Doi: 10.34120/0085-036-142-014

مقارنة طريقتي ألفا وأوميغا في دقة تقدير الثبات في ضوء عدد الأبعاد وطول الإختبار وحجم العينة استناداً لبيانات مولدة

د. عمر سليمان عبيدات

كلية التربية - جامعة الإمام عبدالرحمن بن فيصل

المملكة العربية السعودية

الملخص

تهدف هذه الدراسة إلى تقييم دقة تقدير معامل الثبات باستخدام معاملي كرونباخ ألفا وأوميغا، وذلك باستخدام بيانات مولّدة من خلال الحاسوب، وفقاً لثلاثة متغيرات مستقلة؛ هي: عدد الأبعاد (بُعد واحد، بُعدين، ثلاثة أبعاد) وطول الإختبار 10، 30، 60 فقرة، وحجم العينة 50، 100، 250، 500، 1000؛ حيث ولّد خمسةٌ وأربعون ظرفاً مختلفاً، وتم تقدير قيم الثبات باستخدام معاملي الثبات (ألفا، أوميغا)، وأظهرت نتائج اختبار ت للعينات المستقلة أنه لا يوجد فروق دالة احصائياً بين تقديرات الثبات باختلاف المعاملَين في حال الإختبارات التي تتكون من بُعد واحد على اختلاف أحجام العينات وأطوال الإختبارات. كذلك أظهرت النتائج وجود فروق دالة احصائياً بين تقديرات الثبات باختلاف المعاملين في حال الإختبارات التي تتكون من بُعدين وثلاثة أبعاد على اختلاف أحجام العينات وأطوال الإختبارات، كما أظهرت النتائج أن الفروق في الخطأ المعياري في التقدير قليلة بين طريقتي تقدير الثبات باختلاف ظروف الدراسة. وبذلك توصي الدراسة باستخدام معامل تقدير أوميغا في حال تعدد الأبعاد.

الكلمات المفتاحية: معاملا الثبات كرونباخ ألفا وأوميغا، عدد الأبعاد، حجم العينة، طول الإختبار.

المقدمة

يُعد موضوع "الثبات" من القضايا الرئيسة التي تحظى باهتمام الباحثين الذين يطورون أو يستخدمون المقاييس والاختبارات النفسية والتربوية، ويُعتبر من الأدلة الضرورية اللازم توافرها للاستدلال على الخصائص السيكومترية للمقاييس (Liu & Zumbo, 2007)؛ حيث توفر تقديرات الثبات معلومات أساسية بشأن القياس الصحيح للمتغيرات والسمات. ومن ثم، فإن تحديد معامل الثبات المناسب للاختبار أو للمقياس يُعتبر خاصية يجب أن يشتمل عليها، وهو مؤشر على أنساق القيم التي تم الحصول عليها في ظل الظروف الموحدة (Ercan, Yazici, Sigirli, Ediz & Kan, 2007)، ونجد أن الاختصاصين في القياس والتقويم قدموا معادلات كثيرة لحساب معامل الثبات للمقاييس والإختبارات من خلال طريقة الإختبار وإعادة الإختبار، والصور المتكافئة. ومع تعدد طرائق حساب الثبات إلا أنه يوجد خلاف بين الإختصاصيين حول طرق الإعادة والصور المتكافئة؛ من حيث عدد مرات التطبيق وظروف التطبيق وإعداد الصور المتكافئة والمدة الزمنية بين مرتي التطبيق. وعليه، ظهرت معاملات الإتساق الداخلي (التجزئة النصفية، KR–21، KR–20، كرونباخ ألفا) لحساب معامل الثبات بتطبيق المقياس لمرة واحدة فقط. ويعتبر مؤشر كرونباخ ألفا من أكثر المعاملات استخداما لتقدير الثبات (Raykov, 1997) في جميع مجالات البحث مثل علم النفس والإحصاء والاقتصاد وعلم الاجتماع والطب والتمريض والعلوم السياسية (Cortina, 1993)، ويحسب باستخدام المعادلة التالية:

$$\alpha = \frac{n}{n-1}\left[1 - \frac{\sum_{i=1}^{n} \sigma_{Y_i}^2}{\sigma_x^2}\right] \quad \ldots\ldots\ldots\ldots (1)$$

حيث إن: n عدد فقرات الاختبار، $\sum_{i=1}^{n} \sigma_{Y_i}^2$ مجموع تباين درجات كل فقرة، σ_x^2 تباين الدرجات الكلية للإختبار.

ويُعتبر معامل ألفا لتقدير الثبات الأكثر شيوعاً واستخداماً في العلوم التربوية والنفسية والاجتماعية (Peterson,1994; Hogan, Benjamin & Brezinski, 2000) باعتباره مناسبا لتقدير ثبات التجانس الداخلي عند استخدامه بشكل صحيح (Padilla & Divers, 2013) إلّا أن طريقة كرونباخ ألفا لها افتراضاتها الخاصّة بفقرات

المقياس وهي أن تكون من نموذج «تاو» المتكافئ Tau-equivalent model بحيث تكون الدرجات الحقيقية متماثلة بين الفقرات لنفس الفرد، كما تشترط أن يكون المقياس أحادي البُعد، واستخدام طريقة كرونباخ ألفا في حال عدم تحقق افتراضاتها الخاصة بنموذج تاو المتكافئ فإنها تعطي تقديرا أدنى لقيمة معامل الثبات من القيمة الحقيقية (تيغزة، 2017) (Kamata, Turhan & Darandari, 2003) (Osbourn, 2000)، وكذلك يؤكد كروسر وألجينا (CroKer & Algina, 1986) أنه عند استخدام معامل ألفا في المقاييس متعددة الأبعاد التي لا تحقق افتراضات نموذج تاو فإنها تُعطي تقديراً بالحد الأدنى لمعاملات الثبات، كما أشار (Schmidt & Hunter, 1996) أنه من غير المناسب (غير صالح) استخدام معامل كرونباخ ألفا لتقدير الثبات في المقاييس متعددة الأبعاد. أما دراسة (Cortina, 1993) فقد أشارت إلى أن تقدير قيمة الثبات باستخدام معامل كرونباخ ألفا يميل إلى الإنخفاض بسبب تعدد الأبعاد، ويميل نحو الإرتفاع في حال زيادة الإرتباطات بين الفقرات، كما أشارت مجموعة من الدراسات إلى أن إستخدام معامل كرونباخ ألفا في المقاييس متعددة الأبعاد يؤدي إلى تقديرات متحيزة (Zimmerman, Zumbo & Lalonde, 1993; Komaroff, 1997; Murphy & DeShon, 2000; Osbourn, 2000; Raykov, 2001)، وبذلك قام الباحثون بإقتراح طرق أخرى منها معامل ألفا الطبقي، ومعامل ثيتا، ومعامل أوميغا، ومعامل أوميغا الموزونة، ومعامل بيتا. وطوّر كلٌ من معامل أوميغا كمعامل ثبات للإتساق الداخلي، ويقوم على التحليل العاملي، (Ercan et al., 2007) ويعطي تقديراً للثبات سواءاً كان المقياس أحادي البُعد أو متعدد الأبعاد، ويحسب معامل أوميغا باستخدام المعادلة التالية (Barrett, 2014):

$$\omega = \frac{\left(\sum_{i=1}^{k} \lambda_i\right)^2}{\left(\left(\sum_{i=1}^{k} \lambda_i\right)^2 + \sum_{i=1}^{k}\left(1-\lambda_i^2\right)\right)} \quad \text{.....(2)}$$

حيث تشير λ_i إلى تشبع العامل i ، k عدد الفقرات.

وقام كاماتا وآخرون (Kamata et al., 2003) بمقارنة ألفا مع ثلاث طرائق لتقدير الثبات (ألفا الطبقية، أقصى ثبات، أوميغا) في حال تعدد الأبعاد؛ حيث أشارت النتائج إلى أن الطرائق الثلاث قدّرت الثبات بشكل أفضل بكثير من معامل ألفا في جميع الظروف، حيث قلل معامل ألفا بشكل كبير من الثبات الحقيقي، وكان أكثر وضوحاً عندما كان الإرتباط بين الأبعاد قليلاً وحجم العينة صغيراً 50، أما الفروق في تقدير الثبات حسب حجم العينة كانت قليلة، وأشارت أن معامل أوميغا قد يُبالغ في

تقدير قيمة الثبات، وأوصت بعدم الإعتماد على معامل واحد لتقدير الثبات واستخدام أعلى قيمة في التقدير، كما توصلت دراسة (Segars, 1997) إلى أن تقدير الثبات باستخدام معامل أوميغا في الإختبار أحادي البُعد يساوي تقدير معامل ألفا، وفي المقابل أشار (Yurdugul, 2006) أن معامل أوميغا يعطي تقديرا أعلى في حال تعدد الأبعاد للإختبار. وفي دراسة (Ercan et al., 2007) تم إختبار ملاءمة حجم العينة في معاملات الثبات المختلفة، وأظهرت النتائج أن تقديرات الثبات باستخدام معامل ألفا وثيتا لا ترتبط بحجم العينة حتى في العينات الصغيرة، إلا أن معامل أوميغا يتطلب حجم عينة كبيرة، وأشار (Javali, 2011) إلى أن حجم عينة 50 كاف لحساب أي من معاملات الثبات. وفي دراسة (Green & Yang, 2015) أظهرت النتائج أن معامل ألفا يمكن أن يوفر تقييماً دقيقاً ونسبياً للثبات، ولكن بعض الباحثين قد يسيءُ تفسيره على أنه مؤشر للإتساق الداخلي أو التجانس، كما أن معامل أوميغا يوفر معلومات أكثر حول الثبات إلا أنها قد تكون متحيزة. كما تجدر الإشارة إلى أن معاملي ألفا وأوميغا كلاهما غير متحيزين عندما تتحقق افتراضات نموذج تاو (tau) على الأقل (Padilla & Divers, 2013). وأشار (Kline, 1986) المشار له في (Centre for Academic Success, 2017) أن حجم العينة المناسب لتقدير الثبات يجب أن لا يقل عن 300 فرد، وأشار (Yurdugül, 2008) على أنه من الممكن تقدير الثبات في ظل حجم عينة قليلة في ظروف خاصة مرتبطة بتشبّع العامل الأول بحيث يكون أكبر من 6 فإن حجم عينة 30 كاف، أما إذا كان التشبع أكبر من 3 وأقل من 6 فسيكون الحد الأدنى لحجم العينة 100، كما أفاد إذا كان التشبع أقل من 3، يجب أن لا يقل حجم العينة عن 300، وفي حال أن يكون حجم العينة 500 فأكثر فإن تقدير معامل الثبات لا يرتبط بالتشبّع للعامل الأول، كما اقترح كل من (Bonnett, 2002; Charter, 2003) على أن يكون حجم العينة أكبر من 250 فرد.

مشكلة الدراسة

تتلخص مشكلة الدراسة في أن تقدير الثبات يعتبر جزءاً أساسياً من الخصائص السيكومترية للمقاييس والإختبارات، ويستخدم معامل ثبات كرونباخ ألفا على نطاق واسع في تقدير الثبات. ومن خلال ملاحظة الباحث واطلاعه على الدراسات العربية يكاد يكون هو المعامل المسيطر على جميع الدراسات العربية، ويعود ذلك كون معامل ألفا يتطلب تطبيق الأداة مرةً واحدةً، وكما أنه سهل الحساب، ومتوافر في البرامج

الإحصائية، إلّا أن إستخدامه قد يكون غير مناسب بسبب إنتهاك الإفتراضات الخاصة به؛ مما يؤدي إلى تقدير الثبات بالحد الأدنى، وقد قدم المتخصصون العديد من معاملات تقدير الثبات مثل ألفا الطبقي وأوميغا ومعامل ثيتا. وهذه المعاملات الخاصة بتقدير ثبات الإتساق الداخلي قد تختلف دقة تقديرها بإختلاف عدد الأبعاد، أو اختلاف طول الإختبار، أو اختلاف حجم العينة، وبذلك نجد قلة في الدراسات التي قارنت بين دقة تقدير هذه المعاملات في ضوء هذه المتغيرات، وخاصة الدراسات العربية. ولذلك جاءت هذه الدراسة لتوفير معلومات حول دقة تقدير معاملي الثبات (كرونباخ ألفا، أوميغا)، باختلاف كل من عدد الأبعاد وطول الإختبار، وحجم العينة.

أسئلة الدراسة

السؤال الأول: هل تختلف قيمة الثبات المقدرة باستخدام معاملات الثبات (أوميغا، كرونباخ ألفا) باختلاف عدد الأبعاد (بُعد واحد، بُعدين، ثلاثة أبعاد)، وطول الإختبار 10، 30، 60؟

السؤال الثاني: هل تختلف قيمة الثبات المقدرة باستخدام معاملات الثبات (أوميغا، كرونباخ ألفا) باختلاف عدد الأبعاد (بُعد واحد، بُعدين، ثلاثة أبعاد)، وحجم العينة 50، 100، 250، 500، 1000؟

السؤال الثالث: هل تختلف قيمة الثبات المقدرة باستخدام معاملات الثبات (أوميغا، كرونباخ ألفا) باختلاف عدد الأبعاد (بُعد واحد، بُعدين، ثلاثة أبعاد)، وطول الإختبار 10، 30، 60 وحجم العينة 50، 100، 250، 500، 1000؟

السؤال الرابع: ما قيمة الخطأ المعياري المصاحب في تقدير قيمة الثبات بطريقتي تقدير الثبات (اوميغا، كرونباخ ألفا) باختلاف متغيرات الدراسة (عدد الأبعاد، طول الإختبار، حجم العينة)؟

حدود الدراسة

- استخدمت الدراسة بيانات مولدة باستخدام برنامج Wingen-V3.
- اقتصار البيانات في الدراسة على الإستجابات الثنائية Binary Response (صفر، واحد).

أهمية الدراسة

تكتسب الدراسة أهميتها كونها الدراسة الأولى التي تقارن بين معاملي كرونباخ ألفا وأوميغا في ضوء ثلاثة متغيرات (عدد الأبعاد، وطول الاختبار، وحجم العينة) وتُبين دقة استخدامهما، كما أنها توجه الباحثين إلى استخدام معاملات الثبات المناسبة لتقدير ثبات المقاييس والاختبارات التي يستخدمونها. كما أنه يستفاد من هذا البحث في رفد المؤشرات الإحصائية في التحقق من الخصائص السيكومترية للأدوات والمقاييس المستخدمة في الدراسات المتنوعة في المجالات المعرفية.

تصميم الدراسة

لتحقيق أهداف الدراسة الحالية في الكشف عن دقة تقدير معاملي ثبات الإتساق الداخلي (ألفا وأوميغا) في أدوات القياس حسب النموذج ثلاثي المعلمة وذلك باختلاف عدد الأبعاد وطول الإختبار وحجم العينة، اعتمدت الدراسة على بيانات مولّدة (محاكاة)، وفق المتغيرات المستقلة (عدد الأبعاد وطول الإختبار وحجم العينة)، شملت ثلاثة مستويات للأبعاد (بُعد واحد، بُعدين، ثلاثة أبعاد)، وثلاثة مستويات لطول الإختبار 10، 30، 60 فقرة، وخمسة مستويات من أحجام العينات 50، 100، 250، 500، 1000؛ (Foley, 2010). بذلك تمّ توليد 45 مجموعة بيانات ظرف وكل مجموعة تم تكرارها 30 (replicated) مرة لتمثل الحد الأدنى لاستقرار معاملات الثبات، وفي المرحلة الثانية تمّ حساب معامل الثبات لكل مجموعة بيانات باستخدام معاملي الثبات ألفا وأوميغا، كما حُسبت المتوسطات الحسابية والإنحرافات المعيارية والخطأ المعياري لقيم الثبات على اختلاف طريقة التقدير ومتغيرات الدراسة، وفي حال ظهور فروق تم إجراء اختبار (ت) للعينات المستقلة باختلاف متغيرات الدراسة لتحديد ما إذا كانت معاملات ألفا وأوميغا تتغير وفقها.

توليد البيانات

وُلّدت البيانات باستخدام برنامج Wingen-V3 الخاص بتوليد البيانات من مواقف متعددة ومختلفة قد لا تتحقق في الواقع العملي، وهذا البرنامج من إعداد (Han & Hambbelton, 2007)، وتمّ توليد البيانات بناءً على المحكات الواردة في الدراسات السابقة بشأن محاكاة البيانات للنموذج الثلاثي المعلمة، وهي؛ صعوبة الفقرة، تمييز الفقرة، التخمين، القُدرة (Feinberg & Rubright, 2016; Bulut & Sum-bul, 2017; Mooney, 1997; Mislevy & Stocking, 1989; Harwell & Baker, 1991)،

حيث وُلّدت معلمة الصعوبة للفقرات وفق التوزيع الطبيعي (0.1) a≅N؛ أما معلمة التمييز للفقرات فقد وُلّدت وفق توزيع لوغاريتمي عادي (0.3, 0.2) a~lnN؛ ووُلّدت معلمة التخمين وفق توزيع بيتا (20, 90) c~Beta، أما معالم القُدرة فقد تمّ تقديرها من التوزيع الطبيعي (0,1) θ~، وبناءً على التوصيات من الدراسات السابقة فقد تمّ تكرار (replicated) استجابات المفحوصين (صفر، واحد) بواقع 30 مرة لكل ظرف من ظروف الدراسة وعددها 45.

نتائج الدراسة

هدفت الدراسة إلى مقارنة طريقتين لتقدير الثبات (α، ω) باختلاف عدد من المتغيرات (عدد الأبعاد، حجم العينة، عدد الفقرات)، وذلك عن طريق الإجابة عن أسئلة الدراسة:

أوّلاً- النتائج المتعلقة بسؤال الدراسة الأول: للإجابة عن سؤال الدراسة الأوَّل؛ تم حساب المتوسطات الحسابية والإنحرافات المعيارية لقيم ثبات الإتساق الداخلي باختلاف طريقة التقدير (α، ω)، وذلك حسب متغيري الدراسة عدد الأبعاد (بُعد واحد، بُعدين، ثلاثة ابعاد) وطول الإختبار 10، 30، 60، واجراء إختبار ت، وذلك كما هو مُبيَّن في جدول رقم 1.

جدول رقم 1

الأوساط الحسابية والإنحرافات المعيارية والخطأ المعياري وقيمة اختبار ت لقيم ثبات الإتساق الداخلي باختلاف طريقة تقدير الثبات (α، ω) وفقًا لعدد الأبعاد وطول الإختبار

عدد الأبعاد	طول الإختبار	العدد	طريقة تقدير الثبات						قيمة ت
			أوميغا			ألفا			
			الوسط الحسابي	الإنحراف المعياري	الخطأ المعياري	الوسط الحسابي	الإنحراف المعياري	الخطأ المعياري	
1	10	150	0.622	0.058	0.005	0.612	0.064	0.005	1.361
	30	150	0.830	0.022	0.002	0.829	0.022	0.002	0.599
	60	150	0.904	0.009	0.001	0.902	0.009	0.001	1.822
الكلي		450	0.758	0.125	0.006	0.781	0.129	0.006	0.511

تابع/ جدول رقم 1

الأوساط الحسابية والإنحرافات المعيارية والخطأ المعياري وقيمة اختبار ت لقيم ثبات الإتساق الداخلي باختلاف طريقة تقدير الثبات (α، ω) وفقًا لعدد الأبعاد وطول الإختبار

عدد الأبعاد	طول الإختبار	العدد	طريقة تقدير الثبات						قيمة ت
			أوميغا			ألفا			
			الوسط الحسابي	الإنحراف المعياري	الخطأ المعياري	الوسط الحسابي	الإنحراف المعياري	الخطأ المعياري	
2	10	150	0.765	0.071	0.006	0.725	0.033	0.003	*6.381
	30	150	0.927	0.012	0.001	0.881	0.007	0.001	*38.697
	60	150	0.970	0.012	0.001	0.928	0.004	0.000	*41.891
الكلي		450	0.887	0.098	0.005	0.845	0.891	0.004	*6.853
3	10	150	0.840	0.021	0.002	0.786	0.023	0.002	*21.203
	30	150	0.946	0.012	0.001	0.896	0.012	0.001	*37.446
	60	150	0.990	0.005	0.000	0.940	0.004	0.000	*96.951
الكلي		450	0.925	0.065	0.003	0.874	0.067	0.003	*11.758

يلاحظ من جدول رقم 1 أن معامل الثبات يزيد بزيادة طول الإختبار في كل من طريقتي التقدير، كما يُلاحظ وجود فروق ظاهرة بين المتوسطات الحسابية لقيم ثبات الإتساق الداخلي ناتجة عن اختلاف طريقة التقدير وفقًا لاختلاف عدد الأبعاد وطول الإختبار؛ وللتَحقّق من جوهرية الفروق الظاهرة؛ تمَّ إجراء إختبار ت للعينات المستقلة لقيم ثبات الإتساق الداخلي باختلاف طريقة التقدير (α، ω) وفقًا لعدد الأبعاد وطول الإختبار. كما يلاحظ في جدول رقم 1 عدم وجود فروق دالة إحصائيًا عند مستوى الدلالة $\alpha \leq 0.05$ بين المتوسطات الحسابية لثبات الإتساق الداخلي باختلاف طريقة التقدير (α، ω) عندما يكون عدد الأبعاد بُعدا واحدا؛ باختلاف طول الإختبار 10، 30، 60. كما يتضّح وجود فروق دالة إحصائيًا عند مستوى الدلالة $\alpha \leq 0.05$ بين المتوسطات الحسابية لثبات الإتساق الداخلي باختلاف طريقة التقدير (α، ω) عندما يكون عدد الأبعاد بُعدين وثلاثة أبعاد على إختلاف طول الإختبار 10، 30، 60 لصالح طريقة تقدير الثبات أوميغا.

ثانيًا- النتائج المتعلقة بسؤال الدراسة الثاني: للإجابة عن سؤال الدراسة الثاني؛ تمّ حساب المتوسطات الحسابية والإنحرافات المعيارية لقيم ثبات الإتساق الداخلي باختلاف طريقة التقدير (α، ω)، وذلك حسب متغيري الدراسة عدد الأبعاد (بُعد واحد، بُعدان، ثلاثة أبعاد) وحجم العينة 50، 100، 250، 500، 1000، وإجراء إختبار ت، وذلك كما هو مُبيَّن في جدول رقم 2.

جدول رقم 2

الأوساط الحسابية والإنحرافات المعيارية والخطأ المعياري وقيمة اختبار ت لقيم ثبات الإتساق الداخلي باختلاف طريقة تقدير الثبات (α، ω) وفقًا لعدد الأبعاد وحجم العينة

عدد الأبعاد	حجم العينة	العدد	طريقة تقدير الثبات						قيمة ت
			أوميغا			ألفا			
			الوسط الحسابي	الإنحراف المعياري	الخطأ المعياري	الوسط الحسابي	الإنحراف المعياري	الخطأ المعياري	
1	50	90	.785	.143	.015	.778	.149	.016	.334
	100	90	.775	.138	.015	.770	.144	.015	.249
	250	90	.767	.129	.014	.760	.136	.014	.352
	500	90	.792	.106	.011	.791	.106	.011	.044
	1000	90	.806	.101	.011	.805	.101	.011	.099
2	50	90	.873	.130	.014	.849	.086	.009	1.463
	100	90	.874	.102	.011	.832	.109	.012	*2.644
	250	90	.889	.088	.009	.840	.093	.010	*3.659
	500	90	.895	.085	.009	.847	.081	.009	*3.843
	1000	90	906.	.072	.008	.855	.072	.008	*4.753
3	50	90	.918	.069	.007	.864	.073	.008	*5.114
	100	90	.937	055.	.006	886.	.056	.006	*6.113
	250	90	.927	.063	.007	.875	.064	.007	*5.469
	500	90	.918	.066	.007	.869	.068	.007	*4.869
	1000	90	.926	.069	.007	.874	.070	.007	*5.039

يتضح من جدول رقم 2؛ وجود فروق ظاهرة بين المتوسطات الحسابية لقيم ثبات الاتساق الداخلي ناتجة عن اختلاف طريقة التقدير وفقًا لاختلاف عدد الأبعاد وحجم

العينة؛ وللتَحقُّق من جوهرية الفروق الظاهرة؛ تمَّ إجراء اختبار ت للعينات المستقلة لقيم ثبات الاتساق الداخلي باختلاف طريقة التقدير (α، ω) وفقًا لعدد الأبعاد وحجم العينة. ويتضح كذلك من جدول رقم 2 عدم وجود فروق دالة إحصائيًا عند مستوى الدلالة α=0.05 بين المتوسطات الحسابية لثبات الاتساق الداخلي باختلاف طريقة التقدير (α، ω) عندما يكون عدد الأبعاد بُعدا واحدا باختلاف حجوم العينات 50، 100، 250، 500، 1000، كما لا يوجد فروق جوهرية بين المتوسطات عند عدد الأبعاد (بُعدين) وحجم عينة 50، كما يتضح وجود فروق دالة إحصائيًا عند مستوى الدلالة α=0.05 بين المتوسطات الحسابية لثبات الاتساق الداخلي باختلاف طريقة التقدير (α، ω) عندما يكون عدد الأبعاد بعدين وحجم عينة 100، 250، 500، 1000، وكذلك عند ثلاثة أبعاد وأحجام العينات 50، 100، 250، 500، 1000.

ثالثاً- النتائج المتعلقة بسؤال الدراسة الثالث: للإجابة عن سؤال الدراسة الثالث؛ تمّ حساب المتوسطات الحسابية والانحرافات المعيارية لقيم ثبات الاتساق الداخلي باختلاف طريقة التقدير (α، ω)، وذلك حسب متغيرات الدراسة عدد الأبعاد (بُعد واحد، بُعدين، ثلاثة أبعاد)، وطول الاختبار 10، 30، 60 فقرة، وحجم العينة 50، 100، 250، 500، 1000، وإجراء اختبار ت، وذلك كما هو مُبيَّن في جدول رقم 3.

جدول رقم 3

الأوساط الحسابية والانحرافات المعيارية لقيم ثبات الاتساق الداخلي باختلاف طريقة تقدير الثبات (α، ω) وفقًا لعدد الأبعاد وطول الاختبار وحجم العينة

عدد الأبعاد	عدد الفقرات	حجم العينة	العدد	طريقة التقدير: أوميغا			طريقة التقدير: ألفا			قيمة اختبار ت
				الوسط الحسابي	الانحراف المعياري	الخطأ المعياري	الوسط الحسابي	الانحراف المعياري	الخطأ المعياري	
1	10	50	30	0.591	0.055	0.010	0.575	0.058	0.011	1.038
		100	30	0.590	0.060	0.010	0.577	0.062	0.011	0.808
		250	30	0.596	0.039	0.007	0.579	0.038	0.007	1.695
		500	30	0.653	0.019	0.003	0.653	0.020	0.004	*0.708
		1000	30	0.679	0.044	0.008	0.677	0.044	0.008	0.236

تابع/ جدول رقم 3

الأوساط الحسابية والانحرافات المعيارية لقيم ثبات الاتساق الداخلي باختلاف طريقة تقدير الثبات (α، ω) وفقًا لعدد الأبعاد وطول الاختبار وحجم العينة

عدد الأبعاد	عدد الفقرات	حجم العينة	العدد	طريقة التقدير						قيمة اختبار ت
				أوميغا			ألفا			
				الوسط الحسابي	الانحراف المعياري	الخطأ المعياري	الوسط الحسابي	الانحراف المعياري	الخطأ المعياري	
1	30	50	30	0.862	0.018	0.003	0.859	0.018	0.003	0.602
		100	30	0.839	0.014	0.002	0.837	0.014	0.003	0.325
		250	30	0.809	0.013	0.002	0.807	0.013	0.002	0.673
		500	30	0.815	0.008	0.001	0.815	0.007	0.001	0.273
		1000	30	0.826	0.004	0.001	0.825	0.004	0.001	0.853
	60	50	30	0.903	0.010	0.002	0.900	0.010	0.002	1.322
		100	30	0.897	0.006	0.001	0.895	0.007	0.001	1.003
		250	30	0.897	0.005	0.001	0.895	0.004	0.001	1.574
		500	30	0.908	0.002	0.000	0.906	0.002	0.000	*2.466
		1000	30	0.914	0.002	0.000	0.913	0.002	0.000	1.776
2	10	50	30	0.736	0.142	0.025	0.736	0.029	0.005	−0.009
		100	30	0.734	0.030	0.005	0.682	0.030	0.006	*6.642
		250	30	0.769	0.018	0.003	0.713	0.016	0.003	*12.694
		500	30	0.779	0.014	0.002	0.736	0.010	0.002	*13.876
		1000	30	0.808	0.007	0.001	0.756	0.007	0.001	*29.056
	30	50	30	0.908	0.014	0.002	0.876	0.011	0.002	*9.776
		100	30	0.931	0.009	0.002	0.888	0.007	0.001	*20.251
		250	30	0.931	0.006	0.001	0.880	0.005	0.001	*37.453
		500	30	0.930	0.005	0.001	0.879	0.003	0.001	*46.448
		1000	30	0.933	0.002	0.000	0.881	0.002	0.000	*93.416
	60	50	30	0.976	0.007	0.001	0.936	0.003	0.001	*30.54
		100	30	0.956	0.015	0.003	0.926	0.003	0.001	*10.992
		250	30	0.967	0.010	0.002	0.927	0.002	0.000	*21.981
		500	30	0.975	0.004	0.001	0.926	0.001	0.000	*72.628
		1000	30	0.977	0.001	0.000	0.926	0.001	0.000	*177.756

تابع/ جدول رقم 3

الأوساط الحسابية والانحرافات المعيارية لقيم ثبات الاتساق الداخلي باختلاف طريقة تقدير الثبات (α، ω) وفقًا لعدد الأبعاد وطول الاختبار وحجم العينة

عدد الأبعاد	عدد الفقرات	حجم العينة	العدد	طريقة التقدير: أوميغا			طريقة التقدير: ألفا			قيمة اختبار ت
				الوسط الحسابي	الانحراف المعياري	الخطأ المعياري	الوسط الحسابي	الانحراف المعياري	الخطأ المعياري	
		50	30	0.831	0.030	0.005	0.771	0.032	0.006	*7.506
		100	30	0.865	0.015	0.003	0.813	0.016	0.003	*12.665
	10	250	30	0.841	0.012	0.002	0.788	0.012	0.002	*17.088
		500	30	0.830	0.009	0.002	0.778	0.008	0.002	*22.597
		1000	30	0.831	0.007	0.001	0.778	0.006	0.001	*32.149
		50	30	0.932	0.012	0.002	0.880	0.012	0.002	*16.291
		100	30	0.951	0.006	0.001	0.903	0.006	0.001	*30.135
3	30	250	30	0.951	0.004	0.001	0.900	0.003	0.001	*59.223
		500	30	0.938	0.003	0.001	0.890	0.002	0.000	*67.657
		1000	30	0.958	0.002	0.000	0.907	0.001	0.000	*124.903
		50	30	0.992	0.003	0.000	0.941	0.002	0.000	*80.29
		100	30	0.995	0.001	0.000	0.944	0.001	0.000	*152.257
	60	250	30	0.987	0.001	0.000	0.937	0.001	0.000	*138.759
		500	30	0.986	0.005	0.001	0.939	0.006	0.001	*31.458
		1000	30	0.989	0.005	0.001	0.938	0.004	0.001	*45.966

يلاحظ من جدول رقم 3؛ وجود فروق ظاهرة بين المتوسطات الحسابية لقيم ثبات الاتساق الداخلي ناتجة عن اختلاف طريقة التقدير وفقًا لاختلاف عدد الأبعاد وطول الاختبار وحجم العينة؛ وللتَحقُّق من جوهرية الفروق الظاهرة؛ تمَّ إجراء اختبار ت للعينات المستقلة لقيم ثبات الاتساق الداخلي باختلاف طريقة التقدير (α، ω) وفقًا لعدد الأبعاد وطول الاختبار وحجم العينة. ويتضح -أيضا- من جدول رقم 3 عدم وجود فروق دالة إحصائيًا عند مستوى الدلالة $\alpha \leq 0.05$ بين المتوسطات الحسابية لثبات الاتساق الداخلي باختلاف طريقة التقدير (α، ω) عندما يكون عدد الأبعاد

بُعد واحد وطول الاختبار 10 فقرات باختلاف أحجام العينات -بإستثناء حجم عينة 500 - لصالح طريقة التقدير أوميغا. كما أنه لا يوجد فروق دالة احصائياً بين المتوسطات الحسابية لتقدير الثبات باختلاف طريقة التقدير عندما يكون عدد الأبعاد بُعدا واحدا و30 فقرة؛ باختلاف جميع أحجام العينات، كما أنه لا يوجد فروق دالة احصائياً بين المتوسطات الحسابية لتقدير الثبات باختلاف طريقة التقدير عندما يكون عدد الأبعاد بعدا واحدا و60 فقرة باختلاف جميع أحجام العينات باستثناء حجم عينة 500 لصالح طريقة تقدير أوميغا. كما يتضح أيضاً من الجدول رقم 3؛ وجود فروق دالة إحصائيًا عند مستوى الدلالة $\alpha \leq 0.05$ بين المتوسطات الحسابية لثبات الاتساق الداخلي باختلاف طريقة التقدير (α، ω) عندما يكون عدد الأبعاد بُعدين وطول الاختبار 10 فقرات باختلاف أحجام العينات لصالح طريقة التقدير أوميغا باستثناء حجم العينة 50؛ حيث تكون الفروق غير جوهرية، كما أنه يوجد فروق دالة احصائياً بين المتوسطات الحسابية لتقدير الثبات باختلاف طريقة التقدير عندما يكون عدد الأبعاد بُعدين وطول اختبار 30 فقرة و60 فقرة؛ باختلاف جميع أحجام العينات لصالح طريقة التقدير أوميغا. وكذلك يتضح من الجدول رقم 3؛ وجود فروق دالة إحصائيًا عند مستوى الدلالة $\alpha \leq 0.05$ بين المتوسطات الحسابية لثبات الإتساق الداخلي باختلاف طريقة التقدير (α، ω) عندما يكون عدد الأبعاد ثلاثة أبعاد وطول الإختبار 10 فقرات و30 فقرة و60 فقرة باختلاف أحجام العينات لصالح طريقة التقدير أوميغا.

رابعاً- النتائج المتعلقة بسؤال الدراسة الرابع: للإجابة عن سؤال الدراسة الرابع؛ تم حساب الخطأ المعياري المصاحب لتقدير قيم ثبات الإتساق الداخلي باختلاف طريقة التقدير (α، ω)، وذلك حسب متغيرات الدراسة عدد الأبعاد بُعد واحد، بُعدين، ثلاثة أبعاد، وطول الإختبار 10، 30، 60 فقرة، وحجم العينة 50، 100، 250، 500، 1000، وذلك كما هو مُبيَّن في الجداول السابقة أرقام 1، 2، 3. ويتضح من الجدول رقم 1، الخاص بالخطأ المعياري أن الأخطاء المعيارية متقاربة لتقديرات الثبات باختلاف طريقتي التقدير ونفس طول الإختبار، كما يتضح من الجدول رقم 2 الخاص بالخطأ المعياري أن الأخطاء المعيارية متقاربة عند تقدير الثبات باختلاف عدد الأبعاد ونفس حجم العينة، كما يتضح من جدول رقم 3 أن الخطأ المعياري في تقدير الثبات باختلاف عدد الأبعاد وطول الإختبار وحجم العينة متقارب في ظل نفس الظرف.

مناقشة النتائج

تناولت الدراسة الحالية عددا من المجالات الرئيسة الخاصة بدقة تقدير الثبات باستخدام معاملي ألفا وأوميغا؛ حيث يؤثر استخدام المعامل المناسب في دقة تقدير الثبات للمقياس، كما يجب على مطوري ومستخدمي المقاييس تحديد معامل الثبات المناسب في ضوء عدد الأبعاد وطول الإختبار وحجم العينة. وتعتبر الدراسة الحالية من الدراسات القليلة التي قارنت بين دقة تقدير الثبات باستخدام معاملي ألفا وأوميغا باختلاف عدد من المتغيرات، ممثلة في دراسة تأثير عدد الأبعاد على دقة تقدير الثبات باستخدام معاملي ألفا أوميغا، إضافة إلى دراسة تأثير طول الإختبار على دقة تقدير الثبات باستخدام معاملي ألفا أوميغا، ودراسة تأثير حجم العينة على دقة تقدير الثبات باستخدام معاملي ألفا أوميغا. وتحقق هدف الدراسة بحساب المتوسطات الحسابية والإنحرافات المعيارية لقيم ثبات الإتساق الداخلي باختلاف طريقة تقدير الثبات (α، ω) وفقًا لعدد الأبعاد ولطول الإختبار، وإجراء اختبار ت للعينات المستقلة للحكم على دقة التقدير مما زود بمؤشرات عن الاختلاف في تقدير ثبات الإتساق الداخلي. وأظهرت النتائج أنه لا يوجد فروق جوهرية بين تقديرات الثبات باختلاف طريقة التقدير (α، ω) في حال كان المقياس ببُعد واحد باختلاف طول الإختبار، ووجود فروق جوهرية بين متوسطات ثبات الإتساق الداخلي باختلاف طريقة التقدير (α، ω) باختلاف طول الإختبار في حالة تعدد الأبعاد للمقياس لصالح طريقه أوميغا. وجاءت هذه النتيجة متسقة مع دراسات تناولت طول الاختبار بأن هناك علاقة إيجابية عالية بين طول الإختبار وتقدير معامل الثبات كما هو مذكورفي (Ercan et al., 2007)؛ حيث إن عدد الأبعاد وطول الإختبار يمثلان أهم العوامل التي تفسر تباين دقة تقدير الثبات باختلاف طريقة التقدير، وتتفق النتيجة كذلك مع ما توصل إليه كلٌ من دراسة شميدت وهانتر ودراسة كاماتا وآخرون ودراسة كورتينا (Cortina, 1993; Kamata et al., 2003; Segars, 1997; Yurdugul, 2006; Schmidt & Hunter, 1996)، بأن معامل كرونباخ ألفا يعطي تقديرات غير دقيقة عند استخدامه في الإختبارات والمقاييس التي تتكون من عدة أبعاد. كما اتفقت هذه النتيجة مع دراسة (Padilla & Divers, 2013) بالحصول على تقدير غير متحيز للثبات باستخدام معاملي ألفا وأوميغا عندما تتحقق افتراضات نموذج تاو tau.

وفيما يخص عدد الأبعاد وحجم العينة، أظهرت النتائج أنه لا يوجد فروق في

تقدير الثبات بين معاملي ألفا وأوميغا؛ عندما يكون الإختبار ببُعد واحد باختلاف أحجام العينات. كما أظهرت أن معامل أوميغا يعطي تقديراً للثبات أعلى من معامل ألفا في حال تعدد الأبعاد باختلاف أحجام العينات. ويعزو الباحث هذه النتيجة إلى عدد الأبعاد وليس عدد الأفراد، كما تتفق هذه النتيجة مع دراسة (Yurdugül, 2008; Javali, 2011) في أنه من الممكن تقدير الثبات عندما يكون حجم العينة قليلا. واختلفت مع دراسة (Kamata et al., 2003)؛ حيث جاء تقدير الثبات باستخدام معامل ألفا أقل من الثبات الحقيقي، وخصوصاً عندما يكون حجم العينة صغيراً 50، رغم أن دراسة (Ercan et al., 2007) توصلت إلى أن تقدير معامل الثبات باستخدام معامل ألفا لا ترتبط بحجم العينة حتى في العينات الصغيرة، ويكمن الإختلاف مع هذه الدراسة في كون معامل أوميغا يتطلب حجم عينة كبيرة، كما أن دراستي (Bonnett, 2002; Charter, 2003) أشارت بأن يكون حجم العينة أكبر من 250 فرد لتقدير الثبات، ويكمن الإختلاف بأنه لا يوجد فرق بين تقديرات الثبات باختلاف المعاملين عند تحقق شرط أحادية البُعد، وإنما الإختلاف في حال تعدد الأبعاد، وعلى الرغم من أن دراسة (Terry & Kelley, 2012) أكدت بأن زيادة حجم العينة يؤدي إلى تحسين دقة التقدير في معامل الثبات.

أما فيما يخص تقاطع عدد الأبعاد وطول الإختبار وحجم العينة، فقد أظهرت النتائج بأنه لا يوجد فروق جوهرية بين تقديرات الثبات باختلاف طريقة التقدير (α، ω) عندما يكون الإختبار بُعدا واحدا وطول الإختبار 10، 30، 60 فقرة، باختلاف أحجام العينات باستثناء حجم عينة 500، أما على بقية التقاطعات بين الأبعاد (بُعدان، ثلاثة أبعاد) وطول الإختبار 10، 30، 60 فقرة وحجم العينة 50، 100، 250، 500، 1000 فتوجد فروق جوهرية لصالح طريقة التقدير أوميغا؛ باستثناء بُعدين وطول الإختبار 10 فقرات وحجم عينة 50 فرد فلا يوجد فرق جوهري. ويعزو الباحث الفرق بين قيم الثبات المقدّرة عندما يتكون المقياس من عدة أبعاد يعود لطريقة الإستخدام، وبذلك تتفق هذه النتيجة مع دراسة (Yurdugul, 2006) أن معامل أوميغا يعطي تقديراً أعلى في حال تعدد الأبعاد للإختبار. كما تختلف هذه النتيجة مع دراسة (Ercan et al., 2007) التي أكدت أن استخدام معامل أوميغا لتقدير الثبات يحتاج إلى حجم عينة كبير.

وفيما يخص الخطأ المعياري في التقدير، أظهرت النتائج أن الفروق في تقديرات الثبات باستخدام المعاملين قليلة باختلاف ظروف الدراسة؛ سواءاً باختلاف

عدد الأبعاد في ضوء طول الإختبار، أو عدد الأبعاد في ضوء حجم العينة، أو عدد الأبعاد في ضوء كل من طول الإختبار وحجم العينة. ويعزو الباحث هذه النتيجة كون البيانات التي استخدمت في الدراسة هي بيانات مولّدة حسب شروط وظروف معيارية أهمها التوزيع الطبيعي، وإتفقت هذه النتيجه مع دراسة (Padilla & Divers, 2013) التي أشارت إلى أن معاملات ألفا وأوميغا كلاهما غير متحيزين عندما تتحقق إفتراضات tau على الأقل، وتختلف مع نفس الدراسة في حال عدم تحقق إفتراضات نموذج تاو، وتختلف مع كل من الدراسات التالية (Zimmerman & Zumbo & Lalonde, 1993; Komaroff, 1997; Murphy & DeShon, 2000; Osbourn, 2000; Raykov, 2001)، التي أشارت إلى أن إستخدام معامل كرونباخ ألفا في المقاييس متعددة الأبعاد يؤدي إلى تقديرات متحيزة.

التوصيات

- الاستناد الى مسارات بحثية أخرى في التوصل إلى الاختلافات بين طرق تقدير الثبات.
- استخدام معامل أوميغا لتقدير الثبات في حال عدم تحقق افتراضات نموذج تاو أو أن يكون المقياس مكّونا من أكثر من بُعد واحد.
- إجراء دراسات لتقدير الثبات في نماذج النظرية الحديثة في القياس (النموذج ثنائي المعلمة والنموذج أحادي المعلمة).
- إجراء دراسات لمقارنة طرق تقدير الثبات ببيانات حقيقية، واستخذام معاملات ثبات أخرى.
- إجراء دراسات تقارن بين دقة تقدير الثبات في المعاملات المختلفة باستخدام بيانات مدرجة على تدريج ليكرت.
- إجراء دراسات على البيانات التي لا تتوزع توزيعا طبيعيا لمعرفة أي طرق تقدير الثبات هي الأنسب.

A Comparison of Alpha and Omega Coefficients in Determining the Accuracy of Test Reliability in the Light of Dimensions Number, Test Length and Sample Size; Based on Generated Data

Dr. Omar S. Obeidat

College of Education – Imam Abdulrahman Bin Faisal University
K.S.A

Abstract

This study aims to evaluate the accuracy of reliability coefficient estimation using the Cronbach's alpha and Omega coefficients using generated data through Computerized program according to three independent variables: number of dimensions (one, two, and three), length of the test 10, 30, 60 items, sample size of 50, 100, 250, 500, and 1000. 45 different conditions were generated. The reliability values were estimated using alpha and omega coefficients.

Results of T-test for independent samples indicated that there were no statistically significant difference between the two coefficient estimates of reliability in the case of one-dimension tests with different sample sizes and test-lengths. The results also indicated statistically significant differences between the reliability estimates of the two coefficients in the case of two- and three-dimension tests with different sample sizes and the lengths of the tests. Results also showed slight difference in standard error of estimation for the two methods according to the study conditions. Thus, the study recommends using the Omega estimation coefficient in case of multiple dimensions.

Key words: Reliability Cronbach's alpha and Omega, Number of dimensions, sample size, Length of the test.

المراجع

تيغزة، امحمد (2017). توجهات حديثة في تقدير صدق وثبات درجات أدوات القياس: تحليل نظري تقويمي وتطبيقي. *مجلة العلوم النفسية والتربوية*، 4(1)، 7- 29.

Anderson, R. & Puckett, J. (2003). Assessing Students' Problem Solving Assignment. *New Directions for Teaching and Learning. 95*(9), 81-87.

Barrett, P. (2014). Theta, Alpha, and Omega estimates of reliability for the VO orientation scales. *Cognadev Technical Report Series*,3.

Bonett, D. G. (2002). Sample size requirements for testing and estimating coefficient alpha. *Journal of Educational and Behavioral Statistic***s,** *27,* 335- 340. Doi: 10.3102/10769986027004335.

Bulut, O. & Sunbul, Ö. (2017). Monte Carlo Simulation Studies in Item Response Theory with the R Programming Language. *Journal of Measurement and Evaluation in Education and Psychology*, 8(3), 266-287.

Centre for Academic Success (2017). Statistical Methods – Scale reliability analysis with small samples, Birmingham City University.

Charter, R. A. (2003). Study samples are too small to produce sufficiently precise reliability coefficients. *The Journal of General Psychology***,** 130(2), 117-129. Doi: 10.1080/00221300309601280

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

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory.* New York: Holt, Rinehart & Winston.

Feinberg, R. A., & Rubright, J. D. (2016). Conducting simulation studies in psychometrics. *Educational Measurement: Issues and Practice, 35*(2), 36–49.

Foley, B. P. (2010). "Improving IRT parameter estimates with small sample sizes: Evaluating the efficacy of a new data augmentation technique". Public Access Theses and Dissertations from the College of Education and Human Sciences, 75. http://digitalcommons.unl.edu/cehsdiss/75

Green, S. & Yang, Y. (2015). Evaluation of Dimensionality in the Assessment of Internal Consistency Reliability: Coefficient Alpha and Omega Coefficients**.** *Educational Measurement: Issues and Practice***,** Winter. 34(4), 14–20.

Guadagnoli, E. & Velicer, W. F. (1988). Relation to sample size to the stability of component patterns. *Psychological Bulletin, 103*(2), 265-275.

Han, K. T., & Hambleton, R.K. (2007). User's Manual for WinGen: Windows Software that Generated IRT Model Parameter and Item Response. Center for Educational Assessment Research Report No 642, Amherst, MA: University of Massachusetts Center for Educational Assessment.

Harwell, M. R., & Baker, F. B. (1991). The use of prior distributions in marginalized Bayesian item parameter estimation: A didactic. *Applied Psychological Measurement, 15*(4), 375–389.

Hogan, T. P., Benjamin, A., & Brezinski, K. L. (2000). Reliability methods: A note on the frequency of use of various types. *Educational and Psychological Measurement,* 60, 523-531.

Ercan, I.; Yazici, B.; Sigirli, D.; Ediz, B.; & Kan, I. (2007). Examining Cronbach Alpha, Theta, Omega Reliability Coefficients According to Sample Size. *Journal of Modern Applied Statistical Methods, 6*(1), 27.

Available at: http://digitalcommons.wayne.edu/jmasm/vol6/iss1/27

Javali, S. B.; Gudaganavar. N. V., & Raj, S. M. (2011). *Effect of Varying Sample Size in Estimation of Coefficients of Internal Consistency*. WebmedCentral BIOSTATISTICS, WMC001649. Available at: http://www.webmedcentral.com/article_view/1649

Kamata, A., Turhan, T., & Darandari, D. (2003). *Estimating Reliability for Multidimensional Composite Scale Score*. Paper presented at the annual meeting of American Educational Research Association, Chicago.

Komaroff, E. (1997). Effect of simultaneous violations of essential tau-equivalence and uncorrelated error on coefficient alpha. *Applied Psychological Measurement,* 21, 337-348.

Liu, Y., & Zumbo, B. D. (2007). The Impact of Outliers on Cronbach's Coefficient Alpha Estimate of Reliability: Visual Analogue Scales. *Educational and Psychological Measurement***,** 67, 620-634.

Mislevy, R. J., & Stocking, M. L. (1989). A consumer's guide to LOGIST and BILOG. *Applied Psychological Measurement*, 13, 57–75.

Mooney, C. Z. (1997). Monte Carlo simulations. *Sage University Paper series on Quantitative Applications in the Social Sciences*, 07-116. Thousand Oaks, CA: Sage.

Murphy, K. R. & DeShon, R. P. (2000). Inter-rater correlations do not estimate the reliability of job performance ratings. *Personnel Psychology***,** 53, 873-900.

Osbourn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods*, 5, 343-355.

Padilla, M., & Divers, J. (2013). Coefficient Omega Bootstrap Confidence Intervals: Non-normal Distributions. *Educational and Psychological Measurement, 73*(6), 956–972.

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

Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research*, 32, 329-353.

Raykov, T. (2001). Bias of coefficient alpha for congeneric measures with correlated errors. *Applied Psychological Measurement*, 25, 69-76.

Segars, A. (1997). Assessing the unidimensionality of measurement: a paradigm and illustration within the context of information systems research. *Omega, 25*(1),107-121.

Terry, L. & Kelley, K. (2012). Sample size planning for composite reliability coefficients: Accuracy in parameter estimation via narrow confidence intervals. *British Journal of Mathematical and Statistical Psychology*, 65, 371–401.

Tighezza, M. (2017). New Orientations in the Estimation of Test Score Validity and Reliability: Theoretical and Practical Analysis, (in Arabic). *Journal of Educational and Psychological Sciences*, 4(1),7-29.

Yurdugül, H. (2008) Minimum sample size for Cronbach's coefficient alpha: A Monte-Carlo study. *Hacettepe Universitesi Eğitim Fakültesi Dergisi, 35,* 397-405

Yurdugul, H. (2006). The Comparison of Reliability Coefficients. *Journal of Faculty of Educational Sciences, 39*, 15-37.

Zimmerman, D. W., Zumbo, B. D., & Lalonde, C. (1993). Coefficient alpha as an estimate of test reliability under violation of two assumptions. *Educational and Psychological Measurement***,** *53*, 33-49.