دار المنظومة
DAR ALMANDUMAH
الرواد في قواعد المعلومات العربية

العنوان:	تحري تأثير الدرجات المتطرفة وعدد فئات الاستجابة على تقدير معامل ألفا لكرونباخ
المصدر:	مجلة كلية التربية بأسيوط -مصر
المؤلف الرئيسي:	الضوي، محسوب عبدالقادر
المجلد/العدد:	مج27, ع1
محكمة:	نعم
التاريخ الميلادي:	2011
الشهر:	يناير
الصفحات:	117 - 175
رقم MD:	129860
نوع المحتوى:	بحوث ومقالات
قواعد المعلومات:	EduSearch
مواضيع:	القياس التربوي، الدرجات المتطرفة، درجات الطلاب، درجات الاختبارات، الاختبارات التربوية، معامل ألفا، كرونباخ، الدراسات الإحصائية، نظرية الثبات، النظريات التربوية، علم النفس التربوي، فلسفة التعليم، الدرجات الحقيقية، النماذج التعليمية، الرسائل الجامعية، البحوث التربوية
رابط:	http://search.mandumah.com/Record/129860

© 2016 دار المنظومة. جميع الحقوق محفوظة.
هذه المادة متاحة بناء على الإتفاق الموقع مع أصحاب حقوق النشر، علما أن جميع حقوق النشر محفوظة. يمكنك تحميل أو طباعة هذه المادة للاستخدام الشخصي فقط، ويمنع النسخ أو التحويل أو النشر عبر أي وسيلة (مثل مواقع الانترنت أو البريد الالكتروني) دون تصريح خطي من أصحاب حقوق النشر أو دار المنظومة.

كلية التربية

المجلـــة العلمـــية

تحرى تأثير الدرجات المتطرفة
وعدد فئات الاستجابة على تقدير معامل ألفا لكرونباخ

إعـــــــداد

د/ محسوب عبد القادر الضوى

مدرس علم النفس التربوي

كلية التربية بقنا – جامعة جنوب الوادي

﴿ المجلد السابع والعشرون – العدد الأول – جزء أول – يناير ٢٠١١ ﴾

ملخص الدراسة :

تهدف الدراسة إلى تحرى تأثير الاحتفاظ بالدرجات المتطرفة أو حذفها وعدد فئات الاستجابة عند تقدير معامل ألفا لكرونباخ لأداة قياس مدرجة بطريقة ليكرت ، وكذلك دراسة دلالة الفروق بين قيم معاملات ألفا المستقلة وحجم التأثير المقابل لها . وقد استخدم الباحث مقياس اتجاهات المعلمين نحو تكنولوجيا التعليم ثلاثى التدرج (إعداد : لطفى الخطيب – ٢٠٠٠م) لجمع بيانات البحث الحالى وذلك بإعداد ستة صور من المقياس تبعاً لعدد فئات الاستجابة (٢ ، ٣ ، ٤ ، ٥ ، ٦، ٧) ، وحسبت تقديرات معامل ألفا فى حالة الاحتفاظ بالدرجات الكلية المتطرفة وفى حالة حذف درجتين متطرفتين أو أربعة درجات متطرفة أو ستة درجات متطرفة أو ثمانية درجات متطرفة أو عشرة درجات متطرفة ، وحسبت دلالة الفروق بين أزواج معاملات ألفا المستقلة باستخدام اختبار Feldt's W Test ، وحسبت أحجام التأثير المقابلة للفروق باستخدام اختبار Delta Test . وقد بلغ حجم العينة الكلية (٧٣٢) معلماً بمتوسط عمر قدره (٣١،٥) سنة وانحراف معيارى قدره (١،١) سنة .

وقد بينت نتائج الدراسة الحالية تضخم تقديرات معامل ألفا عند الاحتفاظ بالدرجات المتطرفة ، واختلاف التقديرات باختلاف عدد فئات الاستجابة . وبينت النتائج أيضاً انخفاض تضخم تقديرات معامل ألفا عند حذف أعداد متباينة من الدرجات المتطرفة . وأيضاً بينت النتائج وجود فروق دالة إحصائياً بين بعض أزواج معاملات ألفا المستقلة تقابلها أحجام تأثير صغيرة أو متوسطة فى حالة الاحتفاظ بالدرجات المتطرفة وفى حالة حذف أعداد متباينة من الدرجات المتطرفة ، كما وجدت فروق غير دالة بين قيمة معامل ألفا لصورة المقياس سباعى التدرج وقيمة معامل ألفا لصورة المقياس خماسى التدرج فى حالة الاحتفاظ بالدرجات المتطرفة وفى حالة حذف أعداد متباينة منها .

Abstract

The study aimed at investigating the impact of outliers and the number of response categories on Cronbach's Alpha coefficient estimate the study also aimed at studying differences significance between Alpha coefficients and their effect size. The researcher used teachers attitudes toward educational technology scale with three-point Likert style which prepared by Lotfy Al Khateeb (2000) for collecting the data of the current study by preparing six forms of the scale according to the number of response categories (2, 3, 4, 5, 6, and 7) . The Cronbach's Alpha coefficient was estimated in presence of outliers (2, 4, 6, 8, and 10) and without and in case of deleting (2, 4, 4, 6, 8, and 10 outliers). The significance of the differences was calculated between each pair of independent coefficients by using Feldt's W test. Their effect sizes were calculated by using Delta test. The sample of the current study was de-limited to (732) subjects (mean = 31.5 year, SD=1.1).

The study finding revealed that coefficient Alpha estimates were severely inflated with the presence of outliers and affected by the number of scale points. The findings revealed that coefficient Alpha estimates were downward when there were no outliers, as well and affected by the number of scale points. In addition, the findings indicated that there were statistically significant differences between some pairs of independent Alpha coefficients with small or medium effect sizes in the case of the presence/deleting outliers. There were also nonsignificant differences between Alpha coefficients in the case of using seven and five response categories and in and without outliers.

مقدمة :

بعد مرور أربعة عشر عاماً على ظهور معادلة كيودر-ريتشاردسون Kuder-Richardson للاتساق الداخلى ، قدم كرونباخ عدداً من المعادلات المترادفة أسماها أو رمز لها بالحرف الأغريقى (ألفا) α فى مقاله الموسوعى "معامل ألفا والبنية الداخلية للاختبارات" الذى نشره فى دورية Psychometrika العتيدة سنة ١٩٥١م . مما أعطى دفعاً قوياً لمنهجية تقدير الثبات من منظور الاتساق أو التجانس الداخلى لأداة القياس (Vehkalahti, ٢٠٠٠) .

وحققت مقالة "كرونباخ" حتى عام ٢٠٠٤م نجاحاً عظيماً وتم البحث عنها أكثر من (٥٥٩٠) مرة ، حتى فى السنوات الأخيرة فقد وجد ٣٢٥ بحثاً فى دليل التوثيق Citation Index فى العلوم الاجتماعية فى كل عام تقريباً (Cronbach & Shavelson, ٢٠٠٤) .

وبمراجعة دليل التوثيق فى العلوم الاجتماعية للأدبيات المنشورة من العام ١٩٦٦م إلى العام ١٩٩٠م وجد أن معامل ألفا قد استخدم لتقدير الثبات فى ٦٠% من الأبحاث المنشورة فى العام الواحد فى عدد ٢٧٨ من الدورية العلمية المختلفة (Cortina, ١٩٩٣) .

وعلى الرغم من توافر عدة طرق لحساب معامل الثبات ، يبقى فى الأخير معامل ألفا لكرونباخ بمثابة الطريقة الأكثر شيوعاً لحساب أو تقدير ثبات أدوات القياس ذات الاستجابات من نوع ليكرت (Knapp, ٢٠٠٢: ٧٠) . حتى أن الحزمة الإحصائية فى العلوم الاجتماعية الأكثر شهرة والمعروفة اختصاراً باسم SPSS تضع معامل ألفا باعتباره الطريقة المفترضة فى الاستعمال Default Method ، أى أن الحزمة تستعمل طريقة ألفا تلقائياً إذ لم يحدد المستخدم طريقة أخرى لتقدير الثبات من ضمن الطرق الأخرى التى تقدمها الحزمة .

ويفضل إستخدام معامل ألفا عندما يكون الهدف تقدير معامل ثبات مقاييس الجوانب الوجدانية ، والشخصية نظراً لأنها مقاييس متدرجة الإستجابة لا توجد بها إجابة صحيحة وأخرى خاطئة (Caruso, ٢٠٠٠) .

ويعد معامل ألفا بمثابة مؤشر للاتساق الداخلى لأداة القياس ، ويستخدم على نطاق واسع فى البحوث السيكولوجية والاجتماعية وبحوث الطب الحيوى والبحوث المرتبطة

بالرعاية الصحية (Gilmer & Feldt, ١٩٨٣;Cortina, ١٩٩٣; Graham, ٢٠٠٦; Yan et al., ٢٠١٠) .

ويعتمد معامل ألفا على اتساق أداء الفرد من بند إلى آخر ، وهو يشير إلى قوة الارتباط أو التماسك بين بنود المقياس ، حيث يقوم على حساب متوسط الارتباطات البينية Inter-correlations بين درجات أفراد عينة الثبات على كل بند من بنود المقياس وعندما تكون البنود متجانسة فإن تقدير معامل ألفا يكون الحد الأدنى لمعامل الثبات (Yuan, Guarnaccia & Hayslip, ٢٠٠٣) .

وقد حظى تقدير معامل ألفا لدقته باهتمام عدد كبير من الباحثين ، فقد تم دراسة صيغة بديلة تقريبية له باستخدام قاعدة النسبة الترجيحية (Salih, ١٩٩٤) ، ودراسة الاختبارات الإحصائية لتعادل قيمتين للمعامل (Charter & Feldt, ١٩٩٦) ، ودراسة المعامل من خلال المقارنة بين ثلاثة توزيعات للبيانات هى التوزيع الطبيعى Normal Theory-based Distribution وتوزيع العزوم من الرتبة الرابعة Fourth-Order Moments والتوزيع الامبريقى Bootstrap Empirical Distribution (Yuan, Guarnaccia & Hayslip, ٢٠٠٣) ، ودراسة حجم العينة المتطلب لحساب معامل ألفا (Bonett, ٢٠٠٢, ٢٠٠٣) ، والمقارنة بين صيغتين معدلتين للمعامل باستخدام عينات ذات أحجام صغيرة (Feldt & Kim, ٢٠٠٦) ، وتصحيح المعامل من أخطاء القياس (Rae, ٢٠٠٦) ، ودراسة منعة أو ضلاعة المعامل (Christmann & Aelst, ٢٠٠٦) ، ودراسة الصور الرتبية لمعاملى ألفا وثيتا للمقاييس من نوع ليكرت (Zumbo, Gadermann & Zeisser, ٢٠٠٧) ، ودراسة بنية المعامل ودقته فى ضوء افتراضات نماذج القياس (أمحمد بوزيان ، ٢٠٠٨) ، ودراسة قياس حجم التأثير للفروق بين قيمتين مستقلتين للمعامل (Liu & Weng, ٢٠٠٩) ، ودراسة أخطاء الباحثين فى استخدام المعامل ودراسة عيوبه (Sijtsma, ٢٠٠٩b) ، ودراسة المعادلات البديلة لمعامل ألفا (Sijtsma, ٢٠٠٩a) ، ودراسة اشتقاق صورة جديدة من المعامل تصلح لتقدير ثبات الأدوات المستخدمة فى الدراسات الطولية الكلينيكية وتتغلب على وجود بيانات مفقودة (Yan et al., ٢٠١٠) .

فالاهتمام بدراسة الدرجات المتطرفة أو القيم الشاذة Outliers له تاريخ طويل منذ أن بين "بيرنولى" Bernoulli تأثيرها فى تقدير المعلم أو البارامتر مثل

المتوسط ومعامل الارتباط ومعالم الانحدار ، وبين أن وجود درجة متطرفة واحدة يمكن أن ينتج عنها تحيز شديد فى نتائج الاختبارات الإحصائية الوصفية أو الإستدلالية (Liu, Wu & Zumbo, ٢٠١٠) .

ويؤدى وجود درجات متطرفة إلى الحصول على تقديرات متحيزة للمتوسط الحسابى والانحراف المعيارى ، وهذا بدوره يؤدى إلى تفسيرات مضللة وتعميمات خاطئة فى بعض الأحيان نتيجة لتأثر قيم الاختبارات الإحصائية ومنها اختبار "ت" ، واختبار تحليل التباين ، واختبار تحليل التغاير ، واختبار تحليل الانحدار (Zijlstra, Ark & Sijtsma, ٢٠٠٧) .

ومثلما تتأثر قيم الاختبارات الإحصائية بالدرجات الشاذة ، تتأثر قيمة معامل الثبات باستخدام معامل ألفا لكرونباخ نظراً لاعتماد المعامل على حساب تباين الدرجات الكلية للأفراد وتباين درجات كل بند من بنود أداة القياس (Streiner, ٢٠٠٣a, ٢٠٠٣b) .

وهناك ثمة إتفاق بين الباحثين على أن أدوات القياس المدرجة بأسلوب ليكرت Likert المسمى بأسلوب التقدير الجمعى Summated Rating تعد من أكثر أدوات جمع البيانات استخداماً فى مجال العلوم النفسية والتربوية والاجتماعية ، وتسمى تلك الأدوات بموازين (سلالم) التقدير Rating Scales ، والتى تقوم على استخدام مجموعة استجابات متدرجة على متصل قياس المتغير المستهدف قياسه (Chomeya, ٢٠١٠) .

والأدوات المصممة بطريقة ليكرت تقدم للمفحوصين عبارات أو أسئلة بحيث يستجيبون على متصل الاتفاق أو التفضيل أى متدرجة الاستجابات ، وهذا المتصل يمتد بين طرفى نقيض مثل "تنطبيق تماماً – لا تنطبيق تماماً" أو "أوافق بشدة – أرفض بشدة" ومن أمثلتها مقاييس إتجاهات الطلاب أو الآباء أو المعلمين (Adelson & McCoach, in press) .

وقد ناقش المختصون بالقياس النفسى عبر عقود مختلفة أدوات القياس رتبية الإستجابة وما إذا كانت الاستجابات المقدمة للأفراد عددها فردى أم زوجى ، ويتصل ذلك بأهمية تقديم أو وضع فئة إستجابة وسيطة أو محايدة Neutral Point ، وأيضاً العدد الأمثل من فئات الاستجابة (Gable & Wolf, ١٩٩٣: ٣٤-٣٥) .

مشكلة الدراسة :

لا يخلو بحث علمى فى مجال البحوث النفسية والتربوية والاجتماعية من فحص الخصائص السيكومترية –الثبات والصدق– لأدوات القياس ، وتعد معاملات الثبات والصدق المرتفعة مؤشراً لطمأنة الباحث لأن يتجه نحو جمع بيانات بحثه بمزيد من الثقة .

ونظراً لأن معظم بحوث الماجستير والدكتوراه تستخدم معامل ألفا فى حساب ثبات الأدوات ، فإنه يلاحظ كثيراً من القصور فى فهم منطق المعامل وإمكانياته وحدوده ، كما يلاحظ إهمالاً تاماً للافتراضات التى يقوم عليها والمحددة لشروط استعماله ، والتى قد تسفر عند عدم مراعاتها عن تقديرات غير دقيقة للثبات (Raykov, ١٩٩٧a, ١٩٩٧b) .

كما أن معامل ألفا لكرونباخ يتميز بأنه لا يتمتع بالمنعة أو الضلاعة Robustness لأنه حساس بدرجة شديدة لتجاهل الافتراضات الأساسية لنموذج القياس الخاص به (Christmann & Aelst, ٢٠٠٦) .

ويذكر "ليند وزمبو" (Lind & Zumbo, ١٩٩٣) أن معامل ألفا مثل كل إحصاءات المربعات الصغرى Least-Squares Statistics ربما يتأثر بشكل غير ملائم بالبيانات المتطرفة تلك التى تقع بعيداً عن معظم البيانات الأخرى ، وربما يكون لها تأثير على النماذج المقدرة ومعالمها .

وقد دلت نتائج بعض الأبحاث إلى أن الدرجات المتطرفة تؤدى إلى تضخم شديد فى تقدير معامل ألفا . وأوصت بضرورة فحص البيانات التى يتم جمعها للكشف عن وجود درجات متطرفة ،وأن تحديد هذه الدرجات يعد خطوة هامة فى تحليل البيانات (Liu, Wu & Zumbo, ٢٠١٠; Liu & Zumbo, ٢٠٠٧; Zijlstra, Ark & Sijtsma, ٢٠٠٨) .

وتقع الدرجات المتطرفة عند طرفى التوزيع وبالتالى فإن الدرجات المتطرفة المنخفضة تؤدى إلى انخفاض Underestimation قيم المعاملات الإحصائية مثل المتوسط والتباين ، أما الدرجات المتطرفة المرتفعة فتؤدى إلى تضخم فى قيم المعاملات الإحصائية Overestimation (Liu, ٢٠٠٥) .

ويذكر "ليو" (Liu, ٢٠٠٥) أن الدراسات السابقة فى أدبيات مناهج البحث والإحصاء قدر ركزت على تحديد الدرجات المتطرفة وكيفية التعامل معها ، وأن بعض الدراسات تحرت تأثيرها على تقدير المتوسط الحسابى ومعامل الارتباط ومعالم الانحدار واختبار "ت" واختبار "ف" ، ومع ذلك لا توجد دراسات تصدت لفحص تأثير الدرجات المتطرفة على تقدير معامل ألفا .

وبهذا يكون عدم الدراية الكافية بمصادر الدرجات أو القيم المتطرفة وكيفية التحكم فيها أو حذفها أو تلاشى تأثيرها محل إهتمام الباحثين فى مجال القياس النفسى رغبة فى الحصول على تقديرات نقية لمعاملات الثبات (Zijlstra, Ark & Sijtsma, ٢٠٠٨) .

ومن المشكلات الأخرى المتعلقة باستخدام معامل ألفا لكرونباخ أن العديد من الباحثين يجهلون أن المعامل يعتمد على نموذج القياس "تاو" المترادف بالأساس Essentially Tau-equivalent Model ، وأن تجاهل افتراضات هذا النموذج يؤدى بدوره إلى انخفاض تقدير معامل ألفا ، وما يعمق تلك المشكلة عدم دراية الكثير من الباحثين بالبدائل المناسبة حين يتم تجاهل الفرضيات الأساسية (Miller, ١٩٩٥; Zumbo, ١٩٩٩; Raykov, ١٩٩٧a, ١٩٩٧b; Graham, ٢٠٠٦) .

فوجود بند واحد لا يتبع نموذج "تاو" المترادف بالأساس مقارنة بالبنود الأخرى يؤدى إلى تأثير دراماتيكى على دقة تقدير معامل ألفا . كما أن قصر طول المقياس بشكل ملحوظ يجعله أكثر عرضة لتجاهل افتراضات نموذج القياس الأساسى . ومن العوامل المهمة أيضاً التى تؤثر فى ملاءمة البيانات لنموذج القياس شكل الاستجابات التى تتبع البنود وبخاصة ما يعرف بالبنود ذات الاستجابات متعددة الصياغة ومن أمثلتها بنود الاختبارات التحصيلية التى تحتوى على مزيج من بنود الصواب والخطأ ، وبنود التكملة ، وبنود المقال القصير ، وبنود الاختيار من متعدد Multiple Response Formats (Qualls, ١٩٩٥;Graham, ٢٠٠٦) .

وتعد المقاييس المبنية فى ضوء طريقة ليكرت من أشهر أدوات جمع بيانات البحوث التربوية والنفسية ، والتى ما تزال تحتل مكانة متميزة فى البحث السيكولوجى والتربوى والاجتماعى (Chomeya, ٢٠١٠; Adelson & McCoach, in press) . ومن

المشكلات المتعلقة بها اختيار الباحث لعدد فئات الاستجابة أو شكلها ، فرغبة الباحث فى إنهاء بحثه تجعله يلجأ لتخفيض عدد فئات الاستجابة لضمان سهولة التصحيح وسهولة إدخال البيانات وسهولة تفسير النتائج دون اعتبار لمرجعية علمية فى هذا الاطار .

وعلى الرغم من مضى أكثر من مائة عام على استخدام مقاييس التقدير ، إلا أنه لم يتم التوصل إلى إتفاق بشأن أفضل عدد من بدائل الاستجابة للوصول إلى أفضل الخصائص السيكومترية لأداة القياس (Maydeu-Olivares et al., ٢٠٠٩) .

وقد تباينت نتائج بعض الدراسات السابقة فى تحرى العلاقة بين عدد فئات الاستجابة وتقدير معامل ألفا ، فقد بين "كومريتا وجراهام" (Komorita & Graham,١٩٩٥) ودراسة "ونج" (Weng, ٢٠٠٤) أن المقاييس ذات البنود غير المتجانسة ربما تظهر علاقة موجبة بين تقدير الثبات وعدد فئات الاستجابة ، كما بينت نتائج دراسة "ويزر وشارما ونيدريك" (Weathers, Sharma & Niedrich, ٢٠٠٥) ، ونتائج دراسة "شوميا" (Chomeya, ٢٠١٠; Adelson & McCoach, in press) وجود تأثير لعدد فئات الاستجابة على تقدير معامل ألفا ، ويتعارض ذلك مع نتائج دراسة "ماسترز" (Masters, ١٩٩٤) ، ودراسة "ماكيلفى" (McKelvie, ١٩٩٨) فقد توصلا إلى وجود علاقة موجبة فى حالة وجود اتساق بين الاستجابات .

ويؤكد "بيركيت" (Birkett, ٢٠٠٦) أن النتائج التى يتم الحصول عليها فى دراسات تقدير الثبات باستخدام معامل ألفا على وجه التحديد ليست قابلة للتعميم بصورة مطلقة فى المجتمعات الأخرى ، كما أنه توجد صعوبة فى دراسة التأثيرات الثانوية المحتملة تعديل عدد فئات الاستجابة ، ويوصى بعدم استخدام فئات الاستجابة السداسية .

ونظراً لعدم استقرار الباحثين فى مجال القياس النفسى والتربوى على العدد الأمثل لفئات الإستجابة سواء كان زوجياً أم فردياً ، أوصى بعض الباحثين باستخدام مدى واسع من فئات الاستجابة يتضمن كلاً من البدائل الزوجية والفردية ، فقد اقترح (Oaster, ١٩٨٩; Ko, ١٩٩٤) استخدام مدى من فئات الاستجابة من ستة أو سبعة بدائل ، أما (Preston & Colman, ٢٠٠٠) فقد أوصيا بأن يكون المدى من سبعة إلى عشرة بدائل .

وبناء على ما تقدم ونتيجة لندرة البحوث والدراسات فى هذا المجال على المستوى المحلى ، نبعت فكرة الدراسة الحالية لإلقاء الضوء على تأثير الدرجات الكلية المتطرفة ، وتأثير عدد بدائل أو فئات الإستجابة عند تقدير معامل ثبات درجات مقياس اتجاهات المعلمين نحو تكنولوجيا التعليم باستخدام معادلة ألفا لكرونباخ ، وفحص دلالة الفروق بين قيم معاملات ألفا المستقلة فى حالات مختلفة وحساب حجم التأثير المقابل لها ، وبالتالى تتحدد مشكلة الدراسة الحالية فى الإجابة عن الأسئلة الآتية :-

١- هل يؤثر عدد فئات الاستجابة المقابل للبند فى تقدير معامل ألفا عند الاحتفاظ بالدرجات المتطرفة ؟

٢- هل توجد فروق دالة إحصائياً بين أزواج معاملات ألفا المستقلة عند الاحتفاظ بالدرجات المتطرفة ؟

٣- هل يؤثر عدد فئات الاستجابة المقابل للبند فى تقدير معامل ألفا عند حذف أعداد متباينة من الدرجات المتطرفة ؟

٤- هل توجد فروق دالة إحصائياً بين أزواج معاملات ألفا المستقلة عند حذف أعداد متباينة من الدرجات المتطرفة ؟

أهمية الدراسة :

تظهر أهمية الدراسة الحالية من خلال الآتى :-

- يعد حساب ثبات درجات أدوات القياس أمراً حيوياً للباحثين فى مجال العلوم النفسية والتربوية والاجتماعية ومؤشراً مهماً من المؤشرات الدالة على صلاحية أدوات القياس لاستخدامها فى جمع بيانات عن المتغيرات المستهدفة بالقياس (Graham, ٢٠٠٦) .
- ندرة البحوث العربية المنشورة فى منهجية تقدير الثبات .
- تقييم معامل ألفا لكرونباخ باعتباره من أهم طرق تقدير معامل ثبات أدوات القياس . وذلك بفحص تأثير الدرجات المتطرفة وعدد فئات الاستجابة على تقدير المعامل ، حيث تفيد البحوث بوجود عدد قليل من البحوث المهتمة بذلك الأمر (Zumbo, ١٩٩٩; Liu & Zumbo, ٢٠٠٧; Liu, Wu & Zumbo, ٢٠١٠) .

- عدد البحوث التى تدور حول جودة أدوات جمع البيانات فى ضوء عدد فئات الاستجابة مازال قليلاً (Chomeya, ٢٠١٠) . وبصفة خاصة فى البيئة المصرية وبيئة صعيد مصر على وجه الخصوص .
- يقدم البحث الحالى خلفية نظرية عن معامل ألفا لكرونباخ ونموذج القياس الخاص به والدرجات المتطرفة وعدد بدائل الإستجابة المفضلة ، تفيد الباحثين الراغبين فى إستخدام المعامل عند تقدير الثبات فى دراساتهم .

أهداف الدراسة :

تهدف الدراسة الحالية إلى :-

١- حساب قيم معامل ألفا لكرونباخ عند الاحتفاظ بالدرجات المتطرفة مع تعدد فئات الاستجابة ، وعند حذف أعداد متباينة منها .

٢- دراسة الدلالة الإحصائية للفروق بين معاملات ألفا المستقلة وأحجام التأثير المقابلة لها .

المفاهيم الأساسية فى الدراسة :

نظرية الثبات Reliability Theory :

يعد الثبات خاصية أو ميزة معقدة ، ولكنها تتضمن عامة اتساق القياس . وربما يمكن الحكم على اتساق القياس من خلال الزمن ، والبنود ، والدرجات ، والمختبرين (المفحوصين) ، والمختبرين (الفاحصين) ، أو الدقة فى التصنيف (Payne, ١٩٩٢: ٩٩) . ويعنى الثبات أيضاً اتساق النتائج ، وهذا يكافىء أن أداة القياس مرتفعة الثبات يمكن استخدامها بصورة متكررة فى المواقف غير المتغيرة وينتج عنها نتائج تتميز بالثبات أو نتائج قريبة من الثبات (Ahmann & Glock, ١٩٧١: ٢٨٨-٢٨٩) . وأيضاً يشير الاتساق ، إلى أن الاختبار ينتج عنه نفس تقدير الفرد أو قريباً منه عندما يجرى على الفرد نفسه مرات عديدة مع الافتراض أو التسليم بأن السمة التى يقيسها الاختبار غير قابلة للتغير (Kubiszyn & Borich, ١٩٩٠: ٢٩١) .

وتشير نظرية القياس الكلاسيكى إلىّ أن الدرجة التى يحصل عليها الفرد (X) تنقسم إلى الدرجة الحقيقية T ودرجة الخطأ E كما يلى :

$$X = T + E$$

وتمثل الدرجة الحقيقية تجمع من كل العوامل التى تؤدى إلى الاتساق فى قياس الخاصية المستهدفة، وأحد تلك العوامل هو المستوى الحقيقى للفرد فى الخاصية المستهدفة ، ولكن المتغيرات الموقفية تلعب دوراً . أما درجة الخطأ فتمثل كل العوامل التى تسهم فى عدم اتساق القياس ،ويحكمها إفتراض أن عوامل الخطأ تتوزع عشوائياً(Janda, ١٩٩٨: ٥٨) .

لذا يقع على عاتق الباحث مهمة تعرف مصادر أو عوامل ظهور الخطأ التى تؤدى إلى أخطاء القياس ، والبحث فى علاجها ، وبالتالى يزداد تقدير ثبات أداة القياس ، حيث يقل تأثير أخطاء القياس ليبقى التباين الموجود بين درجات الأفراد يعكس الفروق الفردية الحقيقية فى الصفة أو السمة المراد قياسها لدى هؤلاء الأفراد .

وعند استخدام أى طريقة لتقدير ثبات الاتساق الداخلى ، فإن الاختبار يجب تحليله إلى (k) من الأجزاء ، حيث (X) هى الدرجة الكلية الملاحظة على الاختبار الكلى ، Y١ , Y٢ , Yk ..... هى الدرجات الخاصة بأجزاء الاختبار ، ويمكن التعبير عن ذلك كما يلى :-

$$X = Y١ + Y٢ + Y٣ + \ldots\ldots\ldots\ldots + Yk$$

ويكون أقصى عدد لأجزاء الاختبار هو العدد الكلى لبنوده (Janda, ١٩٩٨: ٥٩) .

وتعتمد العلاقة بين درجات أجزاء أو أقسام المقياس على أربعة نماذج رياضية مختلفة تمثل تقديرات التشابه أو التطابق الإحصائى التى ربما تظهر بين تلك الأقسام بقصد إجراء تحليل لثبات درجات المقياس الكلى ، وتختلف تلك النماذج الرياضية تبعاً للتساوى المفترض لبعض البارامترات الخاصة بأجزاء المقياس ويتصل ذلك بتوزيع الدرجات الحقيقية والدرجات الخطأ والدرجات الملاحظة أو المشاهدة ، وتتمثل تلك النماذج فى : نموذج التوازى الكلاسيكى ، ونموذج "تاو" المترادف ، ونموذج "تاو" المترادف بالأساس ، والنموذج التقاربى (Feldt & Charter, ٢٠٠٣) .

نموذج التوازى الكلاسيكى Classically Parallel Model

يسمى نمـــوذج التوازى التـــام Strict Parallel Model ، و يسمح باستخدام طريقة التجزئة النصفية ، ويتطلـــب تســـاوى متوسطات وتباينات توزيعات الدرجات الحقيقية ودرجات الخطأ للأجزاء المكونة للمقياس ، وبمعنى آخر يجب أن تسهم الأجزاء المكونة للدرجة الكلية بالمقدار نفسه فى تباين الدرجات الحقيقية والدرجات الخطأ للدرجة الكلية (Graham, ٢٠٠٦) .

ويعد نموذج القياس الكلاسيكى الأكثر تقييداً عند استخدامه لتحديد مكونات الدرجة الحقيقية ، ويقوم علـــى افتـــراض أن جميع بنـــود المقياس تقيس متغيراً كامناً وحيداً (أحادية البعد Unidimensionality) يتمثل فى السمة أو المفهوم الذى يشكل القاسم المشترك لبنود المقياس . ويقوم ثانياً على افتراض أن الدرجات الحقيقية لكل بند من بنود المقياس متساوية ، بمعنى أنه يفترض فى بنود المقياس أن تقيس السمة بنفس المقدار ، أو أن مقدار الدرجة الخالية من الخطأ العشوائى متساوية على مستوى البنود التى تشكل المقياس . ويقوم ثالثاً على افتراض أن تباين درجات الخطأ بالنسبة لجميع بنود المقياس يكون متساوياً . وباختصار يدل نموذج التوازى التام على أن بنود المقياس متكافئة أى تقيس كلها ذات المتغير الكامن ، بنفس وحدة القياس The Same Metric ، وبذات درجة الدقة ، وبنفس مقدار الخطأ (Raykov, ١٩٩٧a, ١٩٩٧b; Graham, ٢٠٠٦) .

ويمكن التعبير عن تساوى الدرجة الحقيقية τ بين بنود المقياس للفرد p كما يلى :-

$$\tau_{ip}=\tau_{jp}=.....=\tau_{kp}$$

أى أن الدرجـــة الحقيقيـــة للـــفرد p على البند i مساوية لدرجتـــه الحقيقية على البند j مساوية لدرجته الحقيقية لآخر بند فى المقياس . وكذلك يمكن التعبير عن تساوى تباين درجات الخطأ العشوائى E كما يلى : σ^2

$$\sigma^2(E_{ip})=\sigma^2(E_{jp})=.....=\sigma^2(E_{kp})$$

أى أن مقدار تباين الخطأ للفرد p على البند i ، مساوياً لمقدار تباين الخطأ على البند j مساوياً لتباين خطأ أى بند آخر. وعليه يمكن التعبير عن نموذج التوازى بالمعادلة العامة التالية :-

$$X_{ip} = \tau_p + E_p \ldots\ldots\ldots\ldots\ldots (1)$$

وذلك للدلالة على أن أية درجة ملاحظة للمقياس X_{ip} ذى البنود المتوازية هى محصلة الدرجات الحقيقية المتساوية بين البنود للفرد p ، والدرجات الخطأ المتساوية عبر البنود لذات الفرد (Graham, ٢٠٠٦) .

نموذج "تاو" المترادف⊕ Tau Equivalent Model

يشترك مع نموذج التوازى التام فى مسلمة أو افتراض أحادية البعد ، أى أن بنود المقياس يجب أن تتسم بالتجانس بحيث تشترك كلها فى قياس ذات المتغير أو عامل كامن واحد . كما يشترك معه أيضاً فى مسلمة تساوى الدرجات الحقيقية بين بنود المقياس لذات الفرد ، والتعبير الرياضى لهذا الافتراض هو :

$$\tau_{ip} = \tau_{jp} = \ldots\ldots\ldots = \tau_{kp}$$

ويبقى سارياً على النموذج الحالى . غير أن الاختلاف الوحيد بينهما يكمن فى احتمال عدم تساوى تباين درجات الخطأ عبر بنود المقياس لذات الفرد كما يلى :

$$\sigma^2(E_{ip}) \neq \sigma^2(E_{jp}) \neq \ldots\ldots\ldots \neq \sigma^2(E_{kp})$$

أى أن نموذج "تاو" المترادف يتطلب أن يكون المقياس متجانساً أحادى البعد ، كما يتطلب أن تكون الدرجة الحقيقية للفرد متماثلة بين بنود المقياس ، غير أن درجات الخطأ لأداء الفرد ذاته على المقياس قد تختلف من بند لآخر . ونتيجة لذلك قد تختلف درجة الفرد الملاحظة من بند لآخر (Feldt & Brennan, ١٩٨٩: ١١٠-١١١) .

بالرغم من أن ترجمة مصطلح *Equivalent Model* بالنموذج المتكافىء ترجمة دقيقة ، إلا أن الباحث خشى عند استعمال لفظ "متكافىء" أن يتداخل ذلك مع مصطلح الصور المتكافئة *Equivalent Forms* الذى يشير إلى إحدى طرق تقدير الثبات . ولذلك سيكتفى الباحث باستعمال كلمة "مترادف" أو "متماثل" .

والمعادلة التالية تجمع بين افتراض تساوى الدرجات الحقيقية للبنود ، وافتراض أن كل بند يمتلك مقداراً من الخطأ الخاص به :

$$X_{ip}=\tau_p+E_{ip}.....(2)$$

ويلخص "ريكوف" (Raykov, ١٩٩٧a, ١٩٩٧b) فكرة النموذج المترادف بأنه يدل على أن بنود المقياس تقيس المتغير الكامن نفسه ، بالوحدة نفسها (نفس السلم المتدرج للاجابة لكافة البنود) ، وبنفس الدرجة من الدقة ، ولكن بمقادير مختلفة من الخطأ .

نموذج "تاو" المترادف بالأساس Essentially Tau Equivalent Model

علاوة على عدم اشتراطه لتساوى تباينات الخطأ – فإنه يسمح بإمكان اختلاف قيم الدرجات الحقيقية للبنود بفارق قيمة ثابتة أو عدد ثابت مضاف ، مؤدياً بذلك إلى اختلاف متوسطات الدرجات الحقيقية للبنود ، أو اختلاف فى مستوى الدقة (Raykov, ١٩٩٧a) .

ويقوم هذا النموذج على على افتراض أن كل بند من بنود المقياس يقيس ذات المتغير الكامن (أحادى البعد أو متجانس) ، وباستعمال وحدة القياس نفسها أو سلم القياس ذاته مما ينشأ عن ذلك تساوى تباين الدرجات الحقيقية ، لكن بمستويات أو درجات مختلفة من الدقة المتمثلة فى اختلاف المتوسطات نتيجة تفاوت الدرجات الحقيقية بمقدار ثابت مضاف ، هذا فضلاً عن اختلاف قيم تباين الخطأ (Miller, ١٩٩٥; Raykov, ١٩٩٧a) ، ويمكن تمثيل ذلك رياضياً بالمعادلة التالية :-

$$X_{pi}=(a_i+\tau_p)+E_{pi}.....(3)$$

حيث يدل الرمز a_i على القيمة أو العدد الثابت الذى يمثل مقدار التفاوت فى الدرجة الحقيقية بين كل بند وآخر .

كما أن اختلاف مدى سلم القياس يؤثر فى التباين ، حيث يميل التباين إلى التقلص عندما يكون سلم القياس ضيقاً ، كأن ينطوى علي درجتين فقط (صفر ، ١) ، كما يميل إلى الارتفاع لما تزداد فئات سلم القياس أى تختلف وحدة القياس ، كأن يحتوى السلم المتدرج مثلاً على خمس فئات بأوزان تنطلق من واحد إلى خمسة . أما الدقة التى تتمثل فى اختلاف متوسطات الدرجات الحقيقية للبنود فتدل على اختلاف البنود فى قوتها على الدلالة على نفس

المتغير الكامن رغم قياسها لذات المتغير ورغم تماثل سلم القياس . أما الخطأ فيتمثل فى الفرق بين الدرجة الملاحظة والدرجة الحقيقية (Graham, ٢٠٠٦) .

وفى الأخير يبقى أن نموذج "تاو" المترادف بالأساس يسمح باستخدام معادلة ألفا لكرونباخ ، ويتحرر من تساوى متوسطات وتباينات الدرجات الخطأ ، ولكنه يتطلب تساوى متوسطات وتباينات الدرجات الحقيقية للأجزاء المكونة للدرجة الكلية (Graham, ٢٠٠٦) .

النموذج التقاربى Congeneric Model :

يسمى النموذج المتقارب أو التآلفى ، وهو النموذج الأقل تقييداً من النماذج الأربعة ، وهو النموذج الأكثر عمومية للاستخدام فى تقدير الثبات (Graham, ٢٠٠٦) . ويسلم هذا النموذج بأن كل بند من بنود المقياس يقيس ذات المتغير الكامن أو ذات السمة ، بوحدات أو سلم قياس قد يختلف بين بعض البنود ، وبمستويات مختلفة من الدقة لاختلاف متوسطات الدرجات الحقيقية ، وبمقادير مختلفة من الخطأ (Raykov, ١٩٩٧a) .

ويمكن تلخيص النموذج فى المعادلة الآتية :-

$$X_{pi} = \left[a_i + \beta_i(\tau_p) + E_{pi} \right] \quad \ldots\ldots\ldots\ldots\ldots\ldots (4)$$

وبهذا يفترض هذا النموذج وجود علاقة ارتباطية خطية بين الدرجات الحقيقية للبنود – فيسمح باختلاف الدرجات الحقيقية بين البنود بفارق قيمة ثابتة a_i ، التى تجعل متوسط الدرجات الحقيقية التى تمثل الدقة تختلف بين البنود . ويسمح أيضاً – خلافاً لنموذج "تاو" المترادف – بضرب الدرجة الحقيقية بمعامل ضرب (β_i Multiplicative Term) قد تختلف قيمته من بند لآخر ، مما يترتب عليه اختلاف فى تباين الدرجات الحقيقية للبنود (Raykov, ١٩٩٧a; Graham, ٢٠٠٦) .

وفيما يلى مقارنة بين نماذج القياس الأربعة للثبات تلخص الافتراضات المختلفة ، وتبين حالة الاحتواء بينها :-

جدول (١)

مقارنة بين افتراضات نماذج القياس الأربعة

النموذج	تماثل			
	البنود***	تباين الدرجات الحقيقية*	متوسطات الدرجات الحقيقية**	تباين درجات الخطأ
التام	نعم	نعم	نعم	نعم
المترادف	نعم	نعم	نعم	لا
المترادف بالأساس	نعم	نعم	لا	لا
التقاربى	نعم	لا	لا	لا

(Graham, ٢٠٠٦)

والجدول (١) السابق يمكن تصوره فى شكل هرمى قاعدته النموذج التآلفى وقمته نموذج التوازى التام . فإذا تساوت معاملات الضرب بالواحد الصحيح فى النموذج التآلفى فى المعادلة (٤) فإنه يتحول إلى النموذج المترادف بالأساس ، وإذا آلت القيم الثابتة إلى الصفر فى المعادلة (٣) تحول النموذج المترادف بالأساس إلى النموذج المترادف كما فى المعادلة (٢) ، وإذا ساوت جميع تباينات الخطأ الواحد الصحيح فإن النموذج المترادف يتحول إلى النموذج المتوازى المعبر عنه فى المعادلة (١) (Graham, ٢٠٠٦) .

معامل ألفا لكرونباخ Cronbach's Alpha Coefficient:

اقترح كرونباخ منذ عام ١٩٥١م معادلة تطورت بعد ذلك على يد كايزر وميشيل عام ١٩٧٥م أطلق عليها معامل ألفا (Anastasi & Urbina, ١٩٩٧: ٩٩) . وهذا المعامل يستخدم فى ميدان القياس النفسى لقياس الاتساق الداخلى لأداة القياس عن طريق تحليل البنود وقياس مدى اتساقها مع بعضها البعض (Knapp, ٢٠٠٢: ٧٠) .

*** تعنى تجانس البنود أو الاشتراك فى ذات السمة .

* تعنى تماثل وحدة القياس .

** تعنى ارتفاع مستوى الدقة .

وتقتصر طريقة ألفا لكرونباخ كأحد أهم طرق الاتساق الداخلى على المقاييس ذات البنود المتجانسة وبالتالى فهى تحسب ثبات الدرجات على أنه معامل تجانس بين بنود المقياس الواحد (أحمد الرفاعى غنيم ، ١٩٩١) . ويتأثر الاتساق الداخلى بين البنود بمصدرين من مصادر تباين الخطأ هما أخطاء محتوى البند ، وأخطاء عدم تجانسها ، فكلما كانت البنود متجانسة فيما تقيس كان الاتساق عالياً فيما بينها والعكس صحيح (Anastasi & Urbina, ١٩٩٧: ٩٧) .

ويعبر عن معادلة معامل ألفا بالمعادلة الآتية :-

$$\alpha = \frac{N}{N-1}\left(1 - \frac{\sum_{i=1}^{N}\sigma_{Y_i}^2}{\sigma_X^2}\right)$$

حيث N هى عدد بنود المقياس ، σ_X^2 هو تباين الدرجات الكلية المشاهدة للأفراد ، $\sigma_{Y_i}^2$ هو تباين درجات البنود i (Cronbach & Shavelson, ٢٠٠٤) .

والمعادلة السابقة يمكن التعبير عنها رياضياً بمعادلة ترادفها كالآتى :

$$\alpha = \frac{N^2 \times M(COV)}{SUM(VAR/COV)}$$

حيث N^2 تمثل مربع عدد بنود المقياس ، M(COV) متوسط التغاير بين البنود The Mean Interitem Covariance ، ويمثل المقام مجموع جميع العناصر فى مصفوفة التباين/التغاير Variance/Covariance Matrix (Cortina, ١٩٩٣) .

وتدل دراسات المحاكاة أو المضاهاة الإحصائية Statistical Simulation إلى أن الحد الأكثر تأثيراً فى معادلة ألفا هو تباين درجات المقياس ، بحيث كلما ارتفع تباين درجات المقياس ، أدى ذلك إلى ارتفاع معامل ألفا (Streiner, ٢٠٠٣a, ٢٠٠٣b) . فدراسة "رينهاردت" (Reinhardt, ١٩٩٦) القائمة على المضاهاة أظهرت أن مدى تباين درجات المقياس فسرت مقدار التباين فى معاملات ألفا بنسبة ٦٠% . كما أسفرت دراسة "كاريسو"

(Caruso, ٢٠٠٠) القائمة على استخدام منهج التحليل البعدى للبحوث التى استخدمت مقياس NEO للسمات الكبرى فى الشخصية ، بأن متوسط الثبات لأحد المقاييس الفرعية يساوى (٠،٧٤) عندما طبق فى عدد من الدراسات على الجمهور العام (إذ تتسم العينات بالتباين) ، ولم يرق متوسط الثبات فى الدراسات الأخرى إلا إلى (٠،٦٣) عندما أجرى على عينات إكلينيكية لأنها أكثر تجانساً ، وأقل تبايناً .

وثمة مجموعة من الأساطير حول معامل ألفا ربما تمثل مجموعة من المعتقدات الخاطئة بين بعض الباحثين ، ومنها :

١- ألفا خاصية ثابتة للمقياس

هى الأسطورة الأولى والأساسية التى تحيط بمعامل ألفا وجميع معاملات الثبات الأخرى (Streiner, ٢٠٠٣a) . وقد توصل "ويلكنسون وفريق الإستدلال الإحصائى المشكل من الجمعية الأمريكية لعلم النفس "أنه من المهم تذكر أن الاختبار ليس ثابت أو غير ثابت . فالثبات خاصية لدرجات الاختبار لعينة محددة من المفحوصين" (Wilkinson & The Task Force on Statistical Inference, ١٩٩٩) .

ويشير "تومسون وسيندر" (Thompson & Synder, ١٩٩٨) إلى الأمر ذاته بأن تطبيق المقياس نفسه على عينة أكثر تجانساً ، أو على عينة أكثر تبايناً يؤدى إلى درجات ذات معاملات ثبات مختلفة . وبتعبير آخر ، فالمقياس الذى أظهر مستوى ثبات مرتفع عند تطبيقه على عينة معينة ، قد يظهر مستوى ثبات منخفض عند تطبيقه على عينة أخرى . ومغزى هذا الاستنتاج ، أن الثبات ليس صفة لصيقة أو جوهرية للمقياس ، بل يتوقف على طبيعة الدرجات ومنها التباين ، وبالتالى على العينات• .

• لذا من الخطأ استعمال تعبير "ثبات المقياس" وهو التعبير الذى طالما تردد ومازال يتردد فى بعض البحوث والدراسات ، وبعض مراجع القياس والتقويم ، وبعض كتب مناهج البحث . والأصح أن ، ويتصل بذلك أن نستعمل *Test Scores Reliability* تستعمل عوض ذلك تعبير "ثبات درجات المقياس" كما ورد فى كتاب *Test Items Validity* عوض تعبير "صدق المقياس" تعبير "صدق بنود المقياس" *(Anastasi & Urbina, ١٩٩٧)* .

لـذا يمكن القـول بأنه كلما زاد التباين ، كلمـا ارتفع تقدير معامـل ألفا ، وبالتالى ، فاتساع التباين أو تقلصه يتوقف على مدى تجانس العينة ، أو مدى تباينها . فتباين أفراد العـينة يـؤدى إلى تبايـن الاستجـابات على بنود المقيـاس، وبالتـالى يـؤدى إلى ارتفـاع تبـاين درجـات المقياس ككـل . ومعنى ذلك أن قيم الثبات لذات المقياس قد تختلف باختلاف العينات ، إذ يستتبع اختلاف العينة فى مدى تباين درجات المقياس ككل (Streiner, ٢٠٠٣a, ٢٠٠٣b) .

٢- يقيس معامل ألفا فقط التجانس الداخلى للمقياس

كلما زاد الارتباط بين بنود المقياس زادت قيمة معامل ألفا ، وعلى العكس تعنى القيمة الكبيرة لمعامل ألفا درجة عالية من الاتساق الداخلى ليس صحيحاً دائماً . والسبب أن معامل ألفا أيضاً يتأثر بشدة بطول الاختبار(Streiner, ٢٠٠٣a). وقد توصل "كورتينا" (Cortina, ١٩٩٣) أن اختباراً يتكون من ستة بنود بمتوسط معامل ارتباط بين البنود مقداره (٠،٣٠) بلغ معامل ألفا له (٠،٧٢) ، وبالمحافظة عن نفس قيمة الإرتباط وزيادة عدد البنود لتصبح (١٢ ، ١٨) بنداً وجد أن قيمة معامل ألفا زادت إلى (٠،٨٤ ، ٠،٨٨) على الترتيب .

٣- معامل ألفا الأكبر هو الأفضل دائماً :

تعد معاملات الثبات الكبيرة بأى طريقة من طرق تقدير الثبات هى الأفضل . ولكن معامل ألفا لا يقيس فقط تجانس البنود ، ولكن أيضاً تجانس المتغير المستهدف قياسه . وفى عديد من الحالات حتى المتغيرات التى تبدو أحادية البعد يمكن تصورها أنها تتكون من جوانب مختلفة ، فالقلق يمكن تصوره فى أربعة مكونات معرفية وفسيولوجية وسلوكية ووجدانية . وقيم معامل ألفا المرتفعة ربما تعكس إزدواجية أو تكرار لا لزوم له للمحتوى عبر البنود ، وتشير إلى التكرار أكثر مما تشير إلى التجانس (Streiner, ٢٠٠٣a) .

٤- معامل ألفا يمتد بين الصفر والواحد الصحيح :

حيث أن الثبات هو النسبة تباينين ، فإنه يبدو من الوهلة الأولى عبارة عن عدد محصور بين الصفر والواحد الصحيح كما فى المتباينة الآتية :

$$0 \leq alpha \leq 1.0$$

ولكن أحد المشكلات التى قد تواجه باحث ما- هو تقدير معامل ألفا لأداة القياس والحصول على قيمة سالبة ، وعندئذ تصبح المتباينة السابقة فى الشكل الآتى :

$$c \le alpha \le 1.0, ...with.....c \le 0$$

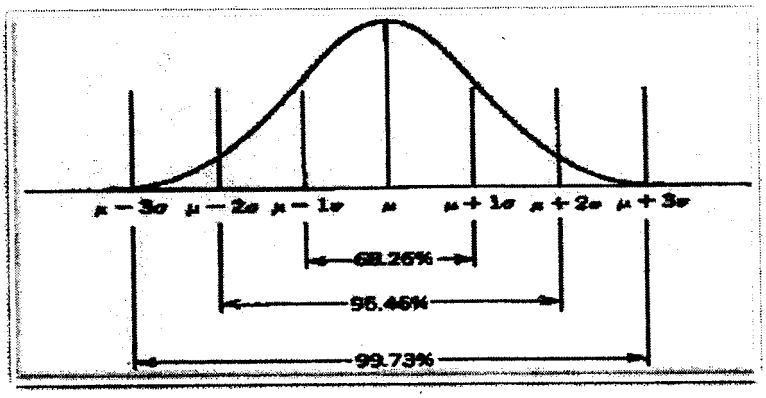
وهذه الحالة ربما يتحكم فيها إتجاه الاستجابة Response Set ، أى يمكن أن يحدث ذلك إذا كانت أداة القياس تتكون من بنود موجبة وسالبة الاتجاه وقام الباحث بالتصحيح وإدخال البيانات على أن البنود أحادية الإتجاه "سالبة فقط أو موجبة فقط" (Sijtsma, ٢٠٠٩b) . أو عندما يكون الارتباط سالب بين بعض البنود فى المقياس والسبب الرئيس لذلك ربما يكون سهو مطور الاختبار ذاته ويمكن تصويب ذلك (Streiner, ٢٠٠٣a) .

الدرجات المتطرفة Outliers :

يطلق مسمى الدرجات المتطرفة أو الشاذة على الابتعاد أو الانحراف العددى عن بقية المشاهدات أو البيانات التى يتم جمعها (Grubbs, ١٩٦٩: ١) . وبتعبير آخر تعرف الدرجة المتطرفة بأنها مشاهدة تنحرف كثيراً عن المشاهدات الأخرى بدرجة تثير المزيد من الشكوك بأنها تم توليدها عن طريق آلية مختلفة (Hawkins, ١٩٨٠: ١٣) .

والدرجات المتطرفة غالباً يتم تحديدها على أنها مشاهدات أو فئات فرعية منها تبدو غير متسقة مع بقية البيانات ، وتلك المشاهدات تستدعى التعامل معها بخاصة عندما تتسبب فى تأثير غير مناسب على مخرجات التحليل الإحصائى للبيانات (Barnett & Lewis, ١٩٩٤: ٧) .

وتعرف الدرجات المتطرفة إحصائياً بأنها الدرجات غير المرغوب فيها ، وتقع بين $(-\infty, \mu - 3\sigma)$ ، وبين $(-\infty, \mu + 3\sigma)$ على طرفى توزيع الدرجات الكلية المشاهدة كما بالشكل التالى :-



شكل (١)

يبين موقع الدرجات المتطرفة على طرفى توزيع الدرجة الكلية المشاهدة

(Hekimoglu, ٢٠٠٥)

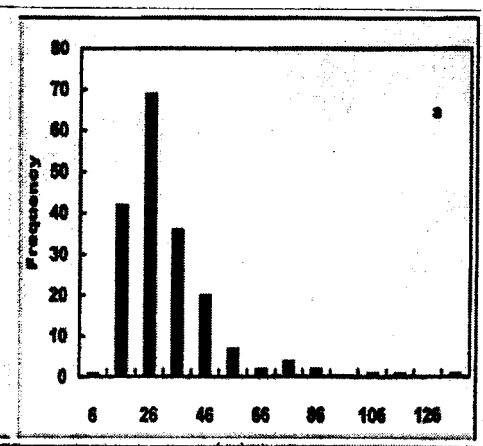
وقد جذبت الدرجات المتطرفة إهتمام العديد من الباحثين فى عدة مجالات مثل علم الفلك والكيمياء والفيزياء وعلم النفس والتربية والاقتصاد (Liu, ٢٠٠٥) .

والقـرار الذى يجب أن يتخذه الباحث بشأن الدرجات المتطرفة هو أحد ثلاثة احتمالات ويتلخص الاحتمـال الأول فـى تحرى تأثيـر الدرجات المتطرفة وذلك بتحليل البيـانات مع الاحتفاظ بتلك الدرجـات وتحليلها بـدون الاحتفـاظ بها ، ويشـار إلى التأثير غير المناسب للدرجات المتطرفة على نتائج الاختبارات الإحصائية بالمشاهدات المؤثرة Influential bservations وعلى سبيل المثال ربما يقرر الباحث حذف تلك المشاهدات المؤثرة من البيانات . والاحتمال الثانى هو المواءمة مع الدرجات المتطرفة ، وهذا يستتبع اختيار اختبارات إحصائية تتميز بالمنعة أو الضلاعة Robustness فيما يتعلق بالدرجات المتطرفة ، أو إجراء تحويل على البيانات . والاحتمال الثالث هو إعتبار الحالات المتطرفة هى ممثل لمجموعة لم تمثل فى العينة ، وربما يقرر عندئذ الباحث أن عينة جديدة يجب جمع بيانات بواسطتها اعتماداً على التكوين الطبقى الملائم أو المناسب Appropriate Stratification ، وبدلاً من ذلك ربما يتم دراسة الدرجات المتطرفة كحالات فردية مثيرة للاهتمام (Zijlstra, Ark & Sijtsma, ٢٠٠٧) .

ويمكن تمييز ثلاثة طرق تنشأ عنها الدرجات المتطرفة فى العينة (Barnett & Lewis, ١٩٩٤: ٣٣-٣٤; Liu & Zumbo, ٢٠٠٧) ، هى :-

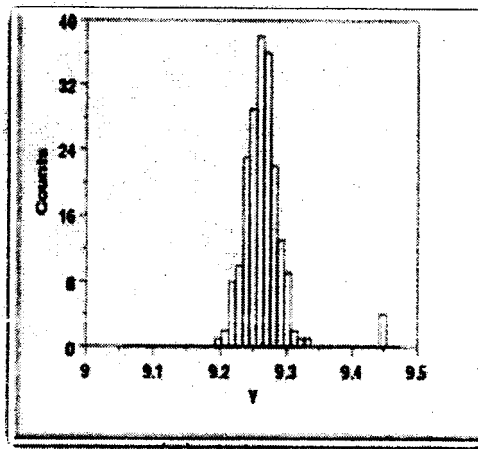
١. خطأ القياس Measurement Error : تنشأ الدرجات المتطرفة لأسباب محددة ، وتعزى إلى خطأ القراءة ، وخطأ التسجيل ، وخطأ الحساب فى البيانات ، وخطأ إدخال البيانات أو أخطاء إعداد وتجهيز البيانات للتحليل مثل الأخطاء المطبعية .
٢. خطأ التنفيذ فى جمع البيانات Execution Error : اختيار أفراد لا ينتمون للمجتمع المستهدف بالدراسة .
٣. الأخطاء غير المتوقعة المتصلة بالقياس من قبل المستجيبين مثل التخمين وعدم اليقظة بسبب التعب وسوء الاستجابة بسبب عدم فهم تعليمات الاستجابة على أداة القياس .

وهناك نوعين من التوزيع يتعلقان بالدرجات المتطرفة ، الأول يمثل توزيع متماثل للدرجات الكلية مع وجود درجات متطرفة على جانبى التوزيع ويسمى التوزيع المتماثل الملوث بالدرجات المتطرفة ، أما الثانى فهو توزيع غير متماثل Symmetric Outlier Contamination Distribution مع وجود درجات متطرفة على أحد جانبى التوزيع ويسمى التوزيع غير المتماثل الملوث بالدرجات المتطرفة Asymmetric Outlier Contamination Distribution كما بالشكلين (٢ ، ٣) الآتيين :



شكل(٢)

توزيع غير متماثل مع وجود درجات المتطرفة نقلاً عن (Liu, ٢٠٠٥)



شكل (٣)

توزيع متماثل مع وجود درجات المتطرفة نقلاً عن (Liu, ٢٠٠٥)

وعلى الرغم من التصدى لبحث الدرجات المتطرفة منذ ما يزيد عن (٢٠٠) عام إلا أن دراسة تأثيرها على تقدير الاختبارات الإحصائية المختلفة ، ودراسة الطرق الإحصائية المناسبة للتعامل معها مازال ميداناً خصباً للبحث كما يظهر من خلال النظرية الإحصائية للدرجات المتطرفة (Liu, ٢٠٠٥) .

فئات الاستجابة Response Categories

يشغل عدد فئات الاستجابة ذهن الباحثين باستمرار ، وقد يكون مصدراً للارتباك أو الحيرة ، لذا يتجه بعض الباحثين إلى اختيار عدد فئات الاستجابة الأقل رغبة فى سهولة التصحيح وإدخال البيانات وتحليلها وتفسيرها . ويفترض بعض الباحثين أن زيادة عدد فئات الاستجابة ربما يؤدى إلى إحساس المستجيبين بالتعب أثناء الاستجابة على البنود أو بالإرتباك والحيرة الذى بدوره قد يفضى إلى تخفيض أو تقليل معدل الاستجابة (Birkett, ٢٠٠٦) .

وبالتالى يضطر بعض الباحثين إلى تحويل سلم التصحيح القائم على أكثر من درجتين إلى سلم تصحيح ثنائى الدرجات ، أى تحويل مجال الدرجات المتصلة أو سلم الدرجات المتصلة المستعمل فى التصحيح إلى سلم ثنائى ، يؤدى إلى تقليص كبير لتباين الدرجات ، وبالتالى إلى انخفاض كبير فى معامل الثبات أو الاتساق الداخلى عند استعمال معادلة كيودر-ريتشاردسون ٢٠-KR مقارنة بقيمة معامل ألفا (أمحمد بوزيان ، ٢٠٠٨) .

وقد ناقش بعض الباحثين أهمية تضمين فئة وسيطة ، بأنه يسمح للمفحوصين أن يستجيبوا استجابة وسيطة أو محايدة ، وأن يكونوا أكثر تمييزاً فى إستجاباتهم ، وتكون درجات المقياس أكثر ثباتاً ، ويقبل المفحوصين أكثر على إكمال الاستجابة ويصبح المقياس أكثر تفضيلاً منهم . ومن ناحية ثانية إهتم بعض الباحثين بفكرة أن وجود إستجابة وسيطة بأنه يجعل المفحوصين أقل تمييزاً ويظهرون أنفسهم أكثر حياداً فى كثير من الأحيان ، بينما يجبر حذف الاستجابة الوسيطة المحايدة المفحوصين لأن يكونوا أكثر عمقاً ويظهر ذلك فى شكل تقديرات دقيقة (Ory & Wise, ١٩٩١) .

ويتفق بعض الباحثين على تفضيل المقياس الذى لا يتضمن الاستجابة الوسيطة لأنه يجبر المفحوصين على اتخاذ خيار واضح ، ولكنهم يثمنون تجاوز هذا التفضيل لو أن نقص الفئة الوسيطة يؤثر فى صدق وثبات الاستجابات . ويرون أن الحيادية تقود إلى بيانات

غامضة وغير محددة وبخاصة بين تلك المجموعات العرقية التى تقدر ثقافتها الاستجابات غير المباشرة (Reid, ١٩٩٠; Garland, ١٩٩١; Busch, ١٩٩٣) .

ومن أوجه الاهتمام الأخرى بأفضل عدد لبدائل الإستجابة أو تضمين فئة وسيطة هو أعمار المفحـوصين ، فصغار المفحوصين يكونون أكثر راحة مع أقل عدد من فئات الاستجابة التـى يمكـن أن تزيد من ثبات الدرجات لأن المفحوصين سوف يستجيبون أكثر إتساقاً ، وأن الصغـار لا يقـدرون على التمييز بدقـة المتطلب فى حالة الاستجابات الخماسية أو الأكثر من الرباعية ، ويميلون إلى الاستجالة المتوسطة حال توافر ذلك (Bourke& Frampton, ١٩٩٢) .

ومنذ أن قدم ليكرت لعمله ، حاول الكثير من الباحثين تحديد أفضل عدد لفئات الاستجابة ، وذلك من خلال اختبار العلاقة بين عدد فئات الاستجابة والاتساق الداخلى ، ولكن وجد تعارض بين النتائج فقد بين البعض ومنهم (Aiken, ١٩٨٣; Wong et al., ١٩٩٣) عدم وجود تأثير لعدد فئات الاستجابة على معامل ألفا ، والبعض بينوا تأثيراً وأوصوا بعدد من البدائل من بديلين إلى ثلاثة ومنهم (Matell & Jacoby, ١٩٧٢; Johnson et al., ١٩٨٢) ، وعدد أكبر من البدائل قد يصل إلى ثمانية عشر بديل (Champney & Marshall, ١٩٧٩) .

كما أن نتائج البحوث بشأن عدد بدائل الاستجابة مازالت حتى حينه متعارضة فستستمر حالة الاشتباك بين الباحثين فى ميدان القياس ، وبهذا يظل العدد الأمثل لفئات الاستجابة محوراً لبحوث ودراسات فى المستقبل .

دراسات وبحوث سابقة :

أولاً : دراسات وبحوث سابقة عن تأثير عدد فئات الإستجابة على تقدير معامل ألفا

هدفت دراسة "حسين العكام" (١٩٩٥) إلى فحص تأثير عدد فئات التدرج لمقياس الاتجاه نحو دراسة الفيزياء ذا بدائل (٣ ، ٤ ، ٥ ، ٦ ، ٧) على خصائصه السيكومترية . ومن بين ما بينته النتائج أن معامل ألفا يزداد بزيادة عدد فئات التدرج حتى الصورة خماسية التدرج ثم يقل فى الصورة سداسية التدرج ويعود للزيادة للصورة سباعية التدرج ، ولم يجد

الباحث فروقاً دالة بين المعاملات مؤكداً أن ثبات المقياس مستقلاً عن عدد البدائل ولا يتأثر بها .

وهدفت دراسة "تورة صالح المحيميد" (١٩٩٩) إلى التعرف على الفروق فى الثبات (طريقة إعادة الاختبار، طريقة معامل ألفا) والصدق (الصدق المرتبط بالمحك،الصدق العاملى) تبعاً لاختلاف تقديرات أوزان ليكرت واختلاف الفئة العمرية ، وقد استخدم مقياس Harris & Piers لمفهوم الذات لدى الطفل وهو ثنائى التقدير، وأصبحت التقديرات بالاضافة إلى التقدير الثنائى ثلاثة ، أربعة ، وخمسة تقديرات . ومن بين ما بينته النتائج زيادة قيم معامل ألفا بزيادة التقديرات من التقدير الثنائى إلى التقدير الخماسى .

أما دراسة "سعيد حسن عبد الفتاح" (٢٠٠٣) فقد هدفت إلى التعرف على مدى تأثير تعدد بدائل الاستجابة لأدوات القياس واختلاف المرحلة الدراسة على معاملات الثبات والصدق ، كما هدفت إلى محاولة التعرف على مدى اختلاف أساليب أو أنماط استجابة الأفراد فى ضوء تغاير عدد بدائل المقياس والمرحلة الدراسية ، ومن بين ما بينته النتائج أن قيم معامل ألفا تزداد بزيادة عدد بدائل الاستجابة وبفروق دالة إحصائية وذلك باستخدام اختبار Hakstain & Walen ، وأوصت الدراسة باستخدام المقياس خماسى التدرج .

هدفت دراسة "أوستر" (Oaster, ٢٠٠٤) إلى حساب معامل ألفا لمقياس Marlowe-Crowne للمرغوبية الاجتماعية لدى طلاب الجامعة ، وذلك بعد تعديل فئات الاستجابة ليكون عددها (٢ ، ٤ ، ٦ ، ٨ ، ١٠) . ومن بين ما بينته نتائج الدراسة تزايد تقدير معامل ألفا بزيادة عدد فئات الإستجابة من (٢) إلى (٨) . أما قيمة معامل ألفا المقابلة لعدد فئات الاستجابة (١٠) فقد كانت أقل من القيمة المقابلة لفئات الاستجابة (٨) .

كما أجرى "ونج" (Weng, ٢٠٠٤) دراسة من أهدافها فحص تأثير عدد بدائل الإستجابة (من ٣ إلى ٩ بدائل) على تقدير معامل ألفا لمقياس إتجاه المعلم نحو مهنة التدريس لدى الطلاب الملتحقين ببرنامج إعداد المعلم بجامعة تايوان الوطنية ، ومن بين ما بينته النتائج تأثير عدد البدائل على الاتساق الداخلى للمقياس ، وأوصت الدراسة بأنها لا تحفز إستخدام الفئات الزوجية دون الفئات الفردية ، وأنه لا داعى لزيادة عدد بدائل الإستجابة عن خمسة بدائل .

وأجرى "ويزر وشارما ونيدريك" (Weathers, Sharma & Niedrich,٢٠٠٥). دراسة هدفت إلى فحص تأثير عدد نقاط ليكرت وخصائص الأفراد على تقدير معامل ألفا ودقة الاستجابة والدور التوسيطى لتكرار الاستجابة على الرغم من اختلاف محتوى البند ، ومن بينته نتائج الدراسة أنه يوجد تأثير دال لعدد نقاط ليكرت على تقدير معامل ألفا ، وأن حجم هذا التأثير صغير ، وأنه قد تعزى الفروق فى معاملات ألفا إلى اختلاف خصائص الأفراد .

وهدفت دراسة "بيركيت" (Birkett, ٢٠٠٦) إلى تقويم تأثير تنوع فئات الاستجابة على تقدير معامل ألفا ومعدل الاستجابة باستخدام استبيان Walston &Walston للصحة النفسية وموضع الضبط متعدد الأبعاد سداسى الاستجابة ، وذلك بعد تعديل الاستجابات ليصبح فى ثلاثة صور فئات الاستجابة لها (٢ ، ٦ ، ١٤) . ومن بين ما بينته النتائج زيادة قيمة معامل ألفا بزيادة عدد فئات الاستجابة من (٢) إلى (٦) ، ووجد أن قيمة معامل الثبات الأعلى تقابل الاستجابة السداسية ، وبفحص الفروق بين معاملات الثبات المستقلة للتأكد من دلالة الزيادة فى معامل ألفا باستخدام اختبار "كرامر" Cramer وجد أنها دالة فقط بين قيمة معامل ألفا المقابلة للاستجابات السداسية وقيمة معامل ألفا المقابلة للاستجابة الثنائية على أحد أبعاد الاستبيان الثلاثة ، وأوصت الدراسة باستخدام فئة الاستجابة الخماسية أو السباعية بدلاً من فئة الاستجابة السداسية .

وهدفت دراسة "محمد حسين سعيد" (٢٠٠٧) إلى التعرف على أثر اتجاه المفردة وعدد بدائل الاستجابة على ثبات أدوات القياس من نوع ليكرت ، ومن بين ما بينته النتائج اختلاف بعض قيم معاملات الثبات باختلاف اتجاه المفردة واختلاف عدد بدائل الاستجابة ، وأن قيم الثبات المرتفعة تحققت فى حالة العبارات الموجبة وفى حالة البدائل السباعية .

وهدفت دراسة "شوميا" (Chomeya, ٢٠١٠) إلى تحرى جودة مجموعة من الاختبارات النفسية (دافعية الإنجاز ، موضع الضبط ، الاتجاه نحو تعاطى الكحوليات) لدى عينة من طلاب الجامعة باستخدام صورتين من كل اختبار الأولى تستخدم فئة الاستجابة الخماسية والثانية تستخدم فئة الاستجابة السداسية ، ومن بين ما بينته النتائج تتجاوز قيمة معامل ألفا لاختبار دافعية الإنجاز واختبار الاتجاه نحو تعاطى الكحوليات فى حالة الاستجابة السداسية قيمة معامل ألفا فى حالة الاستجابة الخماسية ، والعكس تتجاوز قيمة معامل ألفا لاختبار موضع الضبط فى حالة الاستجابة الخماسية قيمة معامل فى حالة الاستجابة السداسية.

بينما هدفت دراسة "أدلسون وماكوتش" (Adelson & McCoach, in press) إلى فحص تأثير فئات الاستجابة الرباعية وفئات الاستجابة الخماسية على تقدير معامل ألفا لمقياس اتجاه تلاميذ المرحلة الابتدائية نحو الرياضيات وذلك بحساب معامل ألفا للصورة الأولى من المقياس والتى تستخدم فئات الاستجابة الرباعية وحساب معامل ألفا للصورة الثانية من المقياس والتى تستخدم فئات الاستجابة الخماسية . ومن بين ما بينته النتائج تجاوز قيمة معامل ألفا للصورة الثانية للمقياس قيمة المعامل للصورة الأولى ، كما ترشح نتائج الدراسة استخدام فئة الاستجابة الخماسية الفردية عن فئة الاستجابة الرباعية .

ثانياً : دراسات وبحوث سابقة عن تأثير الدرجات المتطرفة على تقدير معامل ألفا

أجرى "ليو" (Liu, ٢٠٠٥) دراسة هدفت إلى تحرى تأثير الدرجات المتطرفة على تقدير معامل ألفا باستخدام أسلوب مونت كارلو Monte Carlo Simulation Technique ، وذلك بقياس دقة تقدير المعامل باستخدام مؤشر التحيز Bias Index وقياس قابلية تشتت المعامل باستخدام مؤشر الفاعلية Efficiency Index . ومن بين ما بينته النتائج أن تقدير معامل ألفا لا يتأثر بالتلوث المتماثل للدرجات المتطرفة Symmetric Outlier Contamination ، فى حين أن التلوث غير المتماثل للدرجات المتطرفة يؤثر فى تضخم تقدير المعامل . وأن تقديرات حجم التأثير للفروق بين أزواج المتوسطات للتحيز والفاعلية فى حالة التلوث المتماثل تناظرها أحجام تأثير صغيرة ، أما تقديرات حجم التأثير بين أزواج المتوسطات للتحيز والفاعلية فى حالة التلوث غير المتماثل تناظرها أحجام تأثير كبيرة جداً . والنتائج بعامة تبين أن تقديرات معامل ألفا لا تتأثر بالتلوث المتماثل للدرجات المتطرفة.

وهدفت دراسة "ليو وزمبو" (Liu & Zumbo, ٢٠٠٧) إلى تحرى تأثير الدرجات المتطرفة على تقدير معامل ألفا لمقاييس التماثل البصرى المستخدمة لتقويم جودة الحياة والسعادة والعديد من المتغيرات السيكولوجية الأخرى . ومن بين ما بينته نتائج الدراسة أن تقدير معامل ألفا لا يتأثر بالتلوث المتماثل للدرجات المتطرفة ، فى حين أن التلوث غير المتماثل للدرجات المتطرفة يؤثر فى تضخم تقدير المعامل .

بينما هدفت دراسة "ليو وو وزمبو" (Liu, Wu & Zumbo, ٢٠١٠) إلى دراسة تأثير الدرجات المتطرفة على تقدير معامل ألفا فى حالة المقاييس رتبية الاستجابة ، وقد تم جمع البيانات عن طريقة المحاكاة الإحصائية باستخدام أسلوب مونت كارلو ، ومن بين ما بينته النتائج تضخم تقدير معامل ألفا بسبب وجود درجات متطرفة ، كما تأثر التقدير بعدد فئات الاستجابة المقابلة للبند . وأن كفاءة تقديرات معامل ألفا تضخمت أيضاً وتأثرت بعدد بدائل الإستجابة . وأيضاً وجد أن هناك إنخفاضاً متحيزاً Downward Biased فى تقديرات معامل ألفا عند حذف الدرجات المتطرفة ، وبعامة وجد أن التضخم فى تقدير ألفا فى حالة وجود الدرجات المتطرفة يؤدى إلى تحيز إيجابى .

تعقيب على البحوث والدراسات السابقة :

١- قلة عدد البحوث والدراسات العربية السابقة فى موضوع البحث الحالى .

٢- الندرة الشديدة والملحوظة فى البحوث والدراسات السابقة التى اهتمت بدراسة تأثير الدرجات المتطرفة على الاتساق الداخلى لأدوات القياس على المستوى الدولى .

٣- هدفت بعض الدراسات فقط إلى فحص تأثير عدد فئات الاستجابة الزوجية على تقدير معامل ألفا مثل دراسة (Birkett, ٢٠٠٦); (Oaster, ٢٠٠٤) .

٤- هدفت بعض الدراسات فقط إلى فحص تأثير عدد فئات الاستجابة الفردية على تقدير معامل ألفا مثل دراسة (محمد حسين سعيد ، ٢٠٠٧)

٥- أشارت نتائج بعض الدراسات والبحوث إلى استخدام فئات الاستجابة الفردية وعلى وجه الخصوص فئة الاستجابة الخماسية أو السباعية مثل دراسة (سعيد حسن عبد الفتاح ، ٢٠٠٣) ؛(Adelson & McCoach, in press); (Birkett, ٢٠٠٦).

٦- تتباين النتائج بشأن أفضلية الاستجابات الفردية عن الزوجية (Weng, ٢٠٠٤; Chomeya, ٢٠١٠) .

٧- تتباين النتائج بشأن تأثير عدد البدائل على تقدير الثبات بطريقة ألفا واستقلال تقدير المعامل عن عدد البدائل ، ووتمثل الفئة الأولى من الدراسات فى دراسة (نورة صالح المحيميد ، ١٩٩٩) ، أما الفئة الثانية فمنها دراسة (حسين العكام ، ١٩٩٥).

٨- تتفق نتائج دراسة (Liu, ٢٠٠٥) ونتائج دراسة (Liu & Zumbo, ٢٠٠٧) على أن تقدير معامل ألفا لا يتأثر بالتلوث المتماثل للدرجات المتطرفة ، فى حين أن التلوث غير المتماثل للدرجات المتطرفة يؤثر فى تضخم تقدير المعامل .

٩- يختلف البحث الحالى عن الدراسات والبحوث العربية السابقة فى أنه يتجه نحو دراسة تأثير عدد فئات الاستجابة الفردية والزوجية على تقدير معامل ألفا ، وكذلك دراسة تأثير الدرجات المتطرفة ، ودراسة دلالة الفروق بين قيم معامل ألفا المستقلة وحجم التأثير المقابل لهذه الفروق .

فروض البحث :

فى ضوء المفاهيم الأساسية للبحث والدراسات والبحوث السابقة ، تسعى الدراسة الحالية إلى اختبار الفرض التالية :-

١. تختلف تقديرات معامل ألفا باختلاف عدد فئات الاستجابة عند الاحتفاظ بالقيم المتطرفة .
٢. لا توجد فروق دالة إحصائياً بين كل زوج من معاملات ألفا المستقلة فى حالة الاحتفاظ بالدرجات المتطرفة .
٣. تختلف تقديرات معامل ألفا باختلاف عدد القيم المتطرفة المحذوفة وعدد فئات الاستجابة .
٤. لا توجد فروق دالة إحصائياً بين كل زوج من معاملات ألفا المستقلة فى حالة حذف نفس عدد الدرجات المتطرفة .

إجراءات الدراسة :

أولاً : عينة الدراسة :

تم تحديد حجم عينة الدراسة الحالية فى ضوء نتائج دراسات (Bonett, ٢٠٠٢, ٢٠٠٣) . وقد بلغ حجم العينة الكلية (٧٣٢) معلماً بمتوسط عمر قدره (٣١،٥) سنة

تم حساب الحجم الأمثل لحجم العينة فى البحث الحالى باستخدام المعادلة :-

$$n = \frac{[2k(k-1)]\left[Z_{\alpha/2} + Z_{\beta}\right]^2}{\ln(\tilde{\delta}) + 2}$$

وانحرف معيارى قدره (١،١) سنة ، أخذت منها عينات فرعية حسب مقتضيات تحليل النتائج – عدد فئات الاستجابة– فى البحث الحالى كما بالجدول التالى :

جدول (٢)

توزيع عينة الدراسة

على الصور المختلفة لمقياس اتجاهات المعلمين نحو تكنولوجيا التعليم تبعاً لفئات التدرج

ثنائى التدرج	ثلاثى التدرج	رباعى التدرج	خماسى التدرج	سداسى التدرج	سباعى التدرج
١٤٠	١٣٨	١٥٤	١٥١	١٤٣	١٧٩

ثانياً : أداة القياس : مقياس اتجاهات المعلمين نحو تكنولوجيا التعليم :

أعد المقياس (لطفى الخطيب – ٢٠٠٠م) وهو مقياس ثلاثى التدرج "أوافق – غير متأكد – لا أوافق" ، وقد لجأ الباحث إلى إستخدامه لتوافر شرط تجانس البنود أو أحادية البعد فى بنود المقياس Unidimensionality of Test Items . وقد بلغت عينة إعداد المقياس (١٣٩) معلماً ، ويتكون من (٤٠) بنداً . وحسب معامل الثبات بتطبيق معادلة ألفا كرونباخ فكانت (٠،٨٨) .

أما فى الدراسة الحالية فقد أعاد الباحث وضع المقياس فى ستة صور تختلف فقط فى عدد البدائل مع الاحتفاظ بمحتوى أو مضمون العبارة دون تغيير ، وذلك كالآتى :-

١- الصورة الأولى (ثنائية التدرج أو التقدير) ، ولها بديلين للاجابة هما : موافق ، أرفض .

٢- الصورة الثانية (ثلاثية التدرج) ، ولها ثلاثة بدائل للاجابة هم : موافق ، غير متأكد ، أرفض .

٣- الصورة الثالثة (رباعية التدرج) ، ولها أربعة بدائل هم : موافق بشدة ، موافق ، أرفض ، أرفض بشدة .

حيث k هى عدد بنود الاختبار ، α هى مستوى الدلالة الإحصائية (احتمال الحصول على الخطأ من النوع الأول) ، β احتمال الحصول على الخطأ من النوع الثانى ، $(Z_{\alpha/2}, Z_{\beta})$ هما نقطتين على منحنى التوزيع الطبيعى يناظران القيم الاحتمالية $(\alpha/2, \beta)$ على الترتيب ، وقد بلغ حجم العينة طبقاً للمعادلة $(n = ١٣١)$.

٤- الصورة الرابعة (خماسية التدرج) ، ولها خمسة بدائل هم : موافق بشدة ، موافق ، غير متأكد ، أرفض ، أرفض بشدة .

٥- الصورة الخامسة (سداسية التدرج) ، ولها ستة بدائل هم : موافق بدرجة كبيرة جداً ، موافق بدرجة كبيرة ، موافق ، أرفض ، أرفض بدرجة كبيرة ، أرفض بدرجة كبيرة جداً .

٦- الصورة السادسة (سباعية التدرج) ، ولها سبعة بدائل هم : تنطبق بدرجة كبيرة جداً ، تنطبق بدرجة كبيرة ، تنطبق ، تنطبق بدرجة متوسطة ، لا تنطبق ، لا تنطبق بدرجة كبيرة ، لا تنطبق بدرجة كبيرة جداً .

وللتأكد من تجانس استجابات العينات الفرعية المختلفة قام الباحث بحساب معامل الإلتواء للدرجات الكلية كما يلى :-

جدول (٢)

معاملات الإلتواء لتوزيع درجات العينات الفرعية المختلفة

صور المقياس	معامل الإلتواء
الصورة الثانية	٠،٢٠٩ -
الصورة الثلاثية	٠،١٩١-
الصورة الرباعية	٠،١٨٨ -
الصورة الخماسية	٠،٢٢١-
الصورة السداسية	٠،٢٣٩-
الصورة السباعية	٠،١٩٧-

ويتبين من الجدول (٢) السابق أن معاملات الإلتواء جميعها سالبة وتقع فى الجوار المباشر للقيمة (-٠،٢) ، وهذا يعنى أن توزيعات الدرجة الكلية على مقياس اتجاهات المعلمين نحو تكنولوجيا التعليم متجانسة تبعاً لمعامل الإلتواء .

وقد روعى تقديم شرحاً وافياً لتعليمات المقياس أثناء التطبيق ، كما روعى استبعاد الاستجابات غير المكتملة ، واستبعاد الاستجابات ذات النمط الواحد (اختيار استجابة واحدة فقط على جميع البنود) قبل عملية إدخال البيانات على الحاسب الآلى .

ثالثاً : المعالجة الإحصائية :

١- لتحديد الدرجات المتطرفة استخدمت الحزمة الإحصائية SPSS .

٢- لحساب قيم معامل ألفا لكرونباخ استخدمت الحزمة الإحصائية SPSS .

٣- لحساب دلالة الفروق بين قيم معاملات ألفا المستقلة استخدم اختبار W Test .

٤- لحساب تقدير حجم التأثير المقابل للفروق بين قيم معاملات ألفا المستقلة استخدم اختبار Liu & Weng Test .

نتائج الدراسة وتفسيرها :

نتائج الفرض الأول :

ينص الفرض الأول على " تختلف تقديرات معامل ألفا باختلاف عدد فئات الاستجابة عند الاحتفاظ بالقيم المتطرفة" ، ولاختبار هذا الفرض قام الباحث بحساب تقديرات معامل ألفا فى حالة الاحتفاظ بعدد (١٠ = j) من الدرجات المتطرفة منها خمس قيم منخفضة وخمس قيم مرتفعة ، فى حالة فئات استجابة عددها (٦ =k) كما بالجدول التالى :-

جدول (٣)

تقديرات معامل ألفا عند الاحتفاظ بالدرجات المتطرفة

عدد فئات الاستجابة / عدد الدرجات المتطرفه	K = ٢	K = ٣	K = ٤	K = ٥	K = ٦	K = ٧
J = ١٠	٠،٧٨٠	٠،٨٥٤	٠،٧١٠	٠،٩٣٩	٠،٨٩٦	٠،٩٥٠

ويتبين من الجدول (٣) السابق أن قيم معامل ألفا تزداد باطراد بزيادة عدد فئات الاستجابة عند الاحتفاظ بالدرجات المتطرفة التى عددها (١٠ = J) ، وتعد قيمة معامل ألفا المصورة سداسية التدرج أفضل تقدير للمعامل مقارنة بالصورة ثنائية ورباعية التدرج ، وأيضاً تعد قيمة معامل ألفا للصورة سباعية التدرج أفضل تقدير للمعامل مقارنة بالصورة ذات التدرج الثلاثى والتدرج الخماسى .

ويمكن تفسير هذه النتيجة بأن زيادة عدد بدائل الاستجابة ربما يعطى المستجيب الفرصة للتعبير بدقة أكثر عما يشعر به ، مما يجعله يتخذ قراراً بشأن البديل المناسب الذى

يعبر عما يشعر به أمام كل بند من بنود المقياس ، وهذا بدوره يزيد من التباين الحقيقى على حساب تباين الخطأ ، بينما العدد الصغير من البدائل ربما يجبر المستجيب على اختيار البديل الذى لا يعبر تماماً عن شعوره بشأن محتوى كل بند من البنود ، ويكون اختياره له مضطراً لعدم وجود بديل أكثر دقة ، وهذا بدوره يؤدى إلى الوقوع فى أخطاء مقصودة أو غير مقصودة بسبب عدم تلبية المقياس لاحتياجاته أو انسجامه مع آرائه ، وفى كلتا الحالتين سوف يزداد تباين الخطأ على حساب التباين الحقيقى وهذا يعنى انخفاض تقدير معامل ألفا Underestimation Index .

وتتفق هذه النتيجة مع ما توصلت إليه نتائج دراسة (نورة المحيميد ، ١٩٩٩) ، ودراسة (سعيد حسن عبد الفتاح ، ٢٠٠٣) ، ودراسة (Weathers, Sharma & Niedrich, ٢٠٠٥) ، ودراسة (Chomeya, ٢٠١٠) ، ودراسة (Liu, Wu & Zumbo, ٢٠١٠) ، ودراسة (Adelson & McCoach, in press) . وقد بينت هذه الدراسات التأثير الإيجابى لعدد بدائل الإستجابة على تقدير معامل ألفا .

نتائج الفرض الثانى :

ينص الفرض الثانى على "لا توجد فروق دالة إحصائياً بين كل زوج من معاملات ألفا المستقلة فى حالة الاحتفاظ بالدرجات المتطرفة" . ولاختبار هذا الفرض قام الباحث بحساب الفروق بين كل زوج من معاملات ألفا المستقلة فى حالة حذف نفس عدد الدرجات المتطرفة باستخدام اختبار⊗ Feldt's W Test ، ومن ثم استخدم اختبار⊗⊗ Delta Test لتحديد مستوى حجم التأثير للفروق ذات الدلالة الإحصائية كما بالجداول التالية :-

---

⊗ اختبار Feldt's W Test وضعه "يونارد س. فلت" عام ١٩٦٩م لقياس دلالة الفروق بين زوج من معاملات ألفا كرونباخ المستقلة ، ويتخذ الاختبار المعادلة الاتية :-

$$W=\frac{(1-\alpha_1)}{(1-\alpha_2)},(\alpha_1 \prec \alpha_2)$$

والاختبار يتبع تقاربياً توزيع اختبار "ف" F Central Distribution بدرجة حرية المقام (N_1-1) ، ودرجة حرية البسط (N_2-1) ، حيث (N_1) هو حجم العينة الأولى ، وحجم العينة الثانية (N_2) .

جدول (٤)

دلالة الفروق بين قيم معاملات ألفا المستقلة وحجم التأثير المقابل لها

فى حالة الإحتفاظ بالدرجات المتطرفة

فئات الاستجابة		K = ٢ (α=.780)	K = ٣ (α=.854)	K = ٤ (α=.710)	K = ٥ (α=.939)	K = ٦ (α=.896)	K = ٧ (α=.950)
K = ٢ (α=.780)	W	-					
	Δ	-					
K = ٣ (α=.854)	W	١،٥١	-				
	Δ	٠،١٢	-				
K = ٤ (α=.710)	W	١،٣١	١،٩٩	-			
	Δ	٠،٠٨	٠،٢١	-			
K = ٥ (α=.939)	W	*٣،٦١	*٢،٣٩	*٤،٧٥	-		
	Δ	٠،٤٠	٠،٢٧	٠،٤٧	-		
K = ٦ (α=.896)	W	*٢،١٢	١،٤٠	*٢،٧٩	١،٧٠	-	
	Δ	٠،٢٣	٠،١٠	٠،٣١	٠،١٦	-	
K = ٧ (α=.950)	W	*٤،٤	*٢،٩٢	**٥،٨	١،٢٢	*٢،٠٨	-
	Δ	٠،٤٥	٠،٣٣	٠،٥٣	٠،٠٦	٠،٢٢	-

* قيمة الإحصائى W دالة عند مستوى (٠،٠٥) .

** قيمة الإحصائى W دالة عند مستوى (٠،٠١) .

⊗⊗ اختبار Delta Test وضعه "ليو وونج" (Liu & Weng) عام ٢٠٠٩م "لتقدير حجم التأثير للفروق ذات الدلالة الإحصائية بين أزواج معاملات ألفا المستقلة . ويتخذ الاختبار المعادلة الآتية :-

$$\Delta = \frac{0.5\left[\ln(1-\alpha_1) - \ln(1-\alpha_2)\right]}{\sqrt{\left[\frac{n}{2(n-1)}\right]}} = \frac{0.5\left[\ln(W)\right]}{\sqrt{\left[\frac{n}{2(n-1)}\right]}}, (\alpha_1 \prec \alpha_2)$$

ويحدد الاختبار ثلاثة مستويات لحجم التأثير هى : حجم تأثير كبير $\Delta = 0.8$ ، وحجم تأثير متوسط $\Delta = 0.5$ ، وحجم تأثير صغير $\Delta = 0.2$.

ويتبين من الجدول (٤) السابق وجود فروق دالة إحصائياً بين قيم معاملى ألفا المناظرين لفئات الاستجابة الثلاثية والرباعية لصالح معامل ألفا المناظر لفئة الاستجابة الثلاثية ويقابله حجم تأثير صغير .

ويتبين أيضاً وجود فروق دالة إحصائياً بين قيمة معامل ألفا المناظرة لفئة الاستجابة الخماسية وقيم معامل ألفا المناظرة لفئة الاستجابة الثنائية والثلاثية والرباعية لصالح معامل ألفا المناظر لفئة الاستجابة الخماسية ، وتناظر تلك الفروق أحجام تأثير صغيرة .

ويتبين وجود فروق دالة إحصائياً بين قيمة معامل ألفا المناظرة لفئة الاستجابة السداسية وقيم معامل ألفا المناظرة لفئة الاستجابة الثنائية والرباعية لصالح معامل ألفا المناظر لفئة الاستجابة السداسية ، وتناظر تلك الفروق أحجام تأثير صغيرة . ولا توجد فروق دالة بين قيمة معامل ألفا المناظر لفئة الاستجابة السداسية وقيمة معامل ألفا المناظر لفئات الاستجابة الثلاثية والخماسية .

ويتبين أيضاً وجود فروق دالة إحصائياً بين قيمة معامل ألفا المناظرة لفئة الاستجابة السباعية وقيم معامل ألفا المناظرة لفئة الاستجابة الثنائية والثلاثية والرباعية والسداسية لصالح معامل ألفا المناظر لفئة الاستجابة السباعية ، وتناظر تلك الفروق أحجام تأثير (صغير ، صغير ، متوسط ، صغير) . ولا توجد فروق دالة بين قيمة معامل ألفا المناظر لفئة الاستجابة السباعية وقيمة معامل ألفا المناظر لفئة الاستجابة الخماسية .

وتشير نتائج الجدول (٤) السابق بعامة أن قيم معامل ألفا المقابلة لفئات الاستجابة ذات العدد الفردى أظهرت معاملات دالة إحصائياً مقارنة بتلك المقابلة لفئات الاستجابة ذات العدد الزوجى . وأن أفضل عدد لبدائل الاستجابة هو ٥ . وهذه النتيجة تتفق مع نتيجة دراسة (سعيد حسن عبد الفتاح ، ٢٠٠٣ ؛ Birkett, ٢٠٠٦ ;Weng, ٢٠٠٤) بأن هناك زيادة طفيفة فى تقدير معامل بعد استخدام فئة استجابة أعلى من ٥ . أما قيم أحجام التأثير المناظرة للفروق الدالة إحصائياً فجميعها يقع فى المستوى الصغير ماعدا حجم التأثير المناظر للفرق بين قيمتى معامل ألفا المناظرتين لفئة الإستجابة السباعية وفئة الاستجابة الرباعية حيث يقع فى المستوى المتوسط .

نتائج الفرض الثالث :

ينص الفرض الثالث على " تختلف تقديرات معامل ألفا باختلاف عدد القيم المتطرفة المحذوفة وعدد فئات الاستجابة " ، ولاختبار هذا الفرض قام الباحث بحساب تقديرات معامل ألفا فى حالة أعداد متباينة من القيم المتطرفة (j) (المرتفعة والمنخفضة) ، وعدد فئات استجابة عددها (k = ٦) ، وذلك كما بالجدول التالى :-

جدول (٥)

تقديرات معامل ألفا عند حذف الدرجات المتطرفة

عدد فئات الاستجابة / عدد الدرجات المتطرفة	K = ٢	K = ٣	K = ٤	K = ٥	K = ٦	K = ٧
J = ٢	٠،٧٧٣	٠،٨٤٨	٠،٧١٠	٠،٩٣٦	٠،٨٩٢	٠،٩٤٩
J = ٤	٠،٧٦٠	٠،٨٤٣	٠،٧١٠	٠،٩٣٣	٠،٨٨٩	٠،٩٤٩
J = ٦	٠،٧٤٥	٠،٨٣٧	٠،٧١٠	٠،٩٢٩	٠،٨٨٣	٠،٩٤٨
J = ٨	٠،٧٢٧	٠،٨٢٣	٠،٧١٠	٠،٩٢٦	٠،٨٧٦	٠،٩٤٨
J= ١٠	٠،٧٠٧	٠،٨٠٧	٠،٧١١	٠،٩٢٢	٠،٨٦٨	٠،٩٤٧

ويتبين من الجدول (٥) السابق أن تضخم قيم معامل ألفا يقل كلما زاد عدد القيم المتطرفة المحذوفة فى حالة فئات الاستجابة الثنائية والثلاثية والخماسية والسداسية .

ويتبين أن تضخم قيم معامل ألفا يظل كما هو كلما زاد عدد القيم المتطرفة المحذوفة فى حالة عدد فئات الاستجابة الرباعية ، ماعدا فى حالة حذف جميع القيم المتطرفة فقد زاد تضخم معامل ألفا .

وفى حالة فئة الاستجابة السباعية يتبين من الجدول (٥) أيضاً أن تضخم قيم معامل ألفا يظل كما هو فى حالة حذف قيمتين أو أربعة قيم ، ثم ينخفض ويثبت فى حالة حذف ستة أو ثمانية قيم متطرفة ، وينخفض التضخم أكثر بعد حذف عشرة قيم متطرفة .

وتشير النتائج المبينة فى الجدول (٥) السابق إلى أن أقل قيم لمعامل ألفا ظهرت عند استخدام فئة الإستجابة الرباعية وأعلى قيم للمعامل كانت عند استخدام فئة الإستجابة السباعية . وتعد قيمة معامل ألفا المقابلة لفئة الاستجابة السداسية أفضل تقدير للمعامل مقارنة بالفئات الزوجية الأخرى وهى الثنائية والرباعية ، وأيضاً تعد قيمة معامل ألفا المقابلة لفئة الاستجابة السباعية أفضل تقدير للمعامل مقارنة بالفئات الفردية الأخرى وهى الثلاثية والخماسية .

وبهذا فإن نتيجة اختبار الفرض الثالث تتفق مع نتائج البحوث والدراسات والأدبيات السابقة التى بينت عدم وجود قرار قاطع بشأن العدد الأمثل لفئات الإستجابة التى يمكن إستخدامها ويمكن من خلالها الحصول على إستجابات نقية بدرجة كبيرة وتضمن جودة أداة القياس ومناسبتها لجمع بيانات عن المتغير المستهدف بالقياس ، ومن تلك الدراسات والأدبيات Gable & Wolf, ١٩٩٣: ٣٤-٣٥); (Oaster, ١٩٨٩, ٢٠٠٤); (Preston & Colman, ٢٠٠٠); (Weathers, Sharma & Niedrich, ٢٠٠٥); (Maydeu-Olivares et al., ٢٠٠٩); (Chomeya, ٢٠١٠); (Adelson & McCoach, in press) .

نتائج الفرض الرابع :

ينص الفرض الرابع على "لا توجد فروق دالة إحصائياً بين كل زوج من معاملات ألفا المستقلة فى حالة حذف نفس عدد الدرجات المتطرفة" . ولاختبار هذا الفرض قام الباحث بحساب الفروق بين كل زوج من معاملات ألفا المستقلة فى حالة حذف نفس عدد الدرجات المتطرفة باستخدام اختبار Feldt's W Test ، ومن ثم استخدم اختبار Delta Test لتحديد مستوى حجم التأثير للفروق ذات الدلالة الإحصائية كما بالجداول التالية :-

جدول (٦)
دلالة الفروق بين قيم معاملات ألفا المستقلة وحجم التأثير المقابل لها
(فى حالة الدرجات المحذوفة ٢=j)

فئات الاستجابة		K = ٢ (α=.773)	K = ٣ (α=.848)	K = ٤ (α=.710)	K = ٥ (α=.936)	K = ٦ (α=.892)	K = ٧ (α=.949)
K = ٢ (α=.773)	W	-					
	Δ	-					
K = ٣ (α=.848)	W	١،٤٩	-				
	Δ	٠،١٢	-				
K = ٤ (α=.710)	W	١،٢٨	١،٩١	-			
	Δ	٠،٠٨	٠،٢٠	-			
K = ٥ (α=.936)	W	*٣،٥٥	*٢،٣٨	*٤،٥٣	-		
	Δ	٠،٣٩	٠،٢٦	٠،٤٦	-		
K = ٦ (α=.892)	W	*٢،١	١،٤١	*٢،٦٨	١،٦٩	-	
	Δ	٠،٢٣	٠،١٠	٠،٣٠	٠،١٦	-	
K = ٧ (α=.949)	W	*٤،٤٥	*٢،٩٨	**٥،٦٩	١،٢٥	*٢،١٢	-
	Δ	٠،٤٥	٠،٣٣	٠،٥٣	٠،٠٧	٠،٢٣	-

قيمة الإحصائى W دالة عند مستوى (٠،٠٥) .
قيمة الإحصائى W دالة عند مستوى (٠،٠١) .

ويتبين من الجدول (٦) السابق وجود فروق دالة إحصائياً بين قيم معاملى ألفا المناظرين لفئات الاستجابة الثلاثية والرباعية لصالح معامل ألفا المناظر لفئة الاستجابة الثلاثية ويقابله حجم تأثير صغير .

ويتبين أيضاً وجود فروق دالة إحصائياً بين قيمة معامل ألفا المناظرة لفئة الاستجابة الخماسية وقيم معامل ألفا المناظرة لفئة الاستجابة الثنائية والثلاثية والرباعية لصالح معامل ألفا المناظر لفئة الاستجابة الخماسية ، وتناظر تلك الفروق أحجام تأثير صغيرة .

ويتبين وجود فروق دالة إحصائياً بين قيمة معامل ألفا المناظرة لفئة الاستجابة السداسية وقيم معامل ألفا المناظرة لفئة الاستجابة الثنائية والرباعية لصالح معامل ألفا المناظر لفئة الاستجابة السداسية ، وتناظر تلك الفروق أحجام تأثير صغيرة . ولا توجد فروق دالة بين قيمة معامل ألفا المناظر لفئة الاستجابة السداسية وقيمة معامل ألفا المناظر لفئات الاستجابة الثلاثية والخماسية .

ويتبين أيضاً وجود فروق دالة إحصائياً بين قيمة معامل ألفا المناظرة لفئة الاستجابة السباعية وقيم معامل ألفا المناظرة لفئة الاستجابة الثنائية والثلاثية والرباعية والسداسية لصالح معامل ألفا المناظر لفئة الاستجابة السباعية ، وتناظر تلك الفروق أحجام تأثير (صغير ، صغير ، متوسط ، صغير) . ولا توجد فروق دالة بين قيمة معامل ألفا المناظر لفئة الاستجابة السباعية وقيمة معامل ألفا المناظر لفئة الاستجابة الخماسية .

جدول (٧)

دلالة الفروق بين قيم معاملات ألفا المستقلة وحجم التأثير المقابل لها

(فى حالة الدرجات المحذوفة ٤=j)

فئات الاستجابة		K = ٢ (α=.760)	K = ٣ (α=.843)	K = ٤ (α=.710)	K = ٥ (α=.933)	K = ٦ (α=.889)	K = ٧ (α=.949)
K = ٢ (α=.760)	W	-					
	Δ	-					
K = ٣ (α=.843)	W	١،٥٣	-				
	Δ	٠،١٣	-				
K = ٤ (α=.710)	W	١،٢١	١،٨٥	-			
	Δ	٠،٠٦	٠،١٩	-			
K = ٥ (α=.933)	W	*٣،٥٨	*٢،٣٤	*٤،٣٣	-		
	Δ	٠،٣٩	٠،٢٦	٠،٤٥	-		
K = ٦ (α=.889)	W	*٢،١٦	١،٤١	*٢،٦١	١،٦٦	-	
	Δ	٠،٢٣	٠،١١	٠،٢٩	٠،١٥	-	
K = ٧ (α=.949)	W	*٤،٧١	*٣،٠٨	**٥،٦٩	١،٣١	*٢،١٨	-
	Δ	٠،٤٧	٠،٣٤	٠،٥٣	٠،٠٨	٠،٢٤	-

قيمة الإحصائى W دالة عند مستوى (٠،٠٥) .

قيمة الإحصائى W دالة عند مستوى (٠،٠١) .

ويتبين من الجدول (٧) السابق وجود فروق دالة إحصائياً بين قيمة معامل ألفا المناظرة لفئة الاستجابة الخماسية وقيم معامل ألفا المناظرة لفئة الاستجابة الثنائية والثلاثية والرباعية لصالح معامل ألفا المناظر لفئة الاستجابة الخماسية ، وتناظر تلك الفروق أحجام تأثير صغيرة .

ويتبين وجود فروق دالة إحصائياً بين قيمة معامل ألفا المناظرة لفئة الاستجابة السداسية وقيم معامل ألفا المناظرة لفئة الاستجابة الثنائية والرباعية لصالح معامل ألفا المناظر لفئة الاستجابة السداسية ، وتناظر تلك الفروق أحجام تأثير صغيرة . ولا توجد فروق دالة بين قيمة معامل ألفا المناظر لفئة الاستجابة السداسية وقيمة معامل ألفا المناظر لفئات الاستجابة الثلاثية والخماسية .

ويتبين أيضاً وجود فروق دالة إحصائياً بين قيمة معامل ألفا المناظرة لفئة الاستجابة السباعية وقيم معامل ألفا المناظرة لفئة الاستجابة الثنائية والثلاثية والرباعية والسداسية لصالح معامل ألفا المناظر لفئة الاستجابة السباعية ، وتناظر تلك الفروق أحجام تأثير (صغير ، صغير ، متوسط ، صغير) . ولا توجد فروق دالة بين قيمة معامل ألفا المناظر لفئة الاستجابة السباعية وقيمة معامل ألفا المناظر لفئة الاستجابة الخماسية .

جدول (٨)
دلالة الفروق بين قيم معاملات ألفا المستقلة وحجم التأثير المقابل لها
(فى حالة الدرجات المحذوفة ٦=j)

فئات الاستجابة		K = ٢ (α=.745)	K = ٣ (α=.837)	K = ٤ (α=.710)	K = ٥ (α=.929)	K = ٦ (α=.883)	K = ٧ (α=.948)
K = ٢ (α=.745)	W	-					
	Δ	-					
K = ٣ (α=.837)	W	١،٥٦	-				
	Δ	٠،١٤	-				
K = ٤ (α=.710)	W	١،١٤	١،٧٨	-			
	Δ	٠،٠٤	٠،١٨	-			
K = ٥ (α=.929)	W	*٣،٥٩	*٢،٣٠	*٤،٠٨	-		
	Δ	٠،٣٩	٠،٢٥	٠،٤٣	-		
K = ٦ (α=.883)	W	*٢،١٨	١،٣٩	*٢،٤٨	١،٦٥	-	
	Δ	٠،٢٤	٠،١٠	٠،٢٨	٠،١٥	-	
K = ٧ (α=.948)	W	*٤،٩	*٣،١٣	**٥،٥٨	١،٣٧	*٢،٢٥	-
	Δ	٠،٤٨	٠،٣٥	٠،٥٢	٠،١٠	٠،٢٥	-

* قيمة الإحصائى W دالة عند مستوى (٠،٠٥) .
** قيمة الإحصائى W دالة عند مستوى (٠،٠١) .

ويتبين من الجدول (٨) السابق وجود فروق دالة إحصائياً بين قيمة معامل ألفا المناظرة لفئة الاستجابة الخماسية وقيم معامل ألفا المناظرة لفئة الاستجابة الثنائية والثلاثية والرباعية لصالح معامل ألفا المناظر لفئة الاستجابة الخماسية ، وتناظر تلك الفروق أحجام تأثير صغيرة .

ويتبين وجود فروق دالة إحصائياً بين قيمة معامل ألفا المناظرة لفئة الاستجابة السداسية وقيم معامل ألفا المناظرة لفئة الاستجابة الثنائية والرباعية لصالح معامل ألفا المناظر لفئة الاستجابة السداسية ، وتناظر تلك الفروق أحجام تأثير صغيرة . ولا توجد فروق دالة بين قيمة معامل ألفا المناظر لفئة الاستجابة السداسية وقيمة معامل ألفا المناظر لفئات الاستجابة الثلاثية والخماسية .

ويتبين أيضاً وجود فروق دالة إحصائياً بين قيمة معامل ألفا المناظرة لفئة الاستجابة السباعية وقيم معامل ألفا المناظرة لفئة الاستجابة الثنائية والثلاثية والرباعية والسداسية لصالح معامل ألفا المناظر لفئة الاستجابة السباعية ، وتناظر تلك الفروق أحجام تأثير (صغير ، صغير ، متوسط ، صغير) . ولا توجد فروق دالة بين قيمة معامل ألفا المناظر لفئة الاستجابة السباعية وقيمة معامل ألفا المناظر لفئة الاستجابة الخماسية .

جدول (٩)

دلالة الفروق بين قيم معاملات ألفا المستقلة وحجم التأثير المقابل لها

(فى حالة الدرجات المحذوفة j=٨)

فئات الاستجابة		K = ٢ (α=.727)	K = ٣ (α=.823)	K = ٤ (α=.710)	K = ٥ (α=.926)	K = ٦ (α=.876)	K = ٧ (α=.948)
K = ٢ (α=.727)	W	-					
	Δ	-					
K = ٣ (α=.823)	W	١،٥٤	-				
	Δ	٠،١٣	-				
K = ٤ (α=.710)	W	١،٠٦	١،٦٤	-			
	Δ	٠،٠٢	٠،١٥	-			
K = ٥ (α=.926)	W	*٣،٦٩	*٢،٣٩	*٣،٩٢	-		
	Δ	٠،٤٠	٠،٢٧	٠،٤٢	-		
K = ٦ (α=.876)	W	*٢،٢٠	١،٤٣	*٢،٣٤	١،٦٨	-	
	Δ	٠،٢٤	٠،١١	٠،٢٦	٠،١٦	-	
K = ٧ (α=.948)	W	**٥،٢٥	*٣،٤٠	**٥،٥٨	١،٤٢	*٢،٣٨	-
	Δ	٠،٥٠	٠،٣٧	٠،٥٢	٠،١١	٠،٢٦	-

* قيمة الإحصائى w دالة عند مستوى (٠،٠٥) .

** قيمة الإحصائى w دالة عند مستوى (٠،٠١) .

يتبين من الجدول (٩) السابق وجود فروق دالة إحصائياً بين قيمة معامل ألفا المناظرة لفئة الاستجابة الخماسية وقيم معامل ألفا المناظرة لفئة الاستجابة الثنائية والثلاثية والرباعية لصالح معامل ألفا المناظر لفئة الاستجابة الخماسية ، وتناظر تلك الفروق أحجام تأثير صغيرة.

ويتبين وجود فروق دالة إحصائياً بين قيمة معامل ألفا المناظرة لفئة الاستجابة السداسية وقيم معامل ألفا المناظرة لفئة الاستجابة الثنائية والرباعية لصالح معامل ألفا المناظر لفئة الاستجابة السداسية ، وتناظر تلك الفروق أحجام تأثير صغيرة . ولا توجد فروق دالة بين قيمة معامل ألفا المناظر لفئة الاستجابة السداسية وقيمة معامل ألفا المناظر لفئات الاستجابة الثلاثية والخماسية .

ويتبين أيضاً وجود فروق دالة إحصائياً بين قيمة معامل ألفا المناظرة لفئة الاستجابة السباعية وقيم معامل ألفا المناظرة لفئة الاستجابة الثنائية والثلاثية والرباعية والسداسية لصالح معامل ألفا المناظر لفئة الاستجابة السباعية ، وتناظر تلك الفروق أحجام تأثير (متوسط ، صغير ، متوسط ، صغير) . ولا توجد فروق دالة بين قيمة معامل ألفا المناظر لفئة الاستجابة السباعية وقيمة معامل ألفا المناظر لفئة الاستجابة الخماسية .

جدول (١٠)

دلالة الفروق بين قيم معاملات ألفا المستقلة وحجم التأثير المقابل لها

(فى حالة الدرجات المحذوفة ١٠=j)

فئات الاستجابة		K = ٢ (α=.707)	K =٣ (α=.807)	K = ٤ (α=.711)	K = ٥ (α=.922)	K = ٦ (α=.868)	K = ٧ (α=.947)
K = ٢ (α=.707)	W	-					
	Δ	-					
K = ٣ (α=.807)	W	١،٥٢	-				
	Δ	٠،١٣	-				
K = ٤ (α=.711)	W	١،٠١	١،٥	-			
	Δ	٠،٠٠٤	٠،١٢	-			
K = ٥ (α=.922)	W	*٣،٧٦	*٢،٤٧	*٣،٧١	-		
	Δ	٠،٤٠	٠،٢٨	٠،٤٠	-		
K = ٦ (α=.868)	W	*٢،٢٢	١،٤٦	*٢،١٩	١،٦٩	-	
	Δ	٠،٢٤	٠،١٢	٠،٢٤	٠،١٦	-	
K = ٧ (α=.947)	W	**٥،٥٣	*٣،٦٤	*٥،٤٥ *	١،٤٧	*٢،٤٩	-
	Δ	٠،٥٢	٠،٣٩	٠،٥٢	٠،١٢	٠،٢٨	-

* قيمة الإحصائى W دالة عند مستوى (٠،٠٥) .

** قيمة الإحصائى W دالة عند مستوى (٠،٠١) .

يتبين من الجدول (١٠) السابق وجود فروق دالة إحصائياً بين قيمة معامل ألفا المناظرة لفئة الاستجابة الخماسية وقيم معامل ألفا المناظرة لفئة الاستجابة الثنائية والثلاثية والرباعية لصالح معامل ألفا المناظر لفئة الاستجابة الخماسية ، وتناظر تلك الفروق أحجام تأثير صغيرة .

ويتبين وجود فروق دالة إحصائياً بين قيمة معامل ألفا المناظرة لفئة الاستجابة السداسية وقيم معامل ألفا المناظرة لفئة الاستجابة الثنائية والرباعية لصالح معامل ألفا المناظر لفئة الاستجابة السداسية ، وتناظر تلك الفروق أحجام تأثير صغيرة . ولا توجد فروق دالة بين قيمة معامل ألفا المناظر لفئة الاستجابة السداسية وقيمة معامل ألفا المناظر لفئات الاستجابة الثلاثية والخماسية .

ويتبين أيضاً وجود فروق دالة إحصائياً بين قيمة معامل ألفا المناظرة لفئة الاستجابة السباعية وقيم معامل ألفا المناظرة لفئة الاستجابة الثنائية والثلاثية والرباعية والسداسية لصالح معامل ألفا المناظر لفئة الاستجابة السباعية ، وتناظر تلك الفروق أحجام تأثير (متوسط ، صغير ، متوسط ، صغير) . ولا توجد فروق دالة بين قيمة معامل ألفا المناظر لفئة الاستجابة السباعية وقيمة معامل ألفا المناظر لفئة الاستجابة الخماسية .

وبهذا تشير نتائج الفرض الرابع بعامة إلى أن قيم معامل ألفا المقابلة لفئات الاستجابة ذات العدد الفردى أظهرت معاملات دالة إحصائياً مقارنة بتلك المقابلة لفئات الاستجابة ذات العدد الزوجى . وأن أفضل عدد لبدائل الاستجابة هو ٥ . أما قيم أحجام التأثير المناظرة للفروق الدالة إحصائياً فأكثرها يقع فى المستوى الصغير واقلها يقع فى المستوى المتوسط .

كما يلاحظ أيضاً اتساق النتائج التى تم الحصول عليها من خلال الجداول من (٦ – ١٠) فيما يتعلق بدلالة الفروق بين قيم معاملات ألفا المستقلة ومستويات حجم التأثير المقابلة لتلك الفروق حيث تكون للفروق أكبر دلالة بعد حذف جميع الدرجات المتطرفة وتناظرها أحجام تأثير تقع فى المستوى المتوسط والذى يبتعد بالنتائج عن الأخطاء المصاحبة ، وربما يجعلها عندئذ قابلة للتعميم .

وتتفق النتيجة التى تم التوصل إليها من خلال إختبار الفرض الرابع ، مع الدراسات والأدبيات التى توصى باستخدام فئة إستجابة وسيطة أو محايدة ، وبالتالى يتجه البحث الحالى ليوصى باستخدام مدى واسع من فئات الإستجابة الفردية التى تزيد عن ٣ وعلى وجه التحديد استخدام المقاييس ذات التدريج خماسى الاستجابة فى قياس سمات الشخصية والجوانب الوجدانية ، ومن أمثلة تلك الدراسات والبحوث والأدبيات السابقة (Ory &

Wise, ١٩٩١); (Weng, ٢٠٠٤); (Birkett, ٢٠٠٦); (Chomeya, ٢٠١٠); (Adelson & McCoach, in press) ؛ ودراسة (محمد حسين سعيد ، ٢٠٠٧) .

توصيات الدراسة :

فى ضوء نتائج الدراسة الحالية التى تم استخلاصها ، يورد الباحث التوصيات الآتية :-

١- ضرورة مراجعة ملف أو مصفوفة البيانات بعد إدخالها للتأكد من عدم وجود أخطاء فى رصد الدرجات المناظرة للاستجابات .

٢- ضرورة التأكد من فهم المستجيبين لتعليمات الاجابة على بنود أدوات القياس قبل البدء فى الاستجابة .

٣- ضرورة التأكيد على المستجيبين أن البيانات يتم جمعها لغرض البحث العلمى وهى ليست نوع من التقييم فى العمل وذلك للتخلص مع عامل المرغوبية الاجتماعية وتزييف الاستجابات ، ومن ثم الحصول على استجابات نقية .

٤- اختيار التوقيت المناسب لتطبيق أدوات القياس بحيث يتم تجنيب المستجيبين التعب أو الارهاق .

٥- ضرورة حذف الدرجات الكلية الشاذة أو المتطرفة قبل حساب معامل ألفا لكرونباخ ، وذلك بترتيب الدرجات الكلية للمستجيبين تصاعدياً أو تنازلياً . ويمكن للسهولة استخدام برنامج التحليل الإحصائى SPSS لفرز الدرجات المتطرفة مباشرة .

٦- استخدام المقاييس ذات التدريج خماسى الاستجابة فى قياس سمات الشخصية والجوانب الوجدانية .

٧- ضرورة اختبار الباحث توافر الفرضيات التى يستند إليها معامل ألفا قبل استخدامه ، مثلما يختبر توافر الفرضيات فى بياناته قبل استخدام اختبار "ت" لدراسة دلالة الفروق بين المتوسطات ، وفى حالة عدم توافر الفرضيات يمكن أن يستخدم الصيغة المعدلة من معامل ألفا التى اقترحها (Christmann & Aelst, ٢٠٠٦) لتناسب العينات الصغيرة وتحقق المنعة ، أو يستبدل معامل ألفا بالطرق الحديثة التى قدمها البحث العلمى فى مجال القياس ومنها : نظرية قابلية التعميم Generalizability Theory ، ونمذجة المعادلات البنائية Structural Equation Modeling ، ودالة معلومات الاختبار Test Information Function (Sijtsma, ٢٠٠٩c)

المراجـــع:

أولاً: المراجع العربية:

١- أحمد الرفاعى غنيم (١٩٩١) . تعميم معامل ألفا لحساب ثبات المقاييس ذات المفردات المتجانسة . مجلة كلية التربية . جامعة عين شمس . العدد (١٥) ، ٢٠٧ – ٢٣٨ .

٢- أمحمد بوزيان لزعر تيغزه (٢٠٠٨). البنية المنطقية لمعامل ألفا لكرونباخ ومدى دقته فى تقدير الثبات فى ضوء افتراضات نماذج القياس . مجلة العلوم التربوية والدراسات الإسلامية . كلية التربية . جامعة الملك سعود . ٢١(١) ، ١ – ٥٢ .

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

٤- سعيد حسن عبد الفتاح (٢٠٠٣). مدى اختلاف الخصائص السيكومترية لأدة القياس فى ضوء تغاير عدد بدائل الاستجابة والمرحلة الدراسة : دراسة حالة – مقياس ليكرت . رسالة ماجستير غير منشورة . كلية التربية . جامعة أم القرى .

٥- لطفى الخطيب (٢٠٠٠) . اتجاهات المعلمين فى محافظة إربد نحو تكنولوجيا التعليم . المجلة العربية للتربية . ٢٠(١) ، ١١١-١٣٠ .

٦- 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 .

٧- نورة صالح المحيميد (١٩٩٩) . أثر إختلاف عدد بدائل الاستجابة فى أوزان ليكرت على معاملات الثبات والصدق . رسالة دكتوراه غير منشورة . كلية التربية . جامعة الملك سعود .

ثانياً: المراجع الأجنبية:

8- Adelson, J. L. & McCoach, D. B. (in press). Measuring the Mathematical Attitudes of Elementary Students: The Effects of A 4-Point or 5-Point Likert-Type Scale. Educational and Psychological Measurement.

9- Ahmann, J. S. & Glock, M. D. (1971). Measuring and Evaluating Educational Achievement. Boston: Allyn and Bacon.

10- Aiken, L. R. (1983). Number of Response Categories and Statistics on a Teacher Rating Scale. Educational and Psychological Measurement. 43, 397-401.

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

12- Barnett, V. and Lewis, T. (1994). Outliers in Statistical Data. (3rd ed.). John Wiley & Sons.

13- Birkett, N. J. (2006). Selecting the Number of Response Categories for a Likert-Type Scale. Proceedings of the Survey Research Methods Section, American Statistical Association. 488-492.

14- Bonett, D. J. (2002). Sample Size Requirements for Testing and Estimating Coefficient Alpha. Journal of Educational and Behavioral Statistics. 27(4), 335 – 340 .

15- Bonett, D. J. (2003). Sample Size Requirements for Comparing Two Alpha Coefficient. Applied Psychological Measurement. 27(1), 72- 74 .

16- Bourke, S. & Frampton, J. (1992). Assessing the Quality of School Life: Some Technical Considerations. Paper Presented at the Annual Conference of the Australian Association for Research in Education. Deakin University. Geelong, Australia.

17- Busch, M. (1993). Using Likert Scales in L2 Research: A Researcher Comments. TESOL Quarterly. 27, 733-726.

18- Caruso, J. C. (2000). Reliability Generalization of the NEO Personality Scales. Educational and Psychological Measurement, 60, 236–254.

19- Champney, H. & Marshall, H. (1979). Optimal Refinement of the Rating Scale. Journal of Applied Psychology. 23, 323-331.

20- Charter, R. A. & Feldt, L. S. (1996). Testing the Equality of Two Alpha Coefficients. Perceptual and Motor Skills. 82, 763-768.

21- Chomeya, R. (2010). Quality of Psychology Test Between Likert Scale 5 and 6 Points. Journal of Social sciences. 6(3), 399-403.

22- Christmann, A. & Van Aelst, S. (2006). Robust estimation of Cronbach's alpha. Journal of Multivariate Analysis. 97(7), 1660-1674.

23- Cortina, J. M. (1993). What is Coefficient Alpha? An Examination of Theory and Applications. Journal of Applied Psychology. 78(1), 98-104.

24- Cronbach, L. J. & Shavelson, R. J. (2004). My Current Thoughts on Coefficient Alpha and Successor Procedures. Educational and Psychological Measurement. 64(3), 391 – 418 .

25- Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), Educational Measurement (3rd ed., pp. 105-146). New York: Macmillan.

26- Feldt, L. S. & Charter, R. A. (2003). Estimating the reliability of a test split into two parts of equal or unequal length. Psychological Methods. 8(1), 102-109.

27- Feldt, L. S. & Kim, S. (2006). Testing the Differences between Two Alpha Coefficient with Small Samples of Subjects and Raters. Educational and Psychological Measurement. 66(4), 589 – 600 .

28- Gable, R. K. & Wolf, M. B. (1993).Instrument Development in the Affective Domain. (2nd ed.). Boston, MA: Kluwer Academic.

29- Garland, R. (1991). The Mid-Point on a Likert Rating Scale: Is It Desirable?. Marketing Bulletin. 2, 66-70.

30- Gilmer, J. S. & Feldt, L. S. (1983). Reliability Estimation for A Test with Parts of Unknown Lengths. Psychometrika. 48(1), 99-111.

31- Graham, J. M. (2006). Congeneric and (Essentially) Tau-Equivalent Estimates of Score Reliability: What they are and How to Use Them. Educational and Psychological Measurement. 66(6),930-944.

32- Grubbs, F.E.(1969). Procedures for Detecting Outlying Observations in Samples. Technometrics. 11, 1–21.

33- Hawkins, D. (1980). Identification of Outliers. Chapman and Hall.

34- Janda, L. H. (1998). Psychological Testing: Theory and Applications. Boston: Allyn and Bacon.

35- Johnson, S. M,; Smith, P.C. & Tucker, S. M. (1982). Response Format of the Job Descriptive Index: Assessment of Reliability and Validity by the Multitrait-multimethod Matrix. Journal of Applied Psychology. 67, 500-505.

36- Knapp, T. R. (2002). The Reliability of Measuring Instruments. (2nd ed.). Vancouver, B. C.: Edgeworth Laboratory for Quantitative Educational and Behavioral Science Series.

37- Ko, Y.-H. (1994). A Research for a Better Likert Point-Scale for Mental Health Questionnaire. Psychological Testing. 41, 55-72.

38- Komorita, S. S. & Graham, W. K. (1995). Number of Scale Points and the Reliability of Scales. Educational and Psychological Measurement. 25, 987-995.

39- Kubiszyn, T. & Borich, G. (1990). Educational Testing and Measurement: Classroom Application and Practice. (3rd ed.). London: Scott, Foresman and Company.

40- Lind, J. C. & Zumbo, B. D. (1993). The Continuity Principle in Psychological Research: An Introduction to Robust Statistics. Canadian Psychology. 34, 407-414.

41- Liu, H.-Y. & Weng, L.-J. (2009). An Effect Size Index for Comparing Two Independent Alpha Coefficients. British Journal of Mathematical and Statistical Psychology. 62, 385-400.

42- Liu, Y. (2005). Documenting the Impact of Outliers on Cronbach's Coefficient Alpha Estimate of Reliability: Informing How One Should Interpret the Extant Literature and/or One's Own Research Findings. Unpublished Master Thesis. The University of British Columbia. Canada.

43- Liu, Y.; Wu, A. D. & Zumbo, B.D. (2010). The Impact of Outliers on Cronbach's Coefficient Alpha Estimate of Reliability: Ordinal/Rating Scale Item Responses. Educational and Psychological Measurement. 70(1), 5-21.

44- Liu, Y. & Zumbo, B.D. (2007).The Impact of Outliers on Cronbach's Coefficient Alpha Estimate of Reliability: Visual Analogue Scales. Educational and Psychological Measurement. 67(4), 620-634.

45- Masters, G. R. (1994). The Relationship between Number of Response Categories and Reliability of Likert-Type Questionnaire. Journal of Educational Measurement. 11, 49-53.

46- Matell, M. S. & Jacoby, J. (1972). Is There An Optimal Number of Alternatives for Likert-Scale Items? Effects of Testing Time and Scale Properties. Journal of Applied Psychology. 56(6), 506-509.

47- Maydeu-Olivares, A.; Kramp, U.; Garcia-Forero, C.; Gallardo-Pujol, D. Coffman, D. (2009). The Effect of Varying the Number of Response Alternatives in Rating Scales: Experimental Evidence from Intra-Individual Effects. Behavior Research Methods. 41(2), 295-308.

48- McKelvie, S. J. (1998). Graphic Rating Scale : How Many Categories? . British Journal of Psychology. 69, 185-202.

49- Miller, M. B. (1995). Coefficient Alpha: A Basic Introduction from the Perspectives of Classical Test Theory and Structural Equation Modeling. Structural Equation Modeling. 2, 255-273.

50- Oaster, T. R. F. (1989). Number of Alternatives Per Choice Point and Stability of Likert-Type Scales. Perceptual and Motor Skills. 68, 459-550.

51- Oaster, T. R. F. (2004). Coefficient Alpha Reliability of Likert-Type Scales as a Function of the Number of Alternatives Per Choice Point. Paper Presented at the Annual Meeting of the American Psychological Association (New Orleans, Louisiana, April)

52- Ory, J. C. & Wise, S. L. (1991). Attitude Change Measured by Scales with 4 and 5 Response Options. Paper Presented at the Meeting of the National Council on Measurement in Education, Chicago, IL

53- Payne, D. A. (1992). Measuring and Evaluating Educational Outcomes. New York: Maxwell Macmillan International.

54- 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.

55- Qualls, A. L. (1995). Estimating the Reliability of a Test Containing Multiple Item Formats. Applied Measurement in Education. 8(2), 111 – 120.

56- Rae, G. (2006). Correcting Coefficient Alpha for Correlated Errors: Is ... a Lower Bound to Reliability?. Applied Psychological Measurement. 30(1), 56 - 59.

57- Raykov, T. (1997a). Estimation of Composite Reliability for Congeneric Measures. Applied Psychological Measurement. 21, 173-184.

58- Raykov, T. (1997b). Scale Reliability, Cronbach's Coefficient Alpha, and Violation of Essential Tau-Equivalence with Fixed Congeneric Components. Multivariate Behavioral Research. 32, 329-353.

59- Reid, J. (1990). The Dirty Laundry of ESL Survey Research. TESOL Quarterly. 24, 323-338.

60- Reinhardt, B. (1996). Factors Affecting Coefficient Alpha: A Mmini Monte Carlo Study. In Thompson, B. (ed.), Advances in Social Science Methodology. Texas A&M University. 4, 3–20.

61- Salih, F. A. (1994). An Approximate Likelihood Ratio Test for Coefficient Alpha. The Egyptian Journal of Educational Evaluation. 4(1), 201 -223 .

62- Sijtsma, K. (2009a). Correcting Fallacies in Validity, Reliability, and Classification. International Journal of Testing. 9, 167-194.

63- Sijtsma, K. (2009b). On the Use, Misuse, and the Very Limited Usefulness of Cronbach's Alpha. Psychometrika. 74(1), 107-120.

64- Sijtsma, K. (2009c). Reliability Beyond Theory and Into Practice. Psychometrika. 74(1), 169-173.

65- Streiner, D. L. (2003a). Starting at the Beginning: An Introduction to Coefficient Alpha and Internal Consistency. Journal of Personality Assessment. 80, 99-103.

66- Streiner, D. L. (2003b). Being Inconsistent about Consistency: When Coefficient Alpha Does and Doesn't Matter. Journal of Personality Assessment. 80, 217-222.

67- Thompson, B. & Snyder, P. A. (1998). Statistical Significance and Reliability Analyses in Recent JCD Research Articles. Journal of Counseling and Development. 76, 436-441

68- Vehkalahti, K. (2000). Reliability of Measurement Scales: Tarkkonen's General Method Supersedes Cronbach's Alpha. Unpublished Doctoral Dissertation. Faculty of Social Sciences. University of Helsinki.

69- Weathers, D.; Sharma, S. & Niedrich, R. W. (2005). The impact of the number of scale points, dispositional factors, and the status quo decision heuristic on scale reliability and response accuracy . Journal of Business Research. 58(11), 1516-1524.

71- Weng, L.-J. (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.

70- Wilkinson, L. & The Task Force on Statistical Inference (1999). Statistical Methods in Psychology Journals: Guidelines and Explanations. American Psychologist. 54, 594-604.

71- Wong, C.-S.; Chuen, K.-C. & Fung, M.-Y. (1993). Differences between Odd and Even Number of Response Scales: Some Empirical Evidence. Chinese Journal of Psychology. 35, 75-86.

72- Yan, M.; Alejandro, G. D. V.; Hui, Z. & Tu, X. M. (2010). A U-Statistics-based Approach for Modeling Cronbach Coefficient Alpha Within a Longitudinal Data Setting. Statistics in Medicine. 29(6), 659-670.

73- Yuan, K.-H.; Guarnaccia, C. A. & Hayslip, B., Jr. (2003). A Study of the Distribution of Sample Coefficient Alpha with Hopkins Symptom Checklist: Bootstrap versus Asymptotic. Educational and Psychological Measurement. 63(1), 5-23.

74- Zijlstra, W.P.; Ark, L.A. van der & Sijtsma, K. (2007). Outlier Detection in Test and Questionnaire Data. Multivariate Behavioral Research, 42(3), 531-555.

75- Zijlstra, W.P., Ark, L.A. van der, & Sijtsma, K. (2008). Outlier Detection in the Medical Questionnaire Rising and Sitting Down (QR\&S). In K. Shigemasu, A. Okada, T. Imaizumi, & T. Hoshino (Eds.), New Trends in Psychometrics (pp. 595-604). Tokyo: Universal Academy Press.

76- Zumbo, B. D. (1999). A Glance at Coefficient Alpha with an Eye towards Robustness Studies: Some Mathematical Notes and a Simulation Model (Paper No. ESQBS-99-1). Prince George, B.C.: University of Northern British Columbia. Edgeworth Laboratory for Quantitative Behavioral Science.

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