تأثير عدد فئات الاستجابة وعدد المشاركين على دقة قيم معاملى ألفا وأوميجا في تقدير ثبات درجات المقياس النفسى

إعـــــداد

د/ محمد عبدالهادى عبدالسميع

مدرس علم النفس التربوى

كلية التربية بقنا

جامعة جنوب الوادى

مستخلص

اهتم العديد من الدراسات في مجال القياس النفسى والتربوى بالتحقق من الخصائص السيكومترية للمقياس النفسى وذلك قبل استخدامه في جمع البيانات، ومع ذلك لم تتطرق أى منها إلى دراسة دقة قيم معاملى ألفا وأوميجا في ضوء بعض المتغيرات مثل عدد فئات الاستجابة لليكرت وعدد المشاركين، ولذلك فقد هدفت الدراسة الحالية إلى التعرف على تأثير عدد فئات الاستجابة لليكرت (٣، ٥، ٧) وعدد المشاركين على دقة معاملى ألفا وأوميجا في تقدير ثبات درجات المقياس النفسى، تكّون عدد المشاركين فى الدراسة من ٢٤٣ طالباً وطالبة من طلاب الفرقة الثالثة بكلية التربية بقنا، موزعين على أربع مجموعات طبقاً للنسبة بين عدد بنود مقياس فاعلية الذات الوجدانية وعدد المشاركين: ٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣)، تم تعريب مقياس فاعلية الذات الوجدانية من إعداد Kirk, Schutte, and Hine (2008) حتى يتم الاستفادة منه فى البيئة العربية بعد حساب ثبات درجات ثلاث نسخ منه باستخدام معاملى ألفا وأوميجا، أشارت نتائج الدراسة إلى زيادة دقة قيم معاملى ألفا وأوميجا بزيادة عدد فئات الاستجابة، وخاصة فى حالة العدد الخماسى والسباعى لفئات الاستجابة فى حالة كون عدد المشاركين ٥:١ (١٣٥) بحدود ثقة أفضل، وقد أظهرت النتائج أيضاً أن معامل أوميجا أدى إلى تقدير أدق لقيم معامل ثبات الدرجات مقارنة بمعامل ألفا فى ضوء نسبة التباين المفسر واتساع المسافة بين حدى الثقة، تم مناقشة بعض التوصيات وتقديم بعض البحوث المقترحة.

الكلمات المفتاحية: ثبات درجات المقياس النفسى، ألفا، أوميجا، عدد فئات الاستجابة، عدد المشاركين.

تأثير عدد فئات الاستجابة وعدد المشاركين على دقة قيم معاملى ألفا وأوميجا في تقدير ثبات درجات المقياس النفسى

إعـــــداد

د/ محمد عبدالهادى عبدالسميع

مدرس علم النفس التربوى

كلية التربية بقنا

جامعة جنوب الوادى

مقدمة

تحظى دراسة الخصائص السيكومترية للمقاييس النفسية متمثلة في صدق البنود وثبات الدرجات بأهمية كبيرة بين الباحثين في القياس النفسى والتربوى، كما يُعد حساب معاملات ثبات هذه الدرجات من أهم الإجراءات التي يستطيع الباحث من خلالها الحكم على جودة أدوات بحثه، ومن ثم الثقة في النتائج التي يتم الحصول عليها، ويذكر Johanson and Brooks (2010, 394)[1] أن الباحثين يقومون بإجراء دراسات استطلاعية والتى يتمثل أحد أهم أهدافها فى التحقق من الكفاءة السيكومترية لأدوات البحث قبل استخدامها فى عملية جمع البيانات.

ركز كتاب معايير القياس التربوى والنفسى الذى تصدره أفضل ثلاث جمعيات للبحوث التربوية والنفسية والقياس التربوى فى الولايات المتحدة الأمريكية وهى: الجمعية الأمريكية للبحث التربوى، والجمعية الأمريكية لعلم النفس، والمجلس الوطنى للقياس التربوى AERA, APA, and NCME (2014, 33) على أهمية ثبات درجات أدوات القياس؛ حتى أنه يربط الثبات بالدقة؛ فيفرّق بين مصطلحين للثبات الأول فى النظرية الكلاسيكية للقياس وهو معامل الثبات Reliability coefficient، الثانى فى النظرية الحديثة للقياس وهو ثبات الدرجات/الدقة Scores reliability/precision.

---

[1] تم اتباع دليل الجمعية الأمريكية في علم النفس (الطبعة السادسة) American Psychological Association (2010) في توثيق مراجع البحث سواء في المتن أو قائمة المراجع، وطبقاً لهذا الدليل إذا جاء التوثيق في سياق الفقرة فإنه يتم وضع اسم أو أسماء المؤلفين خارج الأقواس وتبقى سنة النشر وأرقام الصفحات في حالة الاقتباس الحرفى بين قوسين وذلك في متن البحث، كما تم توظيف الدراسات السابقة في عرض المقدمة والمشكلة والإطار النظري بدلاً من تخصيص محور مستقل لها.

وتتعدد الطرق المستخدمة في حساب معامل ثبات درجات أدوات القياس والتي منها طريقة إعادة التطبيق، الصور المتكافئة، معادلات التجزئة النصفية (سبيرمان وبراون – جتمان – رولون)، معامل ألفا لكرونباخ وغيرها من الطرق الشائعة بين الباحثين، وتوجد طرق أخرى تتمثل في معامل بيتا، ثيتا، أوميجا، ونمذجة المعادلة البنائية ولكنها أقل شيوعاً واستخداماً بينهم.

يشير Cronbach (1951, 299) إلى أن تقدير معامل ألفا يتم من خلال حساب تباين البنود والتباين الكلى والتعويض في المعادلة الشهيرة التي قدمها في دراسته المنشورة في مجلة *Psychometrika*، أما بالنسبة لمعامل أوميجا فقدمه McDonald (1978, 1999)، كبديل لألفا للحصول على معامل ثبات أكثر دقة للدرجات.

ويتأثر معامل ثبات درجات المقياس النفسى بعدة عوامل منها: طول المقياس، زمن التطبيق، الفاصل الزمنى بين التطبيقين كما في طريقة إعادة التطبيق، آلية تقسيم المقياس إلى جزءين كما في طرق التجزئة النصفية، وغيرها من العوامل التي تناولتها الأدبيات البحثية المتخصصة في هذا المجال، ولكن توجد عوامل أخرى تؤثر على دقة معامل ثبات الدرجات باستخدام أي الطرق وهى: عدد فئات الاستجابة Number of response categories في أسلوب ليكرت، وعدد المشاركين Number of participants.

ويُعد أسلوب ليكرت المستخدم في تدريج فئات الاستجابة في إعداد المقياس النفسى من أكثر الأساليب استخداماً وانتشاراً؛ نظراً لسهولة استخدامه؛ حيث يشير Zumbo, Gradermann, and Zeisser (2007, 21) إلى أن البيانات من نوعية ليكرت أكثر استخداماً في السياقات التربوية والنفسية وخاصة في حالة المتغيرات المتصلة غير الملاحظة Unobserved continuous variables.

يذكر Shaftel, Nash, and Gillmor (2012, 2) أنه يُستخدم المقياس النفسى لقياس الآراء وسمات الشخصية والظواهر الأخرى في التربية وعلم النفس والصحة والتسويق، ولذلك يكون اختيار عدد فئات الاستجابة أحد أهم الاعتبارات الرئيسة في تصميم هذه المقاييس، ومن ثم يمثل اختيار عدد فئات الاستجابة في ضوء أسلوب ليكرت أهمية كبيرة، وخاصة أن أغلب أدوات البحث في البحوث والدراسات التربوية والنفسية تكون مقاييس يتم تصميمها فى ضوء هذا الأسلوب بفئاته المتعددة سواء فردية أو زوجية، ويرجع ذلك إلى سهولة هذا الأسلوب مقارنة بالأساليب الأخرى.

ويُعد عدد فئات الاستجابة وتأثيرها على الخصائص السيكومترية لأدوات القياس عاملاً

مهماً تناولته دراسات عديدة، وخاصة تأثيرها على معاملات ثبات الدرجات التي حاز ألفا منها على اهتمام أغلب الباحثين، ولذلك يشير Tsang (2012, 121-122) إلى أن العدد الفردى لفئات الاستجابة أفضل من العدد الزوجي؛ لأنه يحتوى على النقطة الوسيطة Midpoint؛ لأنها تتعلق بالمسافة النفسية Psychological distance أثناء الاستجابة، ولا يكون المشارك مضطراً لاختيار اتجاه محدد، ومن ثم الاستجابة الحقيقية التي تعبر عن رأى المشارك بموضوعية من منطلق صدق استجابته من خلال التعبير الشامل عن السمة المقاسة، ولذلك فإن العدد الثلاثى والخماسى والسباعى لفئات الاستجابة هو الأكثر شيوعاً باختلاف الفئات العمرية للمشاركين في الاستجابة على المقياس النفسى.

يرتبط تحديد عدد المشاركين أيضاً بالخصائص السيكومترية لأدوات القياس وخاصة معاملات ثبات درجاتها وكذلك حدود الثقة المرتبطة بهذه المعاملات؛ حيث تناول بعض الباحثين عدد المشاركين المطلوب للدراسة الاستطلاعية التي تهدف أساساً إلى التحقق من الكفاءة السيكومترية للأدوات، فيشير Johanson and Brooks (2010, 399) إلى أن أقل عدد مطلوب من المشاركين في الدراسة الاستطلاعية يمكن أن يكون ٣٠ مشاركاً.

ويضيف Geisinger (2013, 27) أن ما يؤثر على ثبات درجات المقياس النفسى هو التباين الكلى لأداء المشاركين على المقياس؛ لأن قيمة معامل ثبات الدرجات تتأثر باختلاف القدرة الحقيقية لهؤلاء المشاركين، وكلما كان المشاركون غير متجانسين Heterogeneous زادت قيمة معامل ثبات درجات المقياس النفسى وذلك لقياس القدرة الحقيقية للأفراد المجتمع.

ومن المفاهيم الحديثة التي ترتبط بدقة قيم معاملات ثبات درجات المقاييس النفسية مفهوم حدود الثقة Confidence intervals؛ حيث يرى Iacobucci and Duhachek (2003) أنه من الضرورى تدوين حدود الثقة مع معامل ثبات الدرجات، وكلما ضاق اتساع المسافة Interval width بين حدى الثقة الأدنى Lower bound والأعلى Upper bound كانت قيمة معامل ثبات الدرجات أدق، وكلما اتسعت المسافة كانت الدقة أقل، كما تُعد من أهم المعايير المستخدمة للحكم على دقة قيم معامل ثبات الدرجات.

يتضح مما سبق أن هناك أهمية كبيرة لحساب ثبات درجات المقياس النفسى للحصول على نتائج يمكن الثقة فيها، كما توجد عدة طرق لحساب معامل ثبات الدرجات، وأن قيمة ودقة هذا المعامل وحدود الثقة المرتبطة به تتأثر بعدة عوامل منها: عدد فئات الاستجابة وعدد المشاركين.

مشكلة الدراسة:

تتوقف دقة البيانات التي يتم جمعها على دقة أداة القياس التي يتم استخدامها، ويُعد المقياس النفسى الذى يتم تصميمه في ضوء أسلوب ليكرت أحد أشهر أدوات جمع البيانات في الدراسات التربوية والنفسية، ومن هنا تمثل دراسة العوامل المؤثرة على دقة ثبات درجاته: مثل عدد فئات الاستجابة لليكرت وعدد المشاركين أهمية كبيرة، وذلك للحصول على نتائج تتميز بالمنعة الإحصائية Statistical robustness، وبرزت مشكلة البحث من خلال مجموعة من المحاور والتى تتمثل فى التالى:

المحور الأول: يختص بالواقع العملى لطلاب الدراسات العليا، الذى اتضح من خلال تدريس الباحث لمقرر القياس النفسى لطلاب الدبلوم المهنية في التربية؛ حيث تحتوى نواتج التعلم المستهدفة لهذا المقرر على موضوع خطوات تصميم المقياس النفسى – كأحد أهم أدوات القياس في العلوم التربوية والنفسية– وكانت هذه الخطوات تتضمن تحديد عدد فئات الاستجابة أثناء خطوة كتابة البنود، وأيضاً تحديد عدد المشاركين فى التجربة الاستطلاعية اللازمة لحساب الخصائص السيكومترية ومنها ثبات الدرجات؛ حيث أن أهم الأسئلة التي تم مناقشتها هى: ما العدد الأنسب لفئات الاستجابة لأسلوب ليكرت؟ وهل يؤثر ذلك على دقة تقدير معاملات ثبات درجات المقياس النفسى؟

تبيّن عند مراجعة بعض الأدبيات البحثية المتعلقة بهذا السؤال عدم وجود اتفاق واضح بين الباحثين فيما يتعلق بعدد فئات الاستجابة، ومدى ارتباط زيادتها بالزيادة في قيم معاملات ثبات درجات المقياس النفسى، ومن بين هذه الأدبيات دراسة Weng (2004) التي توصلت إلى أن قيمة معامل ثبات درجات المقياس النفسى تزداد بزيادة عدد فئات الاستجابة وذلك بعد دراسته لعدد من (٣–٩) فئة، في حين أشارت نتائج دراسة Chomeya (2010) إلى اختلاف قيمة معامل ألفا في التدريج الخماسى والسداسى باستخدام المقياس الذى تم تطبيقه، حيث زادت قيمة معامل ألفا في التدريج الخماسى في مقياس مركز الضبط، بينما زادت في التدريج السداسى في مقياسى الاتجاه نحو شرب الكحوليات والدافع للتحصيل الدراسى.

وعلى النقيض فإن بعض الدراسات العربية أثبتت أنه لا يزداد معامل ثبات درجات المقياس النفسى بزيادة عدد فئات الاستجابة وخاصة في حالة الأعداد الفردية، ومنها دراسة حجاج غانم أحمد على وياسر عبدالله حفنى حسن (٢٠١١) التي هدفت إلى الكشف عن تأثير عدد بدائل ليكرت (٢، ٣، ٥) على الخصائص السيكومترية للمقياس النفسي متمثلة في معامل ألفا كممثل

لثبات درجات، وصدق التكوين الفرضى كممثل للصدق، حيث تكّون عدد المشاركين من ٥٨٧ طالباً وطالبة من طلاب الجامعة. استجاب المشاركون على ثلاث صور من مقياس مفهوم الذات (إعداد: صلاح الدين أبوناهية، ١٩٩٩)، وقد أشارت نتائج الدراسة-فيما يتعلق بالدراسة الحالية- إلى ارتفاع قيمة معامل ألفا في حالة الصورتين الثلاثية والخماسية لاحتوائهما على النقطة الوسيطة، مع الأخذ في الاعتبار أن معامل ألفا للصورة الخماسية كان أقل من الصورة الثلاثية.

كما قام بعض الباحثين بمراجعة الدراسات المنشورة في مجلات علمية محكمة للتعرف على العدد الأكثر شيوعاً لفئات الاستجابة، حيث وجد Hartley (2013) أن معظم الدراسات استخدمت التدريج الرباعى والخماسى والسباعى لفئات الاستجابة، وذلك بعد مراجعته لعدد من الدراسات المنشورة فى المجلة الدولية لعلم النفس العيادى والصحى *International Journal of Clinical and Health Psychology*.

المحور الثانى: يختص بممارسات الباحثين، وعدم وجود خطوط إرشادية متفق عليها لتحديد عدد فئات الاستجابة وعدد المشاركين وتأثيره على حدود الثقة المرتبطة بدقة قيم معاملات ثبات الدرجات، حيث يختار بعضهم عدداً قليلاً لفئات الاستجابة والمشاركين رغبة منهم في توفير الوقت والجهد أثناء تصميم المقياس النفسى وتطبيقه للحصول على البيانات اللازمة لإجراء الدراسات التربوية والنفسية، وقد يحدث ذلك دون وعى منهم بأن ذلك الاختيار يؤثر على معاملات ثبات الدرجات؛ ولذلك يكون الباحث هو السبب في ضياع وقته وجهده وهو لا يدرى، من خلال بذل مزيد من الجهد في تصميم المقياس النفسى المستخدم في دراسته ولكن لا يختار العدد الأنسب لفئات الاستجابة وكذلك المشاركين مما يؤدى إلى انخفاض قيمة معامل ثبات الدرجات والذى لا يرجع إلى ضعف المقياس الذى تم تصميمه وتطبيقه؛ ولكن للاختيار الخاطئ لعدد فئات الاستجابة والمشاركين.

يؤكد ذلك -فيما يختص بعدد المشاركين- ما أشار إليه Bonett (2002, 335) أن تحديد عدد المشاركين يُعد أحد أهم الخطوات المطلوبة في تصميم الدراسات السيكومترية؛ حيث أنه إذا كان عدد المشاركين صغيراً جداً سوف ينقص من قوة المقياس وتتسع المسافة بين حدى الثقة أيضاً، وعندما يكون عدد المشاركين كبيراً جداً فإن ذلك يكون إهداراً للموارد Wasteful of resources؛ ولذلك قدم بعض المعادلات التي يمكن تطبيقها في حالة تحديد عدد المشاركين المطلوب لحساب معامل ألفا في ضوء مستوى محدد من القوة وحدود الثقة، أما في عام ٢٠٠٣ قدّم أيضاً توصيات لأعداد المشاركين المطلوبة عند مقارنة زوجين من معامل ألفا فقط.

يؤكد ذلك أيضاً-فيما يختص بعدد فئات الاستجابة- ما ذكره Weijtersa, Cabooterb, and Schillewaer (2010, 236) أن بعض الباحثين يجدد عدد فئات الاستجابة لليكرت أثناء تصميم المقياس النفسى بالرغم من عدم وجود خطوط إرشادية واضحة ومحددة تساعدهم في الاختيار الأمثل لعدد هذه الفئات للحصول على نتائج دقيقة.

يضيف Guo, Pohl, and Gerokostopoulos (2013, 1) أن تحديد العدد المناسب للمشاركين أثناء تقدير ثبات درجات أدوات القياس في العلوم التربوية والنفسية أمر مهم؛ لأنه إذا كان عدد المشاركين قليلاً فإن ذلك يؤدى إلى فقدان بعض المعلومات التي يتم الحصول عليها من المقياس، أما إذا كان عدد المشاركين كبيراً فإن ذلك مضيعة للوقت والجهد أثناء عملية التطبيق، ويضيف أنه يوجد مدخلان لتحديد هذا العدد، حيث يتمثل المدخل الأول في حدود الثقة أو ما يُسمى مدخل التقدير، وكلما زاد عدد المشاركين كانت حدود الثقة أضيق، والمدخل الثانى يتمثل في القوة أو تحديد الخطأ من النوعين الأول والثانى، ويجب ألا يقل عدد المشاركين عن ٢٥ مشاركاً.

طالما أن الهدف الأسمى من الدقة فى إعداد المقياس النفسى فى الدراسات التربوية والنفسية هو خفض خطأ القياس الذى يؤثر سلباً على نتائج هذه الدراسات؛ فإن ذلك يتم من خلال تحديد العدد الأمثل لفئات الاستجابة؛ نظراً لأن العدد الأصغر ربما لا يساعد الباحث على إظهار الاستجابة الحقيقية وبالتالي التباين الحقيقى للدرجات مما يؤثر على معامل ثبات الدرجات الناتج.

إذا كانت هناك خطوط إرشادية تساعد الباحثين في تحديد العدد المطلوب من المشاركين أثناء استخدام أسلوب التحليل العاملى والتي منها ٣:١، ٥:١، ١٠:١-أي أن كل بند من بنود المقياس يقابله ثلاثة أو خمسة أو عشرة مشاركين على الترتيب- فماذا عن وضع قواعد محددة Rules of thumb أثناء حساب معاملات ثبات درجات المقياس النفسى، وخاصة أن العديد من الأدبيات البحثية ذكرت أعداداً محددة لا تعتمد على قواعد ثابتة (Ercan, Yazici, Sigirli, Ediz, & Kan, 2007; Feldt & Ankenmann, 1998; Javail, Gudaganavar, & Raj, 2011; Trizano-Hermosilla & Alvarado, 2016) ولعل هذا يُعد مبرراً لاختيار عدد المشاركين في إطار قواعد محددة تعتمد على النسبة بين عدد البنود وعدد المشاركين في إطار الدراسة الحالية.

يشير Padilla and Divers (2013, 956) إلى أن عدد المشاركين والارتباطات البينية Intercorrelations بين البنود التي يشملها المقياس يؤثران على حدود الثقة أثناء حساب معامل ثبات الدرجات باستخدام معامل أوميجا، حيث أنه من الضرورى ألا يقل عدد المشاركين

عن ٥٠ مشاركاً في حالة توافر شرط الاعتدالية، كما أنه لابد ألا يقل عدد المشاركين عن ١٥٠ مشاركاً في حالة عدم التحقق من افتراض الاعتدالية.

المحور الثالث: يختص بمنهجية بعض الدراسات السابقة فيما يتعلق بدراسة عدد فئات الاستجابة، حيث استخدمت هذه الدراسات مجموعات مختلفة من المشاركين للاستجابة على كل صورة من صور المقاييس التي تم تطبيقها، الأمر الذى يجعل احتمالية اختلاف معاملات ثبات الدرجات راجع إلى اختلاف خصائص المشاركين في كل مجموعة وليس عدد فئات الاستجابة التي تم دراستها، كما أن هذه الدراسات استخدمت معامل ألفا بالرغم من عدم دقته في ضوء نتائج الأدبيات البحثية المتخصصة، ومن أمثلة هذه الدراسات في البيئة الأجنبية دراسة Wong, Peng, Shi, and Mao (2011) التي هدفت إلى دراسة تأثير عدد زوجى (رباعى) وفردى (خماسى) من فئات الاستجابة على معاملات ثبات درجات بعض المقاييس النفسية، تكّون عدد المشاركين من ٥٣٢ موظفاً، استجاب ٢٦٦ موظفاً منهم على النموذج الرباعى في حين استجاب ٢٦٦ موظفاً آخرين على النموذج الخماسى، تم تطبيق عدة مقاييس منها السمات الخمس الكبرى في الشخصية، الرضا الوظيفى، الالتزام الوظيفى، الإدراك الوظيفى، سلوك طلب التغذية الراجعة وغيرها، وبالرغم من عدم وجود فروق جوهرية بين التدريج الرباعى والخماسى إلا أن قيمة معامل ألفا كانت أكثر في حالة التدريج الخماسى في معظم المقاييس التي تم تطبيقها.

من أمثلة هذه الدراسات في البيئة العربية دراسة سعيد حسن آل عبد الفتاح الغامدى (٢٠٠٣) التي هدفت إلى التعرف على مدى اختلاف الخصائص السيكومترية لأداة القياس في ضوء تغاير عدد بدائل الإستجابة والمرحلة الدراسية لدى ٤٥١ طالباً من طلاب التعليم العام وذلك باستخدام قائمة القلق لشارلز سبليرجر وآخرين، تم استخدام أربعة نماذج من القائمة تختلف فقط في عدد بدائل الاستجابة (٢، ٣، ٤، ٥)، وقد أشارت نتائج الدراسة إلى ارتفاع قيمة معامل ثبات الدرجات باستخدام التجزئة النصفية وألفا في حالة عدد البدائل الخماسية وأوصت باستخدامها وخاصة مع المراحل الدراسية العليا.

المحور الرابع: يختص بدقة بعض معاملات ثبات درجات المقياس النفسى وخاصة معاملى ألفا وأوميجا، حيث يشير Green and Hershberger (2000, 251) إلى أن معامل ألفا يؤدى إلى انخفاض تقدير معامل ثبات الدرجات، وخاصة في حالة انتهاك افتراضات نموذج تاو المتكافئ في الأساس Essentially –tau equivalent model، أما إذا كانت الأخطاء مرتبطة فإن ذلك يؤدى إلى زيادة تقدير قيمته، وتوصلت نتائج دراسة Raykov (2001a) أيضاً إلى أن ارتباط درجات

الخطأ يؤدى إلى تحيز تقدير معامل ألفا، مما يزيد من محدودية استخدامه في تقدير ثبات درجات المقياس النفسى، ودراسة محمد بوزيان تيغزة (٢٠٠٩) أيضاً التي أشارت نتائجها إلى أن معامل ألفا يؤدى إلى انخفاض تقدير معامل ثبات درجات المقياس النفسى، وتوصى الدراسة باستخدام معامل أوميجا كبديل لألفا في حالة عدم تحقق افتراضات النموذج الخاص به.

وقد توصلت أيضاً دراسة Yang and Green (2011, 378) إلى أن الكثير من الباحثين لا يجيدون تفسير قيمة معامل ألفا بدقة؛ لأنه ليس دليلاً على الاتساق الداخلى للمقياس ولكنه معامل ثبات قائم على الاتساق الداخلى لبنود المقياس، ولذلك يوصى بضرورة أن يجرى الباحثون تحليلات أولية preliminary analyses للدرجات التى يتم الحصول عليها من الدراسة الاستطلاعية؛ وذلك لفهم البنية الداخلية للمقياس النفسى قبل اختيار طريقة حساب معامل ثبات هذه الدرجات، وقد ينتج عن هذه التحليلات اختيار الباحثين لمعامل ألفا أو أوميجا بناءً على توافر الافتراضات.

كما هدفت دراسة Tang and Cui (2012) إلى مقارنة دقة ثلاثة معاملات لحساب ثبات الدرجات (ألفا – لمبادا – أفضل حد أدنى) Greatest Lower Bound (GLB) وذلك من خلال المحاكاة Simulation. باستخدام فنية مونت كارلو Monte-Carlo وأشارت نتائج الدراسة إلى أن تقدير ألفا كان متحيزاً سلباً، كما أن معامل لمبادا أدق من GLB كبدائل لمعامل ألفا.

يشير Deng and Chan (2016, 1) إلى أنه لا تخلو تقريباً أي دراسة من استخدام معامل ألفا وذلك عند قياس أي متغير مكون من مجموعة من البنود في البحوث الاجتماعية والسلوكية؛ لذا من المفضل دراسة مدى دقته فى ضوء متغيرات مثل عدد فئات الاستجابة وأعداد المشاركين.

وقد هدفت دراسة Trizano-Hermosilla and Alvarado (2016) إلى مقارنة أداء بعض معاملات ثبات الدرجات (ألفا – أوميجا – أفضل حد أدنى) وذلك باختلاف أعداد المشاركين (٢٥٠ – ٥٠٠ – ١٠٠٠) والتوزيع الاعتدالى باستخدام فنية مونت كارلو للمحاكاة، وأشارت نتائج الدراسة إلى أفضلية استخدام معامل أوميجا وأفضل حد أدنى حتى في حالة الأعداد الصغيرة للمشاركين وخاصة عندما تكون التوزيعات ملتوية ولا تتوافر افتراضات نموذج تاو المتكافئ في الأساس، كما يشير إلى أن التحيز في قيمة معامل ثبات الدرجات يصل إلى ١٣% عند استخدام ألفا دون توافر افتراضات نموذج تاو المتكافئ في الأساس.

المحور الخامس: يختص بتأثير دقة معاملات ثبات درجات المقياس النفسى على النتائج

الإحصائية، حيث يؤدى عدم الدقة في حساب ثبات درجات المقياس النفسى إلى نتائج لا تتسم بالمنعة الإحصائية Statistical robustness أو مضللة إحصائياً، ومن هنا يشير Geisinger (2013, 23) إلى أن أهمية الدقة في حساب ثبات درجات المقياس النفسى ترجع إلى الإسهام في العلوم النفسية من خلال إظهار العلاقات الحقيقية بين المتغيرات؛ لأنه عندما لا يتسم المقياس النفسى بالدقة في ثبات الدرجات فإن ذلك يؤدى إلى نتائج غير دقيقة، وبالتالي تكون هذه العلاقات غير منطقية، وربما ذلك يكون أحد المبررات المهمة للمشكلة، حيث إن تحديد العدد المناسب لفئات الاستجابة للمقياس النفسى وكذلك عدد المشاركين المستجيبين عليه أمر مهم يؤثر على دقة المقياس النفسى، ومن ثم يحتاج مزيداً من جهد الباحثين في مجال القياس النفسى والتربوى. كما أشارت نتائج دراسة Feldt (2011) إلى أن انخفاض معاملات ثبات درجات أدوات القياس يؤثر سلباً على دقة النتائج الإحصائية المتعلقة باختبار "ف" الخاص بمقارنة المتوسطات.

يتضح من العرض السابق أن بحث تأثير عدد فئات الاستجابة وعدد المشاركين على دقة معاملات ثبات درجات المقياس النفسى ذو أهمية كبيرة وخاصة مع اختلاف الدراسات السابقة التي تناولته واختلاف منهجية هذه الدراسات، بالإضافة إلى عدم وجود خطوط إرشادية واضحة يتبعها الباحثون عند اختيار عدد فئات الاستجابة وعدد المشاركين وتأثيرهما على حدود الثقة المرتبطة بدقة قيم معاملات ثبات درجات المقياس النفسى، علاوة على تناول بعض الأدبيات البحثية لدراسة مدى دقة معاملى ألفا وأوميجا وتأثير هذه الدقة على نتائج الأساليب الإحصائية، والجدير بالذكر أيضاً أن الدراسات التى تناولت تأثير عدد فئات الاستجابة على معامل ألفا ركزت على ارتفاع وانخفاض قيمه باختلاف عدد فئات الاستجابة وليس دقته بانخفاض أو ارتفاع هذه القيم أو حدود الثقة واتساع المسافة بين حديها أو نسبة التباين المفسر، ولذلك فإن الحاجة ماسة لإجراء هذه الدراسة وخاصة في عدم وجود قواعد ثابتة لتحديد عدد المشاركين، ومن هنا تتبلور مشكلة الدراسة في بحث تأثير عدد فئات الاستجابة في ضوء أسلوب ليكرت وعدد المشاركين على دقة قيم معاملى ألفا وأوميجا في تقدير ثبات درجات المقياس النفسى. ولذلك **تحددت مشكلة الدراسة الحالية فى الإجابة عن التساؤلات التالية:**

١. هل يوجد تأثير لعدد فئات الاستجابة (٣، ٥، ٧) وعدد المشاركين: ١:٣ (٨١)، ١:٥ (١٣٥)، ١:٧ (١٨٩)، ١:٩ (٢٤٣) على دقة قيم معامل ألفا في تقدير ثبات درجات مقياس فاعلية الذات الوجدانية؟

٢. هل يوجد تأثير لعدد فئات الاستجابة (٣، ٥، ٧) وعدد المشاركين: ١:٣ (٨١)، ١:٥

(١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣) على دقة قيم معامل أوميجا في تقدير ثبات درجات مقياس فاعلية الذات الوجدانية؟

٣. هل يؤدى استخدام معامل أوميجا إلى تقدير أدق من معامل ألفا لثبات درجات مقياس فاعلية الذات الوجدانية باختلاف عدد فئات الاستجابة (٣، ٥، ٧) وعدد المشاركين: ٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣) فى ضوء نسبة التباين المفسر واتساع المسافة بين حدى الثقة؟

أهداف الدراسة

هدفت الدراسة الحالية إلى بحث تأثير عدد فئات الاستجابة وعدد المشاركين على دقة معاملى ألفا وأوميجا في تقدير ثبات درجات المقياس النفسى، وبشكل أكثر تحديداً فقد هدفت الدراسة إلى التعرف على:

١. تأثير عدد فئات الاستجابة (٣، ٥، ٧) وعدد المشاركين: ٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣) على دقة قيم معامل ألفا في تقدير ثبات درجات مقياس فاعلية الذات الوجدانية.

٢. تأثير عدد فئات الاستجابة (٣، ٥، ٧) وعدد المشاركين: ٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣) على دقة قيم معامل أوميجا في تقدير ثبات درجات مقياس فاعلية الذات الوجدانية.

٣. دقة معاملى ألفا أوميجا فى تقدير ثبات درجات مقياس فاعلية الذات الوجدانية باختلاف عدد فئات الاستجابة (٣، ٥، ٧) وعدد المشاركين: ٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣) فى ضوء نسبة التباين المفسر واتساع المسافة بين حدى الثقة.

أهمية الدراسة

تبعت أهمية الدراسة الحالية من الاعتبارات التالية:

١.تزويد المكتبة العربية بإطارٍ نظرى حديث وواف عن عدد فئات الاستجابة لليكرت وكيفية تحديدها والقواعد المتبعة لتحديد عدد المشاركين وكذلك معاملى ألفا وأوميجا ونماذج القياس النفسى وافتراضاتها، الأمر الذى يجعلها تمثل إضافة لأدبيات البحث في مجال القياس النفسى والتربوى.

٢.رفع وعى الباحثين بأهمية إجراء تحليلات أولية للبيانات المراد حساب ثبات درجاتها،

وذلك قبل اختيار طريقة حساب الثبات لها مما يؤثر على دقة النتائج ومدى الثقة فيها.

٣.تزويد الباحثين بالعدد الأمثل لفئات الاستجابة الذى يؤدى إلى تقدير أدق لمعاملى ألفا وأوميجا وذلك عند تصميم المقياس النفسى وحساب ثبات درجاته.

٤.تزويد الباحثين بدليل عملى على كيفية اختيار العدد الأنسب للمشاركين للاستجابة على المقياس النفسى فى ضوء النسبة بين عدد البنود وعدد المشاركين المقابل لكل بند، الأمر الذى يوفر وقتهم وجهدهم، وكذلك دقة معاملات ثبات الدرجات الناتجة.

٥.تزويد الباحثين بدليل عملى على دقة معامل أوميجا قائم على نسبة التباين المفسر واتساع المسافة بين حدى الثقة، مما يؤثر على شيوع استخدامه، ومن ثم الثقة فى النتائج التى يتم التوصل إليها.

٦.مساعدة الباحثين على استخدام برنامج RStudio التابع للحزمة الحاسوبية R Package وذلك في حساب معاملى ألفا أوميجا، وخاصة أن الحزمة الإحصائية في العلوم الاجتماعية SPSS لا تتضمن كيفية حساب معامل أوميجا وكذلك حدود الثقة المرتبطة به أو بمعامل ألفا.

٧.قدمت الدراسة ثلاث صور (ثلاثية – خماسية – سباعية) من مقياس فاعلية الذات الوجدانية بعد تعريبه وحساب ثبات درجاته باستخدام أدق المعاملات والتي يمكن استخدامها في بحوث ودراسات مستقبلية.

مصطلحات الدراسة

تتمثل مصطلحات الدراسية الحالية في ثبات درجات المقياس النفسى متمثلاً في معاملى ألفا وأوميجا وعدد فئات الاستجابة لليكرت وكذلك عدد المشاركين كالتالى:

١. ثبات درجات المقياس النفسى[٢]

يشير إلى الاتساق فى استجابة المشاركين على أداة القياس، كما أن معامل ثبات درجات المقياس النفسى هو مؤشر يمكّن الباحثين من الثقة في الدرجات التي يتم الحصول عليها ومدى

---

[٢] يرى المتخصصون في القياس النفسى والتربوى أنه من الأفضل استخدام مصطلح ثبات درجات المقياس Scale scores reliability وليس ثبات المقياس لأن الثبات خاصية نسبية للمقياس وليس مطلقة، فالمقياس الذى يكون ثابتاً في مجتمع ما قد لا يكون ثابتاً في آخر، فلو تم تطبيق مقياس ١٠٠ مرة سوف يتم الحصول على ١٠٠ معامل مختلف لثبات الدرجات ولذلك فتمثيل المجتمع أمر مهم أثناء اختيار المشاركين في الدراسة الاستطلاعية (Thompson and Vacha-Haase, 2000).

إمكانية استخدامها في سياقات محددة، ويتمثل في الدراسة الحالية في معاملى ألفا وأوميجا فى تقدير ثبات مقياس فاعلية الذات الوجدانية.

٢. عدد فئات الاستجابة لليكرت

يشير إلى بعض بدائل الاستجابة طبقاً لأسلوب ليكرت والتي تكون أمام كل بند من بنود مقياس فاعلية الذات الوجدانية ويختار المشاركين منها طبقاً لما ينطبق عليهم من محتوى البند، وتقتصر في الدراسة الحالية على أعداد فردية من البدائل: ثلاثية (لا تنطبق – متردد – تنطبق)، وخماسية (لا تنطبق تماماً – لا تنطبق – متردد – تنطبق – تنطبق تماماً) وسباعية (لا تنطبق تماماً – لا تنطبق – لا تنطبق أحياناً – متردد – تنطبق أحياناً – تنطبق – تنطبق تماماً).

٣. عدد المشاركين[٣]

يشير إلى عدد الطلاب المشاركين في الاستجابة على مقياس فاعلية الذات الوجدانية المستخدم في الدراسة الحالية، وقد تم توزيعهم على أربع مجموعات تمثل نسبة عدد بنود المقياس إلى نسبة المشاركين [(٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣)].

الإطار النظرى

يعرض الجزء التالى الإطار النظرى للدراسة الحالية فى ضوء ثلاثة محاور رئيسة تبدأ بالمتغير التابع وهو ثبات درجات المقياس النفسى، نظراً لأنه أكثر عمومية ويتم تناوله من حيث مفهومه وأهميته وطرق حسابه ونماذج القياس المرتبطة به وكذلك معاملى ألفا وأوميجا، ثم متغيران مستقلان يتمثلان فى عدد فئات الاستجابة لليكرت وعدد المشاركين.

أولاً: ثبات درجات المقياس النفسى

تُعد دراسة ثبات درجات أدوات القياس بصفة عامة والمقياس النفسى بصفة خاصة والطرق المختلفة لحسابها أحد الموضوعات المهمة التي تناولتها عديد من أدبيات البحث في القياس النفسى والتربوى، ويذكر Cho (2016, 652) أن البحث فى موضوع ثبات درجات أدوات القياس ليس حديثاً، بل تعود جذوره إلى كتابات Brown في عام ١٩١٠ وكذلك

[٣] يستخدم بعض الباحثين مصطلح حجم العينة بدلاً من المشاركين Participants، ولكن الأحدث مصطلح المشاركين ويُطلق على الأفراد الذين استجابوا على أدوات الدراسة وذلك طبقاً للمراجع الحديثة وتوصيات الجمعية الأمريكية لعلم النفس APA.

Spearman في نفس العام الذّين قدما طريقة لحساب ثبات الدرجات، ولكن هذه الطريقة ينقصها الثبات فكيف يمكنها التعبير عن ثبات الدرجات كما أنها تجعل الباحثين ليس لديهم الصورة الكاملة عن طبيعة البيانات، ولذلك من المفضل أن يكون هناك مدخل منظومى قائم على نظام محدد للاختيار بين معاملات ثبات الدرجات، والذى يستطيع الباحثون اتباعه للحصول على معاملات ثبات للدرجات تتميز بالدقة.

مفهوم ثبات درجات المقياس النفسى

تعددت التعريفات الخاصة بمفهوم ثبات الدرجات في أدبيات القياس النفسى والتربوى، حيث يعرّفه صلاح الدين محمود علام (٢٠٠٠، ١٣٣) إحصائياً بأنه مقدار التباين الحقيقى إلى التباين الكلى وكلما زاد التباين الحقيقى زادت قيمة معامل ثبات الدرجات. ويشير Thorndike (2005) إلى أنه خاصية تقوم في جوهرها على الاتساق في الدرجات التي يحصل عليه الفرد في مقياس ما إذا تم إعادة التطبيق عدد لا نهائى من المرات. كما يعرّفه صفوت فرج (٢٠٠٧، ٢٩٥) بأنه الدرجة الحقيقية التى تعبر عن أداء الفرد لاختبار ما، ومعنى ثبات الدرجة أن المفحوص يحصل عليها فى كل مرة يؤدى فيها الاختبار نفسه أو صورة مكافئة له تقيس نفس الخاصية أو السمة موضع القياس، سواء تم القياس فى الظروف نفسها أو ظروف مختلفة لا تتدخل فيها عوامل عشوائية.

ويشير Geisinger (2013, 26-27) إلى النظرية الكلاسيكية للقياس تفسر الدرجة التي يحصل عليها الفرد (الملاحظة) في ضوء مكونين هما: الدرجة الحقيقة والخطأ العشوائى كما يتضح من المعادلة التالية:

$$X = T + E \qquad (1)$$

وبالتالي فإن تباين الدرجة الملاحظة X يساوى تباين الدرجة الحقيقة T مضافاً إليه تباين درجة الخطأ العشوائى E. ويتضح ذلك من خلال المعالة التالية:

$$\sigma_X^2 = \sigma_T^2 + \sigma_E^2 \qquad (2)$$

يشير ثبات درجات المقياس بناءً على ما سبق إلى نسبة تباين الدرجات الحقيقية إلى تباين الدرجات الملاحظة. ويتضح ذلك من خلال المعادلة التالية:

$$reliability = \frac{\sigma_T^2}{\sigma_X^2} \qquad (3)$$

أو بصيغة أخرى فإنه يساوى ناتج حاصل قسمة تباين الخطأ العشوائى/ تباين الدرجات الملاحظة وذلك مطروحاً من الواحد الصحيح.

$$reliability = 1 - \frac{\sigma_E^2}{\sigma_X^2}$$

(4)

يتضح من المعادلتين (٣) و(٤) أن معامل ثبات درجات المقياس النفسى يتوقف على الأداء الحقيقى للمشاركين، أى أنه كلما زادت قيمة الدرجات الحقيقية زادت قيمة معامل ثبات درجات المقياس، وبالتالي إذا أراد الباحثون الحصول على معامل ثبات مرتفع للدرجات عليهم أن يبذلوا قصارى جهدهم فى مساعدة المشاركين على الأداء الحقيقى من خلال إزالة كل العقبات المؤثرة على ذلك مثل تهيئة الظروف المناسبة أثناء التطبيق وغيرها من العوامل التى ترتبط ببنود المقياس النفسى أو الظروف المحيطة. ويرتبط بذلك تفسير قيمة معامل ثبات الدرجات حيث إنه إذا حصل باحث على قيمة قدرها ٠.٨٥ فإن ذلك يعنى أن ٨٥% من أداء المشاركين في الاستجابة على المقياس ترجع للتباين في الدرجات الحقيقية بينما ١٥% ترجع للتباين في مقدار الخطأ في الأداء أثناء الاستجابة.

أهمية حساب ثبات درجات المقياس النفسى

لعل أكبر دليل على أهمية حساب ثبات درجات أدوات القياس هو أنه لا تكاد تخلو أي دراسة في ميدانى التربية وعلم النفس من حساب ثبات درجات الأدوات المستخدمة فيها، وخاصة المقياس النفسى حتى يستمر الباحث في عملية جمع البيانات بطمأنينة.

تناول بعض الباحثين أيضاً مدى ارتباط ثبات درجات المقياس النفسى بالثقة فى نتائجه وكذلك الاتساق فى أداء المفحوص؛ حيث يشير Kimberlin and Winterstein (2008) إلى أنه يُعد تقدير ثبات درجات أدوات القياس المستخدمة فى جمع البيانات أحد أهم المؤشرات الدالة على خصائص جودة هذه الأدوات والمتمثلة فى الدراسة الحالية فى المقياس النفسى.

يذكر محمود محمد عمر، حصة عبدالرحمن فخرو، تركى السبيعى، وآمنة عبدالله تركى (٢٠١٠، ٢١٥) إلى أن الحصول على معامل ثبات مرتفع للدرجات يؤدى إلى مزيد من الثقة فى نتائجه لأن معامل ثبات الدرجات يُعد مؤشراً كمياً لدقة القياس وكذلك الاتساق فى أداء المفحوص على الاختبار، ويضيف Drost (2011, 105) أن تقدير ثبات الدرجات من أهم الاعتبارات أثناء تكميم السلوك، حيث أن ذلك يعطى مؤشراً على دقة أداة القياس وخلوها من الأخطاء، وخاصة العشوائية.

يتضح من الجزء الخاص بأهمية التأكد من ثبات درجات أدوات القياس قبل استخدامها في

الدراسة الأساسية أنها ليس مجرد خطوة عادية يقوم بها الباحث أثناء إجراء دراساته بل خطوة أساسية يتوقف عليها أمور كثيرة أهمها مدى الثقة في النتائج التي يتم الحصول عليها وكذلك مدى الاتساق فى اتخاذ القرار وغير ذلك من أمور مهمة.

نماذج القياس

يُعد التنويه لنماذج القياس ذا أهمية كبيرة وخاصة عند دراسة دقة معاملات ثبات درجات المقياس النفسى؛ حيث يشير Dunn, Baguley, and Brunsden (2014, 402-403) إلى أنه توجد أربعة نماذج للقياس هى: النموذج المتوازى The parallel model، نموذج "تاو" المتكافئ The tau-equivalent model، نموذج "تاو" المتكافئ في الأساس The essentially tau-equivalent model، النموذج التقاربى The congeneric model. وقد تم اختصار هذه النماذج في ثلاثة نماذج فقط حيث تم ضم النموذجين الثانى والثالث في نموذج واحد وهو النموذج الثالث. وفيما يلى نبذة مختصرة عن كل نموذج وافتراضاته حتى يتيسر للباحثين التحقق من تلك الافتراضات عند اختيار طريقة ثبات الدرجات المناسبة لطبيعة البيانات التي تم جمعها:

أ. النموذج المتوازى

يُسمى هذا النموذج بنموذج التوازى التام، ويُعد أكثر نماذج القياس تقيداً لأن افتراضاته تشمل قياس البنود لمتغيراً (أحادية البعد) Unidimensionality، وتساوى الدرجات الحقيقية أي تساوى تشبعات البنود على نفس العامل الكامن (البعد) بنفس الدرجة، تساوى تباين الخطأ، وهذا يُعد مستحيلاً فى العلوم التربوية والنفسية، وذلك يؤدى إلى عدم دقة معامل ثبات الدرجات المحسوب وتفاوته بدرجة كبيرة سواء بانخفاض التقدير Underestimation أو زيادته Overestimation.

ب. نموذج "تاو" المتكافئ في الأساس

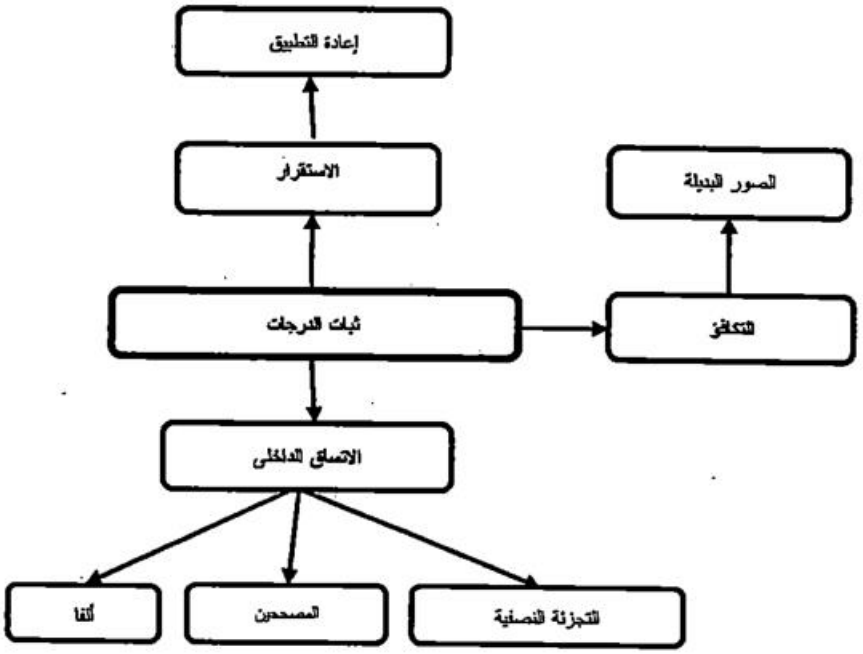
يُعد هذا النموذج أقل تقيداً من النموذج السابق لأن افتراضاته تقوم على ضرورة قياس البنود لمتغير كامن واحد (أحادية البعد)، ويتفق ذلك مع النموذج السابق في تشبع بنود المقياس على العامل الكامن بنفس الدرجة، ويختلف عنه في أنه يسمح بالتباين فى مستوى الخطأ. والجدير بالذكر أن معامل ثبات ألفا يعتمد على هذا النموذج، ولما كان من الصعب جداً فى أدوات القياس أن تتشبع بنود مقياس فرعى بنفس الدرجة على نفس العامل قلل ذلك من دقة استخدام معامل ثبات ألفا، حيث أن قيمة معامل ألفا تتأثر بتحقق افتراضات النموذج المعتمد عليه من عدمه بل وقد يؤدى إلى انخفاض قيمة المعامل المحسوب.

ج. النموذج التقاربى

يُعد هذا النموذج أكثر نماذج القياس مرونة لأن افتراضاته تنص على إمكانية تنوع تشبعات بنود المقياس على العامل الكامن وتنوع مستوى تباين الخطأ، ومن هنا نجد أن هذا النموذج يتميز بالمرونة مقارنة بنماذج القياس الأخرى، والجدير بالذكر أن معامل ثبات أوميجا يعتمد على هذا النموذج من نماذج القياس، مما يجعله أكثر دقة مقارنة بمعاملات ثبات الدرجات.

طرق حساب ثبات درجات المقياس النفسى

تعددت طرق حساب معامات ثبات درجات أدوات القياس ومنها المقياس النفسى، ويوضح شكل (١) أشهر طرق حساب معاملات ثبات الدرجات:



شكل (١) معاملات ثبات درجات أدوات القياس نقلاً عن Drost (2011, p.109)

يتضح من شكل (١) أن معامل أوميجا لا يوجد فى الاتساق الداخلى وذلك يدل على عدم شيوعه بين الباحثين بالرغم من دقته طبقاً للأدبيات البحثية التي تناولته، كما أن مناقشة كل هذه الطرق التي تم عرضها في شكل (١) ليس موضوع الدراسة الحالية لأنها تركز على معاملى ألفا وأوميجا.

معاملى ألفا وأوميجا

يرجع اختيار معاملى ألفا وأوميجا في الدراسة الحالية لثلاثة أسباب: الأول: أنهما يحتاجان تطبيقاً واحداً Single administration، وبذلك يوفران الوقت والجهد بالنسبة لمُعد المقياس والمستجيب عليه، والثانى: أنهما يمثلان نموذجين مختلفين من نماذج القياس فألفا ينتمى إلى نموذج تاو المتكافئ في الأساس، وأوميجا يتبع النموذج التقاربى كما ورد فى الجزء الخاص بنماذج القياس، والثالث: أن ألفا أكثر شيوعاً واستخداماً من أوميجا بالرغم من اختلاف دقة كل منهما فى تقدير ثبات درجات المقياس النفسى، وفيما يلى بعض المعلومات عن كل منهما:

١. معامل ألفا[٤]

يُعد معامل ألفا لحساب ثبات درجات المقياس النفسى أكثر معاملات ثبات الدرجات استخداماً بين الباحثين وتعود جذوره الحقيقية إلى معادلة Kuder and Richardson (1937) والمعروفة باسم KR-20 والمنشورة في المجلد الثانى للمجلة الشهيرة *Psychometrika* والتى تُستخدم في حساب ثبات درجات البنود ثنائية التصحيح Dichotomously scored items، حيث أتى كرونباخ وطور هذه الصيغة عام ١٩٥١م حتى يمكن استخدامها في حالة البنود ذات عدد فئات الاستجابة المتعددة Polytomously scored items. ويمكن حسابه من المعادلة التالية:

$$\alpha = \frac{n}{n-1}\left(1 - \frac{\sum_{i=1}^{n} S_{X_i}^2}{S_X^2}\right) \qquad (5)$$

حيث: n هي عدد البنود، أما بسط المعادلة فيتمثل في مجموع تباين البنود أما مقامها فيتمثل في تباين المجموع الكلى لدرجات المقياس.

يضيف Almehrizi (2013, 438) إلى أن معامل ثبات ألفا يستخدم لحساب الثبات من الدرجات الخام المكونة من عدد n من البنود سواء ثنائية أو متعددة التدريج تحت افتراضات نموذج تاو المتكافئ بالأساس.

---

[٤] تتناول هذه الدراسة معامل ألفا بتوسع، نظراً لشيوع استخدامه بين الباحثين بالرغم من كثرة المعتقدات الخاطئة عنه.

مقارنة معاملات ألفا بين مجموعتين

تعددت جهود الباحثين في مقارنة قيمتى معامل ألفا بين مجموعتين مرتبطتين فقد اقترح Feldt, Woodruff, and Salih (1987, 99) معادلة لمقارنة زوجين لألفا في حالة العينات المرتبطة والتي تظهر في الصورة التالية :

$$T = \frac{(alpha_2 - alpha_1) \times \sqrt{N-2}}{\sqrt{4 \times (1-alpha_2) \times (1-alpha_1) \times (1-r^2)}} \qquad (6)$$

حيث ت هي قيمة الإحصاءة والتي تتبع توزيع اختبار "ت"، حيث يتم مقارنة قيمة "ت" المحسوبة مع قيمة "ت" الجدولية عند درجة حرية ن – ٢، فإذا كانت القيمة المحسوبة أكبر من الجدولية دل ذلك على وجود فروق دالة إحصائياً لصالح القيمة الأعلى، وتشير $alpha_2$ إلى قيمة معامل ألفا للصورة ذات عدد فئات الاستجابة الأكثر، بينما تشير $alpha_1$ إلى قيمة معامل ألفا للصورة ذات عدد فئات الاستجابة الأقل، وN إلى عدد المشاركين، وتشير "ر" إلى معامل ارتباط بيرسون للصورتين موضوع المقارنة.

أسباب شيوع استخدام معامل ألفا بين الباحثين

يُعد معامل ألفا الأكثر شيوعاً لدى الباحثين، حيث يشير Cronbach and Shavelson (2004, 392) إلى أن معامل ألفا تم توثيقه في عديد من البحوث والدراسات المنشورة في دوريات عالمية، ويؤكد ذلك Harzing (2013) أن الدراسة التى نشرها Cronbach عام ١٩٥١م تم توثيقها أكثر من ٢٢٠٠٠ مرة، وفى إحصائية أخرى بلغ عدد مرات توثيقها ٢٨٣٠٧ مرة في عام ٢٠١٦م، مما يجعلها الأكثر توثيقاً بها من أي دراسة أخرى تم نشرها في مجلة *Psychometrika*. ويرجع السبب أيضاً إلى أن البرامج الإحصائية الحديثة والمستخدمة فى القياس النفسى والتربوى مثل الحزمة الإحصائية فى العلوم الاجتماعية SPSS تضع معامل ألفا كاختيار افتراضى أثناء حساب ثبات درجات المقياس النفسى.

يذكر Cho and Kim (2015, 207) أنه بالرغم من تعدد الطرق المنشورة في أدبيات القياس النفسى والتربوى والمستخدمة في حساب معامل ثبات درجات المقياس إلا أن ألفا حاز على اهتمام وتركيز العديد منهم. وهو يقوم على افتراضات نموذج تاو المتكافئ في الأساس التى تم مناقشتها سابقاً وأهمها-بلغة التحليل العاملى- ضرورة أن يكون مقدار التشبع متساو لكل البنود، وأيضاً تباين درجات الخطأ غير مرتبط بين هذه البنود.

القيم المقبولة لمعامل ألفا

لا يوجد اتفاق حول القيم المقبولة لمعامل ألفا لأنها تختلف باختلاف الغرض من الدراسة، حيث يشير Yang and Green (2011, 381) إلى أن القيم المقبولة لمعامل ألفا تتدرج من ٠.٧ في حالة التحليل الأول للخصائص السيكومترية للأداة، ٠.٨ للبحوث الأساسية، ٠.٩ في حالة البحوث الإكلينيكية، كما يقبل بعض الباحثين قيمة ٠.٦٥ لمعامل ألفا في حالة قلة عدد البنود.

بعض المعتقدات الخاطئة عن ألفا

بالرغم أن معامل ثبات ألفا يشتهر بين الباحثين بدرجة كبيرة جداً، إلا أن أغلبيتهم قد يجهل بنيته الداخلية وافتراضات النموذج القائم عليه، حتى أن أحد البحوث الحديثة فى مجال القياس النفسى والتربوى والمنشور عام ٢٠١٥ كان معنوناً "معامل ألفا هو الأكثر شيوعاً ولكن الأقل فهماً" Alpha: Well-known but poorly understood. ويلخص Cho and Kim (2015, 209-214 هذه المعتقدات الخاطئة في التالى:

أ. **يُعد كرونباخ هو أول من قدّم معامل ألفا:** يجد المدقق في الأدبيات البحثية المتخصصة أن معادلة ألفا هي تطوير لمعادلة كيودر وريتشاردسون رقم ٢٠ والمعروفة KR-20 والتي ظهرت عام ١٩٣٧م والتي كانت تستخدم في حساب ثبات درجات البنود ثنائية التصحيح، أما معادلة ألفا فتستخدم في حساب ثبات درجات البنود متعددة التصحيح Polytomously scored items. والجدير بالذكر في هذا السياق أن كرونباخ نفسه لم يطلق عليه اسمه حتى أنه يشعر بالخجل أن معامل ألفا أصبح ملتصقاً باسمه في الأدبيات البحثية طبقاً لم نشره في عام ٢٠٠٤م.

ب. **ألفا يساوى ثبات الدرجات:** اتفق العديد من الباحثين على أن قيمة ألفا تقل عن القيمة الحقيقية لمعامل ثبات الدرجات Lower bound of reliability، وخاصة إذا تم انتهاك افتراضات نموذج تاو المتكافئ في الأساس، وقد تكون قيمة ألفا بالسالب عندما يحتوى المقياس على بند أو أكثر تم صياغتهما بصورة سالبة ولكن تم تصحيحهما بصورة موجبة مثل باقى بنود المقياس، كما تزداد قيمة معامل ألفا عندما يكون هناك ارتباط موجب بين خطأ البنود Positively correlated item errors ومع ذلك تقل القيمة الحقيقية للثبات.

ج. **القيمة المرتفعة لمعامل ألفا تدل على الاتساق الداخلى والتجانس:** يعتقد بعض الباحثين أن ارتفاع قيمة معامل ألفا تدل على الاتساق الداخلى للأداة، ولكن ربما يعود إلى زيادة متوسط معاملات الارتباط بين البنود إلى زيادة عدد البنود المكونة للأداة أو تماثل البنود

بمعنى الاختلاف الظاهرى في الصياغة ولكنها تقيس السمة نفسها.

يشير Yang and Green (2011, 380) إلى أن ألفا لا يُعد مؤشراً للاتساق الداخلى؛ لأنه يعتمد بشكل كبير جداً على عدد البنود التي تتضمنها الأداة، كما أن الاتساق الداخلى هو الارتباطات البينية بين البنود، أما التجانس فيتعلق بأحادية البعد، ويذكر Panayides (2013) أن القيمة المرتفعة لألفا ليست مؤشراً على أحادية البعد كما أنها ليست الأفضل لأنه يمكن الحصول على هذا القيمة في حالة زيادة عدد البنود.

يضيف Davenport, Davison, Liou, and Love (2015, 8) أيضاً أن ثبات الدرجات والاتساق الداخلى والتجانس مفاهيم مرتبطة ولكن تختلف عن بعضها البعض، ويعتمد الاتساق الداخلى على الارتباطات البينية بين البنود وليس على عددها، فهو خاصية للبنود وليس المقياس، ويمكن حسابه عن طريق أخذ متوسط معاملات الارتباط البينية بين هذه البنود، أما التجانس فيعتمد على تشبع البنود على العامل فى التحليل العاملى، وفيما يختص بمعامل ألفا فهو وظيفة لطول المقياس وتشبع البنود، ويؤثر الاتساق الداخلى على ألفا بطريقتين: الأولى أن الاتساق الداخلى لابد أن يكون موجباً وإلا لا يكون المقياس ثابت الدرجات حتى لو كان طويلاً، الثانية: إذا كان موجباً فيحدد مستوى الاتساق الداخلى عند البنود المطلوبة لزيادة قيمة معامل ألفا، وباختصار فالقيمة المرتفعة لمعامل ألفا لا تُعد مؤشراً للاتساق الداخلى للبنود أو التجانس، ويؤكد Sijtsma (2015) أن مفاهيم الاتساق الداخلى وأحادية البعد يرتبطان بالبنية الداخلية لأداة القياس وبالتالي بصدق البنود وليس ثبات الدرجات.

د. **يمكن تحسين ثبات درجات المقياس عن طريق حذف بعض البنود:** تقدم الحزم الإحصائية الحديثة مثل SPSS جدولاً في ملف النواتج يحتوى على قيمة ألفا إذا تم حذف البنود، ولكن لابد ألا يعتمد الباحثون على معلومات هذا الجدول في حالة الرغبة في حذف بعض البنود وتحسين معامل ثبات الدرجات، ولكن من الأفضل فحص مصفوفة معاملات الارتباطات البينية بين البنود وتحديد البنود ذات معاملات الارتباط المنخفض وحذفها.

ه. **لابد أن تساوى أو تزيد قيمة معامل ألفا عن ٠.٧ أو ٠.٨:** يُعد هذا المعتقد من أكثرها انتشاراً بين الباحثين، ولكن تحدد القيمة المقبولة لمعامل ألفا حسب الغرض المحسوب من أجله. فقد تكون القيمة المقبولة ٠.٥ أو ٠.٦ في حالة البحوث الاستكشافية أو ٠.٩ في حالة بحوث الفرز والتصنيف.

و. **معامل ألفا هو الأفضل بين كل معاملات الثبات المنشورة:** ربما يعود هذا المعتقد الخاطئ إلى انتشار استخدامه بين الباحثين، بالرغم من أنه حاز على نقد عديد من الباحثين قديماً

وحديثاً (Green, Lissitz, & Mulaik, 1977; Novick & Lewis, 1967; Osburn, 2000; Raykov, 2001b; Sijtsma, 2009; Zimmerman, Zumbo, & Lalonde, 1993) حتى من قبل كرونباخ نفسه Cronbach and Shavelson (2004, 397)، ومع ذلك مازال يُستخدم فى كثير من الدراسات المنشورة فى دوريات محلية وإقليمية ودولية. ولعل ذلك يرجع إلى غياب الوعى لدى غالبية الباحثين بالطرق البديلة المستخدمة وكذلك توافر الحزم الإحصائية المستخدمة فى حسابها.

يضيف Cho (2016, 654) أيضاً أن الباحثين يستخدمون معامل ألفا دون محاولة التحقق من افتراضاته وذلك لوجود بعض المعتقدات الخاطئة لديهم والتي تتمثل في التالى:

أ. إمكانية استخدامه مع أي نوع من البيانات.
ب. عدم تأثره بانتهاك الافتراضات الخاصة به.
ج. تدل القيمة المرتفعة له على التحقق من افتراضاته.
د. صعوبة التحقق من الافتراضات.
ه. صعوبة استخدام الطرق البديلة في حالة انتهاك افتراضاته.

ومن هنا يمثل التعرف على هذه المعتقدات الخاطئة عن معامل ألفا أهمية كبيرة، كما أن الأهم من ذلك هو أن يضعها الباحثون في الحسبان عند حساب معاملات ثبات درجات المقياس النفسى المستخدم في بحوثهم، ومن ثم تغيير الممارسات الإجرائية المتبعة في ذلك حتى يكون اختيار معاملات ثبات الدرجات قائم على فهم للطبيعة الداخلية لهذه المعاملات.

يتضح من العرض السابق الخاص بمعامل ألفا أن العديد من أدبيات البحث تناولته بالتحليل والنقد وذلك من أكثر من منظور وإن دل ذلك على شئ فإنما يدل على أنه ما زال موضع دراسة ونقاش بين المهتمين بالقياس النفسى والتربوى.

٢. معامل أوميجا

يمثل معامل أوميجا أحد أدق معاملات ثبات درجات المقياس النفسى، ويُعد McDonald (1978, 1999) هو أول من قدم معامل أوميجا كبديل لألفا ويرمز له بالرمز ω، والذى يشير إلى مربع مجموع التشبعات إلى التباين الكلى، وبعبارة أخرى فهو يعنى تباين الدرجات الحقيقية للمقياس والتي ترجع للعوامل المستخلصة من التحليل العاملى مقارنة بالتباين الكلى الذى يرجع لهذه العوامل بالإضافة للخطأ، ويتم حسابه من المعادلة (٧):

$$\omega = [\Sigma (\lambda_j)]^2/([\Sigma(\lambda_j)]^2 + \Sigma \Psi_j^2) \quad (7)$$

حيث تشير $(\Sigma \lambda_j)^2$ إلى مربع مجموع تشبعات البنود بينما تشير $(\Sigma \Psi_j)^2$ إلى مربع مجموع البواقى.

يشير Zinbarg, Revelle, Yovel, and Li (2005, 124) إلى أن معامل ألفا لا يُعد مكافئاً لمعامل أوميجا إلا في شروط صارمة جداً. ويفسر الباحث ذلك في ضوء اختلاف نموذج القياس الذى ينتمى إليه كل منهما، حيث أن معامل ألفا ينتمى إلى نموذج تاو المتكافئ في الأساس المعروف بصرامة افتراضاته مقارنة بالنموذج المتقارب الذى ينتمى إليه معامل أوميجا.

ويذكر Revelle and Zinbarg (2009) أن معامل أوميجا أفضل معامل ثبات لدرجات أدوات القياس وذلك بعد تحليلهما لعدد ١٣ معادلة مختلفة مستخدمة لحساب ثبات درجات هذه الأدوات، ويرجع ذلك أيضاً لانتمائه للنموذج التقاربى الذى يتميز بمرونة افتراضاته، كما يتميز استخدام معامل أوميجا عن ألفا حساب ثبات درجات بأن الباحث يختار النموذج المناسب للبيانات بعد إجراء تحليلات أولية، بينما في الثانية يفترض الباحث نظرياً توافر الافتراضات دون إجراء أي خطوات عملية للتحقق من ذلك مما يؤثر سلباً على معامل الثبات الناتج.

أوصت نتائج دراسة Dunn et al. (2014) أيضاً إلى ضرورة استخدام معامل أوميجا بدلاً من ألفا أثناء حساب الاتساق الداخلى لدرجات المقياس النفسى. ويضيف Zhang and Yuan (2016) أن معامل أوميجا يؤدى إلى تقدير أدق لمعامل ثبات الدرجات إذا لم تتحقق افتراضات نموذج تاو المتكافئ في الأساس.

تناولت دراسات تطبيقية دقة معامل أوميجا مقارنة بألفا مثل دراسة Deng and Chan (2016) التي هدفت إلى اختبار الفرق بين قيمتى ألفا وأوميجا من خلال حساب الفروق في الخطأ المعيارى وحدود الثقة الخاصة بكل منهما وذلك باستخدام الحزمة الحاسوبية R. تم تطبيق أربعة مقاييس نفسية لتحقيق هدف الدراسة تتمثل في الخصائص المعرفية لتلاميذ المدارس المتوسطة، قائمة العوامل الخمسة الكبرى للشخصية، استبيان أساليب الدعابة، مقياس تكيف القدرة التكيفية للأسرة. أشارت النتائج إلى وجود فروق دالة في معظم المقاييس التي تم تطبيقها لصالح معامل أوميجا.

أسباب عدم شيوع استخدام معامل أوميجا بين الباحثين

يرجع عدم شيوع استخدام معامل أوميجا بين الباحثين-بالرغم من دقته بسبب مرونة

افتراضات النموذج التقاربى الذى ينتمى إليه-إلى عدم انتشار برامج الحاسب الآلى المستخدمة فى تقديره وكذلك وعى غالبية الباحثين بها مع أنها متاحة مجاناً على الشبكة العنكبوتية، فمعظم الباحثين يستخدمون الحزمة الحاسوبية فى العلوم الاجتماعية SPSS لسهولتها وانتشارها بينهم وذلك أثناء حساب ثبات درجات المقياس النفسى.

يؤكد ذلك ما أشار إليه Padilla and Divers (2016, 436) إلى أنه بالرغم من عدم تحيز معامل أوميجا فإنه أقل استخداماً وتوثيقاً في الأدبيات المنشورة مقارنة بألفا، وربما يرجع ذلك إلى المعرفة المحدودة بخصائصه الإحصائية من قبل الباحثين وكذلك كيفية حسابه باستخدام البرامج الحاسوبية الحديثة.

يتضح من تناول معامل أوميجا أنه أدق في تقدير معامل ثبات درجات المقياس النفسي مقارنة بألفا وذلك طبقاً لم ورد في الدراسات المنشورة في هذا الميدان، كما أنه أقل شيوعاً واستخداماً، نظراً لعدم وعى غالبية الباحثين ببنيته المنطقية وكذلك كيفية الستخدام البرامج الحاسوبية فى حسابه مثل RStudio.

ثانياً: عدد فئات الاستجابة لليكرت --

يمثل تحديد العدد الأمثل لفئات الاستجابة عند تصميم المقياس النفسى باستخدام أسلوب ليكرت محوراً مهماً والذى يحتاج مزيداً من التفكير من قبل الباحثين، وتبدو حيرة الباحثين في الاختيار بين أمرين: أولهما الاختيار بين العدد الفردى Odd number أو الزوجى Even number لفئات الاستجابة، وثانيهما إذا كان العدد فردياً فأى عدد يتم اختياره (٣، ٥، ٧،٩،..)، أما كان العدد زوجياً فأى عدد يتم اختياره (٢، ٤، ٦، ٨...)؟ ومن هنا تأتى الحيرة بين الباحثين، نظراً لاختلاف ما توصلت إليه نتائج الدراسات السابقة من تأثير عدد فئات الاستجابة على الخصائص السيكومترية للمقياس النفسى لاسيما ثبات درجاته.

يذكر فؤاد أبوحطب وآمال صادق (١٩٩١، ١٤٩-١٥١) أنه توجد أعداد مختلفة من فئات الاستجابة منها ٣، ٥، ٧، ٩، ١١ أو أكثر من ذلك ولكن عدد الحد الأقصى الأكثر شيوعاً هو ٧ فئات ومع كل ذلك فلابد من أن يكون عدد فئات الاستجابة فردياً حتى تتوفر النقطة الوسيطة في بدائل الاستجابة على المقياس النفسى.

تناول بعض الباحثين تأثير النقطة الوسيطة على الخصائص السيكومترية للمقياس النفسى المصاغ باستخدام أسلوب ليكرت؛ حيث هدفت دراسة Weems and Onwuegbuzie (2001) إلى التعرف على تأثير الفئة الوسيطة للاستجابة وكذلك البنود السالبة على معامل ثبات

ألفا، حيث تكّون عدد المشاركين من ١١٦٢ طالباً من طلاب الجامعة والذين استجابوا على استبيان خبرات طلاب الجامعة من إعداد Pace (1994). أشارت نتائج الدراسة فيما يتعلق بالدراسة الحالية إلى أن النقطة الوسيطة كانت أفضل فى التدريج السباعى لفئات الاستجابة وأدت إلى ارتفاع قيمة معامل ألفا.

توصلت نتائج بعض الدراسات العربية أيضاً إلى أن قيمة معامل ثبات الدرجات تزداد في حالة العدد الفردى فقط الذى يحتوى على النقطة الوسيطة مثل دراسة محسوب عبد القادر الضوى (٢٠١١) التي هدفت إلى تحرى تأثير الاحتفاظ بالدرجات المتطرفة أو حذفها وعدد فئات الاستجابة على تقدير معامل ألفا لكرونباخ، تم استخدام مقياس اتجاهات المعلمين نحو تكنولوجيا التعليم وإعداد ست صور منه تبعاً لفئات الاستجابة (٢، ٣، ٤، ٥، ٦، ٧)، حيث تم تطبيقهما على ٧٣٢ معلماً موزعين على خمس مجموعات، وقد أشارت نتائج الدراسة إلى اختلاف تقدير معامل ألفا لكرونباخ باختلاف عدد فئات الاستجابة، حيث كانت قيمة معامل ألفا مرتفعة عند التدريج الثلاثى والخماسى والسباعى مقارنة بالأعداد الزوجية لفئات الاستجابة التي تم تناولها في الدراسة.

يضيف Tsang (2012, 122-124) أن هناك خلافاً بين الباحثين فيما يتعلق بتأثير استخدام النقطة الوسيطة على معامل ثبات درجات المقياس النفسى، وينشأ الخلاف بينهم من مصدرين أولهما منهجى Methodological وثانيهما يختص بمعتقدات المشاركين أثناء الاستجابة Epistemological، ويكاد يكون هناك اتفاق بين الباحثين فيما يتعلق بالجانب المنهجى في عدم التأثير الملحوظ للنقطة الوسيطة على معامل ثبات الدرجات، ولكن فيما يتعلق بجانب المعتقدات فمن المفضل استخدام النقطة الوسيطة.

يرى فريق آخر من الباحثين أن قيمة معامل ثبات الدرجات تزداد بزيادة عدد فئات الاستجابة مثل دراسة Preston and Colman (2000) التي هدفت إلى تحرى اختلاف عدد فئات الاستجابة (٢-١١) على بعض مؤشرات الثبات (ألفا – إعادة التطبيق) وكذلك معاملات الصدق والتمييز، حيث تكّون عدد المشاركين من ١٤٩ مواطناً والذين استجابوا على مقياس الرضا عن الخدمة المقدمة فى بعض الأماكن الخدمية والمكون من ١٠١ بنداً، أشارت نتائج الدراسة إلى أن التدريج الثانى والثلاثى والرباعى أدوا إلى مؤشرات منخفضة، فى حين كانت المؤشرات أعلى فى حالة التدريج الخماسى والسداسى والسباعى.

توصلت دراسة المحاكاة التي أجراها Zumbo et al. (2007) إلى ارتفاع قيمة معامل ثبات الدرجات باستخدام ألفا وثيتا كلما زاد عدد فئات الاستجابة، حيث تم تحديدها في الدراسة من

٢-٧ فئات. كما أشارت النتائج أيضاً إلى أن معامل ألفا يكون أكثر تحيزاً سلباً في حالة عدم تساوى تشبع البنود على العامل وكذلك عدم تحقق شرط أحادية البعد بين البنود.

توصلت بعض الدراسات العربية أيضاً إلى ارتفاع قيمة معامل ألفا بزيادة عدد فئات الاستجابة، ومن بين هذه الدراسات دراسة محمد حسين سعيد (٢٠٠٧) التي هدفت إلى التعرف على تأثير اتجاه المفردة وعدد فئات الاستجابة على ثبات أدوات القياس المصاغة فى ضوء أسلوب ليكرت، تم استخدام مقياس قلق الاختبار لنبيل الزهار ودنيس هوسفر (١٩٨٥) بعد إعادة صياغته في ثلاثة صور مختلفة تبعًا لعدد بدائل الاستجابة (٣، ٥، ٧) وهو عدد فردى لفئات الاستجابة، حيث تكّون عدد المشاركين من ٤٥٠ طالباً من طلاب الصف الأول الثانوى. أشارت نتائج الدراسة أن الصورة ذات البدائل السبعة حققت أعلى قيمة لمعامل ثبات الدرجات.

أجرى Choudhury and Bhattacharjee (2014) دراسة للتعرف على أفضل عدد لفئات الاستجابة (٥-٩) فئة. تم تطبيق أحد المقاييس النفسية على ١٠٠ مشارك. أشارت نتائج الدراسة إلى تفضيل المشاركين للتدريج الخماسى للمقياس، كما أظهرت النتائج أيضاً وجود فروق غير دالة في معامل ألفا لثبات الدرجات، كما زادت قيمة معامل ثبات الدرجات بزيادة عدد فئات الاستجابة.

تسير في نفس الاتجاه دراسة محمد عبدالرحمن إسماعيل (٢٠١٥) التي هدفت إلى التعرف على أثر عدد بدائل الاستجابة في مقياس ليكرت على الخصائص السيكومترية للمقياس وقياس الاتجاهات، ولتحقيق هذا الهدف تم تطوير استبانة خاصة بالبحث تتكون من خمسة بنود مأخوذة من مقياس الالتزام التنظيمي تم تكرارها أربع مرات بعدد خيارات بدائل عدد الاستجابة (٣، ٥، ٧، ٩)، حيث تكّون عدد المشاركين في الدراسة من ٥٠٦ متدربين من مجتمع متدربي معهد الإدارة العامة، المركز الرئيسي بالرياض، وبلغ عدد الاستبانات الصالحة للتحليل ٣٦٩ استبانة. وتوصلت الدراسة إلى عدد من النتائج أهمها: أن معامل ألفا ازداد بزيادة عدد بدائل الاستجابة.

هناك من أشار من الباحثين إلى أن العدد المناسب لفئات الاستجابة هو من ٤-٧ فئة، حيث هدفت دراسة Lozano, García-Cueto, and Muñiz (2008) إلى التعرف على تأثير عدد فئات الاستجابة (٢-٩) على الخصائص السيكومترية (ثبات الدرجات وصدق البنود) وذلك باستخدام فنية مونت كارلو على مقياس مكون من ٣٠ بنداً. وأشارت نتائج الدراسة إلى أن أفضل عدد لفئات الاستجابة ينحصر بين ٤-٧ فئات.

يضيف Wakita, Ueshima, and Noguchi (2012, 534) أن مقياس ليكرت هو أكثر أساليب تصميم مقاييس التقرير الذاتي انتشاراً، كما أن عدد فئات الاستجابة من (٤-٧) فئات هو الأكثر استخداماً بين الباحثين. ويستخدم العدد الفردى لفئات الاستجابة عندما يكون هناك استجابات حيادية في حين يستخدم العدد الزوجى لفئات الاستجابة عندما لا يوجد استجابات حيادية في الأغراض البحثية التي تم تصميم أداة القياس من أجلها، وأشارت نتائج دراسته التي استخدمت ثلاث صور من مقياس السمات الخمس الكبرى في الشخصية إلى أن معاملات ثبات ألفا زادت بزيادة عدد فئات الاستجابة.

يؤكد Bindak (2013, 13) أهمية استخدام الباحثين للتدريج الخماسى لفئات الاستجابة أثناء إجراء الدراسات التي تحتوى على أدوات القياس المتدرجة بأسلوب ليكرت وذلك بناءً على نتائج دراسات المحاكاة. ويرى Sullivan and Artino (2013, 541) أنه عندما قدم ليكرت أسلوب التدريج المعروف باسمه عام ١٩٣٢م لقياس الاتجاهات كان العدد الأمثل لفئات للاستجابة هو (٥) و (٧) فئات.

يشير Clason and Dormody (n.d.) إلى أن أسلوب التقدير الجمعى الذى قدمه ليكرت فى عام ١٩٣٢م يُستخدم فى تقدير استجابات المشاركين فى الدراسات المسحية الخاصة بالاتجاهات وغيرها، كما أن عدد فئات الاستجابة يتغير حسب طبيعة الدراسة وأن عدد خمس فئات للاستجابة هو الأكثر انتشاراً مع أن بعض الباحثين يمكنه استخدام العدد الزوجى لفئات الاستجابة وذلك بحذف الاستجابة المحايدة Neutral.

يوجد من أشار إلى تساوى قيم معامل ألفا باختلاف عدد فئات الاستجابة (٦، ٧، ٩) و (٤، ٥، ٦)، حيث هدفت دراسة Shaftel, Nash, and Gillmor (2012) إلى بحث تأثير عدد خمس فئات للاستجابة (٥، ٦، ٧، ٩، ١٠) على تقدير معامل ثبات ألفا، حيث تكّون عدد المشاركين في الدراسة من ١٥٥٤ مشاركاً من طلاب الجامعة، موزعين على خمس مجموعات طبقاً لعدد فئات الاستجابة والذين استجابوا على مقياس اتجاه الطلاب بين الثقافات والمكون من ٣٥ بنداً. وتوصلت نتائج الدراسة إلى تساوى قيم معامل ثبات ألفا في حالة كون عدد فئات الاستجابة ٦، ٧، ٩ فئات.

هدفت دراسة Lee and Paek (2014) أيضاً إلى التعرف على تأثير عدد فئات الاستجابة (٢-٦) على الخصائص السيكومترية لأدوات القياس، وقد أشارت نتائج الدراسة أنه لا توجد فروق جوهرية بين فئات الاستجابة الرباعية والخماسية والسداسية، ولكن يوجد تأثير أكبر

عندما تكون فئات الاستجابة ثنائية، وخاصة مع قلة عدد بنود أداة القياس كأن يكون ستة بنود فأقل لأن معامل ثبات ألفا كان ٠.٨٢ في حالة التدريج السداسي و٠.٧٣ في حالة التدريج الثنائي، كما يضيف أن الاختلاف في الخصائص السيكومترية نتيجة تغير عدد فئات الاستجابة يكون أكبر عند قلة عدد بنود المقياس. كما أن معامل ألفا وصل ٠.٨٩ في حالة كون عدد البنود ٢٠ بنداً و٠.٦٩ في حالة كون عدد البنود خمسة.

يظل أسلوب ليكرت محور اهتمام الباحثين في الدراسات الحديثة أيضاً؛ حيث هدفت دراسة Subedi (2016) إلى مراجعة الأدبيات البحثية التي تناولت أسلوب التقدير الجمعي لليكرت وتحليلها للخروج منها بتوصيات تفيد الباحثين في الدراسات التطبيقية التي تستخدم هذا الأسلوب. وقد أشارت هذه الدراسة إلى أن هناك جدلاً واسعاً بين الباحثين يتعلق بثبات الدرجات واستخدام النقطة الوسيطة في الاستجابة أثناء استخدام هذا الأسلوب. وتوصى الدراسة باستخدام النقطة الوسيطة وذلك من منظور معتقدات المشاركين، بالرغم من عدم وجود تأثير لاستخدامها من عدمه من منظور منهجي.

الخلاصة أن عملية اختيار عدد فئات الاستجابة تتطلب التأنى والدقة. ويجب أن يتم الاختيار بناءً على أسس علمية فهى ليست عملية عشوائية تتم تبعاً لرغبة الباحثين وأهوائهم بعيداً عن الأدلة العلمية، كما أن أهمية فئات الاستجابة لا تكمن فقط في التأثير على قيم معاملات ثبات درجات المقاييس النفسية بل أيضاً في إتاحة الفرصة أمام المفحوصين للاستجابة بصدق، الأمر الذى يزيد من دقة القياس وقلة الأخطاء. وقد تم اختيار ثلاث فئات فردية للاستجابة في الدراسة الحالية نظراً لاتفاق معظم الدراسات على أهمية العدد الفردى لاحتوائه على النقطة الوسيطة كما أنه الأكثر شيوعاً في الدراسات المنشورة.

ثالثاً: عدد المشاركين

يهتم الباحثون بتحديد العدد المناسب للمشاركين أثناء إجراء الدراسة الاستطلاعية التي يتوقف عليها التحقق من الكفاءة السيكومترية للمقياس النفسى الذى تم تصميمه، فيحاول بعض الباحثين توفير الوقت والجهد في جمع البيانات وتصحيحها وإدخالها على الحساب الآلى وذلك من خلال اختيار عدد قليل من المشاركين في الاستجابة على أدوات القياس الخاصة بدراساتهم، ولكن لابد من وجود أساس علمى يتوقف عليه تحديد عدد المشاركين حتى تؤتى هذه الدراسات ثمارها.

حظى تحديد عدد المشاركين في الاستجابة على المقياس النفسى باهتمام العديد من الباحثين المتخصصين في القياس النفسى والتربوى قديماً وحديثاً، حيث يقترح Fleiss (1984) أنه

عندما يكون عدد المشاركين ١٥-٢٠ مشاركاً فهو كاف لحساب ثبات درجات المقياس النفسى وذلك طبقاً لما ورد في Bonett (2002, 399).

يتفق كل من Kline (1986) و Nunnally and Bernstein (1994) على ضرورة ألا يقل عدد المشاركين عن ٣٠٠ مشارك للحصول على معامل ثبات للدرجات يتسم بالدقة، وذلك طبقاً لما تناوله Yurdugul (2008, 397). ويؤكد ذلك ما أشار إليه Segall (1994, 361) إلى أن ٣٠٠ مشارك يُعد أقل عدد مطلوب لحساب ثبات درجات المقياس النفسى وذلك بناءً على نتائج دراسات المحاكاة.

ويرى Miller (1995) أن عدد المشاركين يمثل أهمية قصوى فى مدى الثقة فى معامل ثبات درجات الاختبار؛ لأنه يمثل المجتمع الأصلى للدراسة، كما أن دقة قيمة معامل ألفا تعتمد على عدد المشاركين مضيفاً أن الدراسات التى تستخدم أسلوب نمذجة المعادلة البنائية لا تتأثر كثيراً بالخطأ فى تحديد المشاركين؛ لأن طبيعة هذه الدراسات تتطلب عدداً كبيراً من المستجيبين على أدوات القياس، كما أن معامل ألفا قد يزداد بطريقة خاطئة نتيجة زيادة المشاركين.

يذكر Charter (1999, 559) أن أقل عدد مطلوب للمشاركين لحساب معامل ثبات الدرجات الذى يمثل المجتمع هو ٤٠٠ مشارك. كما يضيف أيضاً عام (٢٠٠٣) أنه كلما قل عدد المشاركين أدى ذلك إلى عدم استقرار معامل ثبات درجات الاختبار النفسى وبالتالي التأثير السالب على حدود الثقة في القيمة الحقيقية التي تمثل المجتمع، كما لاحظ أيضاً أن عدد المشاركين المدون في الدراسات المنشورة زاد في العقد الحالي عن العقد السابق.

استمرت محاولات الباحثين في تحديد عدد المشاركين في معاملات الارتباط التي يمكن استخدامها في طرق حساب ثبات درجات الاختبار مثل إعادة التطبيق والتجزئة النصفية، حيث قدم Bonett and Wright (2000) معادلات رياضية يمكن الاستفادة منها في تحديد عدد المشاركين المطلوب لحساب معاملات بيرسون، سبيرمان، كندال.

يوجد بعض الباحثين الذين حددوا أعداداً افتراضية للمشاركين لبحث تأثيرها على معاملات ثبات الدرجات، ومن بينها دراسة Ercan et al. (2007) التى هدفت إلى بحث تأثير أعداد المشاركين (١٧٠ - ٢٣٠ - ٢٤٠ - ٣٢٠) على بعض معاملات ثبات الدرجات (ألفا - ثيتا - أوميجا). أشارت نتائج الدراسة إلى أن قيمتى ألفا وثيتا لا يرتبطان بعدد المشاركين، ولكن ترتبط قيمة معامل أوميجا بالعدد الأكبر للمشاركين وخاصة في حالة تمثيل هذه المعاملات لمعالم

المجتمعات التي ينتمى إليها المشاركون في الاستجابة على أدوات القياس. كما أن قيم هذه المعاملات ترتبط أكثر بعدد البنود المكونة لأداة القياس.

هدفت دراسة Javail et al. (2011) أيضاً إلى دراسة تأثير أعداد مختلفة للمشاركين تتمثل في (٣٣ – ٤٣ – ٣٥ – ٣٦ – ٣٧ – ٣٨ – ٣٩ – ٤٠ – ٤٥ – ٥٠ – ٦٠ – ٧٠ – ٨٠ – ٩٠ – ١٠٠) على بعض معاملات الاتساق الداخلى (ألفا – أوميجا – ثيتا – التجزئة النصفية)، وقد أشارت نتائج الدراسة إلى أن قيمة معامل ثبات الدرجات ترتبط أكثر بمتوسط الارتباطات البينية بين البنود أكثر من عدد المشاركين، وأيضاً كانت قيمة معامل ألفا أكبر مقارنة ببقية المعاملات التي تضمنتها الدراسة، كما أوصت الدراسة بضرورة ألا يقل عدد المشاركين عن ٥٠ مشاركاً.

يرى بعض الباحثين – على الجانب الآخر– أن قيمة معامل ثبات الدرجات لا تتأثر بعدد المشاركين فقط بل توجد عوامل أخرى مؤثرة، حيث يضيف Yurdugul (2008, 397) أنه عندما تزداد قيمة أول جذر كامن لبنود المقياس عن ٦ فإنه يقل تحيز معامل ألفا حتى في حالة قلة عدد المشاركين المستجيبين على أداة القياس (٣٠ مثلاً) وذلك من خلال دراسة المحاكاة التي أجراها باستخدام فنية مونت كارلو، كما أظهرت نتائج الدراسة أيضاً أنه عندما يزيد عدد المشاركين عن ٥٠٠ مشارك فإنه لا يوجد تأثير لقيمة الجذر الكامن على معامل ثبات الدرجات لأن هذا العدد كافٍ للحصول على معامل دقيق لثبات للدرجات.

ترتبط حدود الثقة بعدد المشاركين؛ حيث يشير Meeker and Escobar (1998) إلى أنه تُعد حدود الثقة أحد أهم الطرق التى تفيد فى تحديد مدى دقة عدد المشاركين فى تمثيل مجتمع الدراسة، كما أنها مؤشراً على مدى الخطأ فى اختيار المشاركين أو خطأ المعاينة Sampling error.

وتأتى فكرة حدود الثقة من أن تقدير النقطة Point estimation يكون غير دقيق ولذلك نحتاج تقدير مسافى Interval estimation، ومن هنا تُعد حدود الثقة هي أفضل تعبير عن التقدير المسافى لقيم معاملات الثبات التي يرغب الباحثون في التأكد من دقتها، كما يمكن القول في هذا السياق أيضاً أن اتساع المسافة بين حدى الثقة والذى يتم حسابه من طرح الحد الأدنى من الحد الأعلى لحدود الثقة يكون أضيق بزيادة عدد المشاركين، ويُعد مستوى الدلالة ٩٥% هو الأكثر استخداماً بين الباحثين في حالة تدوين حدود الثقة وهو يعنى –في سياق دراسة معاملى ألفا وأوميجا–أنه إذا تم تطبيق المقياس النفسى ١٠٠ مرة على مجموعات مختلفة من المشاركين من

نفس المجتمع، فإن الباحث يكون عنده ثقة بنسبة ٩٥% أن قيم معاملى ألفا وأوميجا للمجتمع تقع بين حدى الثقة الأدنى والأعلى، ومن هنا كلما كانت حدود الثقة أضيق كلما كان الفرق بين القيمة المحسوبة لمعاملى ألفا وأوميجا أقرب إلى القيمة الحقيقية للمجتمع.

يضيف Raykov and Penev (2006) وكذلك Terry and Kelley (2012) أن عدد المشاركين يؤثر على حدود الثقة، فكلما زاد عدد المشاركين أصبحت حدود الثقة أضيق والتي تُعد مؤشراً كمياً على تمثيل معامل ثبات الدرجات للمجتمع الذى اشتق منه المشاركين في الاستجابة على المقياس النفسى، مثال: لو حصل باحث على معامل ثبات ٠.٧٤ نتيجة تطبيق مقياس ما على ٥٠ مشاركاً، وحصل باحث آخر على معامل ثبات ٠.٧٤ نتيجة تطبيق مقياس ما على ٥٠٠ مشارك، فقد يظن البعض أن دقة المعاملين تتساوى وهذا اعتقاد خاطئ؛ لأن الباحث الأول حصل على حدود ثقة عند مستوى ٩٥% تنحصر بين ٦٣-٨٥، في حين حصل الباحث الثانى على حدود ثقة تنحصر بين ٧١-٧٧ وهى أضيق وتدل على دقة معامل ثبات درجات المقياس في تمثيل مجتمع المشاركين.

حاولت بعض الدراسات مقارنة حدود الثقة المستخدمة فى حالة معامل ألفا، حيث هدفت دراسة Romano, Kromrey, and Hibbard (2010) إلى مقارنة ثماني طرق لحساب حدود الثقة حول معامل ألفا لتحديد أيهما أكثر دقة باستخدام فنية مونت كارلو باختلاف أربعة مجموعات للمشاركين:١٠، ٥٠، ١٠٠، ١٠٠٠. أشارت نتائج الدراسة إلى أن القيمة الأعلى لألفا وعدد المشاركين الأكبر أديا إلى أفضل تقدير لحدود الثقة، في حين أن العدد الأقل للبنود أدى إلى عدم دقة حدود الثقة. كما توصلت دراسة Romano, Kromrey, Owens, and Scott (2011) إلى نفس النتائج.

اهتمت أيضاً بعض الدراسات بمقارنة طرق حساب حدود الثقة المرتبطة بمعامل أوميجا أيضاً مثل دراسة Xu (2014) التي هدفت إلى مقارنة ثلاث طرق لحساب حدود الثقة تتمثل في Wald methodوطريقة الاحتمالية likelihood method وكذلك طريقة التحرر من التحيز Bias-corrected and accelerated (Bca) bootstrap في ضوء عدد المشاركين (١٠٠، ٣٠٠، ٥٠٠)، عدد البنود (٦، ١٢)، تشبع العوامل (٠.٣، ٠.٤، ٠.٦، ٠.٧، ٠.٨، ٠.٩)، والاعتدالية (الالتواء، التفرطح). أشارت نتائج الدراسة إلى أنه أدت الطرق الثلاث إلى نتائج متساوية في حالة اعتدالية البيانات، ولكن في حالة عدم الاعتدالية كانت نتائج طريقة (Bca) أفضل من باقى الطرق. وقد أظهرت النتائج أيضاً أن اتساع حدود الثقة يكون أضيق كلما زاد عدد

المشاركين.

يضيف Padilla and Divers (2016, 436) أن معامل أوميجا يتطلب عدد ١٠٠ مشارك أو أكثر لحساب حدود الثقة الخاصة به وذلك طبقاً لنتائج دراسة المحاكاة باستخدام فنية مونت كارلو.

يتضح من عرض الجزء الخاص بعدد المشاركين أنه لا يوجد اتفاق بين الباحثين على تحديد العدد المناسب اللازم لحساب معامل ثبات درجات أدوات القياس، ومن هنا يظل هذا الجزء محور اهتمام العديد من الباحثين في دراسات مستقبلية ومنها الدراسة الحالية.

تعقيب عام على الإطار النظرى والبحوث والدراسات السابقة التي تم توظيفها

يتضح من العرض السابق الخاص بمتغيرات الدراسة أن موضوع دقة ثبات درجات المقياس النفسى يرتبط بعدد فئات الاستجابة وكذلك عدد المشاركين، ويظهر أهم ما تم ملاحظته من خلال القراءة الناقدة للإطار النظرى والدراسات السابقة التي تم توظيفها في النقاط التالية:

١. عدم وجود اتفاق بين الباحثين حول العدد الأمثل لفئات الاستجابة وتأثيره على تقدير ثبات درجات المقياس النفسى، حيث يرى بعض الباحثين أنه لا توجد فروق جوهرية بين العدد الفردى والزوجى لفئات الاستجابة (Wong et al., 2011).

٢. يرى فريق آخر من الباحثين أن العدد الفردى لفئات الاستجابة أفضل وذلك لاحتوائه على النقطة الوسيطة والتي تعطى المستجيب الحرية في الاختيار، الأمر الذى يزيد من موضوعية الاستجابة على المقياس النفسى (Subedi, 2016; Sullivan & Artino, 2013; Tsang, 2012; Weems & Onwuegbuzie, 2011; Weng, 2004). ولذلك تم اختيار أعداد فردية لفئات الاستجابة فى متن الدراسة الحالية.

٣. تناولت معظم الدراسات الأجنبية معامل ألفا فقط أثناء فحص تأثير عدد فئات الاستجابة على تقدير ثبات درجات المقياس النفسى (Shaftel et al., 2012)، كما تناولت كل الدراسات العربية التي تم التوصل إليها أيضاً معامل ألفا كممثل لثبات الدرجات وذلك لمقاييس تم إعدادها من قبل آخرين (حجاج غانم أحمد على، ياسر عبدالله حفنى حسن، ٢٠١١؛ سعيد حسن آل عبد الفتاح الغامدى، ٢٠٠٣؛ محسوب عبدالقادر الضوى، ٢٠١١؛ محمد سعيد حسين، ٢٠٠٧؛ محمد عبدالرحمن إسماعيل، ٢٠١٥)، وذلك على الرغم من عدم دقته في تقدير ثبات درجات المقياس النفسى كما ورد في متن الدراسة، ولذلك تتناول الدراسة الحالية معامل أوميجا بالإضافة إلى ألفا.

٤. استخدمت بعض الدراسات مجموعات مختلفة لبحث تأثير عدد فئات الاستجابة على معاملات ثبات الدرجات (Lozano et al., 2008; Shaftel et al. ; 2012)، ولكن هل هناك ما يضمن أن الاختلاف فى قيم معاملات ثبات الدرجات الناتجة راجعة إلى اختلاف عدد فئات الاستجابة وليس اختلاف خصائص مجموعات المشاركين المستجيبة على المقياس النفسى.

٥. أكدت بعض الدراسات أن التدريج الخماسى لفئات الاستجابة أكثر أعداد فئات الاستجابة مُلاءمة للمشاركين كما أنه يؤدى إلى معاملات ثبات جيدة للدرجات (Bindak, 2013; Clason & Dormody, n.d.)، وكذلك نتائج بعض الدراسات العربية (سعيد حسن آل عبد الفتاح الغامدى، ٢٠٠٣).

٦. توصلت نتائج معظم الدراسات إلى أن قيمة معامل ثبات درجات المقياس النفسى تزداد طردياً مع زيادة عدد فئات الاستجابة (Chomeya, 2010; Choudhury & Bhattacharjee, 2014; Preston & Colman, 2000; Wakita et al., 2012; Weng, 2004)، وكذلك نتائج بعض الدراسات العربية (محسوب عبدالقادر الضوى، ٢٠١١؛ محمد حسين سعيد، ٢٠٠٧؛ محمد عبدالرحمن إسماعيل، ٢٠١٥).

٧. توصلت نتائج بعض الدراسات إلى أن قيمة معامل ثبات درجات المقياس النفسى لا ترتبط بزيادة عدد فئات الاستجابة مثل دراسة Shaftel et al. (2012)، وكذلك دراسة حجاج غانم أحمد على وياسر عبدالله حفنى حسن (٢٠١١).

٨. توصلت نتائج بعض الدراسات إلى أن تأثير عدد فئات الاستجابة على معامل ثبات درجات المقياس النفسى يزداد كلما قل عدد البنود مثل دراسة Lee and Paek . (2014)

٩. راجعت بعض الدراسات أكثر أعداد فئات الاستجابة انتشاراً في الدراسات المنشورة في دوريات علمية محكمة مثل دراسة Hartley (2013).

١٠. عدم وجود معادلات رياضية لحساب عدد المشاركين المطلوب للاستجابة على المقياس النفسى إلا في حالة معامل ألفا أو معامل ارتباط بيرسون المستخدم في طريقة إعادة التطبيق، علاوة على أن هذه المعادلات تحتاج متخصص لتطبيقها حتى أن المعادلات الموجودة المراد حساب معامل ثباتها ركزت على ألفا فقط فى ضوء اختلاف الأدبيات البحثية حول العدد المطلوب (Bonett, 2002, 2003; Bonett & Wright, 2000).

١١. يجد المدقق فى الأدبيات البحثية المنشورة فيما يتعلق بعدد المشاركين أن كلها مجرد اقتراحات لأعداد محددة تتراوح بين ٥٠-٤٠٠ مشارك أو معادلات لمعاملات ثبات مثل

ألفا دون أخرى (Charter, 1999; Ercan et al., 2007; Javail et al., 2011; Kline, 1986; Nunnally & Bernstein; 1994; Segall, 1994;)، وهذا يبرز الحاجة الماسة لوجود قواعد ثابتة لعدد المشاركين قائماً على نتائج تطبيقية تعتمد على النسبة بين عدد بنود المقياس وعدد المشاركين المقابل لكل بند.

١٢. أوصت بعض الدراسات باستخدام معامل أوميجا عندما يكون عدد المشاركين قليلاً مثل دراسة Trizano-Hermosilla and Alvarado (2016)، في حين تعارضت هذه النتيجة مع نتائج دراسة Javail et al. (2011) التي أشارت إلى أن قيمة ألفا كانت أفضل من أوميجا وثيتا باختلاف أعداد المشاركين، وهذا التناقض يدعم الحاجة إلى إجراء الدراسة الحالية.

١٣. يرى العديد من الباحثين أن استخدام معامل ألفا يؤدى إلى تقديرات متحيزة لمعامل ثبات درجات المقياس النفسى، نظراً لصرامة الافتراضات الخاص بالأسلوب الذى ينتمى إليه طبقاً لما تناولته أحدث الدراسات المنشورة في الدوريات العلمية المتخصصة في القياس النفسى والتربوى، وأوصوا باستخدام معامل أوميجا لأنه ينتمى إلى النموذج التقاربى والذى يتميز بمرونة افتراضاته ومن ثم دقة معامل أوميجا القائم عليه (Deng & Chan, 2016; Dunn et al., 2014; Revelle & Zinbarg, 2009; Zhang & Yuan, 2016; Zinbarg et al., 2005)، وكذلك الدراسات العربية (محمد بوزيان تيغزة، ٢٠٠٩).

١٤. استخدمت معظم الدراسات الأجنبية التي تم الاطلاع عليها مقاييس مُعدة ومقننة من قبل آخرين أو قواعد بيانات تم توليدها باستخدام فنية مونت كارلو للمحاكاة سواء أثناء دراسة تأثير عدد فئات الاستجابة أو عدد المشاركين على معاملات ثبات درجات المقياس النفسى (Lozano et al., 2008; Padilla & Divers, 2016; Romano et al., 2010; Trizano-Hermosilla & Alvarado, 2016; Weems & Onwuegbuzie, 2001; Yurdugul, 2008; Zumbo et al., 2007)

يتضح -مما سبق- أن الدراسة الحالية تختلف عن الدراسات السابقة في أنها تسعى لبحث تأثير عدد فئات الاستجابة وعدد المشاركين وحدود الثقة الخاصة واتساع المسافة بين حديها باستخدام قواعد ثابتة وذلك على دقة معاملى ألفا وأوميجا لتقدير ثبات درجات المقياس النفسى، حيث إن كل الدراسات السابقة – وخاصة في البيئة العربية – ركزت على معامل ألفا فقط بالرغم من عدم دقته نظراً لصرامة الافتراضات القائم عليها، كما أنها لم تحدد قواعد واضحة يسهل استخدامها لدى الباحثين لتحديد عدد المشاركين فى الاستجابة على المقياس النفسى فى الدراسات

الاستطلاعية، كما تقدم الدراسة الحالية كوداً لبرنامج RStudio لحساب قيم معاملى ألفا وأوميجا وحدود الثقة المرتبطة بهذه القيم، وتتميز الدراسة الحالية أيضاً بتقديم مقياس تم تعريبه لفاعلية الذات الوجدانية الذى يُفيد في دراسات تطبيقية، ومن هنا تُعد هذه الدراسة إضافة للأدبيات البحثية في مجال القياس النفسى والتربوى وخاصة في البيئة العربية.

فروض الدراسة

تسعى الدراسة الحالية لاختبار الفروض التالية والتي تم صياغتها بصورة تقريرية في ضوء مراجعة الإطار النظرى للدراسة وكذلك الدراسات السابقة التي تم توظيفها:

١. يوجد تأثير لعدد فئات الاستجابة (٣، ٥، ٧) وعدد المشاركين: ٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣) على دقة قيم معامل ألفا في تقدير ثبات درجات مقياس فاعلية الذات الوجدانية.

٢. يوجد تأثير لعدد فئات الاستجابة (٣، ٥، ٧) وعدد المشاركين: ٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣) على دقة قيم معامل أوميجا في تقدير ثبات درجات مقياس فاعلية الذات الوجدانية.

٣. يؤدى استخدام معامل أوميجا إلى تقدير أدق من معامل ألفا لثبات درجات مقياس فاعلية الذات الوجدانية باختلاف عدد فئات الاستجابة (٣، ٥، ٧) وعدد المشاركين: ٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣) فى ضوء نسبة التباين المفسر واتساع المسافة بين حدى الثقة.

إجراءات الدراسة:

أولاً: منهج الدراسة:

استخدمت الدراسة الحالية المنهج الوصفى المقارن، وذلك من خلال وصف تأثير عدد البدائل وعدد المشاركين على دقة معاملى ألفا وأوميجا ومقارنة دقة القيم الناتجة.

ثانياً: المشاركون في الدراسة:

تكّون عدد المشاركين في الدراسة من ٢٤٣ طالباً وطالبة من طلاب الفرقة الثالثة بكلية التربية بقنا للعام الجامعى ٢٠١٦/٢٠١٧م تتراوح أعمارهم بين ٢٠-٢٢ سنة بمتوسط عمرى قدره ٢٠.٢٢ سنة وانحراف معيارى قدره ٠.٤٤ سنة. ويوضح جدول (١) توزيع المشاركين طبقاً للنسبة بين كل بند من بنود مقياس فاعلية الذات الوجدانية وما يقابله من مشاركين:

جدول (١) توزيع المشاركين طبقاً للنسبة بين كل بند من بنود مقياس فاعلية الذات الوجدانية وما يقابله من مشاركين

النسبة بين كل بند وما يقابله من مشاركين	١: ٣	١:٥	١:٧	١:٩
عدد المشاركين	٨١	١٣٥	١٨٩	٢٤٣

ثالثاً: أداة الدراسة: مقياس فاعلية الذات الوجدانية

أعد هذا المقياس في الأصل Kirk, Schutte, and Hine (2008) وذلك لقياس فاعلية الذات لدى الأفراد فيما يرتبط بالنواحى الوجدانية لدى الأفراد، حيث تكّون المقياس في صورته المبدئية من ٣٢ بنداً مصاغة بشكل إيجابى وذات خمس فئات للاستجابة (لا تنطبق تماماً – لا تنطبق – متردد – تنطبق – تنطبق تماماً). تكون عدد المشاركين في تقنين المقياس من ٢٠٧ مشاركين. أظهرت نتائج التحليل العاملى للمقياس وجود عامل واحد والذى فسر٤٤% من التباين الكلى، ومن ثم يتميز المقياس بأحادية البعد وهى متطلب لاستخدام معاملى ألفا وأوميجا، وفيما يرتبط بثبات درجات المقياس فكانت قيمة معامل ألفا مساوية ٠.٩٦ وهى قيمة مرتفعة لمعامل ألفا. وقد أظهرت نتائج الدراسات العاملية على المقياس بعد ذلك أن هناك خمسة بنود لا تتشبع تشبعاً دالاً على المقياس، ومن ثم تكون المقياس في صورته النهائية من ٢٧ بنداً.

أما في سياق الدراسة الحالية فقد تم تعريب المقياس وعرضه على بعض المتخصصين في اللغة الإنجليزية للتأكد من سلامة الترجمة، كما قام أحدهم بترجمة النسخة التي تم تعريبها إلى الإنجليزية مرة أخرى وهو ما يُعرف Back-translation، حيث تطابقت ترجمته مع النسخة الأصلية للمقياس وهو ما يؤكد أيضاً دقة الترجمة.

بالإضافة إلى النسخة الأصلية للمقياس والمكونة من عدد خمس فئات للاستجابة فقد تم إعداد صورتين من المقياس تختلف فقط في عدد فئات الاستجابة كالتالى:

١. الصورة الثلاثية: (لا تنطبق – متردد – تنطبق)
٢. الصورة السباعية: (لا تنطبق تماماً – لا تنطبق – لا تنطبق أحياناً – متردد – تنطبق أحياناً – تنطبق – تنطبق تماماً).

يرجع استخدام هذا المقياس إلى حداثته بالإضافة إلى استخدامه في دراسات تطبيقية منشورة في دوريات علمية دولية محكمة وذات معامل تأثير مرتفع مثل دراسة Pool and Qualter (2012) وكذلك دراسة Alessandri, Vecchione, and Caprara (2015)، الأمر الذى يجعل تعريبه وحساب ثبات درجات ثلاث صور منه في الدراسة الحالية إضافة للمكتبة

العربية، كما أن هذا المقياس يمهد الطريق للباحثين لاستخدامه في دراسات مستقبلية في البيئة العربية.

رابعاً: إجراءات التطبيق والتصحيح وإدخال البيانات:

تضمنت هذه الإجراءات ما يلى:

١.توزيع المقياس على المشاركين وشرح التعليمات بصورة وافية حتى يتسنى لهم الاستجابة بشكل موضوعى يسهم فى إظهار الاستجابة الحقيقية لديهم. والجدير بالذكر فى هذا السياق أن كل المشاركين استجابوا على الصور الثلاث للمقياس وذلك يفيد فى أن التأكد من أن اختلاف قيم معاملى ألفا وأوميجا يرجع إلى اختلاف عدد فئات الاستجابة وليس اختلاف مجموعات التطبيق كما ورد فى التعقيب على الدراسات السابقة.

٢.تم تطبيق المقياس فى جلسة واحدة وذلك بفاصل زمنى قليل بين كل صورة بهدف الراحة، ويرجع اختيار جلسة واحدة للاستجابة على الصور الثلاث هو أن تكون الظروف واحدة لضمان موضوعية الاستجابة وعدم تأثرها بعوامل أخرى.

٣.تحرى الدقة أثناء التطبيق لتجنب أخطاء القياس التي تزيد من ارتباط الخطأ العشوائى الذى يؤثر بدوره سلباً على قيمة معامل ثبات الدرجات الناتج.

٤.تصحيح المقياس فى كل صورة كالتالى:

- الصورة الثلاثية: (١، ٢، ٣).
- الصورة الخماسية: (١، ٢، ٣، ٤، ٥).
- الصورة السباعية: (١، ٢، ٣، ٤، ٥، ٦، ٧). حيث كانت كل البنود موجبة فى كل صور المقياس.

٥.إدخال البيانات على الحاسب الآلى بدقة ومراجعتها حتى يتم الحصول على نتائج دقيقة أيضاً، حيث أن الخطأ فى إدخال البيانات يؤدى إلى قيم لا تعكس الأداء الحقيقى للمشاركين فى الدراسة. ويوضح جدول (٢) عدد الاستجابات التي تم إدخالها على الحاسب الآلى طبقاً لعدد المشاركين:

جدول (٢) عدد الاستجابات في ضوء عدد المشاركين

عدد بنود المقياس × عدد المشاركين	٨١×٢٧	١٣٥×٢٧	١٨٩×٢٧	٢٤٣×٢٧
عدد الاستجابات	٢١٨٧	٣٨٣٥	٥١٠٣	٦٥٦١

تبيّن من جدول (٢) أن عدد هذه الاستجابات يزداد بزيادة عدد المشاركين، الأمر الذى يحتاج إلى أسس علمية لتحديد هذه الأعداد حتى لا يضيع مجهود الباحثين وكذلك تتأثر قيم

المعاملات الناتجة من التحليل إذا كان هذا العدد أقل أو أكثر من المطلوب.

خامساً: المعالجة الإحصائية والحزم الحاسوبية:

تم استخدام برنامج Rstudio من إعداد Rstudio Team (2015) وهو أحد البرامج التابعة للحزمة الحاسوبية R Package وذلك لحساب قيم معاملى ألفا وأوميجا وكذلك حدود الثقة المرتبطة بها. ويمكن الحصول على نسخة مجاناً من هذا البرنامج من على الرابط http://www.rstudio.com/، ويحتوى ملحق (١) على الواجهة الافتتاحية لبرنامج Rstudio وبعض الأكواد المستخدمة فى حساب معاملى ألفا وأوميجا وحدود الثقة المرتبطة بهما. ويمكن حساب معاملى ألفا وأوميجا أيضاً من خلال الرابط https://websem.psychstat.org/apps/alpha

تم أيضاً استخدام برنامج SPSS Version 22 (IBM Corp., 2013)، وذلك لحساب قيم معاملات ارتباط بيرسون بين الصور المختلفة لمقياس فاعلية الذات الوجدانية وذلك لمقارنة قيم معامل ألفا باختلاف عدد فئات الاستجابة في ضوء مجموعات المشاركين وذلك بتوظيف المعادلة (٦) المذكورة في صفحة ١٨.

نتائج الدراسة:

١. نتائج الفرض الأول وتفسيرها:

ينص الفرض الأول على أنه "يوجد تأثير لعدد فئات الاستجابة (٣، ٥، ٧) وعدد المشاركين: ٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣) على دقة قيم معامل ألفا في تقدير ثبات درجات مقياس فاعلية الذات الوجدانية".

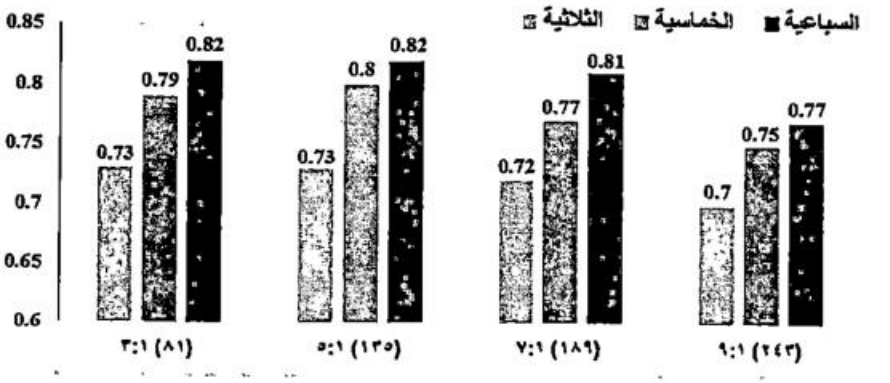
لاختبار صحة هذا الفرض تم حساب قيم معامل ألفا وحدود الثقة* واتساع المسافة المرتبطة بها لثلاث صور من مقياس فاعلية الذات الوجدانية طبقاً لعدد فئات الاستجابة (الثلاثى والخماسى والسباعى) وأربع مجموعات من المشاركين: ٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣). وتعرض الجداول (٣، ٤، ٥، ٦، ٧) والشكلين (٢، ٣) نتائج هذا الفرض:

* تم حساب حدود الثقة في الدراسة الحالية باستخدام طريقة Bias-corrected accelerated(BCa) Bootstrap لأنها تتسم بالدقة مقارنة بالطرق الأخرى كما ورد فى دراسة Xu (2014).

جدول (٣) تأثير عدد فئات الاستجابة وعدد المشاركين على دقة قيم معامل ألفا لثبات درجات مقياس فاعلية الذات الوجدانية وحدود الثقة واتساع المسافة المرتبطة بها

عدد فئات الاستجابة	عدد المشاركين	٣:١ (٨١)	٥:١ (١٣٥)	٧:١ (١٨٩)	٩:١ (٢٤٣)
الثلاثى	القيمة	٠.٧٣	٠.٧٣	٠.٧٢	٠.٧٠
	حدود الثقة	٠.٨١-٠.٦٢	٠.٨٠-٠.٦٥	٠.٧٨-٠.٦٦	٠.٧٥-٠.٦٣
	اتساع المسافة	٠.١٩	٠.١٥	٠.١٢	٠.١٢
الخماسى	القيمة	٠.٧٩	٠.٨٠	٠.٧٧	٠.٧٥
	حدود الثقة	٠.٨٥-٠.٧١	٠.٨٥-٠.٧٣	٠.٨٢-٠.٧١	٠.٨١-٠.٧٠
	اتساع المسافة	٠.١٤	٠.١٢	٠.١١	٠.١١
السباعى	القيمة	٠.٨٢	٠.٨٢	٠.٨١	٠.٧٧
	حدود الثقة	٠.٨٧-٠.٧٦	٠.٨٦-٠.٧٦	٠.٨٤-٠.٧٦	٠.٨١-٠.٧٢
	اتساع المسافة	٠.١١	٠.١٠	٠.٠٨	٠.٠٩

ملحوظة: جميع حدود الثقة المعروضة في نتائج هذا الجدول عند مستوى ٩٥%.



شكل (٢) تأثير عدد فئات الاستجابة وعدد المشاركين على دقة قيم معامل ألفا لثبات درجات مقياس فاعلية الذات الوجدانية

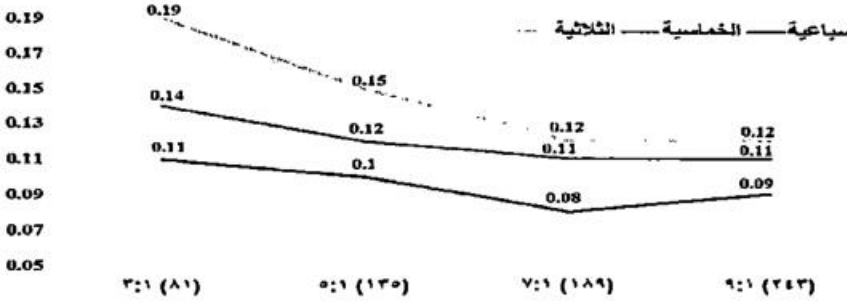
يتضح بصفة عامة من جدول (٣) وشكل (٢) أن قيم معامل ألفا تزداد في جميع مجموعات المشاركين بزيادة عدد فئات الاستجابة كالتالى:

- بلغت قيم معامل ألفا ٠.٧٣، ٠.٧٩، ٠.٨٢، في فى حالة التدريج الثلاثى والخماسى والسباعى على الترتيب، وذلك فى حالة كون عدد المشاركين ٣:١ (٨١)، ويتضح

وجود فرق كبير بين التدريجين الثلاثى والسباعى يصل إلى ٠.٠٩ أي ٩% من التباين الحقيقى.

- تشابهت قيم معامل ألفا عندما كان عدد المشاركين ٥:١ (١٣٥) مع نفس القيم عندما كان عدد المشاركين ٣:١ (٨١)، عدا فى التدريج الخماسى فكانت ٠.٨٠ بدلاً من ٠.٧٩. والجدير بالذكر في هذا السياق أنه حتى إذا تساوت قيم معامل ألفا باختلاف المجموعة الأولى والثانية من المشاركين فيُفضل النظر إلى حدود الثقة واتساع المسافة لمعرفة أيهما أدق، ولذلك تظل القيم المرتبطة بعدد المشاركين ٥:١ (١٣٥) هي الأدق لارتباطها بحدود ثقة (٠.٦٥-٠.٨٠) واتساع مسافة أضيق (٠.١٥) فى حالة التدريج الثلاثى لفئات الاستجابة، وحدود ثقة (٠.٧٦-٠.٨٦) واتساع مسافة (٠.١٠) فى حالة التدريج السباعى لفئات الاستجابة.
- بلغت قيم معامل ألفا ٠.٧٢، ٠.٧٧، ٠.٨١، في فى حالة التدريج الثلاثى والخماسى والسباعى على الترتيب، وذلك فى حالة كون عدد المشاركين ٧:١ (١٨٩)، ويتضح وجود فرق كبير بين التدريجين الثلاثى والسباعى يصل أيضاً إلى ٠.٠٩ أي ٩% من التباين الحقيقى.
- بلغت قيم معامل ألفا ٠.٧٠، ٠.٧٥، ٠.٧٧، في فى حالة التدريج الثلاثى والخماسى والسباعى على الترتيب، وذلك فى حالة كون عدد المشاركين ٩:١ (٢٤٣)، ويتضح وجود فرق كبير بين التدريجين الثلاثى والسباعى يصل إلى ٠.٠٧ أي ٧% من التباين الحقيقى.
- انخفضت قيم معامل ألفا عندما زاد عدد المشاركين عن ٥:١ (١٣٥)، بينما ضاقت المسافة بين حدى الثقة، ولعل ذلك ما تم التنويه إليه في مشكلة البحث أن زيادة عدد المشاركين عن الحد المطلوب قد يؤدى إلى ضياع جهد الباحث في تطبيق وتصحيح الاستجابات المرتبطة بهذا العدد وأيضاً انخفاض قيم معامل ثبات الدرجات.
- زيادة عدد فئات الاستجابة أفضل من زيادة عدد المشاركين، حيث بلغت قيمة معامل ألفا ٠.٨٢ في التدريج السباعى عندما كان عدد المشاركين ٨١ مشاركاً، بينما بلغت ٠.٧٠ في التدريج الثلاثى عندما كان عدد المشاركين ٢٤٣ مشاركاً، مع الأخذ في الاعتبار أيضاً أن دقة حدود الثقة ارتبطت بالتدريج السباعى لفئات الاستجابة.

وقد تم حساب اتساع المسافة بين حدى الثقة في هذا الفرض أيضاً لبيان تأثير عدد فئات الاستجابة وعدد المشاركين على دقة القيم المحسوبة لمعامل ألفا كما ورد فى جدول (٣) ويوضحه شكل (٣):



شكل (٣) تأثير عدد فئات الاستجابة وعدد المشاركين على اتساع المسافة بين حدى الثقة المرتبطة بدقة قيم معامل ألفا لثبات درجات مقياس فاعلية الذات الوجدانية

يتضح من شكل (٣) التالى:

- اتساع المسافة بين حدى الثقة كان أكبر ما يكون في التدريج الثلاثى لفئات الاستجابة، حيث بلغت القيم ٠.١٩، ٠.١٥، ٠.١٢، ٠.١٢، باختلاف مجموعات المشاركين: ٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣) على الترتيب، وهذا يؤكد وجود فروق دالة إحصائياً بين قيم معامل ألفا للصورة الثلاثية من ناحية والصورتين الخماسية والسباعية من ناحية أخرى والتى سترد فى الجداول من (٤-٧).
- اتساع المسافة بين حدى الثقة كان أقل ما يكون في التدريج السباعى لفئات الاستجابة، حيث بلغت القيم ٠.١١، ٠.١٠، ٠.٠٨، ٠.٠٩، باختلاف مجموعات المشاركين: ٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣) على الترتيب، ويُعد ذلك مؤشراً على أن عدد فئات الاستجابة يؤثر على دقة قيم معامل ألفا.
- أن العدد ٣:١ (٨١) كان أكثر مجموعات المشاركين اتساعاً للمسافة، حيث بلغت القيم ٠.١٩، ٠.١٤، ٠.١١ باختلاف فئات الاستجابة الثلاثية والخماسية والسباعية على الترتيب، ومعنى ذلك أن هذا العدد هو الأقل دقة.
- بلغ اتساع المسافة بين حدى الثقة ٠.٠٨ و٠.٠١١ فى التدريجين الخماسى والسباعى لفئات الاستجابة على الترتيب وذلك عندما كان عدد المشاركين ٧:١ (١٨٩)، الأمر

الذى يؤكد وجود فروق دالة إحصائياً بين التدريجين الخماسى والسباعى كما سيرد فى جدول (٦)، مع ملاحظة أن هذه المجموعة من المشاركين هى الوحيدة التى وُجدت فيها فروق دالة إحصائياً بين التدريجين الخماسى والسباعى.

ومجمل نتائج هذا الفرض هي: أنه تزداد دقة قيم معامل ألفا بزيادة عدد فئات الاستجابة، وأن التدريج الخماسى والسباعى لفئات الاستجابة هما الأفضل، كما أن العدد ١:٥ (١٣٥) للمشاركين هو الأفضل لارتفاع قيم معامل ألفا لكنه ليس الأفضل فيما يتعلق باتساع المسافة بين حدى الثقة.

وقد تم أيضاً حساب الفروق بين كل صورتين من صور مقياس فاعلية الذات الوجدانية في ضوء الأربع مجموعات من المشاركين باستخدام معادلة (٦)، وذلك لبيان عما إذا كانت قيم معامل ألفا تختلف اختلافاً دالاً إحصائياً باختلاف عدد فئات الاستجابة، وتعرض جداول (٤-٧) نتيجة هذه الفروق

جدول (٤)

الفروق بين معاملى ألفا لكل صورتين من صور مقياس فاعلية الذات الوجدانية في ضوء عدد المشاركين ١:٣ (٨١)

الإحصاءات المستخدمة / الصورتين موضوع المقارنة	معامل الارتباط (ر)	ت	درجات الحرية	مستوى الدلالة
الثلاثية (٠.٧٣) الخماسية (٠.٧٩)	٠.٩٠	٢.٥٧	٧٩	٠.٠٥
الثلاثية (٠.٧٣) السباعية (٠.٨٢)	٠.٨٢	٣.١٧	٧٩	٠.٠١
الخماسية (٠.٧٩) السباعية (٠.٨٢)	٠.٨٩	١.٥٠	٧٩	غير دال

ملحوظة: قيمة "ت" الجدولية عند درجات حرية ٧٩ = ١.٩٩ عند مستوى ٠.٠٥، و٢.٦٤ عند مستوى ٠.٠١

جدول (٥)

الفروق بين معاملى ألفا لكل صورتين من صور مقياس فاعلية الذات الوجدانية في ضوء عدد المشاركين ١:٥ (١٣٥)

الإحصاءات المستخدمة / الصورتين موضوع المقارنة	معامل الارتباط (ر)	ت	درجات الحرية	مستوى الدلالة
الثلاثية (٠.٧٣) الخماسية (٠.٨٠)	٠.٨٨	٣.٦٦	١٣٣	٠.٠١
الثلاثية (٠.٧٣) السباعية (٠.٨٢)	٠.٨٤	٤.٣٤	١٣٣	٠.٠١
الخماسية (٠.٨٠) السباعية (٠.٨٢)	٠.٩١	١.٤٧	١٣٣	غير دال

ملحوظة: قيمة "ت" الجدولية عند درجات حرية ١٣٣ = ١.٩٦ عند مستوى ٠.٠٥، و٢.٥٨ عند مستوى ٠.٠١

جدول (٦)

الفروق بين معاملى ألفا لكل صورتين من صور مقياس فاعلية الذات الوجدانية في ضوء عدد المشاركين ١:٧ (١٨٩)

الإحصاءات المستخدمة / الصورتين موضوع المقارنة	معامل الارتباط (ر)	ت	درجات الحرية	مستوى الدلالة
الثلاثية (٠.٧٢) الخماسية (٠.٧٧)	٠.٩٠	٣.١٠	١٨٧	٠.٠١
الثلاثية (٠.٧٢) السباعية (٠.٨١)	٠.٨٥	٥.١٠	١٨٧	٠.٠١
الخماسية (٠.٧٧) السباعية (٠.٨١)	٠.٩٠	٣.٠٠	١٨٧	٠.٠١

ملحوظة: قيمة "ت" الجدولية عند درجات حرية ١٨٧ = ١.٩٦ عند مستوى ٠.٠٥، و٢.٥٨ عند مستوى ٠.٠١

جدول (٧)

الفروق بين معاملى ألفا لكل صورتين من صور مقياس فاعلية الذات الوجدانية في ضوء عدد المشاركين ٩:١ (٢٤٣)

الإحصاءات المستخدمة / الصورتين موضوع المقارنة	معامل الارتباط (ر)	ت	درجات الحرية	مستوى الدلالة
الثلاثية (٠.٧٠) الخماسية (٠.٧٥)	٠.٨٨	٢.٩٨	٢٤١	٠.٠١
الثلاثية (٠.٧٠) السباعية (٠.٧٧)	٠.٨٤	٣.٨١	٢٤١	٠.٠١
الخماسية (٠.٧٥) السباعية (٠.٧٧)	٠.٨٨	١.٣٦	٢٤١	غير دال

ملحوظة: قيمة "ت" الجدولية عند درجات حرية ٢٤١ = ١.٩٦ عند مستوى ٠.٠٥، و٢.٥٨ عند مستوى ٠.٠١

يتضح من جداول (٤-٧) أنه توجد فروق دالة إحصائياً عند مستوى ٠.٠١ بين الصورة الثلاثية من ناحية والصورتين الخماسية والسباعية من ناحية أخرى لصالح الصورتين الخماسية والسباعية باختلاف مجموعات المشاركين عدا المجموعة الأولى ٣:١ (٨١) فكان مستوى الدلالة ٠.٠٥ بين الثلاثية والخماسية، كما تبين أيضاً من نفس الجداول عدم وجود فروق دالة إحصائياً بين الصورتين الخماسية والسباعية باختلاف مجموعات المشاركين عدا المجموعة الثالثة ٧:١ (١٨٩) فكانت الفروق دالة عند مستوى ٠.٠١ لصالح الصورة السباعية.

ويمكن تفسير نتائج هذا الفرض - فيما يتعلق بزيادة دقة قيم معامل ألفا لثبات الدرجات بزيادة عدد فئات الاستجابة- أنه كلما زاد عدد فئات الاستجابة احتوى على بدائل من شأنها مساعدة المشاركين على إظهار الاستجابة الموضوعية لديهم، الأمر الذى يزيد من مقدار التباين الحقيقي الذى يؤثر بدوره على دقة قيم معامل ألفا، ويتضح ذلك من أن بدائل الاستجابة في التدريج الثلاثى اقتصرت على ثلاثة اختيارات تمثلت في (لا تنطبق - متردد - تنطبق) وهى لا تعطى الفرصة للمستجيب التي يعبر بها عن رأيه في البند موضوع الاستجابة بموضوعية وبدقة، ومن هنا يأتى التأثير السالب على مقدار التباين الحقيقى الذى يؤثر بدوره على دقة قيم معامل ألفا، أما في حالة التدريجين الخماسى والسباعى لفئات الاستجابة فهما يحتويان على خمسة اختيارات (لا تنطبق تماماً - لا تنطبق - متردد - تنطبق - تنطبق تماماً) وسبعة اختيارات (لا تنطبق تماماً - لا تنطبق - لا تنطبق أحياناً - متردد - تنطبق أحياناً - تنطبق - تنطبق تماماً) واللذان يتيحان

الفرصة للمستجيب للتعبير عن استجاباته بصدق وموضوعية مما يزيد التباين الحقيقى للدرجات، الأمر الذى يؤثر بدوره على ارتفاع قيم معامل ألفا وبالتالي دقتها.

تُعزى نتائج هذا الفرض أيضاً فيما يتعلق بتأثير عدد المشاركين على دقة معامل ألفا إلى أن قيمة معامل ثبات الدرجات تتأثر باختلاف قدرات المشاركين، وكذلك الارتباطات البينية بين البنود المكونة للمقياس وليس العدد، وذلك يفسر عدم زيادة قيم معامل ألفا عند زيادة العدد عن ٥:١ (١٣٥) بسبب تجانس القدرات بين الطلاب المشاركين في الاستجابة على مقياس فاعلية الذات الوجدانية المستخدم في الدراسة الحالية لأنهم من طلاب الفرقة الثالثة بكلية التربية كما ورد فى وصف المشاركين فى الدراسة.

يمكن تفسير دقة حدود الثقة فى حالة التدريج السباعى لفئات الاستجابة أنه عندما تُتاح الفرصة أمام المشاركين للاختيار بين فئات متعددة للاستجابة فإن ذلك يجعل استجاباتهم موضوعية مما يجعلها أكثر تمثيلاً للمجتمع، وهكذا الأمر بالنسبة لضيق اتساع مسافة حدود الثقة عند زيادة عدد المشاركين.

وتتفق نتائج هذا الفرض مع نتائج بعض البحوث والدراسات التي أشارت إلى أن عدد فئات الاستجابة يؤثر على ارتفاع قيم معامل ألفا (Chomeya, 2010; Choudhury & Bhattacharjee, 2014; Preston & Colman, 2000; Weng, 2004; Zumbo et al., 2007)، وكذلك نتائج بعض الدراسات العربية (محسوب عبدالقادر الضوى، ٢٠١١؛ محمد حسين سعيد، ٢٠٠٧؛ محمد عبدالرحمن إسماعيل، ٢٠١٥).

تتفق أيضاً نتائج هذا الفرض مع ما توصلت إليه بعض الأدبيات البحثية فيما يتعلق بزيادة دقة الخصائص السيكومترية للمقياس النفسى عندما يكون عدد فئات الاستجابة من ٤-٧ فئة (Lozano et al, 2008; Sullivan & Artino 2013; Wakita et al., 2012)

اتفقت جزئياً نتائج هذا الفرض من أن الزيادة فى قيمة معامل ألفا لا ترتبط بعدد المشاركين Ercan et al. (2007)، كما اتفقت أيضاً فيما يرتبط بحدود الثقة مع نتائج دراسة Romano et al. (2010) التى أشارت إلى أن حدود الثقة كانت أضيق عند ارتفاع قيمة معامل ألفا وكذلك زيادة عدد المشاركين.

وتختلف نتائج هذا الفرض مع نتائج دراسة Shaftel et al. (2012) وكذلك نتائج دراسة حجاج غانم أحمد على وياسر عبدالله حفنى حسن (٢٠١١) التى توصلت إلى أن قيمة معامل ألفا فى حالة التدريج الثلاثى كانت أفضل منه التدريج الخماسى لمقياس مفهوم الذات.

تختلف جزئياً نتائج هذا الفرض أيضاً مع نتائج بعض الدراسات التى أشارت إلى عدم وجود فروق دالة إحصائياً فى معامل ألفا راجعة إلى الاختلاف بين عدد فئات الاستجابة (Choudhury & Bhattacharjee, 2014; Lee & Paek, 2014)، وخاصة أن الدراسة الحالية توصلت إلى وجود فروق دالة إحصائياً بين التدريج الثلاثى من ناحية والخماسى والسباعة من ناحية أخرى.

تختلف أيضاً نتائج هذا الفرض فيما يتعلق بعدد المشاركين مع ما توصلت إليه دراسة Miller (1995) من أن دقة معامل ألفا تعتمد على عدد المشاركين.

٢. نتائج الفرض الثانى وتفسيرها:

ينص الفرض الثانى على أنه "يوجد تأثير لعدد فئات الاستجابة (٣، ٥، ٧) وعدد المشاركين: ٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣) على دقة قيم معامل أوميجا في تقدير ثبات درجات مقياس فاعلية الذات الوجدانية".

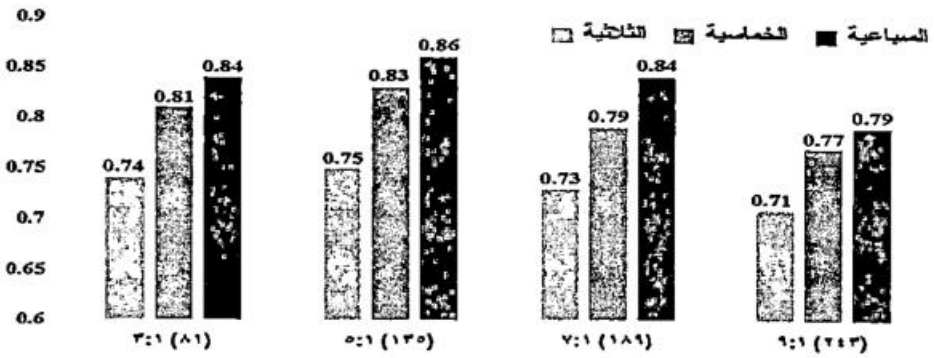
ولاختبار صحة هذا الفرض تم اتباع نفس الخطوات التي تم اتباعها في اختبار صحة الفرض الأول ولكن هذه المرة على معامل أوميجا. ويعرض جدول (٨) وشكل (٤) نتائج هذا الفرض:

جدول (٨)

تأثير عدد فئات الاستجابة وعدد المشاركين على دقة قيم معامل أوميجا لثبات درجات مقياس فاعلية الذات الوجدانية وحدود الثقة واتساع المسافة المرتبطة بها

عدد فئات الاستجابة / عدد المشاركين		٣:١ (٨١)	٥:١ (١٣٥)	٧:١ (١٨٩)	٩:١ (٢٤٣)
الثلاثى	القيمة	٠.٧٤	٠.٧٥	٠.٧٣	٠.٧١
	حدود الثقة	٠.٨٠-٠.٦٤	٠.٨٠-٠.٦٨	٠.٧٩-٠.٦٨	٠.٧٥-٠.٦٥
	اتساع المسافة	٠.١٦	٠.١٢	٠.١١	٠.١٠
الخماسى	القيمة	٠.٨١	٠.٨٣	٠.٧٩	٠.٧٧
	حدود الثقة	٠.٨٦-٠.٧٤	٠.٨٩-٠.٧٨	٠.٨٤-٠.٧٤	٠.٨٢-٠.٧٣
	اتساع المسافة	٠.١٢	٠.١١	٠.١٠	٠.٠٩
السباعى	القيمة	٠.٨٤	٠.٨٦	٠.٨٤	٠.٧٩
	حدود الثقة	٠.٨٩-٠.٧٩	٠.٩٠-٠.٨١	٠.٨٨-٠.٨٠	٠.٨٣-٠.٧٥
	اتساع المسافة	٠.١٠	٠.٠٩	٠.٠٨	٠.٠٨

ملحوظة: جميع حدود الثقة المعروضة في نتائج هذا الجدول عند مستوى ٩٥%.



شكل (٤) تأثير عدد فئات الاستجابة وعدد المشاركين على دقة قيم معامل أوميجا لثبات درجات مقياس فاعلية الذات الوجدانية

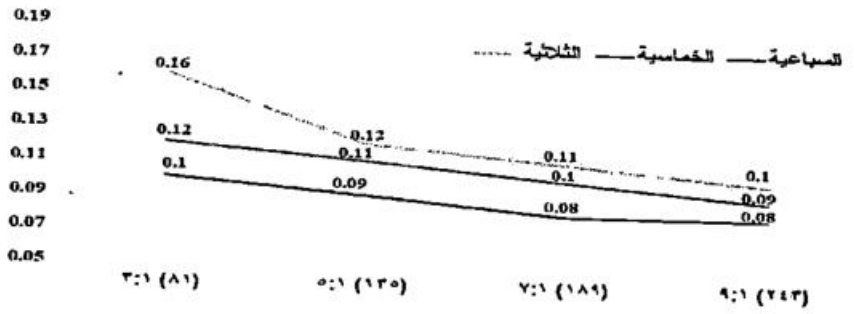
يتضح بصفة عامة من جدول (٨) وشكل (٤) أن قيم معامل أوميجا تزداد بزيادة عدد فئات الاستجابة في جميع مجموعات المشاركين كالتالي:

- بلغت قيم معامل أوميجا ٠.٧٤، ٠.٨١، ٠.٨٤، في فى حالة التدريج الثلاثى والخماسى والسباعى على الترتيب، وذلك فى حالة كون عدد المشاركين ٣:١ (٨١)، ويتضح وجود فرق كبير بين التدريجين الثلاثى والسباعى يصل إلى ٠.١٠ أي ١٠% من التباين الحقيقى.
- بلغت قيم معامل أوميجا ٠.٧٥، ٠.٨٣، ٠.٨٦، في فى حالة التدريج الثلاثى والخماسى والسباعى على الترتيب، وذلك فى حالة كون عدد المشاركين ٥:١ (١٣٥)، ويتضح وجود فرق كبير بين التدريجين الثلاثى والسباعى يصل إلى ٠.١١ أي ١١% من التباين الحقيقى.
- بلغت قيم معامل أوميجا ٠.٧٣، ٠.٧٩، ٠.٨٤، في فى حالة التدريج الثلاثى والخماسى والسباعى على الترتيب، وذلك فى حالة كون عدد المشاركين ٧:١ (١٨٩)، ويتضح وجود فرق كبير بين التدريجين الثلاثى والسباعى يصل إلى ٠.١١ أي ١١% من التباين الحقيقى.
- بلغت قيم معامل أوميجا ٠.٧١، ٠.٧٧، ٠.٧٩، في فى حالة التدريج الثلاثى والخماسى والسباعى على الترتيب، وذلك فى حالة كون عدد المشاركين ٩:١ (٢٤٣)، ويتضح

وجود فرق كبير بين التدريجين الثلاثى والسباعى يصل إلى ٠.٠٨ أي ٨% من التباين الحقيقى.

- يتضح أيضاً أن قيم معامل أوميجا انخفضت عندما زاد عدد المشاركين عن ٥:١ (١٣٥)، بينما ضاقت المسافة بين حدى الثقة.
- زيادة عدد فئات الاستجابة أفضل من زيادة عدد المشاركين، حيث بلغت قيمة معامل أوميجا ٠.٨٤ في التدريج السباعى عندما كان عدد المشاركين ٨١ مشاركاً، بينما بلغت ٠.٧١ في التدريج الثلاثى عندما كان عدد المشاركين ٢٤٣ مشاركاً، مع الأخذ في الاعتبار أيضاً أن دقة حدود الثقة تساوت في التدريجين.

تم حساب اتساع المسافة بين حدى الثقة في هذا الفرض أيضاً لبيان تأثير عدد فئات الاستجابة وعدد المشاركين على دقة القيم المحسوبة لمعامل أوميجا كما يتضح من شكل (٥):



شكل (٥) تأثير عدد فئات الاستجابة وعدد المشاركين على اتساع المسافة بين حدى الثقة المرتبطة بدقة قيم معامل أوميجا لثبات درجات مقياس فاعلية الذات الوجدانية

يتضح من شكل (٥) التالى:

- اتساع المسافة بين حدى الثقة كان أكبر ما يكون في التدريج الثلاثى لفئات الاستجابة، حيث بلغت القيم ٠.١٦، ٠.١٢، ٠.١١، ٠.١٠، باختلاف مجموعات المشاركين: ٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣) على الترتيب.
- اتساع المسافة بين حدى الثقة كان أقل ما يكون في التدريج السباعى لفئات الاستجابة، حيث بلغت القيم ٠.١٠، ٠.٠٩، ٠.٠٨، ٠.٠٨، باختلاف مجموعات المشاركين: ٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣) على الترتيب، ويُعد ذلك مؤشراً على أن عدد فئات الاستجابة يؤثر على دقة قيم معامل ألفا.
- أن العدد ٣:١ (٨١) كان أكثر مجموعات المشاركين اتساعاً للمسافة، حيث بلغت القيم

٠.١٦، ٠.١٢، ٠.١٠ باختلاف فئات الاستجابة الثلاثية والخماسية والسباعية على الترتيب، ومعنى ذلك أن هذا العدد هو الأقل دقة.

- أن العدد ٩:١ (٢٤٣) كان أقل مجموعات المشاركين اتساعاً للمسافة، حيث بلغت القيم ٠.١٠، ٠.٠٩، ٠.٠٨ باختلاف فئات الاستجابة الثلاثية والخماسية والسباعية على الترتيب، مع ملاحظة انخفاض قيم معامل ثبات الدرجات المرتبطة به.

مجمل نتائج هذا الفرض هى أنه تزداد دقة قيم معامل أوميجا بزيادة عدد فئات الاستجابة، وأن التدريج الخماسى والسباعى لفئات الاستجابة هما الأفضل، كما أن العدد ٥:١ (١٣٥) للمشاركين هو الأفضل لارتفاع قيم معامل أوميجا لكنه ليس الأفضل فيما يتعلق باتساع المسافة بين حدى الثقة، كما أن تأثير عدد فئات الاستجابة كان أكبر على قيم معامل أوميجا مقارنة بمعامل ألفا.

يُلاحظ أنه لم يتم فى حالة معامل أوميجا حساب الفروق بين صور مقياس فاعلية الذات الوجدانية كما تم فى حالة معامل ألفا، ولعل ذلك يرجع إلى أنه بمراجعة الأدبيات البحثية لم يتبين أن هناك دراسات تناولت تأثير عدد فئات الاستجابة على معامل أوميجا، ومن ثم لم توجد معادلات رياضية يمكن استخدامها للمقارنة بين قيمتين لمعامل أوميجا مثل معادلة (٦) التى تم استخدامها فى حالة معامل ألفا.

يمكن تفسير نتائج هذا الفرض – فيما يتعلق بزيادة قيم معامل أوميجا بزيادة عدد فئات الاستجابة- أنه كلما زاد عدد فئات الاستجابة احتوى على بدائل والتى من شأنها إتاحة الفرصة أمام المشاركين على إظهار الاستجابة الموضوعية لديهم، الأمر الذى يزيد من مقدار التباين الحقيقي الذى يؤثر بدوره على دقة قيم معامل أوميجا، وخاصة أن تقدير هذا المعامل يعتمد على التشبعات الناتجة من مقدار التباين الحقيقى لاستجابات المشاركين على المقياس النفسى، ويتضح أن تفسير نتائج هذا الفرض تتشابه مع تفسير نتائج الفرض الأول وذلك لارتباط معاملات ثبات الدرجات بالتباين الحقيقى الذى يظهر من خلال عدد فئات الاستجابة، حيث إنه في التدريج الثلاثى اقتصرت على ثلاثة اختيارات تمثلت في (لا تنطبق – متردد – تنطبق) وهى لا تعطى الفرصة للمستجيب للتعبير عن رأيه في البند موضوع الاستجابة بموضوعية وبدقة، ومن هنا التأثير السالب على مقدار التباين الحقيقى الذى يؤثر بدوره على التشبعات، ومن ثم على دقة قيم معامل أوميجا، أما في حالة التدريجين الخماسى والسباعى لفئات الاستجابة فهما يحتويان على خمسة اختيارات (لا تنطبق تماماً – لا تنطبق – متردد – تنطبق – تنطبق تماماً) وسبعة اختيارات

(لا تنطبق تماماً – لا تنطبق – لا تنطبق أحياناً – متردد – تنطبق أحياناً – تنطبق – تنطبق تماماً) واللذان يتيحان الفرصة للمستجيب للتعبير عن استجاباته بصدق وموضوعية مما يزيد التباين الحقيقي للدرجات وتشبعات البنود وقلة تباين الخطأ، الأمر الذى يؤثر بدوره على ارتفاع قيم معامل أوميجا وبالتالي يؤثر على دقتها.

تُعزى نتائج هذا الفرض أيضاً فيما يتعلق بتأثير عدد المشاركين على دقة معامل أوميجا إلى أن قيمة معامل ثبات الدرجات تتأثر باختلاف قدرات المشاركين وكذلك الارتباطات البينية بين البنود المكونة للمقياس وليس العدد كما تم ذكره مع معامل ألفا، وذلك يفسر عدم زيادة قيم معامل أوميجا عند زيادة العدد عن ٥:١ (١٣٥) بسبب تجانس القدرات بين الطلاب المشاركين في الاستجابة على مقياس فاعلية الذات الوجدانية المستخدم في الدراسة الحالية لأنهم من طلاب الفرقة الثالثة بكلية التربية كما ورد فى وصف المشاركين فى الدراسة، مع إضافة أن معامل أوميجا يعتمد على التشبعات التى تتطلب أعداداً محددة للمشاركين منها ٥:١ والذى يُعد أكثرها شيوعاً.

يرجع اتساع المسافة بين حدى الثقة عندما كان عدد المشاركين ٣:١ (٨١) إلى أن الطريقة المستخدمة فى حساب حدود الثقة هي Bias-corrected accelerated (Bca) bootstrap تتطلب عدد ١٠٠ على الأقل حتى تنال الثقة فيها، وخاصة مع معامل أوميجا.

تتفق نتائج هذا الفرض فيما يتعلق بتأثير عدد فئات الاستجابة مع ما توصلت إليه بعض الأدبيات البحثية فيما يتعلق بزيادة دقة الخصائص السيكومترية للمقياس النفسى والمتمثلة فى معامل أوميجا فى متن الدراسة الحالية عندما يكون عدد فئات الاستجابة من ٤–٧ فئة (Lozano et al, 2008; Sullivan & Artino 2013; Wakita et al., 2012)، حيث اتضح ارتفاع قيم معامل أوميجا فى التدريج الخماسى والسباعى عنها فى التدريج الثلاثى لفئات الاستجابة باختلاف مجموعات المشاركين.

وتتفق نتائج هذا الفرض أيضاً فيما يتعلق بعدد المشاركين مع نتائج دراسة Yurdugul (2008) التى توصلت إلى أن هناك عوامل أخرى مؤثرة على قيمة معامل ثبات الدرجات وليس على عدد المشاركين فقط. كما اتفقت النتائج أيضاً مع نتائج دراسة Javail et al. (2011) التى أشارت إلى أن دقة معامل أوميجا لا ترتبط بالزيادة فى عدد المشاركين

وتتفق نتائج هذا الفرض أيضاً فيما يتعلق بالدقة فى حدود الثقة المرتبطة بقيم معامل أوميجا مع نتائج دراسة Raykov and Penev (2006) وكذلك دراسة Terry and Kelley (2012) اللتان أشارتا إلى أن عدد المشاركين يؤثر على حدود الثقة. كما تتفق النتائج مع ما

توصلت إليه دراسة Padilla and Divers (2016) من أن معامل أوميجا يتطلب عدد ١٠٠ مشاركاً أو أكثر لحساب حدود الثقة الخاصة.

تختلف نتائج هذا الفرض فيما يتعلق بعدد المشاركين مع ما توصلت إليه دراسة Ercan et al. (2007) التى أشارت إلى أن قيمة معامل أوميجا تزداد بزيادة عدد المشاركين. ويُلاحظ من عرض ما اتفقت أو اختلفت مع نتائج معامل أوميجا قلة الدراسات مقارنة بألفا ويدل ذلك على ندرة الدراسات التى تناولت معامل اوميجا مقارنة بمعامل ألفا.

٣. نتائج الفرض الثالث وتفسيرها:

ينص الفرض الثالث على أنه "يؤدى استخدام معامل أوميجا إلى تقدير أدق من معامل ألفا لثبات درجات مقياس فاعلية الذات الوجدانية باختلاف عدد فئات الاستجابة (٣، ٥، ٧) وعدد المشاركين: ٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣) فى ضوء نسبة التباين المفسر واتساع المسافة بين حدى الثقة".

ولاختبار صحة هذا الفرض تم استخلاص النتائج التي تم الحصول عليها في الفرضين الأول والثانى الخاصة بقيم معاملى ألفا وأوميجا وطرح قيم معامل ألفا من قيم معامل أوميجا وذلك للحصول على مقدار الفرق وتحويله إلى نسبة مئوية للحصول على نسبة التباين المفسر في ضوء تفسير قيمة معامل ثبات الدرجات بأنه نسبة التباين الحقيقى بالنسبة للتباين الكلى للدرجات، كما تم أيضاً حساب الفروق بين اتساع المسافة بين حدى الثقة. ويعرض جدولى (٩، ١٠) نتائج هذا الفرض:

جدول (٩)

الفروق بين دقة قيم معاملى ألفا وأوميجا باختلاف عدد فئات الاستجابة وعدد المشاركين فى ضوء في نسبة التباين المفسر

عدد فئات الاستجابة / عدد المشاركين		٣:١ (٨١)	٥:١ (١٣٥)	٧:١ (١٨٩)	٩:١ (٢٤٣)
الثلاثى	أوميجا	٠.٧٤	٠.٧٥	٠.٧٣	٠.٧٩
	ألفا	٠.٧٣	٠.٧٣	٠.٧٢	٠.٧٠
	الفروق في نسبة التباين المفسر	%١	%٢	%١	%١
الخماسى	أوميجا	٠.٨١	٠.٨٣	٠.٧٩	٠.٧٧
	ألفا	٠.٧٩	٠.٨٠	٠.٧٧	٠.٧٥
	الفروق في نسبة التباين المفسر	%٢	%٣	%٢	%٢
السباعى	أوميجا	٠.٨٤	٠.٨٦	٠.٨٤	٠.٧٩
	ألفا	٠.٨٢	٠.٨٢	٠.٨١	٠.٧٧
	الفروق في نسبة التباين المفسر	%٢	%٤	%٣	%٢

يتضح من جدول (٩) التالى:

- أن قيم معامل أوميجا أكبر من قيم معامل ألفا باختلاف عدد فئات الاستجابة وعدد المشاركين، مع ملاحظة أن أقل فرق في نسبة التباين المفسر كان في حالة العدد الثلاثى لفئات الاستجابة، حيث كانت النسبة ١%، عدا في حالة كون عدد المشاركين ٥:١ (١٣٥) فكانت النسبة ٢%.
- بلغت نسبة التباين المفسر ٣% و٤% في حالة العدد الخماسى والسباعى لفئات الاستجابة على الترتيب، وذلك أيضاً في حالة عدد المشاركين ٥:١ (١٣٥).
- بلغت نسبة التباين المفسر ٣% في حالة التدريج السباعى لفئات الاستجابة في حالة كون عدد المشاركين ٧:١ (١٨٩).
- بلغت نسبة التباين المفسر ٢% مع بقية عدد فئات الاستجابة وبقية مجموعات المشاركين.

جدول (١٠)

الفروق بين دقة قيم معاملى ألفا وأوميجا باختلاف عدد فئات الاستجابة وعدد المشاركين فى ضوء في اتساع المسافة بين حدى الثقة

عدد فئات الاستجابة / عدد المشاركين		٣:١ (٨١)	٥:١ (١٣٥)	٧:١ (١٨٩)	٩:١ (٢٤٣)
الثلاثى	ألفا	٠.١٩	٠.١٥	٠.١٢	٠.١٢
	أوميجا	٠.١٦	٠.١٢	٠.١١	٠.١٠
	الفروق في اتساع المسافة بين حدى الثقة	٠.٠٣	٠.٠٣	٠.٠١	٠.٠٢
الخماسى	ألفا	٠.١٤	٠.١٢	٠.١١	٠.١١
	أوميجا	٠.١٢	٠.١١	٠.١٠	٠.٠٩
	الفروق في اتساع المسافة بين حدى الثقة	٠.٠٢	٠.٠١	٠.٠١	٠.٠٢
السباعى	ألفا	٠.١١	٠.١٠	٠.٠٨	٠.٠٩
	أوميجا	٠.١٠	٠.٠٩	٠.٠٨	٠.٠٨
	الفروق في اتساع المسافة بين حدى الثقة	٠.٠١	٠.٠١	٠.٠٠	٠.٠١

يتضح من جدول (١٠) أن هناك اتساع المسافة بين حدى الثقة الخاصة بقيم معامل ألفا أكبر من تلك الخاصة بدقة قيم معامل أوميجا مما يعطى مؤشراً كمياً آخر على دقة معمل أوميجا وتمثيله لقيمة المجتمع بشكل أفضل مقارنة بمعامل ألفا كالتالى:

- كان مقدار الفرق فى اتساع المسافة بين حدى الثقة المرتبطة بقيم معاملى ألفا وأوميجا أكبر فى حالة التدريج الثلاثى لفئات الاستجابة ، حيث بلغ هو ٠.٠٣، ٠.٠٣، ٠.٠١، ٠.٠٢، باختلاف مجموعات المشاركين: ٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣) على الترتيب.
- كان مقدار الفرق فى اتساع المسافة بين حدى الثقة المرتبطة بقيم معاملى ألفا وأوميجا أقل فى حالة التدريج السباعى لفئات الاستجابة، حيث بلغ هو ٠.٠١، ٠.٠١، ٠.٠٠، ٠.٠١، باختلاف مجموعات المشاركين: ٣:١ (٨١)، ٥:١ (١٣٥)، ٧:١ (١٨٩)، ٩:١ (٢٤٣) على الترتيب.

مجمل نتائج هذا الفرض أن معامل أوميجا يعطى تقديرات أدق لمعاملات ثبات درجات مقياس فاعلية الذات الوجدانية باختلاف عدد فئات الاستجابة وعدد المشاركين فى ضوء نسبة التباين المفسر واتساع المسافة بين حدى الثقة.

يمكن تفسير نتائج هذه الفرض في ضوء ما أوصت به الأدبيات البحثية فى هذا

السياق بأن معامل ألفا يؤدى إلى انخفاض تقدير قيمة معامل ثبات الدرجات وذلك لصعوبة استيفاء البيانات لافتراضات نموذج تاو المتكافئ بالأساس القائم عليه معامل ألفا، وخاصة صعوبة تساوى القيم الحقيقية للدرجات أو تشبعات البنود بمقدار متساو على العامل الواحد.

تُعزى نتائج هذا الفرض الخاصة بارتفاع قيم معامل أوميجا وأنه يؤدى إلى تقديرات أدق لمعامل ثبات درجات المقياس النفسي، نظراً لمرونة افتراضات النموذج النوعى القائم عليه والتى تجعل تقدير هذه القيم يتم بشكل أدق، كما أن ارتفاع نسبة التباين المفسر في حالة العدد الخماسى والسباعى لفئات الاستجابة ينتج من أنه كلما زاد عدد فئات الاستجابة زاد تباين الدرجات، وبالتالي قيمة معامل ثباتها، ويتسق ذلك مع تعريف معامل أوميجا والوارد فى المعادلة (٧) هو أنه التباين الحقيقى مقسوماً على التباين الحقيقى مضافاً إليه مقدار الخطأ، كما يرجع قلة مقدار الفرق بين حدى الثقة الخاصة بقيم معامل أوميجا مقارنة بتلك الخاصة بمعامل ألفا إلى أن قيم معامل أوميجا تُعد تمثيلاً حقيقياً للمجتمع، ومن ثم ارتباطها بحدود ثقة أضيق.

خلاصة القول فى تفسير نتائج هذا الفرض أنه بصرف النظر عن الفروق فى نسبة التباين المفسر فإن معامل أوميجا يؤدى إلى تقدير أدق لمعامل ثبات درجات مقياس فاعلية الذات الوجدانية لأن قيمة معامل ألفا ربما تزداد ولكن هذه الزيادة غير دقيقة لأنها تنتج من ارتباط درجات الخطأ ومن ثم التحيز بالزيادة فى القيمة الناتجة.

تتفق نتائج هذا الفرض مع نتائج بعض الدراسات التي أشارت إلى أن معامل ألفا يؤدى إلى انخفاض تقدير قيمة معامل ثبات درجات المقياس النفسى (Green & Hershberger, 2000; Raykov, 2001a; Sijtsma, 2009; Tang and Cui, 2012; Trizano-Hermosilla & Alvarado, 2016; Zimmerman et al., 1993)، كما تتفق نتائج هذا الفرض فيما يتعلق بدقة قيم معامل أوميجا مع نتائج بعض الدراسات (Deng & Chan, 2016; Dunn et al., 2014; Revelle & Zinbarg, 2009; Zhang & Yuan, 2016; Zinbarg et al., 2005)، وكذلك أحد الدراسات العربية (محمد بوزيان تيغزة، ٢٠٠٩).

تختلف نتائج هذا الفرض مع نتيجة دراسة Javail et al. (2011) التي التي توصلت إلى أن قيمة معامل ألفا كانت أعلى من معامل أوميجا، مع ضرورة العلم أن ارتفاع قيمة معامل ألفا لا تعنى دقتها.

تعليق عام على نتائج الدراسة الحالية:

يتضح من خلال نتائج الدراسة الحالية أنه يوجد تأثير لعدد فئات الاستجابة وعدد المشاركين على دقة قيم معاملى ألفا وأوميجا، حيث إن العدد الثلاثى لفئات الاستجابة أدى إلى أقل قيم لمعاملى ألفا وأوميجا باختلاف مجموعات المشاركين، كما أن هذا العدد أيضاً ارتبط بأكبر

اتساع للمسافة بين حدى الثقة مما يجعله أقل دقة، ويتشابه فى ذلك مع المجموعة الأولى للمشاركين ٣:١ (٨١) التى أدت إلى أقل قيم وكذلك أكبر اتساع للمسافة بين حدى الثقة باختلاف عدد فئات الاستجابة، وقد أدى التدريجان الخماسى والسباعى لفئات الاستجابة إلى أفضل قيم لمعاملى ألفا وأوميجا، وكذلك أقل مقدار لاتساع المسافة بين حدى الثقة مما يدل على دقته، كما أن معامل أوميجا أكثر دقة من معامل ألفا باختلاف عدد فئات الاستجابة وعدد المشاركين فى ضوء نسبة التباين المفسر واتساع المسافة بين حدى الثقة، كما أن دقة معامل ثبات الدرجات ارتبطت بعدد فئات الاستجابة أكثر منه بعدد المشاركين.

توصيات الدراسة:

توصى الدراسة الحالية فى ضوء نتائجها بضرورة التالى:

١. توجيه نظر الباحثين إلى عدم تفسير القيمة المرتفعة لمعامل ألفا بالاتساق الداخلى للمقياس أو تجانسه، وكذلك التفكير العميق في الأسباب التي أدت إلى ارتفاع قيمته من حيث كونها ناتجة من ارتفاع الارتباطات البينية بين البنود أو عددها حتى يتم الثقة في المقياس النفسى من عدمها، ويتم ذلك من خلال فحص مصفوفة معاملات الارتباط بين البنود والتي تُعد جزءً رئيساً من نواتج الحزم الإحصائية الحديثة.

٢. توجيه نظر الباحثين أيضاً إلى عدم الثقة في معامل ألفا لكونه الأكثر شيوعاً واستخداماً بين معاملات ثبات الدرجات المنشورة، وإنما استخدامه بعد التحقق من افتراضات نموذج تاو المتكافئ في الأساس القائم عليه.

٣. عقد ورش عمل للباحثين لتوعيتهم بكيفية استخدام برنامج RStudio في حساب معاملى ألفا وأوميجا وحدود الثقة المرتبطة بهما، وخاصة أنها متاحة مجاناً.

٤. تضمين مقررات القياس النفسى في الدراسات العليا على كيفية استخدام معامل أوميجا لحساب ثبات درجات المقياس النفسى وذلك باستخدام البرمجيات الحديثة والتي منها RStudio.

٥. تدوين حدود الثقة التى يتم الحصول عليها مع قيم معاملات ثبات الدرجة للتأكد من دقة هذه القيم.

٦. توجيه نظر الأساتذة المشرفين على طلاب الدراسات العليا بضرورة إرشاد طلابهم على استخدام عدد فئات الاستجابة الخماسى أو السباعى عند تصميم أدوات القياس في بحوثهم، وكذلك معامل أوميجا لحساب ثبات درجات المقاييس النفسية المستخدمة في البحوث التي يشرفون عليها، نظراً لدورهم في النهوض بمنظومة البحث العلمى في

الميدان التربوى.

٧.توجيه نظر الباحثين بأن زيادة عدد فئات الاستجابة يرتبط بدقة معاملى ألفا وأوميجا أكثر من زيادة عدد المشاركين وبالتالي توفير الوقت والجهد أثناء عملية التطبيق والتصحيح.

٨.أن يكون اختيار عدد المشاركين في الدراسة الاستطلاعية لتقنين المقاييس النفسية بنسبة ٥:١ أي كل بند من بنود المقياس يقابله خمسة مشاركين.

٩.أن يضع الباحثين في الاعتبار أن قيمة معامل ثبات درجات المقياس النفسى تتأثر بعوامل أخرى مثل عدد متوسط الارتباطات بين البنود وليست فقط عدد فئات الاستجابة وعدد المشاركين.

بحوث ودراسات مقترحة:

تُعد هذه الدراسة استكمالاً لجهود بعض الباحثين في موضوع ثبات درجات المقياس النفسى في ضوء متغيرات مثل عدد فئات الاستجابة وعدد المشاركين، ولكن ما زالت البيئة العربية تحتاج إلى مزيد من الدراسات في هذا الموضوع، كما يمكن القول أنه ستظل دراسة موضوع ثبات درجات المقياس النفسى محور دراسات مستقبلية كثيرة، نظراً لأهميته في القياس النفسى والتربوى، ولذلك فهناك العديد من البحوث والدراسات المقترحة فى هذا المجال والتى يتمثل بعضها فى التالى:

١.تحرى تقدير بعض معاملات ثبات درجات المقياس النفسى في ضوء متغيرات أخرى مثل عدد البنود وكذلك متوسط الارتباطات البينية بينها.

٢.دراسة البنية الداخلية لمعامل ألفا الطبقى Stratified alpha الذى اقترحه بعض الباحثين وتحديد مدى دقته في تقدير معامل ثبات الدرجات في حالة تعدد أبعاد المقياس النفسى.

٣.تأثير عدد البدائل على معاملات أخرى مثل بيتا وثيتا وبعض معادلات التجزئة النصفية.

٤.مقارنة تأثير مدى انتهاك افتراضات نماذج القياس وعدد فئات الاستجابة على معامل ثبات درجات المقياس النفسى.

٥.تأثير عدد فئات الاستجابة وعدد المشاركين على قيم بعض الإحصاءات الوصفية ونتائج بعض الاختبارات الاستدلالية.

٦.تأثير عدد فئات الاستجابة وعدد المشاركين على مؤشرات جودة المطابقة فى التحليل العاملى التوكيدى.

٧.مقارنة طرق مختلفة لحساب حدود الثقة المرتبطة بقيم معاملات ثبات درجات المقياس النفسى.

مراجع الدراسة

حجاج غانم أحمد على، ياسر عبدالله حفنى حسن (٢٠١١). تأثير عدد بدائل ليكرت على الخصائص السيكومترية للمقياس النفسي وافتراضات التصميم العاملي ثنائي الاتجاه. جامعة عين شمس، *مجلة الإرشاد النفسى*، ٢٩ ، ٥٥-١٣١.

سعيد حسن آل عبد الفتاح الغامدى (٢٠٠٣). *مدى اختلاف الخصائص السيكومترية لأداة القياس في ضوء تغاير عدد بدائل الاستجابة والمرحلة الدراسية: دراسة حالة مقياس ليكرت* (رسالة ماجستير). كلية التربية، جامعة أم القرى.

صفوت فرج (٢٠٠٧). *القياس النفسى* (ط٦). القاهرة: الأنجلو المصرية.

صلاح الدين محمود علام (٢٠٠٠). *القياس والتقويم التربوى والنفسى: أساسياته وتطبيقاته وتوجهاته المعاصرة* (ط١). القاهرة: دار الفكر العربى.

فؤاد أبوحطب، آمال صادق (١٩٩١). *مناهج البحث وطرق التحليل الإحصائى في العلوم النفسية والتربوية والاجتماعية*. القاهرة: مكتبة الأنجلو المصرية.

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

محسوب عبدالقادر الضوى (٢٠١١). تحرى تأثير الدرجات المتطرفة وعدد فئات الاستجابة على تقدير معامل ألفا لكرونباخ. جامعة أسيوط، *مجلة كلية التربية*، ٢٧ (١)، ١١٩-١٧٥.

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 عبدالرحمن إسماعيل (٢٠١٥). أثر عدد بدائل الاستجابة في مقياس ليكرت على الخصائص السيكومترية للمقياس وقياس الاتجاهات: دراسة تطبيقية على متدربي معهد الإدارة العامة، المملكة العربية السعودية. معهد الإدارة العامة، *مجلة الإدارة العامة*، ٥٥ (٤)، ٨٣٥- ٨٧٥.

محمود محمد عمر، حصة عبدالرحمن فخرو، تركى السبيعى، آمنة عبدالله تركى (٢٠١٠). *القياس النفسى والتربوى*. عمان: دار المسيرة.

AERA, APA, & NCME (2014). *Standards for educational and psychological testing*. Washington, DC: Author.

Alessandri, G., Vecchione, M., & Caprara, G. V. (2015). Assessment of regulatory emotional self-efficacy beliefs: A review of the status of the art and some suggestions to move the field forward. *Journal of Psychoeducational Assessment, 33*(1), 24-32. doi: 10.1177/0734282914550382

Almehrizi, R. S. (2013). Coefficient alpha and reliability of scale scores. *Applied Psychological Measurement, 37*(6), 438-459. doi: 10.1177/0146621613484983

American Psychological Association (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.

Bindak, R. (2013). Relationship between randomness and coefficient alpha: A Monte Carlo simulation study. *Journal of Data Analysis and Information Processing, 1,* 13-17. doi: 10.4236/jdaip.2013.12003

Bonett, D. J. (2002). Sample size requirements for testing and estimating coefficient alpha. *Journal of Educational and Behavioral Statistics, 27*(4), 335-340. doi: 10.3102/10769986027004335

Bonett, D. J. (2003). Sample size requirements for comparing two alpha coefficients. *Applied Psychological Measurement, 27*(1), 72-74. doi: 10.1177/0146621602239477

Bonett, D. J., & Wright, T. A. (2000). Sample size requirements for estimating Pearson, Spearman, and Kendall correlations. *Psychometrika, 65*(1), 23-28. doi: 10.1007/BF02294183

Charter, R. A. (1999). Sample size requirements for precise estimates of reliability, generalizability, and validity coefficients. *Journal of Clinical and Experimental Neuropsychology, 21*(4), 559-566.

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

Cho, E. (2016). Making reliability reliable: A systematic approach to reliability coefficients. *Organizational Research Methods, 19*(4) 651-682. doi: 10.1177/1094428116656239

Cho, E., & Kim, S. (2015). Cronbach's coefficient alpha: Well known but poorly understood. *Organizational Research Methods, 18*(2), 207-230. doi: 10.1177/1094428114555994

Chomeya, R. (2010). Quality of psychology test between Likert scale 5 and 6 points. *Journal of Social Sciences, 6*(3), 399-403. doi: 10.3844/jssp.2010.399.403

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

Clason, D. L., & Dormody, T. J. (n.d.). Analyzing data measured by individual Likert-type items. *Journal of Agricultural Education, 35*(4), 31-35.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334. doi: 10.1007/BF02310555

Cronbach, L. J., & Shavelson, R. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement, 64*(3), 391-418. doi: 10.1177/0013164404266386

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. doi: 10.1111/emip.12095

Deng, L., & Chan, W. (2016). Testing the difference between reliability coefficients alpha and omega. *Educational and Psychological Measurement, 76*, 1-19. doi: 10.1177/0013164416658325

Drost, E. A. (2011). Validity and reliability in social science research.

Education Research and Perspectives, 38(1), 105-123.

Dunn, T. J., Baguley, T. & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology, 105*(3), 399-412. doi: 10.1111/bjop.12046

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), 291-303.

Feldt, L. S. (2011). Estimating the effect of changes in criterion score reliability on the power of the F test of equality of means. *Educational and Psychological Measurement, 71*(2), 420-430. doi: 10.1177/0013164410375572

Feldt, L. S., Woodruff, D. J, & Salih, F. A. (1987). Statistical inference for coefficient alpha. *Applied Psychological Measurement, 11*(1), 93-103. doi: 10.1177/014662168701100107

Feldt, L. S., & Ankenmann, R. D. (1998). Appropriate sample size for comparing alpha reliabilities. *Applied Psychological Measurement, 22*(2), 170-178. doi: 10.1177/01466216980222006

Geisinger, K. F. (2013). Reliability. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA handbook of testing and assessment in psychology: Vol. I. Test theory and testing and assessment in industrial and organizational psychology* (pp. 21–42). Washington, DC: American Psychological Association. doi: 10.1037/14047-003

Green, S. B., & Hershberger, S. L. (2000). Correlated errors in true score models and their effect on coefficient alpha. *Structural Equation Modeling: A Multidisciplinary Journal, 7*(2), 251-270. doi: 10.1207/ S15328007SEM0702_6

Green, S, B., Lissitz, R. W., & Mulaik, S. A. (1977). Limitations of coefficient alpha as index of test unidimensionality. *Educational and Psychological Measurement, 37*(4), 827-

838. doi: 10.1177/001316447703700403

Guo, H., Pohl, E., & Gerokostopoulos, A. (2013, January). *Determining the right sample size for your test: Theory and application.* Annual Reliability and Maintainability Symposium.

Hartley, J. (2013). Some thoughts on Likert-type scales. *International Journal of Clinical and Health Psychology, 14*(1), 83-86. doi: 10.1016/S1697-2600(14)70040-7

Harzing, A. W. (2013). Publish or perish [Computer software]. Retrieved from http://www.harzing.com/pop.htm

IBM Corp. (2013). IBM SPSS Statistics for Windows, Version 22 [Computer software]. Armonk, NY: IBM Corp.

Iacobucci, D., & Duhachek, A. (2003). Advancing alpha: Measuring reliability with confidence. *Journal of Consumer Psychology, 13*(4), 478–487. doi: 10.1207/S15327663JCP1304_14

Javail, S. B., Gudaganavar, N. V., & Raj, S. M. (2011). Effect of varying sample size in estimation of coefficients of internal consistency. *WebmedCentral Biostatistics, 2*(2), WMC001649.

Johanson, G. A., & Brooks, G. P. (2010). Initial scale development: Sample size for pilot studies. *Educational and Psychological Measurement, 70*(3), 394-400. doi: 10.1177/0013164409355692

Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *Research Fundamentals, 65*(1), 2276-2284. doi: 10.2146/ajhp070364

Kirk, B. A., Schutte, N. S., & Hine, D. W. (2008). Development and preliminary validation of an emotional self-efficacy scale. *Personality and Individual Differences, 45*(5), 432-436. doi: org/10.1016/j.paid.2008.06.010

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika, 2*(3), 151-160. doi:

10.1007/BF02288391

Lee, J, & Paek, I. (2014). In search of the optimal number of response categories in a rating scale. *Journal of Psychoeducational Assessment, 32*(7), 663–673. doi: 10.1177/0734282914522200

Lozano, L., 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. doi: 10.1027/1614-2241.4.2.73

McDonald, R. P. (1978). Generalizability in factorable domains: Domain validity and generalizability. *Educational and Psychological Measurement, 38*(1), 75-79.

McDonald, R. P. (1999). *Test theory: A unified approach*. Mahwah, NJ: Lawrence Erlbaum Associates.

Meeker, W. Q., & Escobar, L. A. (1998). *Statistical methods for reliability data*. New York, NY: John Wiley & Sons, Inc.

Miller, M. B. (1995). Coefficient alpha: A basic introduction from the perspectives of classical test theory and structural equation modeling. *Structural Equation Modeling, 2*(3), 255-273. doi: 10./107055195095400

Novick, M. R., & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurements. *Psychometrika, 32*(1), 1- 13. doi: 10.1007/BF02289400

Osburn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods, 5*(3), 343-355. doi: 10.1037//I082-989X.5.3.343

Padilla, M. A., & Divers, J. (2013). Coefficient omega bootstrap confidence interval: Nonnormal distributions. *Educational and Psychological Measurement, 73*(6), 956–972. doi: 10.1177/0013164413492765

Padilla, M. A., & Divers, J. (2016). A comparison of composite reliability estimators: Coefficient omega confidence intervals in the

current literature. *Educational and Psychological Measurement, 76*(3), 436–453. doi: 10.1177/0013164415593776

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

Pool, L. D., & Qualter, P. (2012). Improving emotional intelligence and emotional self-efficacy through a teaching intervention for university students. *Learning and Individual Differences, 22*(3), 306-312. doi:.org/10.1016/j.lindif.2012.01.010

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica, 104*, 1-15.

Raykov, T. (2001a). Bias of coefficient alpha for congeneric measures with correlated errors. *Applied Psychological Measurement, 25*(1), 69-76. doi: 10.1177/01466216010251005

Raykov, T. (2001b). Estimation of congeneric scale reliability using covariance structure analysis with nonlinear constraints. *British Journal of Mathematical and Statistical Psychology, 54*(2), 315-323. doi: 10.1348/000711001159582

Raykov, T., & Penev, S. (2006). A direct method for obtaining approximate standard error and confidence interval of maximal reliability for composites with congeneric measures. *Multivariate Behavioral Research, 41*(1), 15-28. doi: 10.1207/s15327906mbr4101_2

Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the GLB: Comments on Sijtsma. *Psychometrika, 74*(1), 145-154. doi: 10.1007/S11336-008-9102-Z

Romano, J. L., Kromrey, J. D., & Hibbard, S. T. (2010). A Monte Carlo study of eight confidence interval methods for coefficient alpha. *Educational and Psychological Measurement, 70*(3) 376–393. doi: 10.1177/0013164409355690

Romano, J. L., Kromrey, J. D., Owens, C. M., & Scott, H. M. (2011).

Confidence interval methods for coefficient alpha on the basis of discrete, ordinal response items: Which one, if any, is the best? *The Journal of Experimental Education, 79*(4) 382-403, doi: 10.1080/00220973.2010.510859

RStudio Team (2015). RStudio: Integrated Development for R [Computer software]. Boston, MA: RStudio, Inc.

Segall, D. O. (1994). The reliability of linearly equated tests. *Psychometrika, 59*(3), 361-375. doi: 10.1007/BF02296129

Shaftel, J., Nash, B. L., & Gillmor, S. C. (2012, April). *Effects of the number of response categories on rating scales*. Roundtable presented at the annual conference of the American Educational Research Association, Vancouver, British Columbia.

Sijtsma, K. (2009). Reliability beyond theory and into practice. *Psychometrika, 74*(1), 169- 173. doi: 10.1007/s11336-008-9103-y

Sijtsma, K. (2015). Delimiting coefficient α from internal consistency and unidimensionality. *Educational Measurement: Issues and Practice, 34*(4), 10- 13. doi: 10.1111/emip.12099

Subedi, B. P. (2016). Using Likert type data in social science research: Confusion, issues and challenges. *International Journal of Contemporary Applied Sciences, 3*(2), 36-49.

Sullivan G. M., & Artino A. R. Jr. (2013). Analyzing and interpreting data from Likert-type scales. *Journal of Graduate Medical Education, 5*(4), 541-542. doi: 10.4300/JGME-5-4-18

Tang, W., & Cui, Y. (2012, April). *A simulation study for comparing three lower bounds to reliability.* Paper presented at the annual meeting of the American Educational Research Association, Vancouver, Canada.

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(3), 371–401. doi: 10.1111/j.2044-8317.2011.02030

Thompson, B. L., & Vacha-Haase, T. (2000). Psychometrics is datametrics: The test is not reliable. *Educational and Psychological Measurement, 60*(2), 174–195. doi: 10.1177/0013164400602002

Thorndike, R. M. (2005). *Measurement and evaluation in psychology and education* (7th ed.). Upper Saddle River, NJ: Pearson Education, Inc.

Trizano-Hermosilla, I, & Alvarado, J. M. (2016). Best alternatives to Cronbach's alpha reliability in realistic conditions: Congeneric and asymmetrical measurements. *Frontiers in Psychology, 7*, 1-8. doi: 10.3389/fpsyg.2016.00769

Tsang, K. K. (2012). The use of midpoint on Likert scale: The implications for educational research. *Hong Kong Teachers' Centre Journal, 11*, 121-130.

Wakita, T., Ueshima, N., & Noguchi, H. (2012). Psychological distance between categories in the Likert scale: Comparing different numbers of options. *Educational and Psychological Measurement, 72*(4), 533-546. doi: 10.1177/0013164411431162

Weems, G. H., & Onwuegbuzie, A. J. (2001). The impact of midpoint responses and reverse coding on survey data. *Measurement and Evaluation in Counseling and Development, 34*(3), 166-176.

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. doi: 10.1177/0013164404268674

Weijtersa, B., Cabooterb, E., & Schillewaer, N. (2010). The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing, 27*(3), 236–247. doi:

10.1016/j.ijresmar.2010.02.004

Wong, C., Peng, K. Z., Shi, J., & Mao. Y. (2011). Differences between odd and even number response formats: Evidence from mainland Chinese respondents. *Asian Pacific Journal of Management, 28*(2), 379-399. doi: 10.1007/s10490-009-9143-6

Xu, J. (2014). *A comparison of three approaches to confidence interval estimation for coefficient omega* (Master's thesis). Available from ProQuest Dissertations and Theses database. (UMI No. 1583481)

Yang, Y., & Green, S. B. (2011). Coefficient alpha: A reliability coefficient for the 21st century? *Journal of Psychoeducational Assessment, 29*(4), 377-392. doi: 10.1177/0734282911406668

Yurdugul, H. (2008). Minimal sample size for Cronbach's coefficient alpha: A Monte- Carlo study. *H. U. Journal of Education, 35*, 397-405.

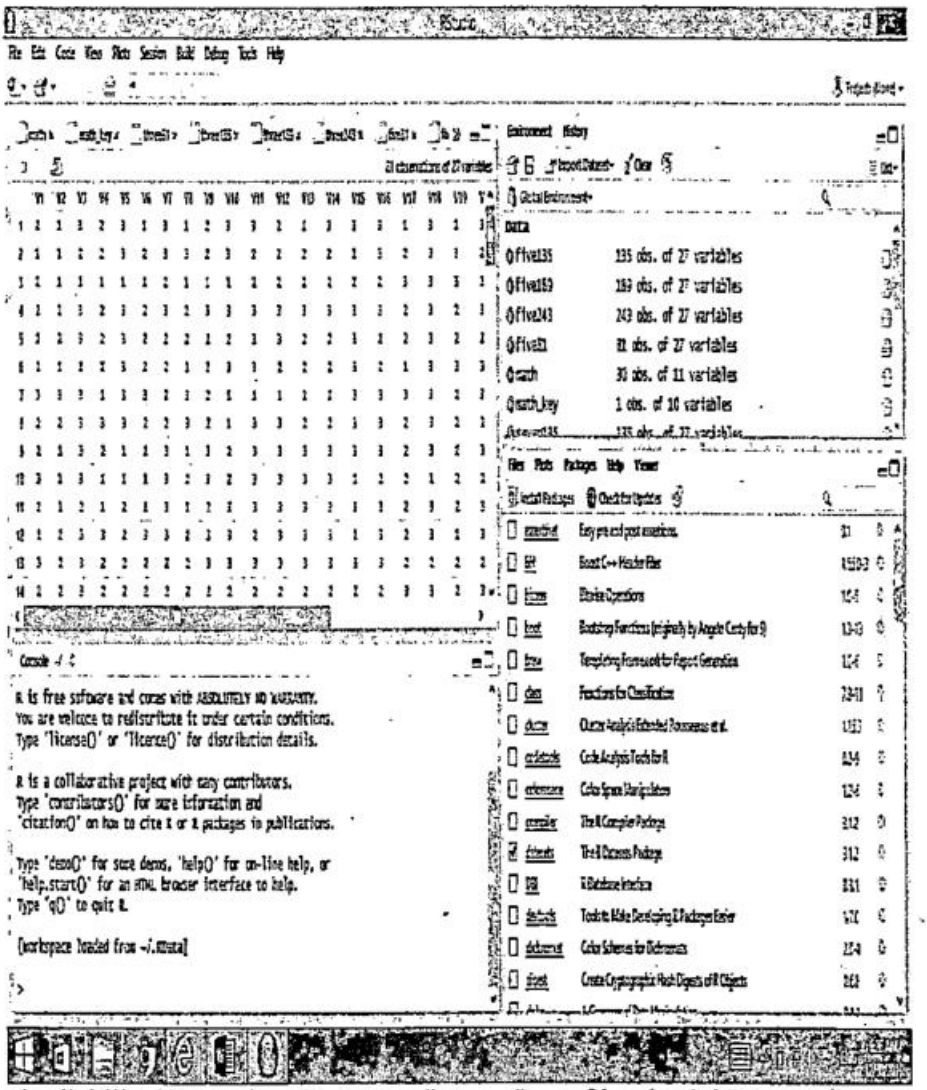
Zhang, Z., & Yuan, K. (2016). Robust coefficients alpha and omega and confidence intervals with outlying observations and missing data: Methods and software. *Educational and Psychological Measurement, 76*(3), 387–411. doi: 10.1177/0013164415594658

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*(1), 33-49. doi: 10.1177/0013164493053001003

Zinbarg, R. E., Revelle, W., Yovel, I, & Li, W. (2005). Cronbach's α, Revelle's β, and McDonald's ω_H: There relations with each other and two alternative conceptualizations of reliability. *Psychometrika, 70*(1), 123-133. doi: 10.1007/s11336-003-0974-7

Zumbo, B. D., Gradermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for Likert rating scale. *Journal of Modern Applied Statistical Methods, 6*(1), 21-29.

ملحق (١) الواجهة الافتتاحية لبرنامج RStudio وبعض الأكواد المستخدمة في حساب قيم معاملى ألفا وأوميجا وحدود الثقة المرتبطة بهما




```
ci.reliability(data = three81, type = "omega", conf.level = 0.95, interval.type = "bca", B=1000)
ci.reliability(data = three135, type = "omega", conf.level = 0.95, interval.type = "bca", B=1000)
```

The Effect of the Number of Response Categories and Number of Participants on the Accuracy of Alpha and Omega Values in Estimating Scores Reliability of the Psychological Scale

Prepared by
Dr. Mohammed Abdelhady Abdelsamea
Lecturer of Educational Psychology
Qena Faculty of Education
South Valley University

Abstract

In the field of educational and psychological testing, many studies have addressed the importance of the psychometric properties of the psychological scale before data collection. However, no study has yet covered the accuracy of alpha and omega coefficients based on some variables such as Likert's number of response categories (3, 5, 7) and number of participants. Accordingly, the current study aimed at identifying the effect of the Likert's number of response categories and number of participants on the accuracy of alpha and omega values in estimating scores reliability of the psychological scale. Participants were 243 third-year students enrolled at Qena Faculty of Education. They were divided into four groups based on the percentage between the items of emotional self-efficacy scale and number of participants: 1:3 (81), 1:5 (135), 1:7 (189), and 1:9 (243). Emotional self-efficacy scale, prepared by Kirk, Schutte, and Hine (2008), was translated and used in the current study. It was used to be useful in the Arab environment especially after estimating scores reliability of three versions using alpha and omega. The results revealed that the accuracy of alpha and omega values increased as the number of response categories increased as well as the number of participants 1:5 (135) with better confidence intervals. It was also found that omega resulted in more accurate values than alpha did as the latter underestimated the scores reliability coefficient values. Some recommendations were discussed and some directions for future research were presented.

Keywords: Scores reliability, alpha, omega, number of response categories, number of participants.