$	يستخدم كتقدير عام لثبات الدرجة
أوميجا Omega	$\omega = \frac{\left(\sum_{i=1}^{I} a_i\right)^2}{\left(\sum_{i=1}^{I} a_i\right)^2 + \sum_{i=1}^{I} \psi_i^2}$	يستخدم في حالة فهم البناء الكامن للمقياس، ونثق في نموذج العامل المناسب للبيانات
ألفا الطبقية Stratified α	$\alpha_{Stratified} = 1 - \frac{\sum_{m=1}^{M} \sigma_m^2 (1-\alpha_m)}{\sigma_X^2}$	يستخدم عند تقدير الثبات لمقياس يتكون من عدة مقاييس فرعية

ويعد معامل ألفا الأكثر شيوعاً بين الباحثين، حيث قام Peterson (1994)في دراسة بعنوان: التحليل البعدي لمعامل ألفا كرونباخ بمراجعة البحوث المنشورة في (١٦) مجلة علمية خلال الأعوام (١٩٦٠ – ١٩٩٢)، وكان إجمالي معاملات ألفا في هذه البحوث (٤٢٨٦) كما يوضحها الشكل التالي:



شكل (١) التوزيع النسبي ل ٤.٢٨٦ معامل ألفا Peterson (1994: 386)

وقام Cho(2016) بمراجعة جميع أعداد مجلتي علم النفس التطبيقي Journal of Applied psychology (JAP)، والإدارة الأكاديمية Academy of Management Journal (AMJ) خلال عامي (٢٠١٣، ٢٠١٤) وعددها (٢٩٢) بحثاً منشوراً وجد أن (١٣٨) بحث وبنسبة (٤٧.٣%) كانت في مجال حساب الثبات، وأن أكثر من (٨٠%) من هذه البحوث استخدمت معامل ألفا من خلال (١٦) صيغة لفظية مرادفة له.

ويشير Morera &Stokes (2016) إلى أنه عند استخدام محرك الباحث العلمي Google Scholar تم الحصول على (٢٥.٠٠٠) مقالة، أو فصل في كتاب، أو بحوث لطلاب رجعت للمقالة الأولى لكرونباخ والتي تناول فيها معامل ألفا. وكذلك الحصول على (١١٩) بحث منشور في المجلة الأمريكية للصحة العامة American Journal of Public Health خلال الأعوام من (٢٠١١ إلى ٢٠١٣) تناولت ثبات درجات المقاييس، منها (١٠٥) بحث بنسبة (٨٨.٢%) استخدمت معامل ألفا، في مقابل (١٢) بحث استخدمت إعادة التطبيق، وبحث واحد استخدم كيودر ريتشادسون ٢٠، وبحث واحد استخدم ألفا المتعدد polychoric α.

ويرجع Yang & Green (2011) في دراسة بعنوان معامل ألفا: معامل الثبات في القرن الحادي والعشرين شيوع استخدام معامل ألفا بين الباحثين لأربعة أسباب، وهي:

١- سهولة تفسير معامل ألفا، وأن قيمته التي تقترب من الواحد الصحيح تعتبر مؤشراً للاتساق الداخلي، (وهنا ينبغي التأكيد على التعامل بحذر وعدم انتهاك المسلمات التي يستند إليها هذا المعامل).

٢- تحقق ميزة في معامل ألفا غير متوفرة في معاملات الثبات الأخرى، وهي الموضوعية وعدم تدخل الباحث، فمثلا في طريقة التجزئة النصفية يتم التدخل في تقسيم بنود الأداة، ويتدخل في طريقة إعادة التطبيق في تحديد الوقت المتاح بين التطبيقين.

٣- معامل ألفا مفيد في إجراء مراجعات للمقاييس في ضوء قيم معامل ألفا بعد حذف درجة البند، حيث يمكن اتخاذ قرارات حول البنود التي يجب تضمينها أو حذفها من الأداة، وأيضا هو مفيد في تقييم الصدق الكامن للأداة (الجذر التربيعي لمعامل ألفا).

٤- أن العديد من الباحثين الذين استخدموا معامل ألفا وضعوا إطاراً عاماً لتفسير قيمته بأنها صغيرة أو متوسطة أو كبيرة.

ويشير McNeish (2018: 415) في دراسة بعنوان: الشكر لمعامل ألفا، سنحصل عليها من هنا إلى أنه عند استخدام محرك الباحث العلمي لثلاث مجلات تصدر عن الجمعية الأمريكية لعلم النفس (APA) وهي: مجلة علم النفس التربوي Journal of Educational Psychology (JEP)، مجلة علم النفس الاجتماعي والشخصية Journal of Personality and Social (JPSP) Psychology، مجلة علم النفس الإكلينيكي Journal of Abnormal Psychology (JAP) وذلك خلال الفترة من (يناير، ٢٠١٤ حتى أكتوبر ٢٠١٦) بإجمالي (٣٦٩) دراسة منشورة في المجلات الثلاث ، منها (١١٨) دراسة تهتم بالثبات والخصائص السيكومترية ، موزعة (١٠٩) دراسة بنسبة (٩٢%) استخدمت معامل ألفا فقط في مقابل (٩) دراسات بنسبة (٨%) استخدمت طرق بديلة بالإضافة إلى ألفا للتحقق من الثبات.

وأجرت سومية شكرى محمود (٢٠١٩) دراسة بهدف التعرف على الأخطاء الشائعة في إجراءات التحقق من ثبات وصدق أدوات القياس المستخدمة في البحوث التربوية العربية، وذلك بمراجعة (٧٢) بحثاً منشوراً في مجلة الجامعة الإسلامية للدراسات التربوية والنفسية خلال الفترة (٢٠١٢ إلى ٢٠١٦)، وأظهرت النتائج أن (٢٣) دراسة بنسبة (٢٥%) استخدمت معامل ألفا فقط، (٢٦) دراسة بنسبة (٢٨%) استخدمت معامل ألفا مع طرق أخرى.

وقام الباحث الحالي بمراجعة أعداد المجلة السعودية للعلوم النفسية الصادرة عن الجمعية السعودية للعلوم التربوية والنفسية (جستن) خلال الفترة من (إبريل ٢٠١٩ حتى إبريل ٢٠٢٠) وعددها ثلاثة أعداد بإجمالي (١٦) بحث في مجالات علم النفس المختلفة اتضح أن جميعها وبنسبة (١٠٠%) استخدمت معامل ألفا كطريقة رئيسية للتحقق من الثبات في جميع الأدوات الواردة في هذه البحوث. وقد شعر الباحث الحالي بمشكلة التعامل مع معامل ألفا من خلال الخبرة التدريسية لمقررات القياس والتقويم والإحصاء، والتحليلات الإحصائية التي أسندت إليه في مشاريع بحثية بمختلف التخصصات العلمية. وإصرار الباحثين المستمر على استخدام معامل ألفا. ولذا تهدف هذه الدراسة إلى التعرف على الحقائق والمعتقدات الخاطئة الشائعة لدى الباحثين حول معامل ألفا.

مشكلة الدراسة: مما سبق تتضح مشكلة الدراسة الحالية فيما يلي:

١. ما المسلمات التي يستند إليها معامل ألفا في التحقق من ثبات درجات أدوات القياس؟.

٢. ما الحقائق والمعتقدات الخاطئة الشائعة لدى الباحثين حول استخدام معامل ألفا في التحقق من ثبات درجات أدوات القياس؟.

٣. ما التوصيات والمقترحات للتغلب على المعتقدات الخاطئة الشائعة لدى الباحثين حول استخدام معامل ألفا في التحقق من ثبات درجات أدوات القياس؟.

أهمية الدراسة: تتضح أهمية الدراسة الحالية فيما يلي:

١- الاستخدام الواسع لمعامل ألفا كطريقة للتحقق من ثبات الأدوات بين الباحثين.

٢- غياب البحوث العربية التي استهدفت تحليل أهم المعتقدات الخاطئة الشائعة لدى الباحثين حول معامل ألفا.

٣- تطوير الوعي المعرفي لدى الباحث العربي حول معامل ألفا وأهم مسلماته.

مصطلحات الدراسة : تتمثل مصطلحات الدراسة الحالية فيما يلى :

١-معامل ألفا : اقترح كرونباخ عام (١٩٥١) عدداً من المعادلات المترادفة رمز لها بالحرف الاغريقي α مستفيداً من جميع صيغ المعادلات التي سبقته ، ويعتبر الأكثر استخداماً بين الباحثين في التحقق من ثبات درجات أدوات القياس النفسي والتربوي.

٢- ثبات درجات الأداة: الاتساق في استجابات الأفراد على أداة القياس ، ومؤشر يمكن الباحثين من استخدام الأداة في مواقف محددة سواء تم القياس في الظروف نفسها أو ظروف مختلفة لا تتدخل فيها عوامل عشوائية.

منهج الدراسة: إن الدراسة الحالية **بحث نظري تحليلي** اعتمد الباحث فيه على عمليات الرصد والاستدلال والتحليل والمقارنة للبحوث والدراسات السابقة التي تناولت معامل ألفا، ومعتقدات الباحثين حوله.

النتائج:

الإجابة عن السؤال الأول: ما المسلمات التى يستند إليها معامل ألفا فى التحقق من ثبات درجات أدوات القياس:

لمعرفة متى يؤدى استخدام معامل ألفا إلى قيمة دقيقة لثبات درجات أدوات القياس، ومتى يترتب على استخدامه تقدير غير دقيق للثبات من حيث تضخم أو انخفاض القيمة. لابد من التعرف على أهم النماذج والمسلمات التي يستند معامل ألفا.

أجرى Schmidt & Hunter (1996) دراسة مسحية وتقويمية لأخطاء القياس في الدراسة النفسي: دروس من ٢٦ سيناريو بحثي، وتم التأكيد على أن اختيار معامل ألفا لتقدير الثبات عندما ينطوي على عدد من الأبعاد (غير متجانس) يعتبر اختياراً غير مناسب، أو غير صالح لتقدير الثبات.

وقدم Schmidt (1996:353) أربع ملاحظات تتعلق بالاستخدام الصحيح لمعامل ألفا، وهي:

١- معامل ألفا ليست مؤشراً مناسباً لأحادية البعد لقياس التجانس.

٢- في حالة عدم تحقق أحادية البعد، يجب استخدام التصحيح لجميع القيم المحسوبة.

٣- لا يوجد نقطة قطع مقبولة أو غير مقبولة لمعامل ألفا.

٤- الاقتصار على معامل ألفا فقط كمؤشر للعلاقات في المقاييس متعددة الأبعاد غير كاف، ويجب توفر الارتباطات الداخلية بين البنود وتصحيحها.

ويشير Christmann & Aelst (2006) إلى أن معامل ألفا حساس بدرجة شديدة جداً لتجاهل المسلمات الأساسية لنموذج الثبات الخاص به، ويتميز بأنه لا يتمتع بالمنعة أو القوة The alpha coefficient is known to be non-robust

وتشير دراسات كل من: Graham (2006: 934)، أمحمد بوزيان تيغزة (٢٠٠٩: ٦٦٢)، Raykov & Marcolides(2013: 123) ، محسوب عبدالقادر الضوي (٢٠١٧: ١٢٩)، محمد عبد الهادي عبدالسميع (٢٠١٧: ٣٣٢) إلى وجود أربعة نماذج للثبات Reliability models، هي: النموذج المتوازى Parallel model، ونموذج تاو المتكافئ Tau-equivalent model، ونموذج تاو المتكافئ في الأساس Essentially tau-equivalent model، والنموذج التقاربي Congeneric model . وأن هذه النماذج تختلف فيما بينها تبعاً للتساوي المفترض لبعض البارامترات الخاصة بأجزاء المقياس من حيث توزيع الدرجات الحقيقية والدرجات الخطأ والدرجات الملاحظة أو المشاهدة. ويقارن Price (2017: 220) بين النماذج الأربعة من حيث افتراضات تساوي الدرجات الملاحظة المتوقعة، وتساوي الانحرافات المعيارية أو التباين، وتساوي مكونات التغاير، وتساوي معاملات التغاير أو الارتباط، وتساوي معاملات الثبات كما يوضحها الجدول التالي:

جدول (٢) مقارنة بين نماذج الثبات

مسلمات النموذج	النماذج			
	المتوازي	تاو المتكافئ	تاو	التقاربي
١-تساوي الدرجات الملاحظة المتوقعة	×	×	-	-
٢-تساوي الانحرافات المعيارية (التباين)	×	-	-	-
٣-تساوي مكونات التغاير	×	×	×	-
٤-تساوي معاملات التغاير أو الارتباط	×	-	-	-
٥-تساوي معاملات الثبات	×	-	-	-

ويشير كل من Dunn, Baguley and Brunsden (2014) إلى دمج النموذجين، نموذج تاو المتكافئ، ونموذج تاو المتكافئ في الأساس في نموذج واحد، وبذلك أصبحت نماذج الثبات ثلاثة نماذج. ويقترحا الباحثين التمييز بين النماذج الثلاثة في ضوء أربعة عناصر رئيسة، وهي:

١- مدى تحقق أحادية البعد Unidimensionality في أداة القياس.

٢- ما إذا كانت الدرجات الحقيقية للبنود المختلفة لها نفس المتوسط (الحساسية Sensitivity).

٣- ما إذا كانت الدرجات الحقيقية للبنود المختلفة لها نفس التباين Variance

٤- ما إذا كان هناك تساوي في تباين الخطأ.

وتشترك النماذج الثلاثة في توفر أحادية للبعد في الأداة، ولكنها تتفاوت في توفر العناصر الثلاثة الأخرى. وتتضح خصائص كل نموذج على النحو التالي:

١- **النموذج المتوازي**: وتشمل مسلماته قياس البنود لمتغير أحادي البعد، وتساوي الدرجات الحقيقية أي تساوي تشبعات البنود على نفس العامل الكامن بنفس الدرجة، وتساوي تباين الخطأ. ومن الواضح أنه من الصعوبة تحقيق هذه المسلمات في أدوات القياس النفسي والتربوي.

٢- **نموذج تاو المتكافئ في الأساس**: ومن مسلمات هذا النموذج تحقق أحادية البعد في البنود، وتشبع البنود على نفس العامل الكامن بنفس الدرجة، وعدم تساوي تباين الخطأ. ويعتمد معامل ألفا على هذا النموذج، ويوضح ذلك الانتهاكات التي يقع فيها المتخصصين في القياس النفسي والتربوي، حيث يصعب تشبع البنود على نفس العامل بنفس الدرجة.

٣- **النموذج التقاربى:** ويعتبر أكثر نماذج الثبات مرونة، ومن مسلماته تنوع تشبعات البنود على العامل الكامن، وكذلك تنوع مستوي تباين الخطأ. ويعتمد معامل أوميجا على هذا النموذج.

الإجابة عن السؤال الثاني: ما الحقائق والمعتقدات الخاطئة حول معامل ألفا لدى الباحثين في التحقق من ثبات درجات أدوات القياس:

على الرغم من التزايد المستمر في استخدام معامل ألفا، إلا أنه يوجد كثير من المعتقدات الخاطئة الشائعة بين الباحثين سواء المتخصصين في القياس أو غير المتخصصين مما دفع الباحث الحالي إلى رصد هذه المعتقدات ومحاولة تفنيدها والتحقق من صحتها من خلال مراجعة عدد كبير من البحوث والدراسات السابقة حول معامل ألفا، وفي دراسة مسحية للبحوث التي استخدمت معامل ألفا للتحقق من الثبات قام بها Cortina (1993: 98) تم رصد خمس حقائق عن معامل ألفا، وهي: (أ) ألفا هي متوسط لجميع معاملات ثبات التجزئة النصفية ، (ب) ألفا هو الحد الأدنى لثبات الاختبار، (ج) ألفا هو مقياس لتشبع العامل الأول، (د) يتساوى معامل ألفا مع الثبات عند توفر شرط نموذج تاو المتكافئ، (ه) يعد معامل ألفا صيغة أكثر عمومية لمعامل كيودر ريتشاردسون للتكافؤ.

ورصد كل من Cho & Kim (2015) في دراسة بعنوان معامل ألفا لكرونباخ هو الأكثر شيوعاً ولكن الأقل فهماً، المعتقدات الخاطئة الشائعة لدى الباحثين حول معامل ألفا، وهي:

١-أن كرونباخ هو أول من قدم معامل ألفا:

1. Common Misconception: Alpha was first developed by Cronbach.

على الرغم من أن الباحثين اللاحقين يقدمون تفسيرات أكثر تفصيلاً، وأدلة أحدث، وتعديلات أكثر فائدة وأهمية على الأعمال السابقة، فإن الأوساط الأكاديمية تعترف بالباحث الذي ابتكر المعادلة. في مجال القياس، من الصعب تخيل عدم استخدام اسم مبتكر المعادلة، وبالتالي ينظر إلى كرونباخ بشكل عام أنه صاحب معادلة ألفا. ولكن المتتبع لتطور معامل ألفا يتضح له يلي:

- أن معادلة سبيرمان بروان Spearman-Brown وجدت منذ أكثر من (٢٠٠) عام تقريباً. وسميت بهذا الاسم لأنها نشرت بشكل منفرد في عام (١٩١٠) في المجلة البريطانية لعلم النفس British Journal of Psychology Spearman(1910) , Brown (1910). وتعتمد هذه المعادلة على افتراض أن نصفي الاختبار متكافئان، ومن ثم تساوي تباينات الدرجة لنصفي الاختبار. ولقد تم اقتراح صيغ أكثر مناسبة في حالة عدم تساوي التباينات للنصفين بواسطة Flanagan (1937), Rulon (1939), and Mosier(1941). وجميع هذه الصيغ تستخدم لحساب الثبات في حالة تجزئة الاختبار إلى نصفين. وقدم كيودر وريتشاردسون Kuder and Richardson(1937) معادلة لحساب الثبات لدرجات الاختبارات الثنائية (صفر، ١) وسميت (KR-20).

- وفي عام (١٩٤٥) قدم جوتمان Guttman معادلة لحساب الثبات في حالة تساوى أو عدم تساوى الانحرافات المعيارية لنصفي الاختبار.

- وفي عام (١٩٥١) وبعد مرور أربعة عشر عاما على ظهور معادلات كيودر ريتشاردسون للاتساق الداخلي، اقترح كرونباخ عددا من المعادلات المترادفة رمز لها بالحرف الاغريقي α مستفيداً من

جميع صيغ المعادلات التي سبقته، ونشر ذلك في المقال الشهير (معامل ألفا واختبارات البنية الداخلية) ، وأكد في مقدمة المقال أن معادلة كيودر رتشاردسون هي حالة خاصة من معادلة ألفا (Cronbach , 1951). وكشفت مراجعة مؤشر اقتباسات العلوم الاجتماعية Social Sciences Citations Index من عام (١٩٦٦) إلى عام (١٩٩٠) أنه تم الاستشهاد بمقال كرونباخ (١٩٥١) حوالي (٦٠) مرة سنويًا في (٢٧٨) مجلة مختلفة(Cortina , 1993) . وحققت هذه المقالة حتى عام (٢٠٠٤) نجاحا كبيراً وتم الاستشهاد بها ما لا يقل عن (٥.٥٩٠) مرة، حتى في السنوات الأخيرة وجد حوالي (٣٢٥) اقتباساً في كل عام وفقا لمؤشر العلوم الاجتماعية. (Cronbach & Shavelson ,2004)

- وعلى الرغم من أن المعامل يطلق عليه اسم ألفا، إلا أن كرونباخ لم يقرن تسمية هذا المعامل باسمه

> Although the coefficient is typically called Cornbrash's alpha, Cronbach never named the coefficient after himself

- وفي عام (٢٠٠٤) أكد على ذلك بقوله أنه من المحرج بالنسبة لي أن المعادلة أصبحت معروفة تقليدياً باسم ألفا
كرونباخ

> It is an embarrassment to me that the formula became conventionally known as Cronbach's α

Cronbach & Shavelson, (2004: 397)

- وبمراجعة دليل البرنامج الإحصائي SPSS يتضح أن معامل ألفا بالنسبة للبيانات الثنائية يكافئ معامل كيودر ريتشاردسون (٢٠)

Reliability Analysis Statistics

You can select various statistics that describe your scale and items. Statistics that are reported by default include the number of cases, the number of items, and reliability estimates as follows:

- **Alpha models.** Coefficient alpha; for dichotomous data, this is equivalent to the Kuder-Richardson 20 (KR20) coefficient.

٢-معامل ألفا يساوي ثبات الدرجات:

2. Common Misconception: Alpha equals reliability.

يرى الكثير أن معامل ألفا يساوي ثبات درجات الاختبار، ولكن قد تكون ألفا أقل أو أكبر من ذلك عندما يتم انتهاك المسلمات التي بنى عليها هذا المعامل؛ وعلى هذا النحو يساء فهم وتفسير القيمة الناتجة. كما تزداد قيمة ألفا عن القيمة الحقيقة للثبات عندما يوجد ارتباط موجب بين خطأ البنود .(Cho , 2015)

ولقد خلص Cortina (1993: 98) في دراسته المسحية التي تمت على البحوث التي استخدمت معامل ألفا للتحقق من الثبات إلى أن ألفا هو الحد الأدنى لثبات الاختبار. ويلخص Raykov(2004:343) الحالات الثلاث لمعامل ألفا، وهي:

أ- بافتراض عدم ارتباط أخطاء القياس، فأن ألفا هو الحد الأدنى لمعامل الثبات بغض النظر عن البنية العاملية للمقياس.

ب- وعند افتراض استقلالية درجات الخطأ، أو عدم ارتباطها، فأن معامل ألفا يكون مساوياً لثبات المقياس عندما تكون بنود المقياس أو مكوناته من نوع مقاييس الدرجة الحقيقة المتكافئة tau equivalent، اما إذا كانت من النوع المتقارب Parellel، فمعامل ألفا لا يساوي ثبات المقياس بل يمثل الحد الأدنى له.

ج- لكن في حالة ارتباط درجات الخطأ، فإن استخدام معامل ألفا يقل عن قيمته أو يؤدي إلى تقدير متضخم لمعامل الثبات.

٣-القيمة المرتفعة لمعامل ألفا تعد مؤشر للاتساق الداخلي والتجانس:

3. Common misconception: A high value of alpha is an indication of internal consistency.

يعتقد بعض الباحثين أن ارتفاع قيمة ألفا لأداة القياس يعد مؤشر للاتساق الداخلي له، ولكن ربما يعود ذلك إلى زيادة متوسط معاملات الارتباطات البينية للبنود، وكذلك عدد البنود التي تتضمنها الأداة أو درجة تشابه هذه البنود Yang & Green (2011:380). فثبات درجات الأداة والاتساق الداخلي والتجانس مفاهيم مرتبطة إلى حد كبير ولكنها تختلف عن بعضها البعض، حيث يعتمد الاتساق الداخلي على متوسط الارتباطات البينية للبنود وليس على عددها، فهو خاصية للبنود وليس خاصية للمقياس، أما التجانس فيعتمد على تشبع البنود على العامل وفقا لنتائج التحليل العاملي، ومعامل ألفا هو دالة لطول الأداة وتشبع البنود، وبالتالي فالقيمة المرتفعة لمعامل ألفا لا تعد مؤشر للاتساق الداخلي أو التجانس (Davenport, Davison Liou, & Love ,2015:8). ويؤكد Sijtsma (2015: 10) أن الاتساق الداخلي وأحادية البعد يرتبطان بالبنية الداخلية لأداة القياس وبالتالي بصدق البنود وليس ثبات الدرجات.

ويقترح كل من Vaske, Beaman & Sponarski (2016) في دراسة بعنوان: إعادة التفكير في الاتساق الداخلي بمعامل ألفا، أن على الباحثين القيام بثلاثة خطوات ضرورية قبل الحكم على الاتساق الداخلي للمقياس، وهي:

أ- التحقق من أحادية البعد Unidimensionality للمقياس باستخدام التحليل العاملي الاستكشافي أو التحليل العاملي التوكيدي.

ب- عند تحقق افتراض أحادية البعد، يجب على الباحث التحقق من أن قيمة معامل ألفا، وجميع مؤشرات المقياس مقبولة والتي تتمثل في: جميع معاملات الارتباط البينية للمفردات تكون موجبة inter-item correlations should all be positive، أن تبلغ معاملات الارتباط المصححة Corrected Item-Total Correlations (٠.٤ فأكثر لجميع البنود)، وأن يكون معامل ألفا لكرونباخ في حالة حذف البند Alpha if Item Deleted أقل من ألفا للمقياس ككل.

ج- التحقق من مقدار ونمط الاتساق باستخدام البرامج الإحصائية SAS, SPSS، وعلى الباحثين توظيف هذه البرامج لتطوير تعريفات بديلة للاتساق الداخلي.

ويرى كل من Dunn, Baguley and Brunsden (2014) أن أهم الصعوبات في استخدام ألفا كمؤشر للاتساق الداخلي تتمثل في: (أ) تعتمد ألفا على مسلمات يصعب تحقيقها، (ب) يؤدى انتهاك هذه المسلمات إلى تضخم ألفا وتخفيض تقديرات الاتساق الداخلي للمقياس ، (ج) "ألفا عند حذف درجة البند" في العينة لا يعكس تأثير حذف البند على ثبات المجتمع ، (د) تقدير نقطة لمعامل ألفا لا يعكس

التباين الموجود في عملية التقدير، مما يؤدى إلى ثقة زائفة في الاتساق الداخلي للمقياس. ويوصي الباحثين باستخدام معامل أوميجا بدلا من ألفا للتحقق من الاتساق الداخلي لدرجات المقياس.

ويوصي أحمد تيغزة (٢٠٠٩: ٦٥٨) بضرورة التريث في الحكم على قيمة معامل ألفا المرتفعة بأنها تدل على اتساق داخلي مرتفع للأداة، إلا بعد التأكد من أن بنود المقياس أو عدد من هذه البنود التي ترتبط فيما بينها ارتباطا مرتفعاً لا تتشابه تماماً في دلالتها رغم اختلاف صياغتها اللغوية.

وأظهرت نتائج دراسة محسوب عبدالقادر الضوي(٢٠١١) تضخم تقديرات معامل ألفا عند الاحتفاظ بالدرجات المتطرفة، واختلاف التقديرات باختلاف عدد فئات الاستجابة. وانخفاض تقديرات معامل ألفا عند حذف أعداد متباينة من الدرجات المتطرفة.

ويؤكد Taber (2018) على أن القيمة المرتفعة لمعامل ألفا تقدم دليلاً محدوداً على ثبات أداة الدراسة، وأن القيم المرتفعة جدا في الواقع غير مرغوب فيها عند تطوير أدوات في المعرفة العلمية وفهم العلوم.

٤-يمكن تحسين ثبات درجات المقياس عن طريق حذف بعض البنود:

4. Common Misconception: Reliability will always be improved by deleting items using "alpha if item deleted."

يرى Kopalle and Lehmann (1997) أن حذف البنود ذات معاملات الارتباط البينية المنخفضة يؤدى إلى مبالغة في قيم ألفا أو تضخم في قيمة معامل ألفا can lead to an overstatement alpha inflation " , or , "of alpha . وأن الحكم على قيمة معامل ألفا لابد أن يكون في ضوء العينة التي طبق عليها المقياس، وكذلك المتغير موضع القياس مع إجراء الصدق عبر العينة Cross-Validation sample وأن حذف البنود يكون مرجعه الأساس النظري والمنطقي للمقياس.

ويضيف Raykov (2007) أن التغير الذي يطرأ على الثبات الكلي للمقياس ضئيل جدا باستخدام " ألفا بعد حذف درجة البند "، ويعتبر هذا الإجراء مضللاً ' Coefficient alpha if item deleted' can be misleading ويقترح Raykov (2008) استخدام مدخل نمذجة المتغير الكامن لتقييم جودة المقياس بعد حذف درجة البند

وأظهرت نتائج دراسة كل من Serbetar & Sedlar (2016) أن إجراء حذف درجة البند يؤدى إلى قيم مرتفعة لمعامل ألفا، وقد تكون في الواقع غير حقيقية، فقد تغيرت قيمة معامل ألفا للمقياس ككل من (٠.٧٩) إلي (٠.٨٣) وهى زيادة طفيفة جداً. ويتضح ذلك في الجدول التالي:

جدول (٣) التغير في قيمة ألفا عند حذف درجة البند Serbetar & Sedlar (2016:193)

الخطوة	رقم البند المحذوف	معامل الارتباط الداخلي للبند Inter-item correlation	ألفا المعدل Adjusted alpha
١	٢٣	٠.١٤	٠.٨٠٦
٢	١٩	٠.١٥	٠.٨١٢
٣	٣	٠.١٥	٠.٨١٥
٤	١٠	٠.١٦	٠.٨١٩
٥	١٢	٠.١٧	٠.٨٢٣
٦	٨	٠.١٨	٠.٨٢٥
٧	٧	٠.١٩	٠.٨٢٨
٨	١١	٠.٢٠	٠.٨٢٩

يتضح من الجدول السابق باستخدام حذف درجة البند من خلال ثمان خطوات متتالية تغيرت قيمة معامل ألفا من (٠.٨٠٦) إلي (٠.٨٢٩) وهى تغير طفيف جداً نتج عنه حذف (٨) بنود من المقياس وقد تكون ذات أهمية لقياس المتغير.

٥-لابد أن تساوي أو تزيد قيمة معامل ألفا عن (٠.٧ أو ٠.٨)

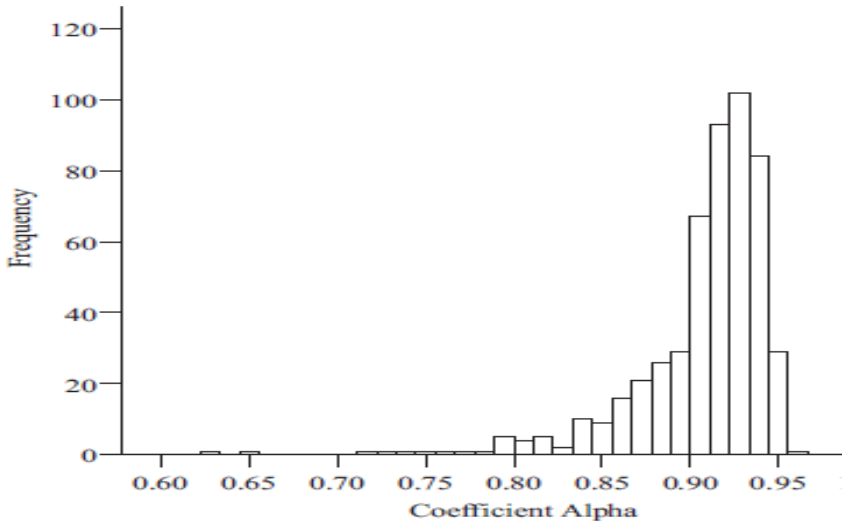
5. Common Misconception: Alpha should be greater than or equal to .7 (or, alternatively, .8).

يعد هذا المعتقد هو الأكثر انتشاراً بين الباحثين، وبرغم وجود آراء كثيرة ومتضاربة حول قيمة ألفا ولكن جميعها يحكمها شيء واحد وهو السعي المستمر لدى الباحثين للحصول على قيم مرتفعة بغض النظر عن حقيقة هذه القيمة. ويؤكد Streiner (2003b) على أنه ليس ضروريا الحصول على قيم مرتفعة لمعامل ألفا، لأن ارتفاع قيم ألفا يرجع إلى أسباب كثيرة متباينة وأهمها تشابه الدلالات اللفظية للبنود لدى المفحوصين على الرغم من اختلاف الصياغة.

ويذكر أحمد عبد الخالق (١٩٩٦: ٥٠) أن معامل الثبات الذي يساوي أو يزيد عن (٠.٧) يعد مقبولاً في مقاييس الشخصية، ومع ذلك فتجدر الإشارة إلى أنه من الأهمية بمكان ألا يسعي مؤلف الاختبار إلى الحصول على معامل اتساق داخلي مرتفع.

ويري رجاء أبو علام (٢٠٠٠: ٤٨٣) أنه يجب أن تكون معاملات الثبات أعلى ما يمكن، فإذا زادت عن قيمة (٠.٨٠) كان هذا مفضلاً عندما يتعلق الأمر بالمقاييس التي تتناول الاتجاهات والميول والجوانب الانفعالية. أما إذا كان معامل الثبات يتعلق باختبارات معرفية كاختبارات التحصيل والاستعداد فهذه يجب أن تكون قريبة جدا من قيمة (٠.٩٠) ويفضل أن تزيد عن هذه القيمة، ويرجع هذا إلى طبيعة الأسئلة التي تتعلق بهذا النوع من الاختبارات.

وقدم كل من (Salvucci, Walter, Conley, Fink, & Saba, 1997: 115). تصنيفا لمعاملات الثبات من حيث قوتها على النحو التالي: أقل من (٠.٥) ثبات منخفض، وبين (٠.٥، ٠.٨) ثبات متوسط، وأكبر من (٠.٨) ثبات مرتفع. وفي دراسة مرجعية قام بها Rodriguez & Maeda (2006) بعنوان: التحليل البعدي لمعامل ألفا لاختبارات الرياضيات في (٥١٢) مدرسة اتضح أن قيم معامل ألفا محصورة بين (٠.٦٠، ٠.٩٥) كما يوضحها الشكل التالي:



شكل (٣) توزيع قيم معامل ألفا Rodriguez & Maeda (2006: 313)

واتفق العديد من الباحثين على أن قيمة معامل ألفا (٠.٧٠) فأعلى تعد مؤشر مقبول لتوفر خاصية ثبات درجات المقياس (Nunnally ,1978). وفي ضوء هذا المحك قام كل من Lance, Butts & Michels (2006: 205) بمراجعة الاستشهادات في مؤشر العلوم الاجتماعية Social Science Index citation خلال الفترة من (٢٠٠٠ إلى٢٠٠٤) وذلك في (١١) مجلة ، تم الحصول على (٩٠) استشهاد، منهم (٤٠) استشهاد بنسبة (٤٤%) استخدمت محك (٠.٧) كمؤشر للثبات، (٥٠) استشهاد بنسبة (٥٦%) أكدوا على ضرورة مراجعة القيمة لتطوير جودة المقياس. وينتقدوا هذه الاستشهادات وتم وصفها بأنها مخلة ومنقوصة وعلى الجميع مراجعة النص الكامل في كتاب (Nunnally ,1978) ، ويتضح ذلك في النص التالي: ذكر Nunnally (1978) أن معامل الثبات التي تبلغ قيمته (٠,٧) يُعد كافيًا، ويتساءل هل هي تلك الأسطورة صحيحة؟ الأساس وراء الحقيقة: خمن Nunnally أنه لربما ينئ الباحث بالمقاييس التي يكون ثباتها متواضع (٠,٧) بعيداً عن بذل مزيد من الجهد والوقت في مجال بحث جديد. الخرافة: على النقيض لما أشار له العديد من الباحثين (ضمنياً أو صريحاً) فإنه يرى وفق قسم "معايره للثباتStandards of " Reliability: (أ) لم يعلن أن (٠,٧) يعد معيارًا عالميًا للثبات، (ب) معامل الثبات البالغ قيمته (٠,٧) لا يكون ملائما لكل الدراسات والبحوث (باستثناء ما ذكر في الجملة السابقة) أو الممارسة، (ج) ليس لديه ما يقوله عن طريقة استثنائية لتقدير الثبات (على سبيل المثال: ألفا لكرونباخ).

> *Summary*. The legend: Nunnally (1978) said that .70 reliability is adequate, right? The kernel of truth: Nunnally conjectured that perhaps one can get away with measures that have only modest reliabilities of .70 or thereabouts if one wants to save time and effort in a new area of research. The myth: Contrary to many researchers' (implicit or explicit) claims, Nunnally's "Standards of Reliability" section (a) did not proclaim .70 as a universal standard of reliability, (b) did not indicate that .70 reliability was adequate for research (except as qual-ified in the previous sentence) or practice, and (c) had nothing to say about any particular method of estimating reliability (e.g., Cronbach's alpha).

٦-معامل ألفا هو الأفضل بين كل معاملات الثبات المنشورة

6. Common Misconception: Alpha is the best choice among all published reliability coefficients.

يرى صلاح علام (٢٠٠٠: ١٧٤) أنه يمكن ترتيب طرق حساب الثبات ترتيباً تنازلياً بحسب حجم القيم التقديرية التي نحصل عليها باستخدامها، وهي كالتالي: التجزئة النصفية، كيودر ريتشاردسون، إعادة تطبيق الاختبار، الصيغ المتكافئة، معامل ألفا. لذلك ينبغي توقع اختلاف هذه القيم باختلاف الطريقة المستخدمة في موقف معين ولعينة أفراد معينة. فهذه القيم ليست مطلقة، وإنما هي قيم نسبية حيث إنها مرتبطة بكل من الموقف وطبيعة العينة المستخدمة. كما أن هذه القيم تزودنا بتقدير لدرجات الاتساق في درجات الاختبار، ولكنها لا توضح أسباب عدم اتساق الدرجات.

ويقترح كل من Zumbo, Gadermann, and Zeisser (2007) معامل ألفا الرتبي Ordinal Alpha، و معامل ثيتا الرتبي Ordinal Theta لتصحيح التحيز السلبي في معامل ألفا، وبالطبع. كيودر ريتشاردسون في حالة الاستجابات الثنائية.

وقارن Revelle and Zinbarg (2009) بين ثلاثة عشر معادلة مختلفة لحساب ثبات درجات المقاييس، حيث أظهرت النتائج أن معامل أوميجا أفضل معامل ثبات كما أنه يتميز عن ألفا في اختيار النموذج المناسب للبيانات بعد التحليلات الأولية، بينما يفترض في معامل ألفا توافر الافتراضات دون وجود خطوات عملية للتحقق من ذلك مما يؤثر سلبياً على قيم معامل الثبات الناتج.

وأظهرت نتائج دراسة محمد عبد السميع (٢٠١٧) أن معامل أوميجا أدق تقدير لقيم معامل ثبات الدرجات مقارنة بمعامل ألفا في ضوء نسبة التباين المفسر واتساع المسافة بين حدي الثقة. وقدم أمحمد تيغزة (٢٠١٧) تصوراً نظرياً لأهم الطرق البديلة لمعامل ألفا، وهي: ألفا الطبقية Stratified alpha، ومعامل الثبات المركبComposite reliability أو "رو" RHO، وأوميجا الموزونة Weighted Omega. وهدفت دراسة حسان غازي العمرى (٢٠١٨) إلى مقارنة ثلاث طرق لتقدير ثبات الاختبارات المركبة التي تتضمن نوعين من الفقرات هي (ألفا، ألفا الطبقي، راجو) د، وأظهرت النتائج أن الوسط الحسابي لتقدير الثبات كان (٠.٧١، ٠.٧٩، ٠.٥٣) حسب ألفا، وألفا الطبقي Stratified alpha، وراجوRAJU على الترتيب، ووجود فروق دالة بين الطرق الثلاثة. وقارن أحمد كريش (٢٠١٨) بين معامل ألفا، ومعامل ألفا الرتبى Ordinal Alpha لتقدير ثبات درجات الاختبارات ذات البيانات الرتبية، وأظهرت النتائج أن معامل ألفا الرتبي أكبر من معامل ألفا وأنه يعطى تقدير أدق لقيمة الثبات الحقيقي بغض النظر عن عدد البدائل في مقياس ليكرت أو توزيع البيانات في مقابل أن معامل ألفا تقل دقته كلما انحرفت البيانات عن التوزيع الطبيعي وقل عدد البديل.

ويضيف Cho (2016: 654) أيضاً أن الباحثين يستخدمون معامل ألفا دون محاولة التحقق من افتراضاته وذلك لوجود بعض **المعتقدات الخاطئة** والتي تتمثل في التالي: (أ) ألفا معامل ثبات يمكن استخدامه مع أي نوع من البيانات، (ب) عدم تأثر ألفا بانتهاك الافتراضات الخاصة به (أو أن تأثير انتهاك الافتراضات يكون ضئيل جدا على تقدير الثبات)، (ج) القيمة المرتفع لمعامل ألفا مؤشر على تحقق افتراضاته، (د) صعوبة التحقق من الافتراضات، (ه)صعوبة استخدام الطرق البديلة في حالة انتهاك افتراضاته. وناقش Streiner (2003A: 101-102) ما أسماه خرافات حول ألفا Myths about Alpha شائعة بين الباحثين، وهي:

٧-ألفا خاصية ثابتة لوصف المقياس Alpha is affixed property of a scale

هي الخرافة الأولى والأساسية التي تحيط بمعامل ألفا. وأكدت دراسة Anastasi & Urbina (1997) على أن كثير من الباحثون يصفون المقياس بكونه ثابت، بينما المصطلح الأكثر مناسبة هو ثبات درجات المقياس وليست ثبات المقياس. وقد أوصي " ويلكنسون وفريق الاستدلال الإحصائي للجمعية الأمريكية لعلم النفس APA" عند وضع دليل إرشادي للطرق الإحصائية في المجلات النفسية بأن الثبات خاصية لدرجات الاختبار يختص بها عينة محددة من المفحوصين، ولذلك من المهم استخدام ثبات درجات الاختبار وليست ثبات الاختبار (Wilkinson and the Task Force on Statistical Inference *APA Board of Scientific Affairs* ,1999)

٨- يقيس معامل ألفا الاتساق الداخلي للمقياس فقط Alpha Measures only the internal consistency of the scale

يتساءل كل من Tang , Cui and Babenko (2014) في دراسة بعنوان: الاتساق الداخلي: هل نعرف حقاً ما هو وكيف يتم قياسه ؟ ، وتأتي الإجابة في ثلاثة مؤشرات تلقي الضوء حول ماهية الاتساق الداخلي، وذلك على النحو التالي:

(أ) الاتساق الداخلي = متوسط الارتباطات الداخلية.

Internal Consistency = Average Interitem Correlations?

(ب) الاتساق الداخلي = التشبع بالعامل العام.

Internal Consistency = General Factor Saturation?

(ج) الاتساق الداخلي = ثبات الاتساق الداخلي.

Internal Consistency = Internal Consistency Reliability?

ويقتصر استخدام معامل ألفا كأحد أهم طرق حساب الاتساق الداخلي على المقاييس ذات البنود المتجانسة وبالتالي فهي تحسب ثبات الدرجات على أنه معامل تجانس بين بنود المقياس الواحد (أحمد الرفاعي غنيم، ١٩٩١). وعلى الرغم من أن ارتفاع معاملات الارتباط بين بنود الاختبار يؤدي إلى زيادة قيمة معامل ألفا، إلا أن زيادة وارتفاع قيمة معامل ثبات ألفا لا تعكس بالضرورة مستويات مرتفعة من الاتساق الداخلي؛ حيث تتأثر قيمة معامل ألفا للاختبار بزيادة طوله Streiner (2003A). وتدعم نتائج Cortina(1993) افتراض تأثر قيمة معامل ثبات الاختبار بطوله؛ حيث أشارت الدراسة إلى أنه بالمحافظة على متوسط قيمة معامل الارتباط بين البنود (R=0.30) مع تغير عدد البنود ما بين (٦، ١٢، ١٨) ستبلغ قيمة معامل ألفا (٠.٧٢، ٠.٨٤، ٠.٨٨) على التوالي ؛ مما يؤكد على تأثر معامل ألفا.

واعتمد كل من BrckaLorenz, , Chiang & Nelson Laird (2013: 1) عدد من المحكات للاتساق الداخلي Internal Consistency Criteria يجب توفره في المقياس الجيد، ومنها أن القيمة المقبولة لمعامل ألفا هي (٠.٧) فأكثر كما يوضحها الجدول التالي:

جدول (٤) محكات المقياس الجيد وفقا لإحصائيات الثبات

إحصائيات الثبات Reliability Statistics	محكات المقياس الجيد Criteria for a Good Scale
ألفا لكرونباخ	تساوي او أكبر من (٠.٧٠)
مدى الارتباط بين البنود	بين (٠.١٥، ٠.٨٥)
متوسط الارتباط بين البنود	بين (٠.١٥، ٠.٥٠)
مدى الارتباطات المصححة للبنود بالمقياس	يساوي أو أكبر من (٠.٥٠)
مدى ألفا لكرونباخ عند حذف البند	سيؤدى حذف أي بند إلى خفض قيمة ألفا

يتضح من الجدول رقم (٤) وجود أكثر من محك لتحقق الاتساق الداخلي للأداة، وعلى الباحثين استخدام هذه المحكات والالتزام بها للحصول على أداة يتحقق لها الجودة والدقة في القياس، وكذلك التأكيد على أن معامل ألفا عند حذف درجة البند ليست مؤشر لصدق المحك كما يدعى بعض الباحثين ولكنه مؤشر للاتساق الداخلي للأداة.

وقدم (Sharma (2016: 273 تصنيف لقيم معامل ألفا كمؤشر للاتساق الداخلي، يوضحها الجدول التالي:

جدول (٥) تصنيف قيم معامل ألفا كمؤشر للاتساق الداخلي

قيم معامل ألفا	الاتساق الداخلي
أكبر من أو تساوي ٠.٩	ممتاز
أكبر من أو تساوي ٠.٨ وأقل من ٠.٩	جيد
أكبر من أو تساوي ٠.٧وأقل من ٠.٨	مقبول
أكبر من أو تساوي ٠.٦ وأقل من ٠.٧	موضع شك
أكبر من أو تساوي ٠.٥ وأقل من ٠.٦	ضعيف
أقل من ٠.٥	غير مقبول

ويري Vaske, Beaman & Sponarski (2017: 171) ضرورة إعادة التفكير في معامل ألفا كمؤشر للاتساق الداخلي، حيث قاما بمراجعة نتائج (٢٩) مشروعاً بحثياً لقياس الاتجاهات باستبيانات مختلفة انحصر عدد بنودها بين (٣، ٦)، واتضح أن متوسط معامل ألفا(٠.٩٠)، وأن قيمة ألفا تزيد بزيادة عدد البنود والتي انحصرت بين (٠.٨٢، ٠.٩٥) ، وأن جميع قيم ألفا كانت أكبر من (٠.٨٠) وهي القيم الموصي بها في البحوث.

٩- معامل ألفا الأكبر هو الأفضل دائماً Bigger is always Better

يشير Streiner (2003A) إلى أنه بالنسبة لمعظم مؤشرات الثبات؛ فكلما زادت قيمتها كلما كان ذلك أفضل، فالباحث يسعى دومًا للوصول لمستويات من الاتفاق بين المحكمين، والتوصل لدرجات جيدة من الاستقرار للدرجات عبر الزمن. ويدعم صفوت فرج (٢٠٠٢: ٣٠٩) نفس الفكرة؛ حيث يشير إلى أنه إذا كانت لدينا ثلاثة أساليب (أ، ب، ج) يمكن استخدام كل منها لتقدير ثبات مقياس معين. وكان (أ) يوفر معامل ثبات (٠.٩٠)، بينما يوفر (ب) معاملا قدره (٠.٨)، ويوفر (ج) معاملاً قدره (٠.٦) فأن (أ) هو أفضل أسلوب لتقدير معامل ثبات هذا المقياس. ويعتمد هذا المحك على منطق أن الأسلوب الذي يؤدي إلى أعلى معامل ثبات يؤدي إلى تقدير أقصي تباين حقيقي للأداء على الاختبار.

ويستطرد Streiner (2003A) ليوضح أن هذا المعتقد صحيح كذلك في حال معامل ثبات ألفا، ولكن ليس في كل الأحوال. فلقد لاحظ أن معامل ألفا لا يقيس فقط تجانس البنود، ولكنه يقيس أيضا تجانس المتغير المستهدف قياسه، وفي كثير من الحالات حتى المتغيرات أحادية البعد Unidimensional يمكن تصورها أنها تتكون من عوامل مختلفة، فمثلا القلق يتم التعامل معه من قبل غالبية الباحثون باعتباره بنية تتكون من ثلاث مكونات هي المكونات المعرفية والفسيولوجية والسلوكية، وعلى الرغم من ذلك يضيف كل من Koksal & Power (1990) مكون رابع هو المكون الانفعالي، فحتى يكون التحسن في قيمة معامل ألفا لمقياس القلق حقيقيًا فلابد وأن تكون البنى الأربعة المكونة لمفهوم القلق متمايزة وفي نفس الوقت متجانسة مع بعضها البعض، فقيم معامل ألفا المرتفعة ربما تعكس ازدواجيه أو تكرار لا داعي له لمحتوي البنود، وتشير إلى التكرار أكثر مما تشير إلى التجانس Streiner (2003A).

١٠-قيمة معامل ألفا تمتد بين الصفر والواحد الصحيح. Alpha Rang Between 0 and 1

يوصي Streiner (2003A) أن معامل ألفا يجب ألا تتجاوز قيمته كحد أقصى (٠.٩٠). ويرى أن قيمه معامل ألفا يمكن أن تكون أقل من الصفر إذا لم يتم تصحيح البنود العكسية.

وأجرى Tan (2009: 108)دراسة بعنوان إساءة استخدام معاملي ثبات كيودر ريتشادسون٢٠ وألفا، وأظهرت النتائج وجود أربعة أخطاء يقع فيها الباحثون، وهي:

(١) افتقار تقرير الدراسة إلى البيانات المتعلقة بثبات درجات الاختبار.

Misuse 1 : Lack of reporting the reliability of scores gathered on the actual study

(٢) عدم وضع البنية العاملية للبيانات في الحسبان.

Misuse 2 : Not considering the factorial structure of the data

(٣) عدم الإشارة إلى الخطأ المعياري للقياس كمؤشر للثبات عند انتهاك مسلمة أحادية البعد.

Misuse 3: Reporting, the standard error of measurement as a reliability indication when the unidimensionality assumption is violated.

(٤) استخدام كيودر ريتشاردسون ٢٠ وألفا كمؤشرين للتجانس.

Misuse 4: Using a high KR-20 or Cronbach's alpha coefficient as an indicator for homogeneity.

وقدم Ritter(2010) في دراسة بعنوان: إحصاء يساء فهمه على نطاق واسع: ألفا لكرونباخ عدد من السيناريوهات لمعامل ألفا، وهي:

(أ) في حال أن تكون جميع درجات البنود غير مرتبطة تماما، وبالتالي فلن يكون هناك ثمة تباين مشترك بين البنود أو اتساق داخلي، وفي هذه الحالة ستكون قيمة معامل ألفا مساوية للصفر.

Scenario #1: all item score correlations are perfectly uncorrelated.

(ب) في حال وجود ارتباط تام بين جميع درجات البنود؛ بمعنى أن يكون معامل الارتباط مساوي للواحد، وهنا سيكون الاتساق الداخلي تام؛ ومن ثم سيساوي معامل ألفا الواحد الصحيح.

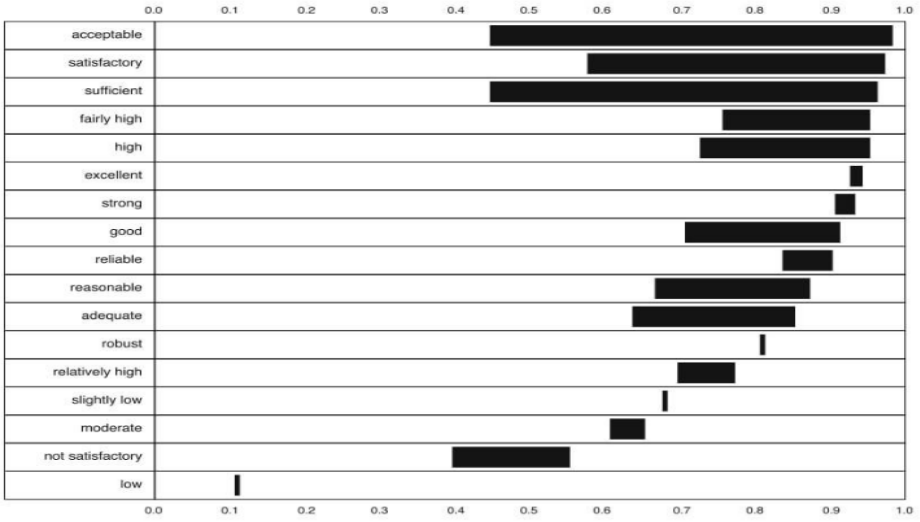
Scenario #2: all item score correlations are perfectly correlated

(ج) في حال أن يكون هناك معاملات ارتباط تامة بين جميع البنود ولكنها بإشارات متباينة، وفي هذه الحالة تكون قيمة معامل ألفا ضعيفة جداً. وبشكل عام، عندما تكون البنود أكثر ارتباطاً، فإن التباين المشترك ستزداد قيمته ومن ثم ستزداد قيمة الاتساق الداخلي وتزداد القيمة المطلقة لمعامل ألفا.

Scenario #3: all item score correlations are perfectly correlated and have mixed signs.

وأجرى Taber (2018)دراسة للتعرف على استخدام الباحثين معامل ألفا عند تطوير أدواتهم في مجال تعليم العلوم من خلال مراجعة (٦٩) بحث منشور في أربع مجلات خلال عام (٢٠١٥)، وهي : المجلة الدولية لتعليم العلوم International Journal of Science Education (IJSE) ، ومجلة الدراسة في تدريس العلوم Journal of Research in Science Teaching (JRST) ، والدراسة في تعليم العلوم Research in Science Education (RISE)، وتعليم العلوم Science Education (SE) . وأظهرت النتائج استخدامهم (١٧) وصفاً لفظياً لقيم معامل ألفا، وهي: (٠.٩٣ – ٠.٩٤) ممتاز،

(٠.٩١-٠.٩٣) قوي، (٠.٨٤-٠.٩٠) ثابت، (٠.٨١) قوي، (٠.٧٦ – ٠.٩٥) مرتفع بوضوح، (٠.٧٣-٠.٩٥) مرتفع، (٠.٧١-٠.٩١) جيد، (٠.٧٠-٠.٧٧) مرتفع نسبياً، (٠.٦٨) منخفض نسبياً، (٠.٦٧ - ٠.٨٧) معقول، (٠.٦٤-٠.٨٥) كاف، (٠.٦١-٠.٦٥) متوسط، (٠.٥٨-٠.٩٧) مرضي، (٠.٤٥-٠.٩٨) مقبول، (٠.٤٥-٠.٩٦) كاف، (٠.٤٠-٠.٥٥) غير مرضي، (٠.١١) منخفض. ويوضح توزيعها الشكل التالي:



شكل (٢) وصف قيم معامل ألفا Taber (2018)

وقدم Hinton , McMurray.& Brownlow (2014: 359) تصنيفاً لقيم معامل ألفا لثبات درجات المقياس يمكن للباحثين الاستناد إليها: (٠.٩٠فأكثر) ممتاز، (٠.٧٠ إلي ٠.٩٠) مرتفع، (٠.٥٠ إلي ٠.٧٠) متوسط، (أقل من ٠.٥٠) منخفض. واقترح George & Mallery (2020: 244) تصنيفاً آخر لقيم معامل ألفا لثبات درجات المقياس: (٠.٩ فأكثر) ممتاز، (٠.٨) جيد، (٠.٧) مقبول، (٠.٦) مشكوك فيه،(٠.٥ فأقل) ضعيف، (أقل من ٠.٥) غير مقبول.

ومن خبرة الباحث في مجال القياس والتقويم، وكثرة الاستفسارات والتساؤلات التي يتم مناقشتها من الباحثين حول معامل ألفا. تم رصد الكثير من المعتقدات الشائعة حول معامل ألفا على الرغم من عدم صحتها وتم حصرها على النحو التالي:

١١-معامل ألفا دال إحصائياً:

درج بعض الباحثين في البيئة العربية على كتابة **أن معامل ألفا دال إحصائيا وهذا غير مقبول علميا**. حيث أن معامل ألفا ليست اختبار إحصائي استدلالي للمقارنة بين المتوسطات أو حساب العلاقات بين المتغيرات يتأخذ فيه قرار بقبول أو رفض الفرض الصفرية، وإنما هدفه التحقق من صلاحية أداة القياس للاستخدام وذلك من خلال مؤشرات أو قيم ألفا. ويشير صلاح أحمد مراد وأمين على سليمان (٢٠٠٥: ٣٦٠) أن البعض يذكر أن معامل الثبات جيد لأنه دال، والحقيقة أن مستوى الدلالة لمعامل الارتباط ليس محكاً لإقرار الثبات من عدمه، لأن مستوى الدلالة يعتمد على حجم العينة. فقد يكون معامل الارتباط (٠.٩٠) وغير دال إذا كانت العينة خمسة أفراد، وقد يكون معامل الارتباط (٠.٢٠) ودالاً إذا كان حجم العينة (١٠٠) فرداً. وعليه لا نستخدم مستوي الدلالة في قبول معامل الثبات.

ويضيف رجاء أبو علام (٢٠٠٦: ٤٨٢) أنه يجب البعد تماماً عن ذكر الدلالات الإحصائية لمعاملات الارتباط التي تستُخدم كمعاملات للثبات. ويرجع ذلك إلى مفهوم أساسي يرتبط بمدلول معامل الارتباط، لنفرض الآن تم الحصول على معامل ارتباط (٠.٢٠) لعينة حجمها (١٠٠) فرد وكان دال إحصائياً عند مستوى (٠.٠١) فهل يمكن القول أن الاختبار ثابت لأن الارتباط دال إحصائياً؟ بالطبع لا، لأن نسبة التباين التي ترجع إلى ارتباط الاختبار مع نفسه لا تزيد (٤%)، وهي نسبة ضئيلة للغاية، وبالتالي لا يمكن اتخاذها قرينة على ثبات الاختبار. ووفقا لنتائج البرنامج الإحصائي SPSS لا يوجد مستوى دلالة إحصائية لمعامل ألفا وإنما تظهر قيمة المعامل وعدد البنود.

١٢-زيادة عدد البنود يؤدى إلى قيم مرتفعة لمعامل ألفا:

تناول Cortina (1993) التغيرات في قيم معامل ألفا وفقا لبنية المقياس وعدد البنود ومتوسط الارتباطات الداخلية بينها باستخدام بيانات إمبيريقية وليست افتراضية كما يوضحها الجدول التالي:

جدول (٦) التغير في قيمة ألفا وفقا لعدد البنود ومتوسط الارتباطات الداخلية Cortina (1993:102)

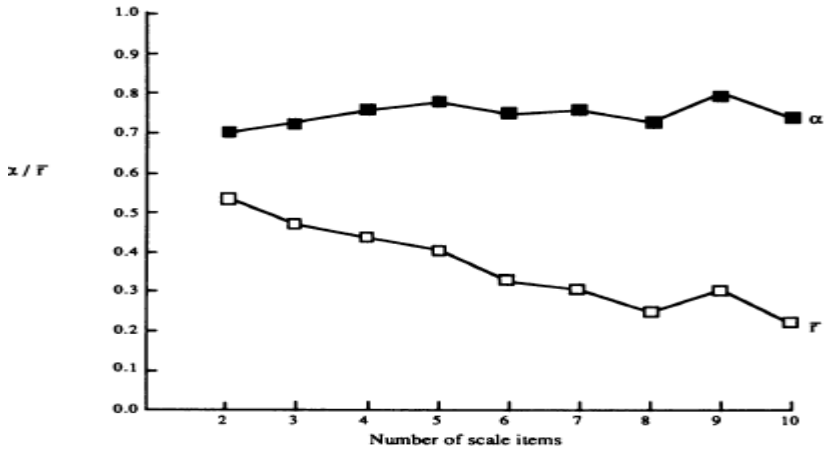
بنية المقياس	عدد البنود	R=0.30 ألفا	R=0.50 ألفا	R=0.70 ألفا
أحادية	٦	٠.٧٢	٠.٨٦	٠.٩٣
	١٢	٠.٨٤	٠.٩٢	٠.٩٦
	١٨	٠.٨٨	٠.٩٥	٠.٩٨
ثنائية	٦	٠.٤٥	٠.٦٠	٠.٧٠
	١٢	٠.٦٥	٠.٧٨	٠.٨٥
	١٨	٠.٧٥	٠.٨٥	٠.٩٠
ثلاثية	٦	٠.٢٨	٠.٤٠	٠.٤٩
	١٢	٠.٥٢	٠.٦٥	٠.٧٤
	١٨	٠.٦٤	٠.٧٦	٠.٨٤

ويتضح من الجدول السابق ما يلي:

أ-زيادة عدد البنود (٦–١٢-١٨) يؤدى إلي زيادة قيم معاملات ألفا سواء أكان المقياس أحادي أو ثنائي أو ثلاثي الأبعاد.

ب-زيادة متوسط الارتباطات (٠.٣٠، ٠.٥٠، ٠.٧٠) يؤدى إلي زيادة قيم معاملات ألفا بغض النظر عن بنية المقياس.

ولقد قام Peterson (1994) بمراجعة (٤٢٨٦) قيمة لمعامل ألفا في ضوء عدد البنود وخلصت النتائج إلى عدم وجود تأثير واضح لطول الاختبار على قيم معاملات ألفا ومتوسط الارتباطات الداخلية، كما يوضحها الشكل التالي:



شكل (٤) العلاقة بين عدد البنود وقيم معاملات ألفا ومتوسط الارتباطات Peterson (1994:390)

بينما يري Schimtt(1996: 351) أن قيمة ألفا لا تتأثر بطول الاختبار وعدد.وقام أمحمد تيغزة (٢٠٠٩: ٦٥٥) باستخدام بيانات فعلية للتحقق من علاقة عدد البنود والاتساق الداخلي بقيم معامل ألفا. ووجد أنه عند تثبيت متوسط الارتباط بين البنود ما بين (٠.١، ٠.٨) مع اختلاف عدد البنود نجد اختلاف قيم ألفا فانحصرت بين (٠.٣٦، ٠.٩٥) لخمس بنود، و (٠.٥٣، ٠.٩٧) لعشرة بنود، و (٠.٦٩، ٠.٩٩) لعشرون بند، و(٠.٧٧، ٠.٩٩) لثلاثون بند، و(٠.٨٢، ٠.٩٩) لأربعون بند. وتم استنتاج ثلاثة نقاط مهمة، وهي: **أولاً:** أنه كلما زاد الاتساق الداخلي للبنود (العلاقات الارتباطية بينها) يزداد معامل ألفا ارتفاعاً. **ثانياً:** عند طول معين للاختبار، فإن الازدياد في قيم ألفا لا يستمر بوتيرة واحدة عند ازدياد متوسط ارتباطات البنود. **ثالثاً:** أنه كلما ازداد حجم الاختبار يتناقص حجم التأثير الذي يمارسه ارتفاع متوسط الارتباطات بين البنود على معامل ألفا.

وأظهرت نتائج دراسة Panayides(2013) تغير قيم ألفا وفقا لبنية المقياس وعدد البنود. كما بالجدول التالى:

جدول (٧) التغير في قيمة ألفا وفقا لعدد البنود وبنية المقياس Panayides(2013: 691)

بنية ثنائية		بنية ثلاثية	
عدد البنود	قيمة ألفا	عدد البنود	قيمة ألفا
٦	٠.٤٩٢	٦	٠.٦٠٣
٨	٠.٦١٣	٩	٠.٧١٤
١٠	٠.٦٩٨	١٢	٠.٧٩٧
١٢	٠.٧٥٧	١٥	٠.٨٢٨
١٤	٠.٧٧٣	١٨	٠.٨٥٦
١٦	٠.٧٩٠	٢٦	٠.٨٧٨
١٨	٠.٧٩٦	٣٦	٠.٩٠٢
٢٠	٠.٨٠٢	٤٦	٠.٩٢٥
٢٤	٠.٨٤٣		
٢٨	٠.٨٦٣		
٣٢	٠.٨٨١		
٣٦	٠.٨٩٧		
٤٠	٠.٩٠٩		

يتضح من الجدول السابق:

أ- عند زيادة عدد البنود من (٦) بنود إلى (٤٠) بند نجد أن معامل ألفا يزيد من (٠.٤٩٢) إلى (٠.٩٠٩) للبنية الثنائية.

ب- عند زيادة عدد البنود من (٦) بنود إلى (٤٦) بند نجد أن معامل ألفا يزيد من (٠.٦٠٣) إلى (٠.٩٢٥) للبنية الثلاثية.

ج- أن قيم معاملات ألفا للبنية الثلاثية أعلى من نظيراتها للبنية الثنائية للمقياس.

ويري Hogan(2015) أن معامل ألفا يتغير وفقا لعدد البنود ومتوسط الارتباط البينية للبنود، كما يتضح في الجدول التالي:

جدول (٨) أثر عدد البنود ومتوسط الارتباطات على معامل ألفا Hogan(2015:134)

عدد البنود	متوسط الارتباطات البينية	قيمة ألفا
٥	٠.١٠	٠.٣٦
٥	٠.٢٥	٠.٦٣
٥	٠.٤٠	٠.٧٧
٢٠	٠.١٠	٠.٦٩
٢٠	٠.٢٥	٠.٨٧
٢٠	٠.٤٠	٠.٩٣
٥٠	٠.١٠	٠.٨٥
٥٠	٠.٢٥	٠.٩٤
٥٠	٠.٤٠	٠.٩٧

يتضح من الجدول السابق:

أ- زيادة عدد البنود يؤدى إلى زيادة قيمة معامل ألفا، حيث تغيرت قيمة ألفا من (٠.٣٦) إلى (٠.٩٧) عند زيادة عدد البنود من (٥) بنود إلى (٥٠) بند.

ب- زيادة متوسط الارتباطات البينية يؤدى إلى زيادة قيمة معامل ألفا، فتغيرت ألفا من (٠.٣٦) إلى (٠.٧٧) لمتوسطات الارتباطات البينية (٠.١٠، ٠.٢٥، ٠.٤٠)، مثلا، في حالة (٥) بنود فقط.

ج- (ج) زيادة عدد البنود إلى (٥٠) بند يؤدى إلى زيادة قيمة معامل ألفا، فتغيرت ألفا من (٠.٨٥) إلى (٠.٩٧) لمتوسطات الارتباطات البينية (٠.١٠، ٠.٢٥، ٠.٤٠).

١٣- زيادة عدد البدائل يؤدى إلى قيم مرتفعة لمعامل ألفا:

يشغل عدد بدائل الاستجابة ذهن الباحثين باستمرار، وقد يكون مصدر للارتباك والحيرة والتردد عند بناء الأداة، وبالتالي تتأثر الاستجابات على الأداة وفقا لعدد البدائل، وقد يتجه بعض الباحثين إلى اختيار عدد أقل من البدائل رغبة في سهولة التصحيح وإدخال البيانات للبرامج الإحصائية، وقد يتجه البعض الآخر إلى زيادة عدد البدائل رغبة في الحصول على قيم مرتفعة من معامل ثبات الدرجات. وقام

Peterson (1994) بمراجعة (٤٢٨٦) قيمة لمعامل ألفا في ضوء عدد البدائل باستخدام منهج التحليل البعدي، كما يوضحها الجدول التالي:

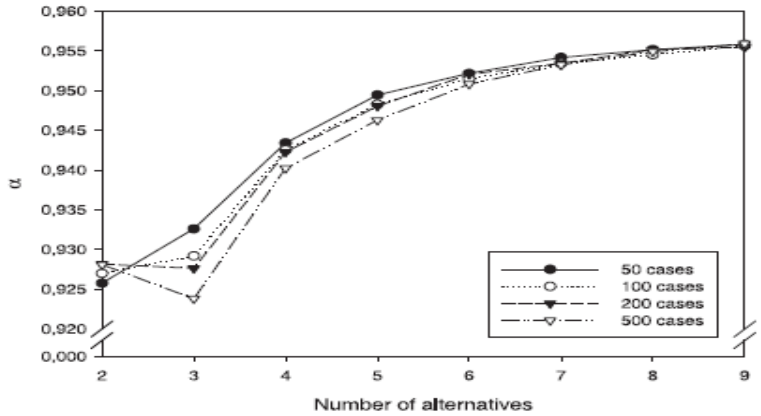
جدول (٩) متوسط ووسيط قيم ألفا وفقا لعدد البدائل Peterson (1994: 387)

عدد البدائل	عدد معاملات ألفا	متوسط ألفا	وسيط ألفا
غير محدد	٦٦٧	٠.٧٦	٠.٧٨
٢	٢٢١	٠.٧٠	٠.٧٤
٣	١٥٩	٠.٧٨	٠.٨٠
٤	٣٠٥	٠.٧٦	٠.٧٨
٥	١٣١٩	٠.٧٧	٠.٧٩
٦	٢٤٩	٠.٧٥	٠.٧٧
٧	٩٩١	٠.٧٨	٠.٨٢
٨فأكثر	٣٧٦	٠.٧٧	٠.٨١

يتضح من الجدول السابق عدم وجود علاقة واضحة بين عدد البدائل وقيم ألفا. فالمتوسط انحصر بين (٠.٧٠، ٠.٧٨) ولكن لا يعطى مؤشر واضح أن زيادة عدد البدائل يؤدى إلى زيادة قيم ألفا. فمثلاً في حالة (٥) بدائل وإجمالي معاملات ألفا (١٣١٩)، وحالة (٨) بدائل فأكثر وإجمالي معاملات ألفا (٣٧٦) يتساوى متوسط ألفا (٠.٧٧). ويتضح نفس الاتجاه في وسيط ألفا فمثلا في حالة عدد بدائل غير محدد وإجمالي معاملات ألفا (٦٦٧)، وحالة (٤) بدائل فأكثر وإجمالي معاملات ألفا (٣٠٥) يتساوى وسيط ألفا (٠.٧٨).

وأظهرت نتائج دراسة Weng (2004) تأثير دال تأثير إحصائياً لعدد البدائل على قيمة معامل ألفا، حيث زادت قيم ألفا من (٠٦٤) إلى (٠.٨١) مع زيادة عدد البدائل من (٣) بدائل إلى (٩) بدائل لاختبار الاتجاه نحو التدريس. وكذلك تؤكد نتائج دراسة Zumbo, Gadermann& Zeisser (2007) زيادة قيم معامل ألفا بزيادة عدد البدائل ، حيث كانت قيم ألفا (٠.٨٢٦، ٠.٨٤٩، ٠.٨٧٢، ٠.٨٨٢، ٠.٨٨٦، ٠.٨٩١) للبدائل من (٢ إلى ٧ بدائل على الترتيب).

بينما أظهرت نتائج دراسة Lozano, Garcia-Cueto & Muniz(2008) أن معامل ألفا يزيد مع زيادة عدد البدائل، بالرغم من أن هذه الزيادة ضعيفة جداً وفي وجود أحجام مختلفة من العينات كما يوضحها الشكل التالي:



شكل (٥) العلاقة بين عدد البدائل ومعامل ألفا Lozano, Garcia-Cueto & Muniz(2008:76)

يتضح من الشكل السابق أن قيم معامل ألفا انحصرت بين (٠.٩٢٥، ٠.٩٥٥) في وجود أحجام عينات (٥٠، ١٠٠، ٢٠٠، ٥٠٠)، وعدد بدائل من (٢ إلى ٩) بدائل.

ويوصي Spiliotopoulou(2009:149) بضرورة التأكد من مناسبة عدد بدائل المقياس. حيث أن استخدام (٤) بدائل فأقل يؤدى إلى أن تكون ألفا أقل من قيمتها Underestimated alpha، وفي المقابل استخدام (٥) بدائل يؤدى إلى ارتفاع قيمة ألفا مقارنة (٦) بدائل.

وأظهرت نتائج دراسة كل من Wong, Pang, Shi, & Mao.(2011) زيادة قيم ألفا وفقا لزيادة عدد البدائل،حيث انحصرت قيمة ألفا بين (٠.٤٢، ٠.٩٥) في حالة (٤ نقاط)، وبين (٠.٥٥، ٠.٩٥) في حالة (٥نقاط) مع مراعاة كل من اختلاف عدد البنود وطبيعة المقاييس ما بين معرفية و نفسية .

وأظهرت نتائج دراسة Choudhury & Bhattacharjee (2014) وجود تأثير دال إحصائياً لعدد البدائل على معامل ألفا، حيث كانت قيم ألفا (٠.٧٨٨، ٠.٨١٥، ٠.٨٣١، ٠.٨٥٨، ٠.٨٦٤) للبدائل من (٥ إلى ٩ بدائل) على الترتيب.

وعلى مستوي البيئة العربية، يوجد الكثير من الدراسات التى تناولت أثر عدد البدائل على معامل ألفا، ومن هذه الدراسات: دراسة حسين أحمد العكام (١٩٩٥)، دراسة نوره صالح المحيميد (١٩٩٩) ، دراسة سعيد عبدالفتاح الغامدي (٢٠٠٣)، دراسة محمد حسين سعيد (٢٠٠٧)، دراسة حجاج غانم علي وياسر عبدالله حسن(٢٠١١)، ودراسة محسوب عبد القادر الضوي (٢٠١١)، دراسة محمد عبدالرحمن إسماعيل (٢٠١٩)، دراسة محمد عبد السميع (٢٠١٧)، والتي أظهرت جميعها زيادة معامل ألفا بزيادة عدد فئات التدريج حتى الصورة خماسية التدريج ثم يقل في الصورة سداسية التدريج، ويعود للزيادة مرة أخرى في الصورة سباعية التدريج، وأن الفروق بين قيم معامل ألفا غير دالة إحصائياً.

١٤ -زيادة حجم العينة تؤدى إلى قيم مرتفعة لمعامل ألفا:

أظهرت نتائج دراسة Petrson(1994) أن التغير في متوسط ووسيط قيم ألفا وفقا لأحجام عينات مختلفة ضعيف جدا، كما يوضحها الجدول التالي:

جدول (١٠) متوسط ووسيط قيم ألفا وفقا لحجم العينة Peterson (1994: 387)

حجم العينة	عدد معاملات ألفا	متوسط ألفا	وسيط ألفا
أقل من ١٠٠	١٠٢٨	٠.٧٦	٠.٨٠
١٠٠-١٩٩	١١٦٩	٠.٧٨	٠.٨٠
٢٠٠-٢٩٩	٦٩٦	٠.٧٨	٠.٨٠
٣٠٠فأكثر	١٢٦٥	٠.٧٥	٠.٧٧
لم تحدد	١٢٨	٠.٧٦	٠.٨٠

يتضح من الجدول السابق عدم وجود علاقة واضحة بين حجم العينة وقيم ألفا. فالمتوسط انحصر بين (٠.٧٥، ٠.٧٨) ولكن لا يعطى مؤشر واضح أن زيادة حجم العينة يؤدى إلى زيادة قيم ألفا. فمثلاً في أحجام عينات من (١٠٠ إلى ١٩٩) وإجمالي معاملات ألفا (١٣١٩)، وحالة (٨) بدائل فأكثر وإجمالي معاملات ألفا (٣٧٦) يتساوى متوسط ألفا (٠.٧٧). ويتضح نفس الاتجاه في وسيط ألفا فمثلاً في حالة عدد

بدائل غير محدد وإجمالي معاملات ألفا (٦٦٧)، وحالة (٤) بدائل فأكثر وإجمالي معاملات ألفا (٣٠٥) يتساوى وسيط ألفا (٠.٧٨).

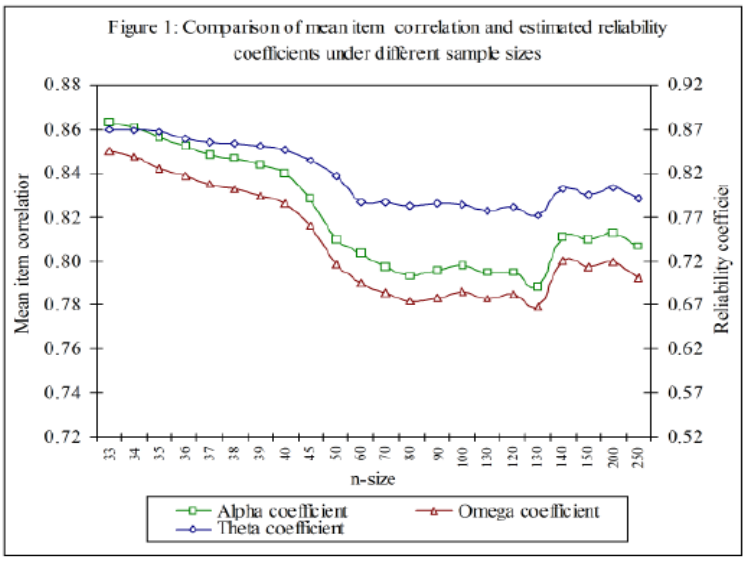
وبحثت دراسة Ercan , Yazici, Sigirli , Ediz , Kan (2007) التغير في معاملات الثبات (ألفا، وثيتا، وأوميجا) في ضوء حجم العينة، وأظهرت النتائج أن قيمتي ألفا وثيتا لا يرتبطان بحجم العينة.

ويؤكد Spiliotopoulou(2009:149) على ضرورة التأكد من مناسبة حجم العينة لاستخدام معامل ألفا، فالعينات ذات الأحجام الكبيرة تؤدى إلى تضخم قيمة ألفا، وينصح الباحثون بإجراء تحليل قوة الاختبار لأحجام العينات المتوقعة، حيث لا توجد قاعدة محددة لاختيار حجم العينة المناسب.

ويضيف أحمد تيغزة (٢٠٠٩: ٦٥٠) أنه كلما زاد التباين، كلما ارتفع مستوي معامل ألفا. وبالتالي، فاتساع التباين أو تقلصه يتوقف على مدى تجانس العينة، أو مدى تباينها. فتباين أفراد العينة يؤدى إلى تباين الاستجابات على بنود المقياس، وبالتالي يفضي إلى ارتفاع تباين درجات المقياس ككل. ومعنى ذلك أن قيم الثبات لذات المقياس قد تختلف باختلاف العينات، إذا يستتبع اختلاف العينة اختلاف في مدى تباين درجات المقياس ككل.

وتري سوسن شاكر مجيد (٢٠١٤: ١٣٧) أن التجانس الشديد في عينة الثبات من حيث الظاهرة المدروسة يؤدى إلى انخفاض واضح في معامل الثبات، لأن التباين داخل هذه العينة المتجانسة يكون منخفضا بقدر لا يسمح بتقدير التباين الحقيقي – أي ثباته – في حين كلما كبر حجم العينة ومن ثم عدم تجانسها كلما كانت أقرب إلى التوزيع الاعتدالي وبالتالي سيصبح تباين الخطأ صغيراً وتؤدى هذه النتيجة إلى ارتفاع ثبات درجات الاختبار.

وأشارت نتائج دراسة Javali , Gudaganavar. and Raj. (2011) إلى أن قيم ثبات الدرجات ترتبط أكثر بمتوسط الارتباط البينية للبنود أكثر من حجم العينة، كما يوضحها الشكل التالي:



شكل (٦) مقارنة متوسط الارتباط بين البنود وتقدير الثبات في ضوء حجم العينة Javali , Gudaganavar. and Raj. (2011: 6)

يتضح من الشكل السابق أن قيم معامل ألفا متقاربة مع باقي قيم ثبات الدرجات لأحجام عينات انحصرت من (٣٣ إلى ٢٥٠ فرد)، وأن متوسط الارتباط بين البنود يتدخل في حجم العينة في التأثير على قيم ثبات الدرجات ارتفاعاً وانخفاضاً.

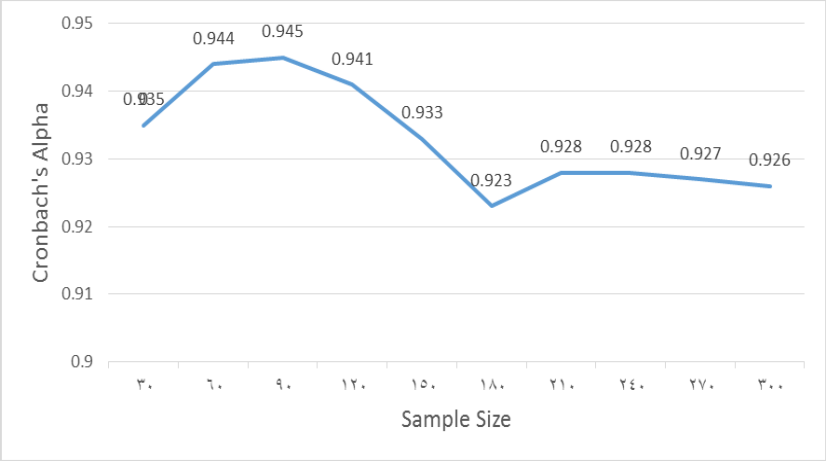
ويري Bonett & Wright (2014) أن تحديد حجم العينة من أهم الخطوات في تصميم الدراسات النفسية، فإذا كان حجم العينة صغير جداً يضعف من جودة المقياس وتتسع المسافة بين حدى الثقة، وعندما يكون العدد كبير جداً فإن هذا إهدارا للموارد.

كما أظهرت نتائج دراسة محمود على موسي (٢٠٢٠أ) تأثر معامل ألفا بحجم العينة، حيث كانت قيم ألفا (٠.٤٨) لحجم عينة (٣٢) فرد في مقابل (٠.٨١) لحجم عينة (٨٤) فرد وأن الفروق بين معاملات ألفا عبر المجموعات المستقلة كانت دالة إحصائياً. بالإضافة إلى أن الفروق بين معاملات ألفا عبر العينات المرتبطة في مواقف القياس المختلفة كانت دالة إحصائياً (محمود على موسي، ٢٠٢٠ب).

مثال عملي: تم استخدام بيانات حقيقية لعينة مكونة (٣٠٠) طالب وطالبة بكلية التربية جامعة الملك سعود طبق عليهم مقياس مكون من (١٢) بند لتقييم مشرف التدريب الميداني، وللتحقق فعليا هل يؤدى اختلاف حجم العينة إلى زيادة معامل ألفا؟ تم اتباع الخطوات التالية:

(١) اختيار عشر عينات بأحجام مختلفة بالإرجاع انحصرت بين (٣٠ إلى ٣٠٠ طالب وطالبة).

(٢) حساب معامل ألفا لكرونباخ لكل عينة على حدة، وجاءت القيم كما يوضحها الشكل التالي:



شكل (٧) العلاقة بين حجم العينة وقيم معامل ألفا

يتضح من الشكل السابق

- أن قيم معامل ألفا انحصرت بين (٠.٩٢٣، ٠.٩٤٥) وفقا لأحجام عينات من (٣٠) حتى (٣٠٠) فرد.

- ارتفاع قيم معامل ألفا لأحجام العينات (٣٠، ٦٠، ٩٠، ١٢٠) على الترتيب.

- تتسم قيم معامل ألفا بالاستقرار إلى حد ما لأحجام العينات (٢١٠، ٢٤٠، ٢٧٠، ٣٠٠).

١٥-اختلاف اتجاه البند يؤدي إلى قيم مرتفعة لمعامل ألفا

أظهرت نتائج دراسة نضال كمال الشريفين (١٩٩٥) أن الترتيب العشوائي للبنود الموجبة والسالبة كان أقل ثباتاً من البنود الموجبة أو السالبة. كما أن البنود الموجبة كانت أفضل ثباتا من البنود السالبة أو الصور التي تحتوي على نصف البنود موجب والنصف الآخر سالب. وتوصل محمد حسين سعيد (٢٠٠٧) إلى اختلاف بعض قيم معامل ألفا للثبات باختلاف اتجاه البند، حيث كانت أفضل صورة من صور الاختبار والتي حققت قيم ثبات مرتفعة متسقة هي الصورة التي كانت جميع بنودها موجبة. وطبق شاهر خالد سليمان (٢٠١٥) خمس نماذج من مقياس الاتجاهات نحو الرياضيات، حيث كان النموذج الأول كل العبارات موجبة بنسبة (١٠٠%)، والنموذج الثاني (٧٠% موجبة، ٣٠% سالبة)، والنموذج الثالث (٥٠% موجبة، ٥٠ % سالبة)، والنموذج الرابع (٣٠% موجبة، ٧٠ %سالبة)، والنموذج الخامس جميع العبارات سالبة بنسبة (١٠٠%) وجاءت قيم معامل ألفا للنماذج الخمسة على الترتيب (٠.٩١، ٠.٩٠، ٠.٨٧، ٠.٨٩، ٠.٨٨)، وكانت الفروق بين هذه القيم غير دالة إحصائياً.

الإجابة عن السؤال الثالث: ما التوصيات والمقترحات للتغلب على المعتقدات الخاطئة الشائعة لدى الباحثين حول معامل ألفا:

بعد تناول أهم المسلمات والنماذج التي بنيت عليها طرق التحقق من ثبات درجات أدوات القياس، وكذلك الحقائق والمعتقدات الخاطئة حول معامل ألفا، يمكن تقديم التوصيات التالية:

١- ضرورة استخدام مصطلح معامل ألفا لكرونباخ وليست معامل ألفا كرونباخ أو كرونباخ ألفا.

٢- ضرورة التحقق من المسلمات التي يتطلبها معامل ألفا قبل استخدامه في بيانات أدوات القياس النفسي والتربوي.

٣- ضرورة استخدام مصطلح معامل ألفا لثبات درجات الأداة (الاختبار، المقياس، الاستبانة) بدلاً من معامل ألفا لثبات الأداة.

٤- ضرورة مراجعة ملف البيانات والتأكد من عدم وجود بيانات مفقودة أو شاذة أو متطرفة .

٥- ضرورة اختيار معامل ألفا الأكثر مناسبة ونوع البيانات ، فمثلاً استخدام معامل ألفا الرتبي مع البيانات الرتبية.

٦- إن ارتفاع قيمة معامل ألفا ليست مؤشر لأحادية البعد واتساق بنود المقياس، فقد تكون القيمة المرتفعة نتيجة عدد البنود أو معاملات الارتباط البينية.

٧- في حالة عدم تحقق مسلمات معامل ألفا، فيتحتم على الباحث استخدام الطرق البديلة للتحقق من ثبات درجات بنود الأدوات، ومنها: معامل ألفا الرتبي، ومعامل ألفا الطبقي، ومعامل أوميجا، ومعامل ثيتا، ومعامل الثبات المركب.

٨- ضرورة إضافة حدود الثقة إلى جانب قيم معاملات الثبات للتأكد من دقة هذه القيم.

٩- أن قيمة معامل ألفا هي متوسط قيم طرق التجزئة النصفية للمقياس عند تقدير الثبات، ولذلك فهو أكثر دقة واستقرارا منها.

١٠- أن قيمة معامل ألفا بعد حذف درجة البند والمستخدم لتحديد تأثير حذف البند على ثبات المقياس ككل لا يعكس هذا التأثير واقعياً، حيث يمكن أن تزاد قيمة ألفا نتيجة عوامل كثيرة منها، حجم العينة ومتوسط الارتباطات البينية للبنود.

١١- يقترح الباحث إجراء البحوث والدراسات التالية في البيئة العربية :

- أثر حجم العينة وعدد البدائل على معامل ألفا لثبات درجات أدوات القياس.
- أثر حجم العينة وبنية الأداة (أحادية – متعددة) الأبعاد على معامل ألفا لثبات درجات أدوات القياس.
- أثر نوع الصفة (معرفية – وجدانية – اجتماعية) وعدد العبارات على معامل ألفا لثبات درجات أدوات القياس.
- المقارنة بين معامل ألفا والطرق البديلة للتحقق من ثبات درجات أدوات القياس.
- مؤشرات التحليل البعدي لمعامل ألفا في الدراسات والبحوث العربية.

المراجع العربية:

أحمد الرفاعي غنيم (١٩٩١). تعميم معامل ألفا لحساب معامل ثبات المقاييس ذات المفردات غير المتجانسة. *مجلة كلية التربية*، جامعة عين شمس، (١٥)، ٢٠٧-٢٣٨.

أحمد كريش(٢٠١٨). معامل (ألفا) الرتبي: تقدير معامل ثبات درجات الاختبار باستخدام البيانات الرتبية. *مجلة العلوم النفسية والتربوية*، جامعة الوادي، الجزائر، ٦ (١)، ١٠-٢٣.

أحمد محمد عبد الخالق (١٩٩٦). قياس الشخصية، جامعة الكويت، مجلس النشر العلمي.

أمحمد بوزيان تيغزة (٢٠٠٩). البنية المنطقية لمعامل ألفا كرونباخ، ومدى دقته في تقدير الثبات في ضوء افتراضات نماذج القياس. *مجلة العلوم التربوية والدراسات الاسلامية*، جامعة الملك سعود، ٢١(٣)، ٦٣٧ – ٦٨٨.

أمحمد بوزيان تيغزة (٢٠١٧). توجهات حديثة في تقدير صدق وثبات درجات أدوات القياس: تحليل نظري تطبيقي وتقويمي. *مجلة العلوم النفسية والتربوية*، جامعة الوادى، الجزائر، ٤ (١)، ٧-٢٩.

حجاج غانم علي وياسر عبدالله حسن(٢٠١١). تأثير عدد بدائل ليكرت على الخصائص السيكومترية للمقياس النفسي وافتراضات التصميم العاملي ثنائي الاتجاه. *مجلة الإرشاد النفسي*، ٢٩، ٥٦-١٣١.

حسان غازي العمري (٢٠١٨). المقارنة بين ثلاث طرق في تقدير ثبات الاختبارات المركبة التي تتضمن نوعين من الفقرات (ألفا، ألفا الطبقي، راجو). *مجلة اتحاد الجامعات العربية للتربية وعلم النفس*، ١٦(٢)، ٨٥-١٠٢.

حسين أحمد العكام (١٩٩٥). أثر فئات تدريج ليكرت لمقياس اتجاه على خصائصه السيكومترية. رسالة *ماجستير غير منشورة*، كلية التربية، جامعة اليرموك.

رجاء أبو علام (٢٠٠٦). *مناهج الدراسة في العلوم النفسية والتربوية*. القاهرة، دار النشر للجامعات.

سعيد عبد الفتاح الغامدي(٢٠٠٣). مدى اختلاف الخصائص السيكومترية لأداة القياس في ضوء تغاير عدد بدائل الاستجابة والمرحلة الدراسية – دراسة حالة: مقياس ليكرت. رسالة ماجستير غير منشورة، كلية التربية، جامعة أم القري.

سوسن شاكر مجيد (٢٠١٤). *أسس بناء الاختبارات والمقاييس النفسية والتربوية*. ط٣، عمان، مركز ديبونو لتعليم التفكير.

سومية شكرى محمود (٢٠١٩). دراسة بهدف التعرف على الأخطاء الشائعة في إجراءات التحقق من ثبات وصدق أدوات القياس المستخدمة في البحوث التربوية العربية. *مجلة كلية التربية*، جامعة أسيوط، ٣٥(٧)، ٦٧١-٦٩٥.

شاهر خالد سليمان (٢٠١٥). أثر اتجاه الفقرة في مقياس اتجاهات نحو الرياضيات على الخصائص السيكومترية للمقياس وفقراته في ضوء بعض المتغيرات، مجلة رسالة الخليج العربي،٣٦(١٣٨)، ٣١-٤٨.

صفوت فرج (٢٠٠٢). القياس النفسي. ط٦، القاهرة، الانجلو المصرية.

صلاح أحمد مراد وأمين على سليمان (٢٠٠٥). *الاختبارات والمقاييس في العلوم النفسية والتربوية. خطوات إعدادها وخصائصها*. ط٢، القاهرة، دار الكتب الحديث.

صلاح الدين محمود علام (٢٠٠٠). *القياس والتقويم التربوي والنفسي. أساسياته وتطبيقاته وتوجهاته المعاصرة. القاهرة*، دار الفكر العربي

محسوب عبد القادر الضوي (٢٠١١). تحري تأثير الدرجات المتطرفة وعدد فئات الاستجابة على تقدير معامل ألفا لكرونباخ. *مجلة كلية التربية*، جامعة أسيوط، ٢٧(١)، ١١٧-١٧٥.

محمد حسين سعيد (٢٠٠٧). أثر اتجاه المفردة وعدد بدائل الاستجابة على ثبات أدوات القياس من نوع "ليكرت". *المجلة المصرية للدراسات النفسية*، ١٧(٥٦)، ٣٥٣ -٣٩٢.

محمد عبد الهادي عبد السميع (٢٠١٧). تأثير عدد فئات الاستجابة وعدد المشاركين علي دقة قيم معاملي ألفا وأوميجا في تقدير ثبات درجات المقياس النفسي. *المجلة المصرية للدراسات النفسية*، ٢٧(٩٦)، ٣١٧-٣٨٤.

محمد عبدالرحمن إسماعيل (٢٠١٩). أثر استخدام نقطة المنتصف "محايد" مقياس ليكرت في الخصائص السيكومترية للمقياس وقياس الاتجاهات. *مجلة الإدارة العامة*، معهد الإدارة العامة بالرياض،٥٩(٣)، ٥٨٦-٦٤٢.

محمود على موسي (٢٠٢٠أ). دلالة الفروق بين معاملات ألفا عبر المجموعات المستقلة لآراء طلاب الثانوية العامة المتقدمين لاختبار القدرات الموحد ٢٠٢٠/٢٠١٩ لوزارة التعليم العالي المصرية. مجلة *كلية التربية بالإسماعيلية*، (٤٦) يناير، ١-١٨.

محمود على موسي (٢٠٢٠ب).اختبار الفروق بين معاملات ألفا عبر قياسات التقرير الذاتي والمهام المعرفية لنظرية العبء المعرفي. *المجلة الدولية للبحوث في العلوم التربوية*، ٣(٢)، ٥٦٣-٦٠٥.

نضال كمال الشريفين (١٩٩٥). أثر توزيع الفقرات الموجبة والسالبة في مقياس اتجاه على خصائصه السيكومترية وأداء الطلبة عليه. *رسالة ماجستير غير منشورة*، كلية التربية، جامعة اليرموك.

نورة صالح المحيميد (١٩٩٩). أثر اختلاف عدد بدائل الاستجابة في أوزان ليكرت على معاملات الثبات والصدق. *رسالة ماجستير غير منشورة*، كلية التربية، جامعة الملك سعود.

Anastasi, A. & Urbina, S. (1997). Psychological Testing. (7th ed.). Upper Saddle River, New Jersey: Prentice Hall.

Bonett, D. & Wright,T.(2014). Cronbach's alpha reliability: Interval estimation, hypothesis testing, and sample size planning. *Journal of Organizational Behavior*, 36(1), 3-15.

BrckaLorenz, A., Chiang, Y. & Nelson Laird, T. (2013). Internal Consistency Reliability, *FSSE Psychometric Portfolio*, 160, 1-4

Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology*, 3(3), 296-322.

Christmann, A. & Aelst, S. (2006). Robust estimation of Cronbach's alpha. *Journal of Multivariate Analysis*, 97, 1660 – 1674.

Cho, E., & Kim, S. (2015). Cronbach's Coefficient Alpha: Well Known but Poorly Understood. *Organizational Research Methods*, *18*(2), 207-230.

Cho, E., (2016) "Making reliability reliable: A systematic approach to reliability coefficients," *Organizational Research Methods*, 19 (4), 651–682.

Choudhury, S. & Bhattacharjee , D. (2014). Optimal number of scale points in likert type scales for quantifying compulsive buying behaviour. *Asian Journal of Management Research*, 4(3), 432 -440.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78(1), 98-104.

Cronbach, L. J & Shavelson, R. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement*, 64(3), 391-418.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.

Davenport, E. C., Davison, M. L., Liou, P. Y., & Love, Q. U. (2015). Reliability, Dimensionality, and Internal Consistency as Defined by Cronbach: Distinct Albeit Related Concepts. *Educational Measurement: Issues and Practice*, *34*(4), 4-9.

Drost, E. (2011). Validity and Reliability in Social Science Research. *Education Research and Perspectives*, 38(1):105-124.

Dunn,T. Baguley, T. and Brunsden, W. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Psychological Society*, 105(3), 399-412.

Ercan ,I. Yazici,B. Sigirli ,D. Ediz ,B. and Kan,I. (2007).Examining Cronbach Alpha, Theta, Omega Reliability Coefficients According to the Sample Size. *Journal of Modern Applied Statistical Methods*. 6(1), 291-303.

Finch, W.& French, B. (2019). *Educational and Psychological Measurement*. New York, Taylor & Francis.

George, D. & Mallery, P (2020). *IBM SPSS Statistics 26 Step by Step: A Simple Guide and Reference*. 6th ed. Taylor & Francis.

Graham, J. (2006). Congeneric and (Essentially) Tau-Equivalent Estimates of Score Reliability What They Are and How to Use Them. *Educational and Psychological Measurement*, 66(6) ,930-944.

Hinton, P., McMurray,I.& Brownlow, C.(2014). *SPSS Explained. East* Sussex, Second edition, England, Routledge Inc.

Hogan, T. (2015*). Psychological Testing a Practical Introduction*. Third Edition. John Wiley & Sons.

Javali, S., Gudaganavar, N. and Raj, S. (2011) Effect of Varying Sample Size in Estimation of Coefficients of Internal Consistency. *Downloaded from http://www.webmedcentral.com on 22-Dec-2011*.

Kopalle, P. K., & Lehmann, D. R. (1997). Alpha inflation? The impact of eliminating scale items on Cronbach's alpha. *Organizational Behavior and Human Decision Processes*, 70(3), 189- 197.

Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria. *Organizational Research Methods*, 9(2), 202-220.

Lozano, L. M., García-Cueto, E., & Muñiz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 4*(2), 73–79.

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods, 23*(3), 412–433.

Morera, O. &Stokes, S. (2016). Coefficient α as a Measure of Test Score Reliability: Review of 3 Popular Misconceptions. *Statistics*,106(3),458-461.

Nunnally , J. & Berstein , I.(1994). *Psychometric Theory* (3rd ed.). New York, McGraw Hill.

Panayides,P.(2013). Coefficient Alpha Interpret with Caution. *Europe's Journal of Psychology*, 9(4), 687-696.

Peterson, R. A. (1994). A meta-analysis of Cronbach's coefficient alpha. *Journal of Consumer Research*, 21, 381-391.

Price, L. (2017). *Psychometric Methods Theory into Practice*. New York, Guilford Press

Raykov, T. & Marcoulides, G. (2011). *Introduction to Psychometric Theory*. New York, Taylor & Francis Group.

Raykov, T. (2004). Point and interval estimation of reliability for multiple components measuring instruments via linear constraint covariance Structure modeling. *Structural Equation Modeling*, 11(3), 342-356.

Raykov, T. (2007). Reliability if deleted, not "alpha if deleted": Evaluation of scale reliability following component deletion. *British Journal of Mathematical and Statistical Psychology*, 60(2), 201-216.

Raykov, T. (2008). Alpha if item deleted: A note on criterion validity loss in scale revision if maximizing coefficient alpha. *British Journal of Mathematical and Statistical Psychology*, 61(2), 275-285.

Revelle,W.and Zinbarg, R(2009). Coefficients alpha, beta, omega and the GLB: Comments on Sijtsma. *Psychometrical*, 74(1), 145-154.

Ritter, N.L. (2010). Understanding a Widely Misunderstood Statistic: Cronbach's α. Paper presented at the annual meeting of the Southwest Educational Research Association, New Orleans, February 18, 2010.

Rodriguez, M and Maeda, Y. (2006). Meta-Analysis of Coefficient Alpha. *Psychological Methods*, 11(3), 306-322.

Salvucci, S., Walter, E., Conley, V., Fink, S., & Saba, M. (1997). *Measurement error studies at the National Center for Education Statistics (NCES).* Washington D. C.: U. S. Department of Education.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, 8(4), 350-353.

Schmidt, F. L.; & Hunter, J. E., (1996). Measurement error in psychological research: Lessons from 26 research scenario. *Psychological Methods*, 1(2), 199-223.

Serbetar I., & SedlarI. (2016). Assessing Reliability of a Multi-Dimensional Scale by Coefficient Alpha. *Journal of Elementary Education*, *9*(1/2), 189-196.

Sharma, B (2016). A focus on reliability in developmental research through Cronbach's Alpha among medical, dental and paramedical professionals, *Asian Pac. J. Health Sci.*, 3 (4):271-278.

Sijtsma, K. (2015). Delimiting coefficient alpha from internal consistency and unidimensionality. *Educational Measurement: Issues and Practices*, *34*(4), 10–13.

Souza, A., Alexandre, N. & Guirardello, E. (2017). Psychometric properties in instruments evaluation of reliability and validity. *Epidemiologia e Servicos de* Saude, *Brasília*, 26(3),649-659.

Spearman, C. (1910). Correlation calculated from faulty data. *British Journal Psychology*, 3(3), 271-295.

Spiliotopoulou, G. (2009). Reliability reconsidered: Cronbach's alpha and paediatric assessment in occupational therapy. *Australian Occupational Therapy*, 56, 150- 155.

Streiner, D. L. (2003A). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, 80(1), 99-103.

Streiner, D. L. (2003b). Being inconsistent about consistency: When coefficient alpha does and doesn't matter. *Journal of Personality Assessment, 80*(3), 217–222.

Taber, K. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*. 48, 1273-1296.

Tan, S. (2009). Misuses of KR-20 and Cronbach's Alpha Reliability Coefficients. *Education and Science*, 34(152), 101- 112.

Tang, W, Cui, Y, Babenko , O (2014).Internal Consistency: Do We Really Know What It Is and How to Assess It? *Journal of Psychology and Behavioral Science,* 2(2) ,205-220.

Vaske, J. J., Beaman, J., & Sponarski, C. (2016). Rethinking internal consistency in Cronbach's alpha. *Leisure Sciences, 39*(2), 1-11.

Weng, L. (2004). Impact of the Number of Response Categories and Anchor Labels on Coefficient Alpha and Test-Retest Reliability. *Educational and Psychological Measurement*, 64(6),956-972.

Wilkinson, L. and the Task Force on Statistical Inference *APA Board of Scientific Affairs* (1999). Statistical Methods in Psychology Journals Guidelines and Explanations. *American Psychological*, 54(8), 594-604.

Wong,C. , Pang, K., Shi, J. & Mao , Y.(2011). Differences between odd number and even number response formats: Evidence from mainland Chinese respondents. *Asian Pacific Journal of Management*, 28(6), 379–399.

Yang, Y., Green, S. (2011). Coefficient alpha: A reliability coefficient for the 21 century? *Journal of Psychoeducational Assessment*, 29(4), 377–392.

Zumbo, B. Gadermann, A. and Zeisser, C. (2007). Ordinal Versions of Coefficients Alpha and Theta for Likert Rating Scales. *Journal of Modern Applied Statistical Methods*.6 (1),21-29.

Alpha Coefficient for Validating Scales Scores' Reliability: Facts and Misconceptions

Dr. Elsayed Abohashem

Kong Saud Unoversity

Abstract

The current research aims to document and analyze the most common misconceptions about the alpha coefficient among researchers, and the most important of these beliefs: (1) Alpha was first developed by Cronbach , (2) the use of alpha without checking its Assumption, (3) the Alpha equals reliability , (4) the high value of the alpha coefficient is an indication of internal consistency and homogeneity, (5) the reliability of the scale grades can be improved by deleting some items, (6) the alpha coefficient value must be equal to or greater than (0.7 or 0.8), (7) the alpha coefficient It is the best among all published reliability coefficients, (8) the alpha is a fixed property to describe the scale, (9) the alpha coefficient measures the internal consistency of the gauge, (10) the largest alpha coefficient is always the best, (11) the alpha coefficient value runs from zero to one, (12) alpha coefficient statistically significant , (13) the increase in the number of items leads to high values for the alpha coefficient, (14) the increase in the number of alternatives leads to high values of the alpha coefficient, (15) the difference in the direction of the item leads to high values of the alpha coefficient, (16) the increase in size The sample leads to high values for the alpha coefficient. Lastly, some recommendations are offered to correct and address these popular misconceptions held by researchers about Alpha coefficient.

Key words: Reliability, Alpha coefficient, Internal Consistency, Misconceptions.