مجلة جامعة الملك سعود، م٢١، العلوم التربوية والدراسات الإسلامية (٣)، ص ص ٦٣٧-٦٨٨، الرياض (٢٠٠٩م/١٤٣٠هـ)

البنية المنطقية لمعامل ألفا لكرونباخ، ومدى دقته في تقدير الثبات في ضوء افتراضات نماذج القياس

أمحمد تيغزة

أستاذ، قسم علم النفس، كلية التربية، جامعة الملك سعود،

الرياض، المملكة العربية السعودية

(قدم للنشر في ٢٧/ ٤/ ١٤٢٨هـ؛ قبل للنشر في ٢٩/١٠/ ١٤٢٨هـ)

ملخص البحث. تستهدف الدراسة الحالية معالجة ثلاثة أسئلة أساسية: ١- ما هي البنية المنطقية التي ينطوي عليها معامل ألفا لكرونباخ، وما متضمناتها؟ ٢- ما هي نماذج القياس أو الحالات التي يؤدي فيها استعمال معامل ألفا إلى تقدير دقيق للثبات، وتلك التي يسفر استخدامه فيها عن تقدير متدن أو تضخيمي لمعامل الثبات الحقيقي؟ ٣- ما هي معاملات الاتساق البديلة التي تعطي تقديرا أدق لمعامل الثبات عندما لا تتوفر الافتراضات assumptions التي يتطلبها استعمال معامل ألفا في بيانات القياس؟

فيما يتعلق بالاجابة عن السؤال الأول، لقد اتضح أن تباين الدرجات الكلية للمقياس (بالمقارنة مع مجموع تباين الفقرات)، وطول الاختبار أو عدد الفقرات، يمارسان تأثيرا قويا في قيم معامل ألفا. ومن متضمنات ذلك أن العينة المستعملة ينبغي أن تكون غير متجانسة في السمة المقاسة، وأن معامل ألفا ليس مؤشرا نقيا لاتساق فقرات المقياس.

ولمعالجة السؤال الثاني، تم التطرق إلى أربعة نماذج للقياس وهي: ١- النموذج المتوازي The parallel model؛ ٢- نموذج "طاو"- المترادف أو المتكافيء The tau-equivalent model؛ ٣- نموذج "طاو" المترادف أو المتكافيء في الأساس The essentially tau-equivalent model؛ ٤- وأخيرا النموذج المتقارب The congeneric model.

إن الافتراضات التي يقوم عليها معامل ألفا هي ذاتها افتراضات نموذج القياس المترادف في الأساس. ولذلك يعطي معامل ألفا تقديرا دقيقا للثبات عندما تتوفر في بيانات القياس مسلمات هذا النموذج (وأيضا في ظل النموذج الأول والثاني السابقين). أما إذا انتهكت إحدى افتراضات نموذج "طاو" المترادف في الأساس أو بعضها، فإن استعمال معامل ألفا يسفر عن تقدير منخفض لمعامل الثبات الحقيقي، وعن تقدير متضخم للثبات عند غياب شرط استقلال الارتباطات بين درجات الخطأ. وبتعبير آخر، عندما تستجيب بيانات القياس للنموذج المتقارب الذي يبدو أكثر تحررا وواقعية في افتراضاته، فإن استعمال معامل ألفا يتمخض عن الحد الأدنى فقط لتقدير الثبات الحقيقي لدرجات للمقياس.

أما بخصوص الإجابة عن السؤال الثالث ، فقد تم التطرق إلى بعض معاملات الاتساق الأخرى التي تعطي تقديرا أدق للثبات من معامل ألفا ، عند توفر افتراضات النموذج المتقارب في بيانات القياس ، أوعند غياب بعض مسلمات نموذج "طاو" المترادف في الأساس. وفي هذا السياق . تم التطرق باقتضاب إلى معامل ثيتا (Theta Coefficient (θ ، ومعامل أوميجا (Omega Coefficient (Ω ، ومعامل ثبات المفهوم (CR: Construct Reliability) ، ومعامل أوميجا الموزونة Weighted Omega. Ω_w ؛ كما تم توضيح كيفية حساب معامل ألفا ، ومعامل ثبات المفهوم ، ومعامل أوميجا الموزونة بتوظيف بيانات مثال واقعي.

وفي الخاتمة ، تم التطرق إلى الاستنتاجات الختامية ، كما تم اقتراح جملة من التوصيات.

المقدمـــــة

الحمد لله رب العالمين والصلاة والسلام على سيدنا محمد وعلى آله وصحبه أجمعين.

إن الخوض في دراسة معامل ألفا لكرونباخ ، كطريقة من طرق تقدير الثبات من منظور الاتساق الداخلي ، يقتضي منطقيا توضيح دلالة مصطلح "الثبات" ، لإبراز أهمية مفهوم الاتساق Consistency في تعريف "الثبات" ، وكذلك لأن مفهوم الاتساق يشكل لحمة وسدى معامل ألفا لتقدير الثبات.

يقضي التعريف المتداول لمفهوم الثبات في كتب التقويم والقياس التربوي ، وكتب مناهج البحث في المكتبة العربية ، بأن الثبات يتمثل في الحصول على نفس النتائج عند إعادة تطبيق مقياس أو اختبار أو أداة مرتين (أو أكثر) في ظروف متماثلة . على نفس الأفراد أو العينة. فمثلا ، يعرف جابر عبد الحميد جابر واحمد خيري كاظم (١٩٩٠ ، ص: ٢٧٦_٢٧٧) في كتابهما عن مناهج البحث في التربية وعلم النفس ، "ثبات" الاختبار بأن " يعطي نفس النتائج إذا ما استخدم الاختبار أكثر من مرة تحت ظروف متماثلة" .

ويعرف ماهر إسماعيل صبري ومحب محمود كامل الرافعي (٢٠٠١ ، ص: ٢٨٥) " الثبات" بأن " يعطي الاختبار نفس النتائج عند تكرار تطبيقه في قياس نفس الشيء أكثر من مرة وفي ظروف تطبيقية مشابهة". كما يعرف علي أحمد سيد وأحمد محمد سالم (٢٠٠٤ ، ص: ١٧٩) الثبات تعريفا مماثلا لكن بصياغة أكثر إجرائية. ففي تصورهما أن "الاختبار الثابت هو الاختبار الذي لو طبق على مجموعة معينة ثم أعيد تطبيقه على مجموعة أخرى متماثلة يعطي نفس النتائج".[١]

وتنطوي هذه الطريقة الشائعة في تعريف الثبات على مشكلة منطقية تقوم على فكرة أن الخاص يعرف العام ، أو على توضيح الأصل بالاحتكام إلى الفرع. فتعريفات الثبات السابقة هي ذاتها تعريفات نوع من أنواع الثبات أو طريقة من طرق تقديره وهو تعريف مفهوم الثبات بطريقة الإعادة ، غير أن هذا النوع من الثبات يستهدف قياس أخطاء القياس (الأخطاء العشوائية) المنبثقة عن عدم استقرار السمة المقاسة ، أو

التي مصدرها البعد الزمني المتبنى عند الإعادة. لكنها تقصر دون تقدير أخطاء القياس الناجمة عن الافتقار للتكافؤ (كما هو الشأن في الثبات بطريقة الصور المتكافئة)، أو تلك الناجمة عن مدى الافتقار إلى الاتساق (كما هو الشأن في الثبات عن طريق التجزئة النصفية، والثبات عن طريق التجانس أو الاتساق الداخلي). يمثل التعريف المتداول لمفهوم الثبات - إذن - طريقة واحدة من طرق تقدير الثبات دون الطرق الأخرى، مع العلم بأن الطرق الأخرى التي لا يمثلها هذا التعريف الشائع تركز على تقدير الأخطاء العشوائية الناجمة عن الافتقار للتكافؤ والافتقار للاتساق؛ وهي مصادر الأخطاء التي لا تقوى على تقديرها طريقة الإعادة التي شكلت أساسا أو أرضية للتعريفات المتداولة عن الثبات.

كما أن تعريف الثبات بناء على تشابه النتائج عند الإعادة غير دقيق، إذ يكفي أن يحتفظ الأفراد برتبهم، رغم اختلاف درجاتهم بين التطبيق الأول والإعادة، ليتحقق مستوى ثبات مرتفع أو تام. فمعامل الارتباط بين درجات التطبيق الأول والتطبيق الثاني للاختبار يتأثر إذا اختلف موقع الفرد أو رتبته بالنسبة لبقية درجات الأفراد، ولا يتأثر باختلاف الدرجات بين التطبيقين إلا إذا أثر ذلك في رتب الأفراد. أما إذا اختلفت الدرجات (ارتفعت مثلا في التطبيق الثاني)، وبقيت رتب الأفراد ثابتة فإن معامل الارتباط لا يتغير. وبالتالي لكي يكون تعريف الثبات القائم على نموذج الإعادة دقيقا فنيا، فلا بد من التركيز على مدى تماثل رتب الأفراد أو مواقعهم بين التطبيقين للاختبار بدلا من التركيز على تشابه الدرجات ذاتها.

ولعل التعريف الذي يقترب من طبيعة مفهوم الثبات هو التعريف الذي يركز على خاصية اتساق درجات أداة القياس، سواء أكان ذلك عن طريق اتساق درجات المقياس عبر الإعادة، أو من صورة إلى أخرى، أو اتساق درجات نصفي المقياس أو أجزائه، أو اتساق درجات فقرات المقياس. إذ تشير الكتابات المرجعية المتخصصة في هذا الشأن:

(eg .Cronbach. 2004: Thompson & Vacha-Haase, 2000; Onwuegbuzie & Daniel, 2002: Sawilowsky, 2000).

بأن مفهوم الثبات يدل على مدى اتساق درجات المقياس أو دقتها [٢]. وتوضح أنازتزي (Anastasi & Urbina, 1997) بأن مفهوم الثبات استعمل استعمالا متنوعا ليغطي جوانب أو أبعاد مختلفة لاتساق الدرجات. فالثبات في مفهومه العام، وفقا لأنازتزي، يدل على مدى دلالة تباين درجات الأفراد الذين أجري عليهم المقياس على الفروق الحقيقية في السمة أو الصفة المقاسة، أي مدى الاتساق في درجات المقياس؛ وأيضا على مدى دلالة تباين هذه الدرجات على أخطاء الصدفة أو الأخطاء العشوائية أي مدى افتقار الدرجات للاتساق.[٣]؛ ص ٣٨]

ولما كان الثبات يستهدف تقدير مدى التباين المتسق أو المنتظم في الدرجات، لذلك نجد أن الدليل

الإرشادي أو المعياري للقياس التربوي والنفسي الذي اضطلعت بإصداره لجنة مشتركة ينتسب أفرادها المختصون إلى ثلاث منظمات علمية متخصصة[١] قد عرف مفهوم الثبات بمقدار أو مدى خلو درجات المقياس من أخطاء القياس.(N.C.M.E: 1985 & A.E.R.A & A.P.A) (٤). واعتمادا على هذا التعريف يوضح صلاح الدين محمود علام (٢٠٠٠) بأن مفهوم ثبات درجات الاختبارات يقصد بها مدى خلوها من الأخطاء غير المنتظمة التي تشوب القياس، أي مدى قياس الاختبار للمقدار الحقيقي للسمة التي يهدف لقياسها. (٥ ؛ ص ١٣١).

نستخلص مما سبق. أن الثبات يعنى بالتباين المتسق للدرجات Systematic Score Variance، أو نسبة هذا التباين المتسق إلى التباين الذي تنطوي عليه درجة الاختبار بشقيه التباين المتسق أو المنتظم (الدرجة الحقيقية[٢] True Score) والتباين غير المتسق أو غير المنتظم Unsystematic/Random Variance (الخطأ العشوائي للقياس Random Errors). أو بتعبير مرادف ومقتضب، يدل الثبات على مدى خلو درجات أداة القياس أو التقدير من الأخطاء العشوائية أو غير المنتظمة.

ويتم تقدير الثبات بناء على أربع طرق أو نماذج: طريقة الإعادة التي تستهدف تقدير استقرار الدرجات، وطريقة الصور المتكافئة التي تستهدف تقدير التكافؤ، وطريقة التجزئة النصفية التي تستهدف تقدير الاتساق، وأخيرا طريقة الاتساق الداخلي التي تستهدف تقدير التجانس والاتساق. ولقد كانت طريقة التجزئة النصفية باستعمال تصحيح سبيرمان – براون، وإلى عهد قريب، أكثر طرق تقدير الثبات استعمالا لأسباب عملية في الغالب إذ تقوم على إجراء واحد للأداة، ولا تتطلب جهدا حسابيا كبيرا. ومع تطور

---

(١) عنوان هذه المطبوعة الإرشادية في القياس التي صدرت سنة ١٩٨٥ كالتالي: Standards for educational and psychological testing (الدليل المعياري للقياس التربوي والنفسي). واشتركت في وضعه ثلاث منظمات علمية وهي: الرابطة الأمريكية للبحث التربوي American Educational Research Association; A.E.R.A والرابطة الأمريكية لعلم النفس American Psychological Association: A.P.A والمجلس القومي للقياس في التربية National Council on Measurement in Education: N.C.M.E.
وأول دليل صدر كان بعنوان: Technical Recommendations for Psychological Tests and Diagnostic Techniques (التوصيات أو الإرشادات التقنية أو الفنية المتعلقة بالمقاييس النفسية أدوات التشخيص). وذلك سنة ١٩٥٤ . وأحدث دليل صدر لحد الآن كان سنة ١٩٩٩. ويحمل نفس عنوان الدليل السابق الذي صدر سنة ١٩٨٥.

(٢) "الدرجة الحقيقية" مفهوم افتراضي، يدل على خصائص لا يمكن ملاحظتها ولا يمكن قياسها قياسا مباشرا. وتعرف إجرائيا بأنها متوسط درجات الفرد المتوقعة على سمة أو متغير معين، عند قياس السمة عددا لا نهائي من المرات. ومن المتعذر طبعا أن يعاد تطبيق الاختبار على الفرد إلى ما لانهاية، ولذلك تبقى إمكانية التحديد التام للدرجة الحقيقية أمرا افتراضيا أو تقريبيا.

الحزم الإحصائية وانتشارها ولا سيما حزمة SPSS ، أضحت طريقة الاتساق الداخلي ممثلة في المعادلة العامة المعروفة بمعامل ألفا كرونباخ / Coefficient Alpha Cronbach's Alpha أكثر شيوعاً واستعمالا بدون منازع في البحوث.

وعلى الرغم من الإقبال المستمر على استعمال معامل ألفا ، بحيث لا يكاد يخلو بحث أو رسالة ماجستير أو دكتوراه من التطرق إلى طريقة التجانس أو الاتساق الداخلي عند تقدير ثبات الأدوات، ممثلة في معامل ألفا لكرونباخ، فإننا نلاحظ كثيرا من القصور في فهم منطق معامل ألفا وإمكانياته وحدوده، كما نلاحظ إهمالا تاما للافتراضات Assumptions التي يقوم عليها المحددة لشروط استعماله، والتي قد تسفر عند عدم مراعاتها عن تقديرات غير دقيقة للثبات.

وقد تتعدى الظاهرة الباحثين غير المتخصصين في القياس إلى مدرسي القياس النفسي والتربوي. ففي إحدى جلسات استعراض خطط بحوث الماجستير لإبداء الملاحظات عليها من طرف الأساتذة الحاضرين، ثار جدل حول رأي أبداه أحد أساتذة مدرسي القياس والإحصاء يجزم فيه بأن قيمة معامل الاتساق الداخلي يجب ألا تزيد عن (٠,٧)، مدعما رأيه بصورة استنسخها من كتاب : "قياس الشخصية" لمؤلفه أحمد عبد الخالق؛ إذ يذكر المؤلف في كتابه السابق تحت عنوان: " في وجوب عدم ارتفاع معامل الاتساق الداخلي"، ما يلي: "يسعى معظم مؤلفي الاستخبارات إلى الحصول على معاملات ثبات قسمة نصفية أو اتساق داخلي مرتفعة. ولكن النظرية السيكومترية الحديثة تؤكد ضرورة الاحتفاظ بتجانس البنود (كما تقاس بالاتساق الداخلي) عند مستوى متوسط بحيث لا يزيد على ٠.٧ تقريباً. وذلك حتى يضيف كل بند جانبا جديدا من المعلومات، مما يرفع من تنوع عينة السلوك المسحوبة واتساعها."(٦؛ ص٥٠)

هذا الموقف يبرز مدى التأثير الكبير الذي تمارسه بعض المراجع الواسعة الانتشار ككتاب "قياس الشخصية" على القراء سواء أكانوا من المتخصصين أم من غير المتخصصين من الباحثين؛ لا سيما إذا كانت بعض الأفكار غير دقيقة كما هو الأمر في الاستشهاد السابق. فذكر موقف "النظرية السيكومترية الحديثة" كلام عام ومبهم يغري القاريء بقبول الرأي على عواهنه، رغم افتقاره إلى التحديد والتوثيق. وتعليل وجوب انخفاض معامل الاتساق الداخلي عن قيمة (٠,٧) وألا يتعداها، لكي يساهم كل بند بجانب جديد من المعلومات، حجة تتعلق أساسا بالصدق بدلا من الثبات. فمساهمة كل بند في إثراء المفهوم موضوع القياس عملية متسقة تتخذ منحى منتظما وغير عشوائي، ولذلك تساهم في الرفع من التباين المنتظم أو المتسق الذي يفترض فيه أنه يعكس السمة المقاسة أو جوانب منها. وتقدير مدى انطواء أو "تلوث" التباين المنتظم الذي يفترض فيه أنه يمثل السمة المقاسة بالخطأ المنتظم الذي يستعصي تمييزه عن التباين المنتظم

الحقيقي ، أمر يضطلع به الصدق. أما الثبات فلا يقوى على الكشف عن الأخطاء المنتظمة وإنما يعنى أساسا بتقدير مدى انطواء درجات المقياس على الأخطاء العشوائية. والأخطاء العشوائية تختلف عن الأخطاء المنتظمة لأنها لا ترتبط بالقوام النظري ، أو بالمحتوى الدلالي للمفهوم ، وإنما ترتبط بمصادر خارجية متذبذبة كالحالة الصحية والمزاجية للمفحوص ، وظروف تطبيق المقياس وغيرها.

إن ما أوردته يمثل مثالا فقط من أمثلة عديدة عن الأفكار النمطية المنتشرة التي تفتقر إلى الدقة عن معاملات الاتساق ، والتي لا يعدم القارئ مصادفتها في بعض كتب القياس وكتب مناهج البحث. ولذلك ، فإن التطرق إلى بعض قضايا معامل ألفا لتقدير الاتساق الداخلي وتوضيحها ، ومنها توضيح البنية المنطقية لمعامل ألفا ، للوقوف على افتراضاته ومسلماته والعوامل التي تؤثر في قيمته ، وتناول نماذج القياس وافتراضاتها ذات العلاقة بمدى دقة معامل ألفا ، وتوضيح متى يكون معامل ألفا دقيقا في تقدير الثبات ، ومتى يفتقر إلى الدقة المنشودة ، والتطرق إلى بعض المعاملات البديلة الأخرى عند عدم توفر الافتراضات التي يتطلبها معامل ألفا في بيانات القياس ؛ قد يساهم في توضيح الرؤى والتصورات ، وفي تغيير الأفكار النمطية غير الدقيقة ، وفي تشجيع التناول التقويمي الناقد كلما تعلق الأمر بتوظيف معامل ألفا لكرونباخ في تقدير الثبات وتأويله.

أهمية الدراسة

تتجلى أهمية الدراسة الحالية فيما يلي :

١ـ غياب البحوث العربية التي استهدفت تحليل البنية المنطقية والإحصائية لمعامل ألفا ، ومعالجة الافتراضات التي تقوم عليها ، وتحليل إشكالية تذبذب دقة معامل ألفا في تقدير الثبات.

٢ـ الاستعمال الواسع لمعامل ألفا في البحوث ورسائل الماجستير والدكتوراه.

٣ـ عدم مراعاة الافتراضات التي يتطلبها الاستعمال السليم لمعامل ألفا في تقدير الثبات في عديد من البحوث ، رغم شيوع استعماله.

٤ـ توضيح الكيفية التي تمكننا من فهم دقيق لمعامل ألفا بالاحتكام إلى التحليل المنطقي (الاستنتاجات القائمة على التحليل والاستدلال والمحاكمة Reasoning) بدون التركيز المطلق على التحليل الرياضي له.

٥ـ تطوير وعي القاري بمواطن دقة معامل ألفا ، ومواطن قصوره.

٦ـ تطوير مهارة الدارس أو الباحث في قراءة نتائج معامل ألفا قراءة ناقدة وتقويمية. مثلا ، هل معامل ألفا المرتفع يدل بالضرورة على ارتفاع الاتساق الداخلي ، أم يدل على تأثير طول الاختبار ، أم هو نتيجة التشابه الكبير بين الأسئلة التي تتفق في الدلالة رغم اختلاف الصياغة؟

٧ـ التعرف على بعض الطرق الأخرى

(المعاملات الأخرى) البديلة التي لا تتطلب توفر افتراضات صارمة في واقع بيانات القياس، مقارنة ببعض افتراضات معامل ألفا الصارمة، والتي تعطي تقديرا دقيقا نسبيا للثبات.

مشكلة الدراسة

تتلخص مشكلة الدراسة في تبيان منطق معامل ألفا لتقدير الثبات، وفي توضيح الحالات التي يسفر فيها استعمال معامل ألفا عن تقدير دقيق للثبات، والحالات التي يؤدي استخدامه فيها إلى تقدير متحيز بتقليص القيمة الحقيقية للثبات أو بتضخيمها؛ وفي التطرق إلى بعض المعاملات البديلة لمعامل ألفا.

وتتجلى المشكلة تفصيلا في أسئلة البحث الآتية:

١ ـ ما البنية المنطقية والإحصائية التي ينطوي عليها معامل ألفا، وما الاستنتاجات التي تنبثق عن التناول المنطقي لمعامل ألفا القائم على التحليل والاستدلال والمحاكمة reasoning؛ فضلا عن التحليل الإحصائي القائم على الاشتقاق الرياضي؟

٢ ـ ما طبيعة نماذج القياس، وما الافتراضات التي تقوم عليها ذات العلاقة باستخدام معامل ألفا؟

٣ ـ ما هي نماذج القياس أو الحالات التي يؤدي فيها استعمال معامل ألفا إلى تقدير دقيق للثبات الحقيقي، والحالات التي يسفر استخدامه فيها إلى تقدير متدن أو تضخيمي لمعامل الثبات الحقيقي؟

٤ ـ ما هي المعاملات البديلة التي تعطي تقديرا أدق لمعامل الثبات عندما لا تتوفر الافتراضات التي يتطلبها استعمال معامل ألفا في بيانات القياس؟

أهداف الدراسة

تتمثل أهداف الدراسة فيما يلي:

١ ـ دراسة البنية المنطقية والرياضية لمعامل ألفا للتوصل إلى فهم دقيق لافتراضاته ومسلماته Assumptions، واستنتاج بعض العوامل التي تؤثر في قيمته.

٢ـ استعراض نماذج القياس التي تفيد في معرفة متى يكون معامل ألفا دقيقا في تقدير الثبات، ومتى يفتقر إلى الدقة المنشودة.

٣ـ توضيح الحالات التي يسفر فيها استعمال معامل ألفا عن تقدير دقيق للثبات، والحالات التي يؤدي فيها توظيف معامل ألفا إلى توفير الحد الأدنى فقط لتقدير الثبات lower bound، الذي قد يختلف اختلافا بينا عن معامل الثبات الحقيقي؛ أو الحالات التي يتمخض فيها استعمال معامل ألفا عن تقدير متضخم للثبات.

٤ـ التطرق إلى بعض المعاملات أو الصيغ الأخرى البديلة التي تمد الباحث بتقدير أدق للثبات، عندما لا تتوفر بعض الافتراضات الصارمة التي يقتضيها معامل ألفا في بيانات القياس.

منهج الدراسة

إن البحث الحالي بحث نظري تحليلي ، ويختلف طبيعة عن البحث التجريبي ، والبحث شبه التجريبي ، والبحث المسحي. وقد وظف الباحث في انجازه عدة طرق وإستراتيجيات ، ومنها المنهج التاريخي وتحليل المضمون لرصد تطور تقدير الثبات من منظور الاتساق ، انطلاقا من طريقة التجزئة النصفية وانتهاء بمعامل ألفا. واستخدم الباحث أيضا طريقة أو استراتيجية التحليل المنطقي لمعادلات معامل ألفا. ونقصد بالتحليل المنطقي استخدام عمليات الاستدلال والتحليل والمقارنة والمحاكمة Reasoning في توضيح معادلات معامل ألفا ، والكشف عن افتراضاته. كما استعمل الباحث الطريقة الوصفية لشرح ظاهر معادلات معامل ألفا ، وتوضيح كيفية حساب بعض معاملات الاتساق البديلة.

وتبنى الباحث أيضا التناول النقدي عند مناقشة مفهوم الثبات ، وتحليل بنية معامل ألفا ، وتناول نماذج القياس ، وتحليل مواطن قوة معامل ألفا ومواطن قصوره. ولعل أهمية المنحى النقدي في معالجة مواضيع معامل ألفا تكمن في لفت انتباه القارئ بأن جل قضايا معامل ألفا قضايا إشكالية ، ومثار تباين في الرؤى والأفكار. وبأن أدب القياس النفسي والتربوي لا يقدم معرفة يقينية تنأى عن النقد ، ولا تقبل النقاش ؛ بل يقدم معرفة منفتحة على المراجعة الناقدة ، وقابلة للدحض Refutability .

ولما كانت الدراسة تستهدف المختصين في القياس ، وغير المختصين من الباحثين والأساتذة وطلاب الدراسات العليا ، لجأنا أحيانا إلى الاستطراد بغية توضيح الجوانب الفنية المتخصصة لمعامل ألفا ، وإلى تحليل المنطق التي تتضمنه معادلات معامل ألفا للتخفيف من الاشتقاق الرياضي من جهة ، وللكشف عن الدلالات النظرية والافتراضات التي تكمن وراء الصيغ الرياضية من جهة أخرى.

معامل ألفا: خصائص بنيته الإحصائية والمنطقية ومتضمناتها

لعل العمل الذي استقطب انتباه الباحثين إلى التركيز على اتساق البنية الداخلية للأداة كمؤشر للثبات ، أو الاتساق أو التجانس الداخلي ، يتمثل في المقالة التي كتبها كيودر Kuder بمعية زميله رتشاردسن Richardson سنة ١٩٣٧م [٧]. وتضمنت المقالة التي شكلت حينئذ مرجعا في تقدير الثبات عددا من المعادلات التي تستهدف تقدير الاتساق أو التجانس الداخلي لأداة القياس. لعل أكثرها شهرة وانتشارا واستعمالا ، المعادلة التي تحمل رقم عشرين ، ويشار إليها عادة بالاختصار التالي : KR-20 ، وهي كما يلي :

$$KR\text{-}20 = \frac{k}{k-1}\left[1 - \frac{\sum p_i q_i}{\sigma_T^2}\right] \quad (١)$$

والمعادلة رقم (٢١) (KR-21) . وهي كما يلي :

$$(٢) \quad \text{KR-21} = \frac{k}{k-1}\left[1 - \frac{\overline{X}\left(k - \overline{X}\right)}{k \times \sigma_T^2}\right]$$

حيث أن k تدل على عدد فقرات أو بنود المقياس. وتشير p_i إلى نسبة الإجابات الصحيحة على الفقرة i . أو نسبة المجيبين الذين حصلوا على الدرجة ١ ، بحيث تدل الدرجة ١ على أن الإجابة صحيحة. وتدل q_i على نسبة الإجابات الخاطئة على الفقرة i . أو نسبة المجيبين الذين حصلوا على الدرجة صفر ، بحيث تدل الدرجة صفر على أن الإجابة خاطئة. ويدل التعبير الرياضي : $\sum p_i q_i$ على حساب حاصل الضرب بين نسبة الإجابات الصحيحة ونسبة الإجابات الخاطئة المكملة لها لكل فقرة، (التي تمثل تباين الفقرة) ، ثم يتم جمع حواصل ضرب النسبتين لجميع الفقرات للحصول على مجموع تباين درجات فقرات المقياس، أي مجموع تباين بنود أو فقرات أداة القياس، ولذلك يمكن أن يتخذ التعبير المرادف المباشر التالي: $\sum p_i\ (1-p_i)$. ويدل σ_T^2 على تباين درجات الاختبار أو المقياس (ككل)، أي تجمع درجات كل فرد على فقرات المقياس لتمثل درجة المقياس (ككل)، ثم يحسب التباين لدرجات المقياس لجميع الأفراد أو المجيبين، فيدعى تباين درجات المقياس (ككل). وتدل $\overline{X}$ في المعادلة رقم (٢) على متوسط درجات المقياس : تحسب عدد الأجوبة الصحيحة للأفراد على مستوى جميع فقرات المقياس، ثم يقسم المجموع على عدد المجيبين.

لقد اقترح سبيرمان Spearman وبراون Brown (كل منهما على انفراد) سنة ١٩١٠ (٨)، طريقة التجزئة النصفية Split- half method للتخلص من إعادة إجراء المقياس، وما يرافق الإعادة من أثر التدريب، وذلك بإجراء تصحيح على معامل الارتباط بين نصفي الاختبار لتعديله عند مضاعفة طول المقياس (أي إضافة فقرات النصف الثاني إلى فقرات النصف الأول للمقياس لأن قيمة معامل الارتباط تزداد عند ازدياد عدد الفقرات) ، والتي عرفت بمعادلة سبيرمان- براون للتجزئة النصفية [٣]. والتي أضحت الطريقة المفضلة للباحثين والممارسين

---

(٣) عند تنصيف الاختبار ، وتصحيح معامل الارتباط عند مضاعفة طول الاختبار ، فيتوقع أن يكون معامل الارتباط للاختبار الكامل بعد إضافة النصفين وفقا لما تسفر عنه معادلة سبيرمان، وبراون المختصرة التي تتخذ الشكل التالي:

$r_{xx} = \frac{2\, r_{1/2\,1/2}}{1 + r_{1/2\,1/2}}$ وهي حالة خاصة (عند تجزئة الاختبار إلى نصفين) من معادلتهما العامة التالية:

$r_{kk} = \frac{k\, r_{xx}}{1 + (k-1) r_{xx}}$ حيث أن r_{xx} يدل على معامل ثبات المقياس، ويدل $r_{1/2\,1/2}$ على معامل ارتباط درجات نصفي الاختبار؛ ويدل r_{kk} على معامل الثبات التقديري للاختبار الذي سيمدد أو يقصر وفقا لمعامل k، ويدل k على معامل (عدد مرات) إطالة أو نسبة تقصير الاختبار.

والدارسين في تقدير الثبات، والتي لم تتقهقر في وتيرة استعمالها إلا في السنوات الأخيرة أمام معامل ألفا لكرونباخ، وإلى حد ما أمام المعادلة رقم عشرين (المذكورة أعلاه) لكيودر-رتشاردسن، بسبب تطور الحزم الإحصائية وانتشارها الواسع، وبالتالي التغلب على مشقة الحساب.

غير أن طريقة التجزئة النصفية لتقدير الثبات قامت على مسلمات أو افتراضات صارمة قلما التفت إليها عند استعمالها. إن التجزئة النصفية قامت على افتراض أو مسلمة التوازي التام Strict Parallel assumptions. بمعنى أنه يفترض في فقرات النصفين أن تتساوى في درجاتها الحقيقية، وأن تتساوى أيضا في قيم تباين خطأ القياس التي تنطوي عليها الدرجات. ومن المعروف أن نظرية القياس الكلاسيكي تقوم على تصور جوهري بأن الدرجة التي يحصل عليها الفرد على الاختبار والتي تدعى بالدرجة الملاحظة تنطوي بالضرورة على جزء أو قدر منها يدل على التباين المتسق أو المنتظم والذي يفترض أن له علاقة وثيقة بالسمة التي يقيسها الاختبار، ولذلك تدعى بالدرجة الحقيقية، وقسم آخر يدل على التباين (وبتعبير غير فني اختلافات أو فروق) غير المتسق أو المنتظم لا علاقة له بموضوع القياس، يدعى إجمالا بالدرجة الخطأ (أو الأخطاء العشوائية غير المنتظمة إذا توخينا التعبير الدقيق لها). فمثلا، إذا أضفنا فقرات النصف الثاني إلى فقرات النصف الأول للاختبار، فلا بد أن تحقق فقرات النصف المضاف إلى فقرات النصف الأول خاصية التوازي، بحيث تكون الدرجات الحقيقية للنصفين متساوية، ويكون تباين الخطأ لفقرات النصفين متساويا أيضاً. وبناء على ذلك، تكون معاملات الارتباط بين كل نصفين من الأنصاف الممكنة للاختبار متساوية، وبالتالي أن كل نصفين يمكن أن يمثل الأنصاف الأخرى الممكنة لتساوي ارتباطاتها ودرجاتها الحقيقية وتباين خطئها. وإذا لم تتحقق بعض شروط التوازي التام، فإن استعمال صيغة سبيرمان-براون لتصحيح معامل ارتباط النصفين يؤدي إلى تقدير متحيز لمعامل الثبات.

لقد انشغل كيودر ورتشاردسن بالتفكير في عملية التجزئة النصفية في حد ذاتها. إذ أن كل طريقة في تجزئة الاختبار تسفر عن ارتباط يختلف عن طرق التجزئة الأخرى الممكنة، وبالتالي تسفر عن معاملات ثبات مختلفة. وبناء عليه، لا يمكن اعتبار أن فقرات كل نصف من النصفين عينة ممثلة للمجال الأوسع الذي يحتوي على جميع الفقرات ذات الصلة بالسمة المقاسة Universe of items؛ طالما أن كل طريقة من طرق التجزئة النصفية تتمخض عن تقدير مختلف لمعامل الثبات. ثم إن اختلاف نتائج الثبات باختلاف طرق التجزئة النصفية الممكنة للمقياس لا يتوافق مع الافتراضات التي قامت عليها هذه الطريقة.

فكيف يمكن التوصل إلى تجزئة موحدة أو وحيدة لذات الاختبار بدون انتهاك افتراضات التوازي؟ لماذا لا يكون الجزء الذي يجزأ على أساسه

الاختبار متمثلا في أبسط وحدة أو الوحدة الأولية التي يحتوي عليها المقياس؟ ولما كانت الوحدة الأولية التي تشكل أداة القياس متمثلة في السؤال أو الفقرة أو البند، لذلك تصور كيودر ورتشاردسن أن تقسيم الاختبار بحسب عدد فقراته يوحد معضلة التقسيم. فالاختبار الذي ينطوي على عشر فقرات يجزء دائما إلى عشرة أجزاء إذا اتخذت الفقرة كأساس للتقسيم، أي إذا احتوى كل قسم على فقرة واحدة من فقرات المقياس. ولذلك فالبيانات التي نحتاجها لتقدير الثبات تقوم أساسا على بيانات الفقرات وليس على بيانات أنصاف الاختبار. غير أنه لتطوير معادلتهما، اضطر كيودر ورتشاردسن أن ينطلقا من ذات الافتراضات أو المسلمات التي قامت عليها طريقة سبيرمان وبراون، أي افتراض التوازي التام (تساوي الدرجات الحقيقية وتساوي تباين الخطأ)، فالفقرات تعامل وكأنها اختبارات متوازية لكي يتسنى تقسيم الاختبار وفقا لعدد فقراته باتخاذ الفقرة أساس التقسيم ، كما افترضا أيضا أن تكون الفقرات مهما كان شكلها (فقرات الصحيح والخطأ أو فقرات الاختيار من متعدد أو غيرها) ثنائية الدرجة Dichotomously scored items، إذ تخصص الدرجة واحد للإجابة الصحيحة والدرجة صفر للإجابة الخاطئة[٤] .

(٤) في الواقع يمكن استعمال أي زوج من القيم أو الدرجات، لكن جرت العادة استعمال القيمتين التاليتين: الصفر ليرمز إلى الإجابة الخاطئة، والواحد ليرمز إلى الإجابة الصحيحة.

ما معنى افتراض التوازي التام بين المكونات الأولية للمقياس التي تتمثل في الفقرات؟ توازي المكونات الأولية للمقياس أو توازي فقراته معناه أن فقرات المقياس أو بنوده التي تشكل المقياس هي بمثابة اختبارات متوازية، فالمقياس الذي يتكون من عشرين فقرة ، يتم تصوره من منظور التوازي كأنه يتكون من عشرين اختبارا أو مقياسا متوازيا، بحيث تتساوى في درجاتها الحقيقية وفي تباين خطئها. ولما كانت الفقرات التي ينظر إليها على أنها اختبارات متوازية متماثلة في درجتها الحقيقية، التي تمثل القاسم المشترك أو المدى أو النطاق الذي تشترك فيه الفقرات، أو اللحمة أو النسيج الذي يجمع فقرات المقياس، لذلك فإن هذه الافتراضات التي تقوم عليها معادلة كيودر- رتشاردسن تتعلق بخاصية اتساق الفقرات وتجانسها.

إن معادلة كيودر- رتشاردسن (KR-20) المذكورة أعلاه (المعادلة رقم ١) تضطلع بتقدير الثبات عندما تصحح كل فقرة من فقرات المقياس على أساس ثنائي باستعمال درجتين فقط. صفر أو واحد أو أي زوج من الدرجات. أما التعبير التالي: $\left[1 - \frac{\sum p_i q_i}{\sigma_T^2}\right]$ فيعكس منطق تعريف الثبات reliability rationale باعتبار الثبات يدل على مقدار الدرجة الحقيقية ، أي نسبة الدرجة الحقيقية إلى الدرجة الملاحظة التي تنطوي على الخطأ. عند توحيد مقام الحد السابق فإنها تؤول إلى التعبير المرادف الآتي:

$\left[\frac{\sigma_T^2 - \sum p_i q_i}{\sigma_T^2}\right]$. ونلاحـــظ أن البســـط :

$\sigma_T^2 - \sum p_i q_i$ يـدل علـى مجمـوع التغـاير بـين فقرات المقياس (أو العلاقات بين الفقرات التي لا تتقيد بوحـدة قيـاس معينـة)، لأن مجمـوع تبـاين درجـات الاختبـار ككـل σ_T^2 تـساوي مجمـوع تبـاين الفقـرات $\sum p_i q_i$ مضاف إليه ضعف قيم تغاير كل فقرة مع أخـرى (العلاقـات البينيـة بـين الفقـرات إذا توخينـا التبسيط) أو كـل زوج من فقرات المقياس. وبنـاء على ذلـك عنـد حـذف مجمـوع تبـاين درجـات الفقـرات: $\sum p_i q_i$ مـن مجموع تباين درجـات الاختبـار σ_T^2 تبقى قيم التغاير بـين الفقـرات الـتي تـدل عـلى القاسـم المشترك بـين الفقـرات، وبالتـالي تعكـس في أساسـها الدرجـة الحقيقيـة للاختبـار. وبمـا أن البـسط يـدل عـلى تباين الدرجة الحقيقية، والمقام يدل على التباين الكلـي لـدرجات المقيـاس الـتي تـضم تبـاين الدرجـة الحقيقيـة وتباين الدرجة الخطأ. فإنه يعكـس نـسبة تبـاين الدرجـة الحقيقية إلى تباين الدرجـة الملاحظـة، أو يعكـس مقـدار التبـاين المتسق الخـالي مـن الخطـأ العـشوائي، أي بعـد تنقيتـه مـن التبـاين العـشوائي. وذاك هـو منطـق الثبـات الذي يستشف من المعادلة كما يـتردد في جنبـات النظريـة السكومترية العريقة.

ويـضطلع الكـسر التـالي: $\frac{k}{k-1}$ بجعـل الحـد الأقـصى لـتصل قـيم نتـائج المعادلـة يـصل إلى الواحـد الـصحيح. إذ بـدون اسـتعمال هـذه النـسبة يـستحيل أن تـصل معادلـة KR-20 إلى القيمـة القـصوى أو سقفها الأعلى المتمثـل في الواحـد الـصحيح عنـد تحقـق الثبات التام. وكلما قل عدد فقرات المقياس كلما ازداد تأثير هذه النسبة التصحيحية على نتيجة معادلة KR-20 للاتساق الداخلي. ذلك أن الحد التالي $\sum p_i q_i$ لا يساوي الصفر عند تحقق الثبات التام وبالتالي فإن الحد الأيـــسر $\left[1 - \frac{\sum p_i q_i}{\sigma_T^2}\right]$ أو المـــرادف لــــه $\left[\frac{\sigma_T^2 - \sum p_i q_i}{\sigma_T^2}\right]$ لا يـصل إلى أقـصاه أبـدا عنـد تحقـق الثبـات التـام أي لا يـصل إلى الواحـد الـصحيح، لذلك يضطلع الحد $\frac{k}{k-1}$ برأب هذا النقص، بضبط الطـرف الأقـصى للمـدى النظـري للثبـات إلى الواحـد الصحيح عند تحقق الثبات التام.

إضافة إلى مـسلمة التـوازي الـصرف، فإنـه عـلى الرغم مـن أن المعادلـة: KR-20 لم تقـم عـلى افـتراض تـساوي معـاملات الارتبـاط بـين فقـرات المقيـاس، بـل يكفـي تقاربهـا، لكنهـا اشـتقت بنـاء عـلى افـتراض أن مـصفوفة الارتباطـات بـين الفقـرات مـن طـراز مرتبـة الوحدة unit rank. ومعنى ذلك، أنه يفترض في جميع فقرات المقيـاس أن تقيس بعـدا واحـدا فقـط. أو عـاملا واحدا مشتركا. أي يفترض في جميع فقرات المقياس أن تكون متجانسة بحيـث تنتمي إلى بعـد واحـد فقـط، أو تـتشبع عـلى عامـل عـام واحـد عنـد إجـراء التحليـل العاملي(٩ ؛ ص ٢٥٦).

أما المعادلة KR-21 فأوردناها لتواتر التطرق إليها في المراجع العربية. وللتنبيه على الافتراضات الأكثر صرامة التي تقوم عليها مقارنة بمعادلة KR-20. لقد اقترحت المعادلة في الربعينيات من القرن الماضي، وكان الضابط أحيانا للمفاضلة بين المعادلات المترادفة يسر الاستعمال والتخفيف من مشقة الحساب لأن كل العمليات الحسابية كانت تجرى يدويا، أو باستعمال حاسبات يدوية غير متطورة في الخمسينيات والستينيات. ومع ذلك ما زالت بعض المراجع الحديثة تتطرق إليها بحجة سهولة حسابها مقارنة بمعادلة KR-20 رغم التطور الكبير في مجال الكمبيوتر، وفي مجال الرزم الإحصائية العديدة.

إن المعادلة KR-21 قائمة على افتراض تساوي مستوى الصعوبة لكافة فقرات المقياس. بمعنى أن النسبة التالية : p_i التي تدل على نسبة الإجابات الصحيحة (عدد الإجابات الصحيحة على سؤال معين: i مقسوما على الإجابات الصحيحة والخاطئة على ذات السؤال) والتي تعكس في ذات الوقت معامل صعوبة السؤال أو مستواه يجب أن تكون متساوية عبر جميع فقرات الاختبار أو بنوده. وعند عدم توفر هذا الشرط الذي يستحيل تحقيقه في الواقع، فإن معامل الاتساق (الثبات) الناجم عند استعمال هذه المعادلة يتسم بعدم الدقة وبالانخفاض مقارنة بمعامل الاتساق الناتج عن استعمال المعادلة : KR-20 .

وبعد مرور أربعة عشر عاما على ظهور معادلة كيودر ورتشاردسن للاتساق الداخلي، قدم كرونباخ عددا من المعادلات المترادفة أسماها أو رمز لها بالحرف الأغريقي الصغير الحجم : α (أي ألفا)، في مقاله الموسوعي الذي نشره في مجلة سيكومتريكا العتيدة سنة ١٩٥١ميلادي(١٠) . مما أعطى دفعا قويا لمنهجية تقدير الثبات من منظور الاتساق أو التجانس الداخلي لأداة القياس. ولقد استقطبت معادلة ألفا لكرونباخ اهتمام الباحثين أكثر مما استقطبته معادلة كيودر ورتشاردسن، على الرغم من قواسمهما المشتركة، ذلك أن ألفا أعم من KR-20 لأنها تستعمل لتقدير التناسق الداخلي سواء أكانت درجات التصحيح ثنائية، أم متصلة ، وبالتالي لا يضطر الباحث إلى تحويل سلم التصحيح القائم على أكثر من درجتين إلى سلم تصحيح ثنائي الدرجات ، أي استعمال الدرجة صفر والدرجة واحد مثلا. ذلك أن تحويل مجال الدرجات المتصله أو سلم الدرجات المتصلة المستعمل في التصحيح (مثل تخصيص أوزان تتراوح من واحد إلى خمسة لفقرات الاتجاه صيغت فئات أجوبتها المتدرجة على شاكلة سلم ليكرت الخماسي الفئات الذي قد يتراوح من "موافق تماما" إلى "غير موافق إطلاقا") إلى سلم ثنائي الدرجات، يؤدي إلى تقليص كبير لتباين الدرجات، وبالتالي إلى انخفاض كبير في معامل الثبات أو الاتساق الداخلي عند استعمال KR-20 مقارنة بقيمة معامل ألفا. فمعادلة KR-20 ، إذن، حالة خاصة من معادلة ألفا التي تستعمل في حالة استعمال درجتين فقط في

التصحيح أو في حالة استعمال مجال من الدرجات. ولذلك لا نجد ذكرا لطريقة كيودر ورتشاردسن في الحزمة الإحصائية : SPSS في حين وجدت طريقة ألفا لتقدير الثبات ، بل واعتبرت ألفا الطريقة الافتراضية في الاستعمال Default method . أي أن الحزمة تستعمل طريقة ألفا تلقائيا إذا لم يحدد المستخدم طريقة أخرى لتقدير الثبات من ضمن الطرق التي تعرضها الحزمة.

لمعامل ألفا عدة صيغ مترادفة ، والصيغة الأكثر ألفة وورودا في كتب القياس وكتب مناهج البحث هي كما يلي :

$$(٣) \quad Cronbach\,Alpha(\alpha) = \frac{k}{k-1}\left(1 - \frac{\sum \sigma_i^2}{\sigma_T^2}\right)$$

التعبير المستخدم في معامل ألفا (معادلة ألفا) ، والذي يتمثل في : $\sum \sigma_i^2$ يدل على مجموع تباين درجات الفقرات. أي يتم حساب تباين درجات الأفراد لكل فقرة ، ثم تجمع قيم التباين المحسوبة لكافة فقرات المقياس. إن معادلة ألفا كرونباخ لا تختلف في جوهرها عن KR-20 ، للدلالة على مدى خلو الدرجات من الأخطاء العشوائية ، أو للدلالة على نسبة تباين الدرجة الحقيقية إلى الدرجة الملاحظة . إذ أنه يمكن إعادة ترتيب حدود التعبير $\left(1-\frac{\sum \sigma_i^2}{\sigma_T^2}\right)$ وتوحيد مقام الحدين في كسر واحد لتتخذ الشكل التالي $\left(\frac{\sigma_T^2 \quad \sum \sigma_i^2}{\sigma_T^2}\right)$ ولذلك ، تظهر المعادلة كالتالي :

$$(٤) \quad Cronbach's\ Alpha(\alpha) = \frac{k}{k-1}\left(\frac{\sigma_T^2 - \sum \sigma_i^2}{\sigma_T^2}\right)$$

يدل البسط $\sigma_T^2 - \sum \sigma_i^2$ على التباين الباقي لدرجات المقياس ككل σ_T^2 بعد حذف مجموع تباين الفقرات $\sum \sigma_i^2$ منها. لكن ماذا يمثل هذا التباين الباقي ؟ إنه يمثل التغاير بين فقرات المقياس مثنى مثنى ، أي يمثل التباين المنتظم غير العشوائي وهو ما يشار إليه بتباين الدرجة الحقيقية. وبالتالي يدل الكسر على مقدار تباين الدرجة الحقيقية إلى تباين الدرجة الكلية.

لم يتعد ما سبق من شرح لمعامل ألفا وظيفة التوضيح الوصفي ، غير أنه لتحقيق فهم أعمق لمعامل ألفا ، من الضروري أن نحلل بنية معادلة ألفا ، لنقف على أبعاد المعادلة الأكثر تأثيرا ، ونستنتج العوامل التكوينية للمعادلة التي تساهم في تحديد قيمها.

عند فحص معادلة الفا رقم (٣) ، ندرك أن الكسر داخل القوس : $\frac{\sum \sigma_i^2}{\sigma_T^2}$ يلعب دورا كبيرا في تحديد معامل ألفا ، ذلك لأن الحد الأيسر خارج القوس $\frac{k}{k-1}$ قليل التأثير في قيمة معامل ألفا لا سيما عند ازدياد عدد الفقرات. يدل الكسر داخل القوس ـ كما سبق أن أوضحنا ـ على مجموع تباين درجات الفقرات مقسوما على تباين درجات الاختبار ككل. إن نتيجة

الكسر بعد حذفها من الواحد الصحيح تقرر قيمة معامل ألفا. إذ يرتفع معامل ألفا كلما كانت قيمة نتيجة الكسر منخفضة، وينخفض معامل ألفا كلما كانت قيمة نتيجة الكسر مرتفعة. ومنطقيا، تكون نتيجة الكسر منخفضة كلما صغرت قيم البسط وارتفعت قيم المقام، وبالتالي يرتفع معامل الثبات (معامل ألفا) كلما انخفضت قيم مجموع تباين الفقرات، وارتفعت قيم تباين الاختبار ككل. يوحي هذا الوصف للقارئ، بأنه للحصول على معامل ثبات مرتفع، لا بد من الإبقاء على تباين فقرات المقياس منخفضة، أو العمل على رفع تباين درجات المقياس.

لكن أيهما أكثر تأثيرا في قيمة معامل ألفا: مجموع تباين الفقرات ببسط الكسر، أم تباين درجات المقياس ككل بالمقام؟

تدل دراسات المضاهاة الإحصائية statistical simulation studies أن الحد الأكثر تأثيرا في معادلة ألفا هو تباين درجات المقياس، بحيث كلما ارتفع تباين درجات المقياس، أدى ذلك إلى ارتفاع معامل ألفا (١١). فدراسة رينهاردت (Reinhardt, 1996) القائمة على المضاهاة أظهرت أن مدى تباين درجات المقياس فسرت مقدار التباين في معاملات ألفا بنسبة ٦٠٪. كما أسفرت دراسة كاريسو (Caruso, 2000) البعدية meta-analysis للبحوث التي وظفت إحدى مقاييس الشخصية، بأن متوسط الثبات لأحد المقاييس الفرعية يساوي ٠,٧٩ عندما طبق في عدد من الدراسات على الجمهور العام (إذ تتسم العينات بالتباين). ولم يرق متوسط الثبات في الدراسات الأخرى إلا إلى ٠,٦٢ عندما أجري على عينات إكلينيكية لأنها أكثر تجانسا، وأقل تباينا. (١٢)

يمكن القول، إذن، كلما قوي التباين، كلما ارتفع مستوى معامل ألفا. وبالتالي، فاتساع التباين أو تقلصه يتوقف على مدى تجانس العينة، أو مدى تباينها. فتباين أفراد العينة يؤدي إلى تباين الاستجابات على فقرات المقياس، وبالتالي يفضي إلى ارتفاع تباين درجات المقياس ككل. ومعنى ذلك، أن قيم الثبات لذات المقياس قد تختلف باختلاف العينات، إذ يستتبع اختلاف العينة اختلاف في مدى تباين درجات المقياس ككل. ولذلك لاحظ طومسن (Thompson, 1994) بأن تطبيق المقياس نفسه على عينة أكثر تجانسا، أو على عينة أكثر تباينا يؤدي إلى درجات ذات معاملات ثبات مختلفة (١٣). وبتعبير آخر، فالمقياس الذي أظهر مستوى ثبات مرتفع عند تطبيقه على عينة معينة، قد يظهر مستوى ثبات منخفض عند تطبيقه على عينة أخرى. ومغزى هذه الاستنتاج، أن الثبات ليس صفة لصيقة أو جوهرية للمقياس، بل يتوقف على طبيعة (منها تباين) الدرجات، وبالتالي على المجموعات أو العينات. ولذا من الخطأ استعمال تعبير "ثبات المقياس" وهو التعبير الذي طالما تردد وما زال يتردد في البحوث والدراسات، ومراجع التقويم والقياس، وكتب مناهج البحث. والأصح أن نستعمل عوض ذلك تعبير: "ثبات

درجات المقياس[٥]. فالثبات ليس صفة ملازمة للمقياس، بحيث إذا أظهر بحث (أو بعض البحوث) توفرها في المقياس، يتخذ ذلك دليلا على أن الثبات قد حسم أمره بالنسبة للمقياس، وبالتالي ليس من الضروري إعادة تقديره من طرف باحثين آخرين الذين يوظفون ذات المقياس في أبحاثهم على عينات أخرى من نفس المجتمع. فكون الثبات خاصية ملازمة للمقياس، معناه أنه يتأثر بالمقياس ذاته فقط ولا يتأثر باختلاف العينات وتباينها، وبالتالي يبقى الاختبار محتفظا بمستوى ثباته عند تقديره باستعمال عينات مختلفة أخرى.[٦]

أما القضية الثانية الحرجة التي يمكن استنتاجها عند فحص بنية معادلات ألفا لكرونباخ فحواها: هل تتوقف قيمة معامل ألفا على مدى الاتساق الداخلي لفقرات المقياس فقط (مستوى الارتباطات فيما بينها، أو متوسط معاملات ارتباطاتها)، أم تعكس أيضا تأثير عدد الفقرات أو طول الاختبار، بحيث أن ارتفاع عدد الفقرات قد يكون عاملا حاسما في ارتفاع معامل ألفا، حتى في حالة انخفاض الاتساق الداخلي لفقرات المقياس انخفاضا كبيرا؟

لقد أورد كرونباخ (Cronbach, 1951) عدة صيغ مترادفة لمعامل ألفا، ومن بينها الصيغة التالية القائمة على حساب التغاير بين فقرات المقياس:

$$Cronbach\ Alpha(\alpha) = \frac{k}{k-1}\left(\frac{\sum\sum CV_{ij}}{\sigma_T^2}\right) \quad (٥)$$

يدل بسط الكسر على مجموع التغاير covariance بين أزواج فقرات المقياس. إن الجدول الذي يرصد العلاقات بين جميع فقرات المقياس وذلك بوضعها في بداية الصفوف، ووضعها أيضا في بداية الأعمدة يسمى بمصفوفة التباين والتغاير variance-

(٥) تنسحب نفس الملاحظة على الصدق. فالتعبير الشائع "صدق المقياس" يوحي بأن الصدق خاصية أو صفة ملتحمة بالمقياس. وما دامت صفة في المقياس فهي ملازمة له وجودا وعدما. بينما الصدق يتمثل في تأويل درجات الاختبار وتفسيرها، وفي الاستدلالات المترتبة عليها. وبالتالي فالصدق شأنه شأن الثبات عملية (أوجملة عمليات)، أو إجراء (جملة إجراءات)، وليس صفة ملازمة لموصوفها. ولما كان الصدق يتمركز حول عملية تأويل الدرجات وتفسيرها للتوصل إلى استدلالات أو استنتاجات معينة، فإنه من قبيل تحري الدقة استعمال تعبير "صدق درجات المقياس"، عوضا عن التعبير التالي: "صدق المقياس".

(٦) وليس من قبيل المصادفة، أن تحتفظ الدورية المتخصصة في القياس النفسي والتربوي: Educational and Psychological Measurement في استعمال تعبير "Test Reliability" أو تعبير "Test Validity"، بحيث أن عناوين البحوث التي =

- تستهدف تقدير الثبات والصدق لدرجات المقاييس المختلفة، والتي تخصص لها المجلة حيزاً واسعاً لها في كل عدد، خلت تماماً في العشريتين الأخيرتين من هذين التعبيرين، واستبدلا بشكل شامل بالتعبيرين الدقيقين التاليين المسحمين، وتطوير نظرية القياس: "Test score reliability" للدلالة على "ثبات درجات المقياس"؛ وكذلك: "test score validity" للدلالة على "صدق درجات المقياس".

covariance matrice أو مصفوفة التغاير اختصارا (covariance matrice). والتغاير يعبر عن العلاقة بين متغيرين باستعمال وحدة القياس الأصلية للمتغيرين، في حين أن الارتباط يقدر العلاقة بين المتغيرين ولكن بتقييد وحدات قياس المتغيرين بتحويلها إلى درجات معيارية بمتوسط يساوي صفرا، وانحراف معياري يساوي الواحد. ومعنى ذلك أن التغاير يتحول إلى معامل ارتباط بين متغيرين عند تحويل درجاتهما الخام إلى درجات معيارية.[٧] تحتوي مصفوفة التغاير بين فقرات المقياس في خلاياها القطرية على قيم التباين للفقرات (تغاير الفقرة مع نفسها، إن صح التعبير)، وفي خلاياها اللاقطرية (الخلايا الأخرى) على التغاير بين كل فقرة وأخرى. فبسط المقام للمعادلة السابقة رقم (٥) يدل على جمع قيم التغاير كافة التي توجد بالخلايا غير القطرية للمصفوفة، أما المقام الذي يدل على تباين درجات الاختبار ككل فيحتوي على مجموع قيم التباين للفقرات الموجودة بالخلايا القطرية للمصفوفة، وأيضا على مجموع قيم التغاير بين فقرات المقياس الموجودة في الخلايا غير القطرية لمصفوفة التغاير. أو بتعبير آخر تمثل تباين درجات المقياس ككل جميع قيم المصفوفة. فإذا احتوى الاختبار على ١٠ فقرات مثلا، فإن عدد الخلايا القطرية عشرة وبالتالي توجد عشرة قيم للتباين، وأن عدد الخلايا غير القطرية تسعون خلية أو تسعون قيمة تغاير. وبالتعويض في المعادلة رقم (٥)، فإن مجموع قيم التغاير التسعين يدل على البسط الذي يمثل مجموع تغاير الفقرات كلها، أما المقام الذي يمثل تباين الاختبار فيحسب بإضافة مجموع قيم التغاير التسعين إلى مجموع قيم التباين العشرة.

لقد تم التركيز حتى الآن على كيفية قراءة المعادلة رقم (٥)، وفهمها، ولكي نجيب عن السؤال السابق، سنركز انتباهنا على الكسر لأنه الحد الحاسم في المعادلة. يدل الكسر على نسبة مجموع تغاير الفقرات

---

(٧) يمكن التعبير عن التغاير بدلالة الارتباط كما هو مبين في الصيغة التالية: $cov_{ij} = r_{ij}\,\sigma_i\,\sigma_j$ أي أن معامل التغاير بين فقرتين (فقرة i، وفقرة j) يساوي معامل الارتباط بين درجات الفقرتين مضروبا في الانحراف المعياري لكليهما. وتظهر المعادلة العلاقة القوية بين مفهوم الارتباط وبين مفهوم التغاير، فإذا انعدم الارتباط بين الفقرتين، انعدم أيضا تغايرهما. وأيضا كلما ارتفع معامل الارتباط بين فقرتين أو متغيرين، ارتفع مستوى تغايرهما. ولقد أشرنا إلى أن التغاير يوظف وحدة القياس الأصلية لكل متغير في حين أن الارتباط يعبر عن العلاقة بتحويل وحدات القياس الأصلية للمتغيرين إلى درجات معيارية، بمتوسط يساوي صفرا، وانحراف معياري يساوي الواحد الصحيح. وتبرز العلاقة القوية بين الارتباط والتغاير عند التعبير عن درجات المتغيرين، أو الفقرتين اللتين يراد تقدير تغايرهما بالدرجات المعيارية، وبالتعويض في الانحراف المعياري لإحدى الفقرتين بالواحد الصحيح، والقيام بنفس الشيء بالنسبة للفقرة الأخرى، سنجد أن التغاير بين الفقرتين يساوي معامل الارتباط بينهما، وبناء عليه= يمكن القول أن معامل الارتباط هو التغاير بين الدرجات المعيارية للفقرتين أو المتغيرين.

إلى مجموع تباينها وتغايرها. فأي ارتفاع في قيم التغاير أو أي ارتفاع في العلاقات الارتباطية بين الفقرات يؤدي إلى ارتفاع في قيمة بسط الكسر ولا يؤدي إلى ارتفاع المقام بنفس الوتيرة ، الأمر الذي يتمخض عنه ارتفاع في قيمة معامل ألفا. إن ارتفاع معامل ألفا بارتفاع متوسط الارتباطات بين الفقرات (أي ازدياد مستويات التغاير بين الفقرات) لأمر يثلج الصدر ، ويجعلنا نطمئن على أن معامل ألفا يعكس مستوى العلاقات الارتباطية أو التغايرية بين فقرات المقياس . ويمثل بناء على ذلك الاتساق الداخلي لفقرات المقياس أو يعكس متوسط العلاقات الارتباطية البينية للفقرات. وبالتالي هل نستطيع أن تستخلص من ذلك أن ارتفاع معامل ألفا دليل على ارتفاع متوسط العلاقات الارتباطية بين فقرات المقياس. أو بتعبير آخر ، هل ارتفاع معامل ألفا دليل على ارتفاع مستوى الاتساق الداخلي للمقياس؟

إن ارتفاع قيمة معامل ألفا لا يدل بالضرورة على ارتفاع مستوى الاتساق الداخلي للمقياس. وفي الواقع ، أن ثمة عاملا حاسما آخر يمارس تأثيرا كبيرا على نتائج معامل ألفا ويتمثل هذا العامل في طول الاختبار ، أو عدد الفقرات.

ولكي يتم توضيح التأثير الكبير الذي يمارسه طول الاختبار على نتائج معامل ألفا ، سيتم اختيار إحدى الصيغ المرادفة لمعامل ألفا ، القائمة أساسا على متوسط الارتباطات بين الفقرات ، وعلى الدرجات المعيارية بدلا من الدرجات الخام لفقرات المقياس ، والمعادلة هي الآتي :[٨]

$$Cronbach's\ \ Alpha\ (\alpha) \ = \ \frac{k\,\overline{r_{ij}}}{1 + (k-1)\,\overline{r_{ij}}} \qquad (٦)$$

يدل التعبير $\overline{r_{ij}}$ على متوسط معاملات الارتباط الممكنة بين فقرات المقياس ، أما الحد k فصادفناه في المعادلات السابقة ، ويدل على عدد الفقرات أو طول الاختبار (١٤ : ص ١٠٠).

لكن كيف يسلك معامل ألفا عند تفاوت عدد

---

(٨) هذه المعادلة في الواقع إحدى الصيغ المرادفة للمعادلة العامة لسبيرمان براون التي أشرنا إليها في الهامش رقم(٣)، ورغم تماثل المعادلتين إلا أنهما تقومان على افتراضات مختلفة. فالمعادلة العامة لسبيرمان-براون تقوم على افتراض التوازي الصارم أو المحض، أي افتراض تساوي الدرجات الحقيقية، وتساوي تباين الخطأ لأجزاء الاختبار أو المقياس. بينما المعادلة الحالية (المعادلة رقم ٦) فتمثل حالة خاصة لمعادلة ألفا لكرونباخ العامة (أي تقدير التجانس الداخلي عند تحويل درجات الفقرات إلى درجات معيارية). ولما كانت حالة خاصة لمعامل ألفا، فهي تقوم على ذات الافتراضات التي يقوم عليها معامل ألفا،(والتي هي أقل تشددا من افتراضات سبيرمان وبروان) أي افتراض تساوي الدرجات الحقيقية بين الفقرات (او قد تختلف بمقدار عدد ثابت كما سنرى في المعالجات القادمة)، لكن لا تشترط كما هو الحال في التوازي الصارم الذي أخذ به سبيرمان وبراون تساوي تباين الخطأ. نستنتج مما تقدم، أنه من الممكن أن نصادف معادلتين متماثلتين في القياس، رغم قيامهما على افتراضات متباينة.

الفقرات وعند تثبيت متوسط معامل الارتباط بين الفقرات؟ وكيف يسلك أيضا في حالة الارتفاع التدريجي لمتوسط معاملات الارتباط بين الفقرات عند تثبيت طول الاختبار؟ لنتمكن من معاينة ذلك أنشأنا الجدول رقم (١)، الذي يظهر قيم معامل ألفا عندما يحتوي الاختبار على خمس وعشر وعشرين وثلاثين فقرة، وكذلك عند مستويات مختلفة من متوسط ارتباطات الفقرات. لنركز أولا على قيم معامل ألفا عند ازدياد متوسط معاملات الارتباط عند طول معين للمقياس.

يلاحظ أن معامل ألفا يزداد كلما ارتفع متوسط معاملات الارتباط بين الفقرات أي ارتفاع الاتساق الداخلي. فعند الابقاء على طول الاختبار ثابتا عند خمس فقرات مثلا، فإن قيمة معامل ألفا ارتفعت من ٠.٣٦ إلى ٠.٥٦ عندما ارتفع متوسط ارتباط الفقرات من ٠.١ إلى ٠.٢، ويزداد معامل ألفا ارتفاعا كلما ارتفع متوسط الارتباط بين الفقرات. ولكن بوتيرة متناقصة. بمعنى لا يرتفع بنفس المقدار عند الازدياد التدريجي في متوسط ارتباطات الفقرات. غير أن حجم تأثير متوسط ارتباط الفقرات على معامل ألفا يتناقص في الشدة كلما ازداد طول الاختبار. فعندما يكون طول الاختبار ٣٠ فقرة، فإن ارتفاع متوسط ارتباط الفقرات من ٠.١ إلى ٠.٢ أدى إلى ارتفاع معامل ألفا من ٠.٧٧ إلى ٠.٨٨. علما بأن معامل ألفا قفز من مستوى ٠.٣٦ إلى مستوى ٠.٥٦ عند ازدياد متوسط ارتباط الفقرات من ٠.١ إلى ٠.٢ لما كان طول الاختبار خمس فقرات فقط.

ومما سبق، يمكن استنتاج ما يلي: **أولاً**، أنه كلما ازداد الاتساق الداخلي للفقرات (العلاقات الارتباطية بينها) يزداد معامل ألفا ارتفاعا. **وثانياً**، عند طول معين للاختبار، فإن الازدياد في قيم ألفا لا يستمر بوتيرة واحدة عند ازدياد متوسط ارتباطات الفقرات. **وثالثاً**، أنه كلما ازداد حجم الاختبار يتناقص حجم التأثير الذي يمارسه ارتفاع متوسط الارتباطات بين الفقرات على قيم الفا.

غير أن التأثير الواضح والبين، فيتمثل في الارتفاع الكبير في قيم معامل ألفا عند ازدياد طول الاختبار. ففقرات المقياس التي ترتبط فيما بينها بمتوسط منخفض جدا لا يتجاوز الواحد من عشرة يقفز معامل ألفا لها من ٠.٣٦ عندما يكون طول الاختبار خمس فقرات إلى ٠.٧٧ عندما يصبح طول الاختبار ٣٠ فقرة. ومتوسط ارتباطات الفقرات الذي مقداره ثلاثة من عشرة (أي ما تشترك فيه الفقرات لا يتجاوز التسعة بالمئة من تباين المقياس، وهذه قراءة أخرى بتحويل متوسط الارتباط إلى معامل تحديد بتربيعه) والذي يبقى مع ذلك متوسط ارتباط منخفض نسبيا، نجد معامل ألفا له يبدأ بمستوى مرتفع قدره ٠.٨١ عندما يكون طول الاختبار ١٠ فقرات فقط، ويرتفع إلى ٠.٩٠ عندما يكون طول الاختبار ٢٠ فقرة، ويزداد ارتفاعا ليصل إلى ٠.٩٥ وهو مستوى ثبات مرتفع جدا عندما يكون طول الاختبار ٤٠ فقرة.

الجدول رقم (١) قيم معاملات ألفا عند اختلاف عدد الفقرات أو طول الاختبار، وتفاوت قيم متوسطات الارتباطات (مستويات التجانس الداخلي) بين الفقرات، عند توظيف معادلة رقم (٦) لمعامل ألفا: $\frac{k\,\overline{r_{ij}}}{1+(k-1)\overline{r_{ij}}}$.

عدد الفقرات أو طول المقياس: k	متوسط الارتباطات بين الفقرات $\overline{r}_{ij}$	معامل ألفا: $Alpha\ (\alpha)$	عدد الفقرات أو طول المقياس: k	متوسط الارتباطات بين الفقرات $\overline{r}_{ij}$	معامل ألفا: $Alpha\ (\alpha)$
خمس فقرات	٠.١	٠.٣٦	عشرون فقرة	٠.١	٠.٦٩
	٠.٢	٠.٥٦		٠.٢	٠.٨٣
	٠.٣	٠.٦٨		٠.٣	٠.٩٠
	٠.٤	٠.٧٧		٠.٤	٠.٩٣
	٠.٥	٠.٨٣		٠.٥	٠.٩٥
	٠.٦	٠.٨٨		٠.٦	٠.٩٧
	٠.٧	٠.٩٢		٠.٧	٠.٩٨
	٠.٨	٠.٩٥		٠.٨	٠.٩٩
عشرة فقرات	٠.١	٠.٥٣	ثلاثون فقرة	٠.١	٠.٧٧
	٠.٢	٠.٧١		٠.٢	٠.٨٨
	٠.٣	٠.٨١		٠.٣	٠.٩٢
	٠.٤	٠.٨٧		٠.٤	٠.٩٥
	٠.٥	٠.٩٠		٠.٥	٠.٩٦
	٠.٦	٠.٩٣		٠.٦	٠.٩٨
	٠.٧	٠.٩٦		٠.٧	٠.٩٨
	٠.٨	٠.٩٧		٠.٨	٠.٩٩
أربعون فقرة	٠.١	٠.٨٢	أربعون فقرة	٠.٦	٠.٩٨
	٠.٢	٠.٩١		٠.٧	٠.٩٩
	٠.٣	٠.٩٥		٠.٨	٠.٩٩
	٠.٤	٠.٩٦		٠.٩	٠.٩٩
	٠.٥	٠.٩٧			

بل وأكثر من ذلك، يمكن أن يحتوي المقياس على عوامل (أبعاد) مستقلة تماما (وبتعبير غير فني شتات من المكونات كالمفاهيم أو السمات التي تكون المقياس)، ومع ذلك يبدي المقياس مستوى ثبات مقبول أو مرتفع عند ازدياد طول الاختبار. لقد أوضح كورتينا (Cortina, 1993) أن إنشاء اختبار يحتوي على مجموعتين من الفقرات (عاملين) بحيث أن إحدى المجموعتين مستقلة عن الأخرى رغم تقارب الارتباطات بين الفقرات داخل كل منهما، فإن معامل ألفا لهذا الاختبار - إذا كان طوله ٦ فقرات (ثلاث

فقرات لكل مجموعة المشكلة للمقياس) — يساوي ٠,٤٥ ؛ ويرتفع إلى ٠,٦٥ عندما يكون طوله ١٢ فقرة، ويبلغ مستوى ٠,٧٥ عندما يكون طوله ١٨ فقرة. إن المقياس الذي قوامه ١٨ فقرة، إذا انطوى على ثلاثة أبعاد مستقلة، أو ثلاث مجموعات (أو محاور أو ثلاث فروع من الفقرات) لا ترتبط فيما بينها إطلاقا، يسفر عن معامل ألفا قدره ٠,٦٤ ؛ رغم استقلالية مجموعات الفقرات التي تكونه. يقول كورتينا: "**إذا كان طول المقياس أكثر من ١٤ فقرة، فإن معامل ألفا يمكن أن يساوي ٠,٧٠ أو يتعدى ذلك، حتى في حالة احتواء المقياس على ثلاث مجموعات أومقاييس فرعية مستقلة، وفي حالة ارتباط الفقرات داخل كل مجموعة ارتباطا منخفضا نسبيا (تحوم الارتباطات البينية حول مستوى ٠,٣٠)**. (١٥ ؛ ص١٠٢)

ما الدرس الذي يمكن استقاؤه من دراسة أثر طول الاختبار في معامل ألفا؟ كيف ينبغي أن نقرأ قيم معاملات ألفا المرتفعة في البحوث؟

الدرس الذي نستقيه هو أن ارتفاع معامل ألفا قد يكون ناتجا عن ارتفاع قيم الارتباطات بين الفقرات، وهذا ما نبحث عنه في الاختبار لكونه يمثل مدى الاتساق الداخلي للمقياس، كما أن ارتفاع معامل ألفا قد يكون بسبب طول الاختبار لا سيما إذا كان الارتباط بين فقرات المقياس (الاتساق الداخلي) ضعيفا وغير كاف لتبرير ارتفاع معامل ألفا. كما أن ارتفاع معامل ألفا قد يكون بسبب عامل آخر لم نتطرق إليه لحد الآن وهو تماثل الفقرات وتشابهها بحيث تكرّر بعضها بعضا item over-redundancy فيما تقيسه رغم اختلاف الصياغة الظاهرية لها، الأمر الذي يساهم في تضخيم حجم معامل ألفا. (١٦)

ولذلك ينبغي التريث عند الحكم على ثبات درجات المقياس بناء على معامل ألفا، ولا سيما إذا كانت قيمته مرتفعة، إلا بعد تفكير وتدبر في السؤالين التاليين:

أولاً: هل قيمة معامل ألفا التي تبدو مرتفعة للمقياس الذي استعملته في بحثي تدل على مستوى كاف من الارتباطات البينية للفقرات، أم أن هذا الارتفاع في معامل ألفا سببه طول الاختبار وليس اتساق الفقرات أو العلاقات الارتباطية فيما بينها التي تبدوا منخفضة؟

إن طرح هذا السؤال الهام ينبه الباحث بأن السعادة الغامرة التي تستبد بالباحث وهو يذكر بانشراح نفسي قيمة معامل ألفا المرتفعة (كأن تكون ثمانية من عشرة أو أعلى من ذلك) يفتقر إلى أساس أو مبرر منطقي ما لم يتدبر الباحث سبب هذا الارتفاع، هل هو راجع إلى ارتباط الفقرات فيما بينها أي إلى اتساقها، أم راجع إلى طول الاختبار، أم إلى كليهما: اتساق الفقرات وطول الاختبار. ونتائج تقدير الثبات باستعمال حزمة SPSS تمد الباحث بعناصر للإجابة إن هو أحسن استغلالها. فالحزمة يمكن أن تزوده بمصفوفة الارتباطات بين الفقرات. وهذه المصفوفة ذات أهمية كبيرة لأنها تبين ما إذا كانت الارتباطات بين الفقرات مقبولة في مستوياتها، وبالتالي يعتبر أن معامل ألفا

المرتفع يعكس الاتساق الداخلي للفقرات. أما إذا لاحظ أن عددا كبيرا من الارتباطات بين الفقرات ضعيفة فإن ذلك يجعله يقف على حقيقة ارتفاع قيمة ألفا باعتبارها لا تعكس تأثير الاتساق الداخلي للفقرات لضعف ارتباطاتها، وإنما تعكس التأثير الكبير لعدد الفقرات التي تفتقر إلى قاسم مشترك بينها. كما تزود الحزمة الباحث بمعلومة هامة جدا لم تستغل إطلاقا من طرف الباحثين في عرضهم لنتائج ألفا. وتتجلى في مربع معامل الارتباط المتعدد squared multiple correlation coefficient: R^2 بين كل فقرة وباقي فقرات المقياس. ويسمى أيضا بمعامل التحديد المتعدد multiple coefficient of determination بين الفقرة وباقي فقرات المقياس. ويدل معامل التحديد على نسبة التباين التي تفسرها (أو تشترك بها) فقرة معينة مع كافة فقرات المقياس، وبالتالي تدل على مقدار التباين المشترك (وبتعبير غير فني مقدار العلاقة) بين الفقرة وباقي الفقرات. أو مقدار المساهمة التي تقدمها الفقرة للمعلومات التي يستهدف المقياس الحصول عليها عن السمة المقاسة.

أما بالنسبة للتساؤل الإشكالي الذي يتعلق بحجم معامل الارتباط الذي يعتبر مقبولا، فاختلفت الآراء حول الموضوع لكونها لاتعتمد على أساس رياضي وانما ترجع إلى تقدير المتخصصين. فمثلا يرى روبنسون (Robinson, 1991) أنه إذا كان معامل ألفا مقداره ٠,٨٠ أو أعلى من ذلك فيكفي أن يكون الحد الأدنى لمتوسط الارتباطات بين الفقرات قدره ٠,٣٠ ؛ فإذا كان الأمر كذلك، أي إذا كان متوسط الارتباطات ٠,٣٠ أو أكثر فإن معامل ألفا المرتفع يدل على وجود تجانس داخلي كاف بين الفقرات. أما كلارك وواطسون (Clark & Watson, 1995) فيميزان بين قياس المفاهيم المتعددة الأوجه ذات النطاق الواسع التي تنطوي على مفاهيم فرعية، وبين قياس المفاهيم الضيقة المجال. فإذا استهدف المقياس قياس مفهوم واسع النطاق ومتباين فيقترح أن تتراوح متوسطات الارتباطات بين الفقرات من ٠,١٥ إلى ٠,٢٠ ؛ وأن تتراوح من ٠,٤٠ إلى ٠,٥٠ عند قياس المفاهيم الضيقة المجال. (١٧)

ثانياً: هل معاملات ألفا المرتفعة جدا أي تلك التي تتجاوز التسعة من عشرة ($\alpha \geq 0.9$) . تدل على اتساق داخلي مرتفع للفقرات (عند التأكد من أن متوسط الارتباطات بين الفقرات مرتفع بالفعل) أم أنها صنيع التشابه الكبير بين عدد من الفقرات، بحيث أنها تكرر بعضها بعضا، ولا تضيف شيئا للمقياس؟

إن الارتفاع الكبير في معامل ألفا قد لا يعكس الاتساق الداخلي للفقرات بقدر ما يعبر عن التشابه الكبير بين فقرات المقياس التي تكرر بعضها بعضا مع اختلاف طفيف في الصياغة اللغوية [١٨]. وقد نتحفظ في قبول اقتراح بعض الثقاة فيما يتعلق بالمستوى الذي ينبغي أن يبلغه معامل الثبات بعامة ومعامل ألفا بخاصة. فنانلي (Nunnally, 1967; 1978 ; Nunnally & Bernstein, 1994) مثلا رغم أنه اقترح في بداية الأمر مستوى ٠,٧٠

كحد أدنى مقبول لمعامل الثبات لأغراض البحث، ثم رفع هذا المستوى الأدنى إلى ٠.٨٠ في الطبعات اللاحقة لكتابه الشهير في نظرية القياس النفسي. غير أنه أوصى بألا يقل معامل الثبات عن ٠,٩٠ عند توظيف المقياس لأغراض إكلينيكية. ومن الأفضل في نظره أن تحوم قيمة معامل الثبات أو معامل ألفا حول ٠.٩٥ [١٩]. ونعتقد أن اشتراط مستوى ثبات لا يقل عن ٠,٩٠ ينطوي على مغالاة حتى ولو استعمل المقياس لأغراض إكلينيكية. إذ من المحتمل جدا أن فقرات المقياس الإكلينيكي الذي تتجاوز مستوى ثبات درجات فقراته ٠,٩٠ تعكس التماثل فيما تقيسه الفقرات بنسبة لا يستهان بها، ولا تمثل فقط الاتساق الداخلي للفقرات. وفي هذا السياق يرى استرينر (Streiner, 2003) أن معامل ألفا الذي تتجاوز قيمته ٠.٩٠ من المحتمل جدا أن يدل على وجود تكرار redundancy لا مبرر له فيما تقيسه الفقرات بدلا من الدلالة على الاتساق الداخلي. ولذلك يوصي ألا يتجاوز معامل ألفا قيمة ٠.٩٠ كحد أقصى [٢٠]. قد نوافق هذا الباحث فيما يتعلق بالتأثير الكبير الذي يمارسه تشابه محتوى ودلالة الفقرات في تضخيم قيمة معامل ألفا، ومع ذلك لا نشاطره رأيه الذي يقضي بتحديد سقف لقيمة معامل ألفا الذي يجب ألا يتعدى ٠.٩٠ ؛ ذلك أن ارتفاع معامل ألفا يرتد إلى أسباب كثيرة متباينة بما في ذلك سبب تشابه دلالة الفقرات. ولذلك ينبغي التوصية بالتريث في الحكم على قيمة معامل ألفا المرتفعة بأنها تدل على اتساق داخلي مرتفع للأداة، إلا بعد التأكد من أن فقرات المقياس أو عددا من هذه الفقرات التي ترتبط فيما بينها ارتباطا مرتفعا لا تتشابه تماما في دلالتها رغم اختلاف صياغتها اللغوية.

نماذج القياس

لمعرفة متى يؤدي استعمال معامل ألفا إلى نتائج دقيقة، أي تتطابق قيمته مع قيمة معامل الثبات، ومتى يترتب على استعماله تقدير غير دقيق للثبات الحقيقي، أي يمد الباحث عند استعماله بالحد الأدنى لمعامل الثبات، بدلا من تزويده بالقيمة التي تتطابق ومعامل الثبات، لا بد من التعرف على نماذج القياس وافتراضاتها، للوقوف على نموذج القياس الذي يعكس الافتراضات أو المسلمات التي قامت عليها عملية اشتقاق معامل ألفا بمختلف صيغه أو معادلاته.

في نظرية القياس، يميز عادة بين أربعة نماذج من القياس Measurement Models التي تتمايز في مدى تقيدها أو تحررها من الافتراضات والمسلمات المؤسسة لأنواع مفهوم التوازي. وهذه النماذج هي:

أولاً: نموذج التوازي التام Strict Parallel Model، أو يسمى بالنموذج المتوازي اختصارا: Parallel Model.

ثانياً: نموذج الدرجات الحقيقية المترادفة أو نموذج التماثل في الدرجة الحقيقية (طاو)، أو نموذج "طاو" المتكافيء(٩): Tau-equivalent model / τ-

---

(٩) على الرغم من أن ترجمة مصطلح "equivalent model" بالنموذج المتكافيء ترجمة دقيقة، إلا أننا نخشى عند استعمال-

equivalent model . علما بأن الرمز "طاو" يشير إلى الحرف اليوناني τ الذي عادة ما يستعمل للدلالة على الدرجة الحقيقية .

ثالثاً: نموذج التماثل في الدرجة الحقيقية أساسا ، أو نموذج "طاو" المتماثل أو المترادف في الأساس : Essentially essentially τ-equivalent model tau- equivalent model .

رابعاً: النموذج المتقارب أوالتقاربي أو التآلفي Congeneric Model .

إن نموذج التوازي التام ، أو نموذج التوازي اختصارا ، يقوم **أولاً** على افتراض أن بنود المقياس تقيس متغيرا كامنا واحدا single latent variable ، أو بعدا واحدا unidimensional يتمثل في السمة أو المفهوم الذي يشكل القاسم المشترك لبنود المقياس. ويقوم **ثانيا** على مسلمة أن الدرجات الحقيقية لكل بند من بنود المقياس متساوية ، بمعنى أنه يفترض في بنود المقياس أن تقيس السمة بنفس المقدار. أو أن مقدار الدرجة الخالية من الخطأ العشوائي متساوية على مستوى الفقرات التي تشكل المقياس. ويقوم ثالثا على افتراض أن تباين درجات الخطأ بالنسبة لجميع فقرات المقياس يكون متساويا. وباختصار ، يدل نموذج التوازي على أن فقرات المقياس تقيس كلهـا ذات المتغير الكامن (السمة) ، بنفس وحدة القياس the same metric ، وبذات الدرجة أو المستوى من الدقة ، وبنفس المقدار من الخطأ (Raykov, 1997; Graham, 2006) (٢١).

فإذا رمزنا لكل فقرة من الفقرات التي تقيس نفس البعد بالرمز i ، وإلى فقرة أخرى بالرمز j ، ونرمز إلى الفقرة الأخيرة للمقياس بالرمز k ، ورمزنا إلى الفرد الذي يجيب عن الفقرات أو المقياس بالرمز p ، لأمكن التعبير عن خاصية تساوي الدرجات الحقيقية (τ) بين الفقرات لفرد معين كما يلي :

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

أي أن الدرجة الحقيقية للفرد p على الفقرة i مساوية لدرجته الحقيقية على الفقرة الأخرى j مساوية لدرجته الحقيقية لآخر فقرة في المقياس (أو أي فقرة من فقرات المقياس). ولأمكن التعبير أيضا على تساوي تباين o^2 درجات الخطأ العشوائي (ε) كما يلي :

$$o^2(\varepsilon_{ip}) = o^2(\varepsilon_{jp}) = \ldots\ldots\ldots\ldots = o^2(\varepsilon_{kp})$$

أي أن مقدار تباين الخطأ للفرد p على الفقرة i ، مساويا لمقدار تباين الخطأ على الفقرة j مساويا لتباين خطأ أي فقرة أخرى للمقياس ، أو لتباين درجة آخر فقرة في المقياس. وعليه يمكن التعبير عن نموذج التوازي بالصيغة العامة التالية :

= لفظ "متكافئ" أن يتداخل ذلك بمصطلح "الصور المتكافئة" التي تدل على إحدى طرق تقدير الثبات، أو إحدى نماذج الثبات. ولذلك سنكتفي باستعمال كلمة "مترادف" أو "متماثل".

$$X_{pi} = \tau_p + \varepsilon_p \quad (٧)$$

وذلك للدلالة على أن أية درجة ملاحظة X_{pi} للمقياس ذي الفقرات المتوازية هي محصلة الدرجات الحقيقية المتساوية بين الفقرات للفرد p ، والدرجات الخطأ المتساوية عبر الفقرات لذات الفرد .

أما نموذج "طاو" المتماثل أو المترادف: Tau-equivalent model / τ-equivalent فيشترك مع نموذج التوازي في مسلمة أحادية البعد unidimensionality ، أي أن فقرات المقياس يجب أن تتسم بالتجانس بحيث تشترك كلها في قياس ذات البعد أو عامل كامن واحد. كما يشترك معه أيضا في مسلمة تساوي الدرجات الحقيقية بين فقرات المقياس لذات الفرد. وعليه، فإن التعبير الرياضي التوضيحي لهذا الافتراض ، أي:

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

يبقى ساريا على النموذج الحالي. غير أن الاختلاف الوحيد بينهما يكمن في احتمال عدم تساوي تباين درجات الخطأ عبر فقرات المقياس لذات الفرد.

$$o^2(\varepsilon_{ip}) \neq o^2(\varepsilon_{jp}) \neq \ldots\ldots\ldots \neq o^2(\varepsilon_{kp})$$

يظهر جليا من هذه المقارنة أن نموذج "طاو" المتماثل أو المترادف يتطلب أن يكون المقياس متجانسا أحادي البعد أو العامل، كما يتطلب أن تكون الدرجة الحقيقية للفرد متماثلة بين فقرات المقياس، غير أن درجات الخطأ لأداء الفرد ذاته على المقياس قد تختلف من فقرة لأخرى. ونتيجة لذلك قد تختلف درجة الفرد الملاحظة من فقرة لأخرى (Feldt & Brennan, 1989) (٢٢).

ويمكن أن نلخص افتراض تساوي الدرجات الحقيقية للفقرات، وفكرة أن كل فقرة تمتلك مقدارا من الخطأ الخاص بها (ε_{pi})، في المعادلة التالية :

$$X_{pi} = \tau_p + \varepsilon_{pi} \quad (٨)$$

ويلخص ريكوف (Raykov,1997a,1997b) بدقة وتركيز هذا النموذج بقوله أن نموذج "طاو" المترادف يدل على أن مفردات المقياس أو فقراته تقيس المتغير الكامن نفسه، بالوحدة نفسها on the same scale or metric (نفس السلم المتدرج للإجابة لكافة الفقرات)، وبنفس الدرجة من الدقة، لكن بمقادير مختلفة من الخطأ.

وإذا كان نموذج "طاو" المترادف اتسم بتماثل الدرجات الحقيقية ، وباختلاف قيم تباين الخطأ، فإن نموذج "طاو" المتماثل أو المترادف في الأساس: Essentially tau-equivalent model/ essentially τ-equivalent model . علاوة على عدم اشتراطه لتساوي تباينات الخطأ. فإنه يسمح بإمكان اختلاف قيم الدرجات الحقيقية للفقرات بفارق قيمة ثابتة أو عدد ثابت مضاف، مؤديا بذلك إلى اختلاف متوسطات الدرجات الحقيقية للفقرات، أو اختلاف في مستوى الدقة. (٢٣).

نموذج "طاو" المتماثل في الأساس يقوم إذن على افتراض أن كل فقرة من فقرات المقياس تقيس ذات المتغير الكامن (أحادي البعد أو متجانس)، وباستعمال وحدة القياس نفسها أو سلم القياس ذاته مما ينشأ عن ذلك تساوي تباين الدرجات الحقيقية ؛ لكن بمستويات أو درجات مختلفة من الدقة المتمثلة في اختلاف المتوسطات نتيجة تفاوت الدرجات الحقيقية بمقدار ثابت مضاف ، هذا فضلا عن اختلاف قيم تباين الخطأ. [٢٤]

ويمكن تمثيل ذلك بالمعادلة الآتية :

$$X_{pi} = (a_i + \tau_p) + \varepsilon_{pi} \qquad (٩)$$

حيث يدل الرمز a_i على القيمة الثابتة أو العدد الثابت الذي يمثل مقدار التفاوت في الدرجة الحقيقية بين كل فقرة وأخرى.

وإن اختلاف وحدة القياس أو سلم القياس يؤثر في التباين ، بحيث يميل التباين إلى التقلص عندما يكون سلم القياس ضيقا ، كأن ينطوي على درجتن فقط : صفر (للإجابة الخاطئة)، وواحد(للإجابة الصحيحة)، كما يميل إلى الارتفاع لما تزداد فئات سلم القياس أي تختلف وحدة القياس ،كأن يحتوي السلم المتدرج مثلا على خمس فئات بأوزان تنطلق من واحد إلى خمسة. أما الدقة التي تتمثل في اختلاف متوسطات الدرجات الحقيقية للفقرات فتدل على اختلاف الفقرات في قوتها على الدلالة على نفس المتغير الكامن رغم قياسها لذات المتغير ورغم تماثل سلم القياس أو وحدات القياس. أما الخطأ فيتمثل في الفرق بين الدرجة الملاحظة والدرجة الحقيقية.

ولئن كان نموذج "طاو" المترادف في الأساس ، لا يشترط تساوي متوسطات الدرجات الحقيقية التي تختلف بمقدار ثابت ، رغم تساوي قيم تباينها ، ولا يشترط أيضا تساوي تباين درجات الخطأ ، فإن النموذج الرابع والأخير: النموذج المتقارب Congeneric Model ، فيحتفظ ـ أسوة بالنماذج الثلاث التي تطرقنا إليها ـ بمسلمة تجانس مفردات المقياس أو مسلمة البعد الأحادي أو العامل العام الممثل لكافة بنود المقياس ، لكنه في المقابل يتحرر من مسلمة تساوي الدرجات الحقيقية بحيث لا تختلف في متوسطها فحسب (كما هو الشأن في نموذج "طاو" المترادف في الأساس)، بل تختلف أيضا في تباينها ، ويتحرر أيضا من قيد مسلمة تساوي تباين درجات الخطأ. ويمكن تلخيصه في المعادلة التالية :

$$X_{pi} = [a_i + \beta_i (\tau_p) + \varepsilon_{pi}] \qquad (١٠)$$

وواضح من المعادلة ، أن النموذج المتقارب ـ إذ يفترض وجود علاقة ارتباطية خطية linear correlation بين الدرجات الحقيقية للفقرات – فإنه يسمح باختلاف الدرجات الحقيقية بين البنود بفارق قيمة ثابتة a_i ؛ التي تجعل متوسط الدرجات الحقيقية التي تمثل الدقة تختلف بين الفقرات. ويسمح أيضا ـ خلافا لنموذج "طاو" المترادف ـ بضرب الدرجة الحقيقية

بمعامل ضرب (β_i : multiplicative term) قد تختلف قيمته من فقرة لأخرى؛ مما يترتب عليه اختلاف في تباين الدرجات الحقيقية للبنود.

وقصارى القول أن النموذج الرابع التقاربي يسلم بأن كل مفردة من مفردات المقياس تقيس ذات المتغير الكامن أو ذات السمة، بوحدات قياس أو سلم قياس قد يختلف بين بعض الفقرات، وبمستويات مختلفة من الدقة لاختلاف متوسطات الدرجات الحقيقية، وبمقادير مختلفة من الخطأ. [٢٥]

ولكي يتسنى مقارنة النماذج الأربعة، ولمزيد من التوضيح والتركيز، لخصنا الخصائص الأساسية لهذه النماذج في الجدول رقم (٢).

الجدول رقم (٢). مقارنة بين افتراضات أو مسلمات نماذج القياس الأربعة.

نماذج القياس الأربعة	Measurement models	أوجه المقارنة بين كل فقرة وأخرى للمقياس أو بين أجزائه: الاشتراك في بعد واحد (التجانس)	تماثل تباين الدرجات الحقيقية (وحدة القياس متماثلة)	تماثل متوسطات الدرجات الحقيقية (ارتفاع مستوى الدقة)	تماثل تباين درجات الخطأ
نموذج التوازي	Parallel model	نعم ❖	نعم	نعم	نعم
نموذج "طاو" المترادف	Tau(τ)-equivalent model	نعم	نعم	نعم	لا
نموذج "طاو" المترادف في الأساس	Essentially tau(τ)-equivalent model	نعم	نعم	لا	لا
النموذج التقاربي	Congeneric model	نعم	لا	لا	لا

❖ تدل "نعم" على أن المسلمة المنصوص عليها في عنوان العمود ضرورية. وبالمقابل تدل "لا" على أن المسلمة غير ضرورية.

وتجمع النماذج الأربع علاقة الاحتواء بحيث أن نموذج التوازي يعد حالة خاصة من نموذج "طاو" المترادف، و نموذج "طاو" المترادف يعتبر حالة خاصة من نموذج "طاو" المترادف في الأساس، وهذا الأخير هو حالة خاصة من النموذج المتقارب. والنموذج المتقارب هو أعمها، وأكثرها تحررا من الافتراضات المقيدة، وأكثرها انسجاما مع واقع المقاييس والاختبارات.

مواطن قصور معامل ألفا

لقد ثار جدل واسع حول مدى دقة معامل ألفا في تقدير الثبات على مدى أربعة عقود خلت، وذلك

نتيجة للصيت العريض الذي حظيت به، والاستعمال الواسع الذي عرفته وما زالت تعرفه، بل ازدادت وتيرة استعمالها في كافة البحوث التي توظف الأدوات (الاستبيانات، سلالم التقدير، الاختبارات، القوائم inventories ، أدوات الملاحظة، وغيرها) والتي تسعى إلى تقدير ثبات درجات الأدوات المستعملة وصدقها ضمن إجراءات البحث. لقد استهدفت الدراسات التقويمية النقدية لأداء معامل ألفا الإجابة على الإشكالية التالية: **متى يكون معامل ألفا دقيقا في تقديره للثبات، ومتى يفتقر إلى الدقة؟** وإذا رمنا التحديد الدقيق للسؤال الذي شغل مقومي أداء معامل ألفا، لكون السؤال السابق - رغم بساطته - سؤال فضفاض، ومطاط، ويفتقر إلى التحديد، لأمكن تفريع صياغته إلى السؤالين المحددين التاليين:

أولاً: في ظل أي نموذج قياس من النماذج الأربعة التي تطرقنا إليها آنفا، يكون معامل ألفا دقيقا في تقدير الثبات لكون النموذج يستجيب وافتراضات طريقة معامل ألفا ومسلماته؟ وفي ظل أي نموذج تفتقر فيه طريقة معامل ألفا إلى الدقة في تقدير الثبات، لكون النموذج لا يتقيد بعدد من المسلمات أو الافتراضات التي يتطلبها؟

ثانياً: هل تحتفظ طريقة معامل ألفا بدقتها في تقدير الثبات رغم تباين خصائص بنية المقاييس من حيث البنية العاملية للمقياس (أي في حالة تعدد العوامل المفسرة لتباين فقرات المقياس، بدلا من عامل واحد فقط، أو تعدد أبعاد المقياس بدلا من بعد وحيد unidimensionality، علما بأن انتظام الفقرات في عدة أبعاد multidimensionality يمثل خاصية أغلب المقاييس المستعملة)؟

لقد لاحظ نوفيك ولويس - في وقت مبكر - أن معامل ألفا يتطلب أن تكون الدرجات الحقيقية للفقرات متساوية، وحتى إذا اختلفت، فإنها تختلف فيما بينها بقيمة ثابتة، لكن لا يشترط تساوي تباين الخطأ بالضرورة. وأطلقا على افتراض قياس الفقرات لنفس المفهوم، وافتراض تساوي الدرجات الحقيقية بين الفقرات، أو اختلافها بمقدار ثابت فقط، مع احتمال تباين درجات الخطأ، بنموذج أو مقياس "طاو" المترادف في الأساس essentially tau-equivalent model/measure ، إذ يرجع الفضل إليهما في نحت هذا المصطلح الواسع الانتشار في نظرية القياس كما أسلفنا. الاستنتاج الهام الذي توصلا إليه هو أن معامل ألفا يقوم على ذات الافتراضات التي تأسس عليها نموذج "طاو" المترادف أساسا. **غير أن هذا النموذج بافتراضه ضرورة أن تقيس كل فقرة البعد ذاته الذي تقيسه الفقرات الأخرى للمقياس، وبنفس المقدار أو الدرجة، نموذج مثالي قائم على افتراضات مقيدة للغاية تتجافى وواقع استعمال المقاييس في الواقع.** وبالتالي، إذا لم تنسجم فقرات المقياس مع افتراضات نموذج "طاو" المترادف في الأساس، فإن استعمال معامل ألفا يؤدي إلى تقدير محافظ conservative estimate لمعامل الثبات، أي يترتب عن تبني معامل ألفا في ظل هذا النموذج تقدير للثبات بقيمة أدنى من قيمة معامل الثبات الحقيقي (٢٦).

لقد أوضحت كثير من الدراسات (e.g.,Graham, 2006; Green & Hershberger, 2000; Kamaroff, 1972 McDonald, 1999; Novick & Lowis, 1967; Zimmerman, 1972; Raykov,1997a,b; 1998, 2001a,b; 2004a,b;) على مدى العقود الثلاثة الماضية، أن معامل ألفا لكرونباخ يسفر عن تقدير غير دقيق للثبات، إذا لم يتحقق في درجات بنود المقياس شروط (مسلمات أو افتراضات) نموذج "طاو" المترادف في الأساس. وهي افتراضات وإن كانت أقل صرامة وتقييدا من نموذج التوازي التام ونموذج "طاو" المترادف، إلا أنها تبقى مع ذلك شروطا صارمة قلما تتوفر في بيانات القياس في الواقع. (٢٧).

ولذلك نجد مثلا كورتنا (cortina, 1993) - في دراسة مسحية للبحوث التي وظفت معامل ألفا في قياس الثبات - يخلص إلى النتيجة التي مفادها أن معامل ألفا لكرونباخ يعطي تقديرا تمثل قيمته الحد الأدنى للثبات lower- bound estimate of reliability ؛ عندما لا تكون فقرات أو مكونات المقاييس من نوع "طاو" المترادف في الأساس. وإذا علمنا أن الاختبارات من نوع "طاو" المترادف في الأساس، يندر وجودها لاستعصاء إنجاز أدوات قياس تحقق افتراضاته ومسلماته ، فإن تقدير معامل ألفا لكرونباخ يبقى يمثل الحد الأدنى للثبات الحقيقي لأدوات القياس المستعملة. (٢٨)

ويشير آلن وين (Allen & Yen, 1979) بصريح العبارة في كتابهما في نظرية القياس بأن " استعمال طريقة معامل ألفا وكذلك طريقة كيودر- رتشاردسن تنتجان الحد الأدنى عند تقدير ثبات المقياس" [٢٩ ؛ ص ٨٨]. أما كروكر، وألجينا (Crocker & Algina, 1986)، فيوضحان بأنه في المقاييس المتعددة الفقرات التي لا ينطبق عليها افتراضات النموذج المتوازي، أو نموذج "طاو" المترادف، أونموذج "طاو" المترادف في الأساس، فإن استعمال معامل ألفا يعطي الحد الأدنى لمعاملات ثباتها. (٣٠ ؛ ص ١٢١)

ويلخص ريكوف (Raykov, 2004) وضعية معامل ألفا بدقة وتركيز في ثلاثة جوانب، وهي:

١- بافتراض عدم ارتباط درجات الخطأ، تعكس قيم معامل ألفا الحد الأدنى لمعامل الثبات للمقاييس المتعددة البنود، بغض النظر عن البنية العاملية للمقياس. بمعنى سواء أكانت بنود المقياس وحيدة البعد أو متعددة الأبعاد، فإن استعمال معامل ألفا لكرونباخ يسفر عن تقدير منخفض للثبات. ويميل إلى تخفيض معامل الثبات الحقيقي إلى الحد الأدنى.

٢- ودائما عند افتراض استقلالية درجات الخطأ، أو عدم ارتباطها، فإن التقدير الذي يتمخض عنه استعمال معامل ألفا، يكون مساويا لثبات المقياس، عندما تكون بنود المقياس أو مكوناته من نوع "طاو" المترادف في الأساس، على الأقل (أو من نوع "طاو" المترادف، أو من النوع المتوازي. أما إذا كانت البنود من النوع المتقارب، فمعامل ألفا لا يساوي ثبات المقياس بل يمثل الحد الأدنى له) . غير أن المتطلبات الصارمة للافتراضات التي يقوم عليه نموذج "طاو" لا

تتوفر في واقع القياس النفسي والتربوي والاجتماعي.

٣- لكن في حالة ارتباط درجات الخطأ، فإن استعمال معامل ألفا يؤدي إما إلى تقدير متضخم للثبات أو إلى تقدير منخفض له. (٣١)

ومما يجدر الانتباه له في المعالجة السابقة، أن توظيف معامل ألفا يؤدي إلى تقليص معامل الثبات الحقيقي. إذا كانت بنود المقياس تمت بصلة إلى النموذج المتقارب لتحرره من القيود الصارمة (الافتراضات المقيدة) التي تقوم عليها النماذج الثلاثة الأخرى. هذا مع افتراض عدم ارتباط الأخطاء.

أما في حالة ارتباط الأخطاء، فإن استعمال معامل ألفا قد لا يؤدي إلى التقليص من القيمة الحقيقية لثبات المقياس فحسب، بل يؤدي - في الغالب - إلى تضخيم الثبات.(٣٢) إذ يوضح روزنبوم (Rozenboom, 1966) أن تطبيق فقرات المقياس في جلسة واحدة، يؤدي إلى ارتباط أخطاء القياس بين الفقرات. وارتباط أخطاء القياس تؤدي إلى تضخيم قيمة معامل ألفا. (٣٣)

ولقد قام زيمرمن وآخرون (Zimmerman et al., 1993) بفحص شامل لمشكلة تأثير أخطاء القياس المرتبطة في معامل ألفا، واستخلصوا أن معامل ألفا يقدر الثبات تقديرا تضخيميا إذا كانت معاملات الارتباط بين أخطاء القياس موجبة. ويدللون على ذلك بأنه إذا كان معامل الثبات لمقياس ينطوي على ثماني فقرات يساوي ٠.٥٠، وكان ارتباط أخطاء القياس بين أزواج الفقرات يساوي ٠.١٠، فإن معامل ألفا يعطي تقديرا للثبات يساوي ٠,٧٢ مؤديا إلى تضخيم قيمة ثبات المقياس بمقدار ٠.٢٢ (٣٤).

لقد سبق أن أوضحنا مرارا أن معامل ألفا يتطلب أن يكون المقياس أحادي البعد، أو بتعبير آخر أن تكون فقراته متجانسه. وبلغة التحليل العاملي أن تتشبع جميع الفقرات (أو أغلبها على الأقل) تشبعا مرتفعا على عامل واحد، بحيث أن هذا العامل يستوعب أو يفسر نسبة عالية من تباين فقرات المقياس أو مكوناته. **لكن ماذا لو استعملنا معامل ألفا لتقدير ثبات درجات مقياس متعدد الأبعاد أو العوامل (أي ينطوي على عاملين أو أكثر)؟ هل تعطي طريقة معامل ألفا تقديرا دقيقا للثبات عند خرق مسلمة التجانس، أو اختلال شرط البعد الواحد للمقياس؛ أم تؤدي إلى تقدير تقليصي أو تضخيمي لمعامل الثبات؟ هل يمكن أن نعتبر معامل ألفا مؤشرا دقيقا لتجانس المقياس أي لأحادية بعد المقياس؟ أو بتعبير آخر، هل ارتفاع قيمة معامل ألفا، يعتبر دليلا على أن المقياس مرتفع التجانس، وعلى أنه ينطوي على بعد واحد فقط؟**

لنبدأ بمعالجة السؤالين الأخيرين المتعلقين بإمكانية اعتبار معامل ألفا مؤشرا دقيقا لأحادية المقياس أو تجانسه. تفيد أغلب الدراسات التي عالجت هذه القضية (eg. Cortina, 1993; Feldt & Qualls, 1996;Green, Lissitz, & Mulaik, 1977; Hattie, 1985; Miller, 1995; Raju, 1982; Terwilliger & Lele, 1979 أن معامل ألفا يعتبر مؤشرا (دليلاً) غير دقيق على

تجانس فقرات المقياس أو مكوناته، أو على أحادية البعد Unidimensionality للمقياس. ولذلك، تم التخلي عنه من طرف السيكومتريين كمؤشر لتجانس الفقرات [٣٥] . وإذا كان الأمر كذلك لدى المتخصصين، فإن اعتبار معامل ألفا المرتفع دليلا على تجانس فقرات المقياس أو اشتراكها في بعد واحد، ما زال اعتقادا مستحكما ومتفشيا لدى الباحثين. وبذلك يقعون فريسة المغالطة التي يسميها المناطقة بمغالطة تأكيد النتائج أو المترتبات Fallacy of affirming the consequences . إذ تقوم المغالطة على الاعتقاد بأنه: إذا كانت س، فإذن ص. وبالتالي إذا كانت ص، فإذن س. أي "ارتفاع التجانس إذن ارتفاع معامل ألفا"، وبالتالي. فإن "ارتفاع معامل ألفا، إذن ارتفاع التجانس". فإذا كان شرط تجانس المقياس قائما، فإذن يرتفع مستوى معامل ألفا، ويستدل من ذلك بالتالي، أن ارتفاع معامل ألفا دليل على تجانس فقرات المقياس.

أما فيما يتعلق بسلوك معامل ألفا عند اختلال شرط البعد الواحد للمقياس، أي اختلال شرط التجانس، بحيث ينطوي المقياس على بعدين أو عدة أبعاد تمثل فقرات المقياس، فإن استعمال معامل ألفا يسفر عن تقدير أدنى من مستوى معامل الثبات الحقيقي للمقياس ككل. وفي هذا السياق، يرى شميدت وهانتر (Schmidt & Hunter, 1996) – عقب دراسة مسحية وتقويمية لأخطاء القياس ـ أن انتقاء معامل ألفا لتقدير الثبات عندما ينطوي المقياس على عدد من الأبعاد (غير متجانس) يعتبر اختيارا غير مناسب، أو غير صالح لتقدير الثبات. (٣٦)

ولقد اجتهد كورتين (Cortina, 1993) باستعمال فنية المضاهاة الإحصائية statistical simulation technique في إثبات ميل معامل ألفا إلى الانخفاض بسبب تعدد أبعاد المقياس (الافتقار إلى التجانس)، والارتفاع بسبب العلاقات الارتباطية البينية للفقرات (الاتساق). ولكنه لاحظ أيضا أن طول الاختبار (عدد الفقرات) قد يسفر عن قيمة مرتفعة لمعامل ألفا عند انتظام الفقرات في ثلاث مجموعات أو أبعاد أو عوامل مستقلة فيما بينها (غياب التجانس والاتساق معا في بنية الاختبار). ولذلك يخلص من دراسته إلى القول بأن استعمال معامل ألف مفيد عندما تتوفر مسلماته في البيانات، أي عندما يصنف التباين الخاص للفقرة في زمرة أخطاء القياس. (٣٧)

ولعل الفكرة الأخيرة، أي التباين الخاص، تقتضي توضيحا لأن هذا الشرح يمدنا بتعليل لافتقار معامل ألفا للدقة في غياب تجانس المقياس. إن كرونباخ (Cronbach, 1947) نفسه يقسم التباين الكلي للمقياس المتعدد الفقرات إلى تباين العامل العام، وهو التباين المشترك بين جميع فقرات المقياس؛ وتباين فقرات مجموعة العوامل، أو العوامل الطائفية، أي مقدار التباين الذي تشترك فيه مجموعة من الفقرات؛ والتباين الخاص بكل فقرة، وتباين الخطأ العشوائي Error Variance (٣٨). تقوم طريقة معامل ألفا لكرونباخ على منطق نسبة الدرجة الحقيقية التي تمثل كل التباين المنتظم الذي تشترك فيه كل الفقرات

(مستبعدة التباين المشترك بين فقرات المجموعات أو العوامل) إلى التباين الكلي للمقياس المنتظم وغير المنتظم أو العشوائي. ومعنى ذلك أنه عند استعمال معامل ألفا، فإن التباين المشترك بين فقرات كل مجموعة من مجموعات الفقرات المكونة للمقياس، والتباين الخاص بكل فقرة، يصنف في زمرة الأخطاء العشوائية رغم أنها ليست كذلك. و ينتج عن هذا السلوك تقليص في مقدار تباين الدرجة الحقيقية وتضخيم في مقدار تباين الخطأ، مما يؤدي إلى تقدير منخفض للثبات الحقيقي للمقياس. أما الطرق البديلة التي سنأتي على بعضها التي تراعي تعدد البنية العاملية أو تعدد أبعاد المقياس، فلا تقوم على اعتبار تباين فقرات كل عامل من العوامل المشكلة للمقياس من قبيل الخطأ العشوائي، بل تعتبر أن تباين فقرات العوامل تصب في مجرى تباين الدرجة الحقيقية، لكونها تباينا منتظما على مستوى كل عامل وليس تباينا عشوائيا، وبالتالي ليست من الأخطاء العشوائية. ولذلك تعطي هذه الطرق تقديرا دقيقا للثبات مقارنة بتقدير معامل ألفا.

ولقد حاول كرونباخ وزملاؤه (Cronbach, Schonemann, & McKie, 1965) تدارك الأمر باقتراح تعديل لصيغة معامل ألفا، تأخذ بعين الاعتبار المجموعات الفرعية التي يتكون منها المقياس، وأسموا هذه الصيغة المعدلة بمعامل ألفا الطبقي Stratified Alpha ؛ وفيما يلي معادلتها:

$$\text{Strat } \alpha = 1 - \frac{\sum_{j=1}^{c} \sigma^2_{x_j} \left(1 - {}_{\alpha}\rho_{x_j x'_j}\right)}{\sigma^2_x} \qquad (١١)$$

يدل التعبير ${}_{\alpha}\rho_{x_j x'_j}$ على تقدير معامل ألفا لفقرات مجموعة من مجموعات (طبقة من طبقات) المقياس، ويدل التعبير $\sigma^2_{x_j}$ على تباين هذه المجموعة (الطبقة)، وتدل سيجما $\sum_{j=1}^{c}$ على أن عملية الجمع تشمل حواصل الضرب لتباين درجات كل مجموعة في نتيجة طرح معامل ألفا لكل مجموعة من الواحد الصحيح. بحيث يبدأ الجمع انطلاقا من المجموعة الأولى j=1 (الطبقة الأولى) إلى آخر مجموعة c (آخر طبقة). أما التعبير σ^2_x فيدل على تباين المقياس ككل. (٣٩)

غير أنه يشترط أن تكون كل مجموعة من مجموعات الفقرات، أو كل طبقة من طبقات المقياس من نوع "طاو" المترادف في الأساس على الأقل (٤٠). بمعنى يشترط أن تقيس فقرات كل مجموعة بعدا واحدا، وأن يكون تباين الدرجات الحقيقية متساويا، ولا تختلف هذه الدرجات إلا بمقدار ثابت، أما تباين الخطأ فلا يشترط أن يكون متماثلا بين الفقرات. معنى ذلك، إذا كانت مجموعات الفقرات المكونة للمقياس أو بعضها من نوع النموذج المتقارب Congeneric model، فمن المحتمل أن ينحو معامل ألفا الطبقي نحو تقدير منخفض لثبات المقياس.

بعض الصيغ البديلة

لقد تطرقنا إلى الحالات التي تحتفظ فيها معامل ألفا بدقتها في تقدير الثبات والحالات التي تفتقر فيها إلى الدقة في تقدير الثبات، ورأينا أن هذا التقدير يكون دقيقا في ظل توفر مسلمات نموذج "طاو" المترادف في الأساس في البيانات، وفي ظل النماذج التي تمثل حالات خاصة له: نموذج "طاو" المترادف ونموذج التوازي. ورأينا أن هذا التقدير يكون متحيزا يفتقر إلى الدقة في ظل النموذج الرابع: النموذج المتقارب. **فإذا ارتبطت فقرات المقياس ببعد واحد، لكن مع اختلاف تباين درجاتها الحقيقية، واختلاف تباين خطئها؛ يعكس معامل ألفا الحد الأدنى للثبات الحقيقي للمقياس، كما يمكن أن يعكس تقديرا تضخيميا في حالة ارتباط درجات الخطأ.**

أما إذا انطوى المقياس على بنية عاملية متعددة الأبعاد، فإن استعمال معامل ألفا يسفر عن تقدير أدنى من مستوى معامل الثبات الحقيقي.

ونتيجة لمواطن القصور هذه، اقترحت عدة معادلات بديلة تعطي تقديرا أدق من معامل ألفا، عندما لا تتوفر مسلمات نموذج "طاو" المترادف في الأساس، على الأقل؛ أي تنسجم هذه المعادلات ومسلمات النموذج الأكثر تحررا من المسلمات المقيدة: النموذج المتقارب، وبعضها يأخذ بعين الاعتبار استقلالية الأخطاء أو ارتباطها، وأحادية البنية العاملية للمقياس أو تعددها. ولذلك، سنتعرض لبعض هذه الطرق البديلة مع أمثلة تطبيقية لبعضها.

معامل ثيتا: (θ) Theta Coefficient

لقد قام ارمور (Armor, 1974) باقتراح هذه الطريقة لتقدير الثبات، لأنها تعطي تقديرا أدق لمعامل الثبات عندما تتجاوز بيانات القياس نموذج طاو المترادف في الأساس، إلى النموذج المتقارب الذي لا يشترط في الدرجات التساوي في الدرجات الحقيقية، ولا التساوي في تباين درجات الخطأ. كما تستعمل أيضا عند انتهاك المقياس شرط واحدية البعد أو التجانس، لينطوي على عوامل أو أبعاد، شريطة أن يستأثر أحد العوامل بتفسير نسبة كافية من تباين فقرات المقياس، كأن يفسر مثلا نسبة لا تقل عن ٤٠٪ من التباين الكلي للمقياس(٤١). وبتعبير آخر، في وجود أكثر من عامل واحد مشكل للمقياس، فيجب أن ينفرد أحدها بالتفوق على العوامل الأخرى في تمثيل تباين المقياس. وكلما تفوق هذا العامل عن غيره في تمثيل تباين المقياس (معلومات المقياس بتعبير غير فني)، بحيث تزداد نسبة التباين الذي يقوى على تفسيره عن ٤٠٪، كلما كان الوضع أفضل، وأكثر انسجاما مع منطق معادلة ثيتا.

لعل التوضيح المبسط السابق يجعل القارئ يستنتج أن معامل ثيتا يستعمل بعد إجراء التحليل العاملي على بيانات المقياس. كما أن القارئ الذي يتوفر على خلفية ولو كانت بسيطة عن التحليل العاملي يستنتج أن الوصف السابق يشير إلى أسلوب

معين شائع الاستعمال من أساليب التحليل العاملي، وهو أسلوب التحليل إلى المكونات الأساسية أو الرئيسية Principal Components Analysis(PC).

تنطلق طريقة التحليل إلى المكونات الرئيسية - شأنها في ذلك شأن طرق التحليل العاملي الأخرى - من تحليل مصفوفة الارتباطات بين المتغيرات، أو بين جميع الفقرات المكونة للمقياس، للكشف عن المجموعات التي ترتبط عناصرها أو فقراتها ارتباطا عاليا. فيشرع أسلوب المكونات الأساسية في استخراج المجموعة ذات الارتباط المرتفع بين فقراتها بحيث يكوّن من هذه الفقرات (المتغيرات) تشكلة أو تركيبة خطية linear combination لتحقيق أقصى مجموع ممكن لمربعات الارتباطات بين هذا التشكلة أو التركيبة الخطية وبين الفقرات الأصلية. أو بتعبير آخر، للتوصل إلى تحديد أوزان للمتغيرات الأصلية لتحقيق تركيبة خطية من المتغيرات بأوزانها تفسر أقصى نسبة من تباين متغيرات المقياس أو فقراته. وتدعى هذه التركيبة الخطية الكامنة (البنية التحتية) التي تمثل أقصى تباين في فقرات المقياس بالعامل، ومربع معامل الارتباط بين هذه البنية الخطية وبين الفقرات الأصلية يدعى بالتشبع. عند استخراج العامل الأول بهذه الطريقة، يستخرج العامل الثاني وذلك بتكوين أفضل تشكلة خطية ثانية من المتغيرات بأوزان أخرى من التشكيلات الخطية العديدة الممكنة تفسر أقصى نسبة من التباين الباقي عند استخراج العامل الأول. أي أن العامل الثاني مستقل عن العامل الأول بحيث ما يمثله من تباين في المقياس ليس تكرارا أو استنساخا للتباين الذي يمثله العامل الأول وذلك لعدم ارتباطهما. ويمثل العامل الثاني أفضل تشكلة خطية لأوزان المتغيرات أو الفقرات الأصلية بعد التشكلة الخطية التي يمثلها العامل الأول لكونها تحقق أقصى مجموع مربع ارتباط بالفقرات الأصلية، أي أعلى قيمة مميزة أو جذر كامن eigenvalue . وتدل القيمة المميزة عل المساحة المشتركة (نسبة التباين المشترك) بين العامل وبين الفقرات. ويسهل فهمه عندما نتصوره بأنه يدل على قوة العلاقة التي تربط العامل بعدد من فقرات المقياس.

نستخلص مما سبق أن أسلوب التحليل إلى المكونات الأساسية يستخرج عوامل متدرجة من حيث أهميتها بدءا بالعامل الأول ونزلا إلى العامل الأخير. إذ يتسم العامل الأول بتمثيل أكبر نسبة من التباين، كما يحتوى على أعلى قيمة مميزة، أي أن علاقته بفقرات الاختبار التي يمثلها أقوى من علاقة العوامل الأخرى المستخرجة بفقرات المقياس التي يمثلها.

بعد هذا التقديم لمنطق التحليل إلى العوامل الأساسية، نعود إلى معامل ثيتا. لقد اقترح أرمور (Armor, 1974) عدة صيغ مترادفة لمعامل ثيتا لعل أكثرها شهرة واستعمالا المعادلة التالية:

$$\theta = \frac{k}{k-1}\left(1 - \frac{1}{Eigen_1}\right) \qquad (١٢)$$

ويدل k على عدد الفقرات (أو المتغيرات).

أما اللفظ $Eigen_1$ فيدل على الجذر الكامن للعامل الأول المستخرج بطريقة المكونات الأساسية[٤٢]. وتتوفر هذه المعلومات في نتائج التحليل العاملي عند استعمال أية حزمة إحصائية. ويبقى مجرد التعويض في المعادلة أعلاه لتقدير ثبات درجات المقياس.

يفضل استخدام هذه الطريقة إذا اتسمت مصفوفة التشبعات على العوامل قبل التدوير بالمواصفات التالية: ١ـ يجب أن يفسر العامل الأول (المكون الرئيسي الأول) المستخرج أعلى نسبة من التباين بحيث ينبغي ألا تقل هذه النسبة عن ٤٠ ٪ من تباين فقرات المقياس. ٢ـ ينبغي أن تفسر المكونات الأخرى المرتبة تنازليا نسبا متقاربة ومتناقصة من التباين المتبقي. ٣ـ كل الفقرات أو معظمها ينبغي أن تتشبع تشبعا كافيا (لا تقل عن ٠.٣٠) على المكون أو العامل الأول. ٤ـ كل الفقرات أو معظمها ينبغي أن يكون تشبعها على العامل أو المكون الأول أعلى من تشبعها على باقي المكونات أو العوامل. (٤٣ ؛ ص ٦٠)

معامل أوميجا (١٠) Omega Coefficient (Ω) للمقياس كله.

لا يقل معامل أوميجا أهمية ودقة في تقدير ثبات المقاييس عن معامل ثيتا. وإذا كان معامل ثيتا يستعمل عقب التحليل إلى المكونات الأساسية، فإن أوميجا تستعمل بعد التحليل العاملي بالطرق الأخرى غير المكونات الأساسية، أي عند استعمال طريقة المحاور الأساسية Principal Axis Factoring (PAF)، أوطريقة الاحتمال الأقصى Maximum Likelihood (ML)، أو طريقة التحليل العاملي الصوري Image Factoring (IMAGE)، أو طريقة المربعات الصغرى غير الموزونة Unweighted Least Square(ULS) ، أو طريقة المربعات الصغرى المعممة Generalized Least Squares (GLS)، أو أي طريقة أخرى.

لقد طور كل من هيس، وبوهرنستيدت (Heise & Bohrnstedt, 1970) معاملا لتقدير الثبات بناء على نتائج التحليل العاملي، وفضلا تسميته بأوميجا. إن الشكل العام لمعامل أوميجا القائم على توظيف تباين كل الفقرات وتغايرها، وقيم شيوعها أو اشتراكياتها communalities هو:

$$Omega\ coefficient\ (\Omega) = 1 - \frac{\sum \sigma_i^2 + \sum \sigma_i^2\, h_i^2}{\sum_i \sum_j \operatorname{cov}(x_i, x_j)} \quad (١٣)$$

حيث يدل التعبير $\sum \sigma_i^2$ على مجموع تباين فقرات المقياس، ويدل التعبير $\sum \sigma_i^2\ h_i^2$ على مجموع حواصل ضرب تباين كل فقرة من فقرات المقياس في قيمة شيوعها، علما بأن لكل فقرة قيمة شيوع معينة، والتي تساوي مجموع مربعات تشبع كل فقرة على العوامل المستخرجة في حالة التدوير المتعامد

---

(١٠) أحيانا يرمز لها بالحرف الإغريقي الكبير: Ω ، وأحيانا يرمز لها بالحرف الإغريقي الصغير: ω

الذي يقوم على استقلال العوامل المستخرجة أوعدم ترابطها. أما المقام $\sum \sum \text{cov}(x_i x_j)$ فيدل على تغاير بين الفقرات الموجودة في الأعمدة والفقرات الموجودة بالصفوف لمصفوفة التغاير بين فقرات المقياس. وبتعبير آخر يدل التعبير السابق على مجموع جميع قيم التغاير بين فقرات المقياس [٤٤]. وتجدر الإشارة إلى أن كل الحزم الإحصائية تزود المستعمل للتحليل العاملي بالإحصاءات الوصفية لجميع الفقرات بما في ذلك التباين والانحراف المعياري لجميع الفقرات، وبمصفوفة الارتباطات ومصفوفة قيم التغاير بين الفقرات، والقيمة المميزة لكل عامل، وقيم الشيوع لكل فقرة أو متغير، علاوة على مصفوفة العوامل بتشبعات الفقرات عليها قبل التدوير وبعده. وبالتالي يقوم المستعمل لأية حزمة بالحسابات البسيطة السهلة التي تتطلبها المعادلة رقم (١٣).

وتختصر المعادلة السابقة إلى الشكل التالي إذا أردنا الاشتغال على الارتباطات بدلا من التباين والتغاير:

$$Omega coefficient(\Omega) = 1 - \frac{k - \sum h_i^2}{k + 2\sum r_{x_i x_j}} \quad (١٤)$$

تدل k على عدد الفقرات، وتدل $\sum h_i^2$ على مجموع قيم الشيوع لفقرات المقياس. أما التعبير الجديد $\sum r_{x_i x_j}$ فيدل على مجموع الارتباطات بين فقرات المقياس التي تظهرها الحزم الإحصائية بشكل مصفوفة معاملات الارتباط بين الفقرات.(٤٥، ص ٦٥)

ومن الأهمية بمكان، الإشارة إلى أن معامل أوميجا يستعمل لتقدير ثبات درجات المقياس سواء أكانت بنية المقياس بسيطة أي تحتوي على بعد أو عامل واحد، أم كانت بنيته مركبة أو معقدة بحيث يحتوى على عدد من الأبعاد أو العوامل. فإذا احتوى المقياس على بعدين أو عاملين أو أكثر باستعمال طرق التحليل العاملي (باستثناء طريقة المجموعات الرئيسية) فإن استعمال معامل أوميجا يعطي تقديرا للثبات على مستوى المقياس ككل بغض النظر عن تعدد عوامله، ولا يعطي تقديرا للثبات على مستوى كل بعد في المقياس.

ولعل المقارنة بين المعاملات الثلاثة: معامل ألفا لكرونباخ، ومعامل ثيتا، ومعامل أوميجا، من حيث مدى دقتها، أو اقترابها من معامل الثبات الحقيقي، تفيد في توضيح وإبراز مدى قدرة كل منها على تقدير الثبات.

يتمخض استعمال المعاملات الثلاثة عن نتائج متكافئة إن لم تكن متماثلة إذا كانت بيانات القياس تنسجم ومتطلبات النموذج المتوازي، أو تتماشى ونموذج "طاو" المترادف، أو توافق على الأقل نموذج "طاو" المترادف في الأساس لأن النموذج الأخير - كما أوضحنا سالفا (وببعض الإسهاب) - أقلها تشددا في صرامة الافتراضات التي يجب أن تتوفر في بيانات الفقرات.

أما إذا كانت بيانات القياس لا تستجيب لمتطلبات نموذج "طاو" المترادف في الأساس (وهي لا

تحقق بالضرورة افتراضات نموذج التوازي، وافتراضات نموذج "طاو" المترادف)، كأن تكون درجات الفقرات تنسجم والنموذج الأكثر تحررا من صرامة الافتراضات التي تقوم عليها النماذج الثلاثة السابقة، (أي عند اختلاف الدرجات الحقيقية وتباين الخطأ)، وعند انطواء المقياس على بنية متعددة العوامل أو الأبعاد، فإن معامل ألفا يعطي أدنى تقدير لمعامل ثبات درجات المقياس، في حين أن معامل أوميجا يعطي أفضل تقدير بحيث تقترب قيمته اقترابا كبيرا من معامل الثبات الحقيقي. أما ثيتا فتحتل موقعا وسطا في هذه الحالة، بحيث تتفوق على معامل ألفا وتتخلف عن معامل أوميجا في دقة تقدير الثبات.(٤٦)

مؤشرات ثبات المفهوم أو التكوين الفرضي Construct reliability (CR)

يدل ثبات المفهوم على مدى دقة أو اتساق المؤشرات التي قد تتمثل في الفقرات، أو قد تكون مقاييس جزئية أو فرعية، أو قد تتخذ أشكال أخرى، في قياس مفهوم معين. إذ يتوقع في المؤشرات (فقرات مقياس مثلا) التي يفترض الباحث أنها تنتمي إلى مفهوم معين أن تنطوي على نسبة من التباين المشترك بينها وبين المفهوم. فإن كانت هذه النسبة مثلا تساوي ٦٠٪، فمعنى ذلك أن المفهوم الذي يراد قياسه يفسر أو يحدد نسبة تباين في المؤشر او الفقرة قدرها ٦٠٪ من مجمل تباين المؤشر أو الفقرة. وبالتالي فباقي التباين الذي لا يحدده المفهوم في الفقرة يدعى بالخطأ، ومقداره في هذه الحالة ٤٠٪.

يقدر ثبات المفهوم في الغالب باستعمال طريقة التحليل العاملي التوكيدي CFA: Confimatory Factor Analysis. والتحليل العاملي التوكيدي، على خلاف التحليل العاملي الاستكشافي EFA: Exploratory Factor Analysis، ينطلق من خلفية تنظيرية على أساسها يفترض الباحث سلفا وقبل إجراء التحليل العاملي، أي المتغيرات الملاحظة أو المقاسة (المؤشرات، أو الفقرات، أو المقاييس الفرعية أو غيرها) تنتمي إلى أي العوامل الكامنة. فإذا كانت المؤشرات تتكون من فقرات مقياس معين، واعتقد الباحث - وفقا لتنظيره - بأن المفهوم ينطوي على عاملين أو بعدين، فإن الباحث سيقوم بتعيين الفقرات التي تتشبع على العامل الأول، والفقرات التي تتشبع على العامل الثاني سلفا، أي قبل إجراء التحليل العاملي. وبالتالي تكون وظيفة التحليل العاملي اختبار حسن مطابقة البنية العاملية المفترضة (انتماء مجموعة معينة من الفقرات إلى البعد الأول، وانتماء مجموعة الفقرات الأخرى إلى البعد الثاني)، مع البيانات.

إن إحدى المعادلات التي تمكن من تقدير مدى دقة عدد من المؤشرات في قياس بعد أو مفهوم (عامل كامن)، في سياق التحليل العاملي التوكيدي بخاصة، أو التحليل العاملي بعامة، هي المعادلة التالية (٤٧):

$$CR = \frac{\left(\sum_{i=1}^{p}\lambda_i\right)^2}{\left(\sum_{i=1}^{p}\lambda_i\right)^2 + \sum_{i=1}^{p}\varepsilon_{ii}} \qquad (١٥)$$

يدل التعبير $\sum_{i=1}^{p}\lambda_i$ على مجموع تشبعات المؤشرات (الفقرات) على العامل ؛ كما يدل التعبير $\sum_{i=1}^{p}\varepsilon_{ii}$ على مجموع تباين الخطأ للمؤشرات أو الفقرات (أو مجموع خطأ القياس للمؤشرات). وخطأ القياس للمؤشرات ، في سياق التحليل العاملي ولا سيما التحليل العاملي التوكيدي ، يمكن تقديره بتربيع تشبع كل مؤشر على العامل ، وحذف ناتج تربيع التشبع من الواحد الصحيح. ومعنى ذلك أن التعبير الدال على خطأ القياس للمؤشرات: $\sum_{i=1}^{p}\varepsilon_{ii}$ يمكن استعاضته بالتعبير المرادف التالي: $\sum_{i=1}^{p}\left(1-\lambda_i^2\right)$ ، وبالتالي تتخذ المعادلة رقم (١٥) ، الشكل المرادف الآتي :

$$CR = \frac{\left(\sum_{i=1}^{p}\lambda_i\right)^2}{\left(\sum_{i=1}^{p}\lambda_i\right)^2 + \sum_{i=1}^{p}\left(1-\lambda_i^2\right)} \qquad (١٦)$$

وعلى الرغم من الأهمية التي تكتسيها معادلة ثبات المفهوم رقم (١٥) أو رقم (١٦) ، والإقبال المتزايد على استعمالها ، إلا أنها كانت موضع انتقاد لسببين : يتمثل السبب الأول في نزعة معامل ثبات المفهوم إلى الانخفاض إذا كانت إشارة تشبع أو تشبعين أو عدد قليل من التشبعات سالبة. وهذا السلوك يضاد طبيعة معامل الثبات الذي يقوم على القيم المطلقة للتشبعات ، ولا يتأثر بالإشارات الموجبة أو السالبة للتشبعات. ويتجلى السبب الثاني في أن معامل ثبات المفهوم قد تنخفض قيمته إذا أضيف متغير (فقرة مثلا) إلى فقرات المقياس. ويفترض في معامل الثبات أن تزداد قيمته بدلا من انخفاضها ، أو أن يحتفظ بمستواه على الأقل عند إضافة مؤشر أو متغير إلى المقياس.

ففي سياق النمذجة باستعمال المعادلات البنائية Structural Equation Modeling ، حيث يلجأ إلى التحليل العاملي التوكيدي لاختبار صدق النموذج القياسي المفترض ، أو اختبار صلاحية المؤشرات الملاحظة أو المقاسة في الدلالة على مفهوم كامن (أو مفاهيم أو تكوينات فرضية كامنة Latent constructs) وذلك بتقدير صدق هذه المؤشرات المقاسة (مجموعات الفقرات مثلا بحيث كل مجموعة تشكل مقياسا لمفهوم كامن) أو ثباتها ، تميل بعض الدراسات المتخصصة في تقدير الثبات إلى تفضيل المعادلة التي تدعى بأوميجا الموزونة (Weighted Omega: Ω_w) ، كما تدعى أحيانا بـ : Coefficient-H [٤٨]. وهي كما يلي :

$$\Omega_w = \frac{\sum \frac{\lambda_i^2}{1-\lambda_i^2}}{1+\sum \frac{\lambda_i^2}{1-\lambda_i^2}} \qquad (١٧)$$

وإذا كانت دلتا مربع λ_i^2 تدل على مربع التشبع النمطي pattern loading لمؤشر معين أو فقرة معينة على عامل معين، فإن سر اختلاف معادلة أوميجا عما سواها يكمن في التعبير التالي $\frac{\lambda_i^2}{1-\lambda_i^2}$ ، إذ يدل مربع تشبع الفقرة على العامل على المساحة المشتركة (التباين المشترك) بين الفقرة والعامل أو بين الفقرة والمفهوم المقاس. ولذلك يدل على الدرجة الحقيقية للفقرة . وحذف مربع تشبع الفقرة (الدرجة الحقيقية) من الواحد الصحيح كما هو مبين في المقام يدل على تباين الخطأ أو على الخطأ اختصارا. فالكسر إذن يعبر عن نسبة الدرجة الحقيقية إلى الخطأ، ويمثل وزن فقرة معينة أو مدى أهميتها في تحديد المفهوم. فإذا كان تشبع فقرة معينة ٠.٨ مثلا، فإن نسبة الدرجة الحقيقية إلى الخطأ (الكسر السابق) أي وزن الفقرة يساوي ١.٧٧ ، أما إذا كان تشبع الفقرة أقل من ذلك كأن يكون مثلا ٠,٥ ؛ فإن وزن الفقرة يقل إلى ٠.٣٣ ومن الواضح أن المعادلة تراعي مدى إسهام كل فقرة في تحديد المفهوم الكامن بدلالة الأوزان التي اشتقت من نسبة الدرجة الحقيقية إلى الخطأ. وهذا ما افتقدناه في المعادلة التي قامت على جمع التشبعات كما هي بدون اشتقاق أوزان لها تعكس تفاوتها في الدلالة على المفهوم، ولم تربع التشبعات إلا بعد أن تم دمجها في مجموع كما يدل على ذلك التعبير الجوهري التالي في بناء المعادلة رقم ١٥ أو رقم ١٦ :

$$\left(\sum_{i=1}^{p}\lambda_i\right)^2$$

ولذلك نجد معادلة أوميجا الموزونة تتمتع بخصائص يفتقر إليها معامل ثبات المفهوم، ومنها أن استعمالها يسفر عن أقصى تقدير لمعامل الثبات الحقيقي، إذ تعتبر أكثر دقة من معامل ألفا ومن معامل ثبات المفهوم. وأن قيمتها لا تقل عن مربع أعلى تشبع لأهم فقرة أو مؤشر من مؤشرات المفهوم أو المقياس، بينما قد تكون نتائج استعمال المعادلات الأخرى دون ذلك. كما أن معامل أوميجا الموزونة لا تتأثر بوجود تشبع أو تشبعات نمطية سالبة، ولا تنخفض قيمتها إطلاقا عند إضافة فقرة أو فقرات إلى مجموعة الفقرات التي تقيس المفهوم الكامن. (٤٩)

ولعلنا نحتاج إلى مثال توضيحي نوظف فيه معامل المعادلات السابقة. لتقدير ثبات المقاييس الأربعة لأبعاد المعرفة وفقا لتصنيف بلوم المعدل الذي تمت مراجعته وتعديله حديثا [٥٠]. انطوى مقياس بعد المعرفة الوقائعية أو الحقائقية factual dimension على فقرتين، ومقياس بعد التصور الذهني Conceptual knowledge على ثلاث فقرات، ومقياس بعد المعرفة الإجرائية Procedural knowledge على ثلاث فقرات، ومقياس بعد المعرفة الديناميكية أو ما وراء المعرفة dynamic knowledge / metacognition على ثلاث فقرات. (الشكل رقم ١)

وعقب استعمال التحليل العاملي التوكيدي

(بطريقة الاحتمال الأقصى). والتحقق من مطابقة النموذج المفترض الموضح في الشكل رقم (١) للبيانات. رصدت في الجدول رقم (٢) قيم تشبع الفقرات أو المؤشرات على المفهوم أو العامل الذي تنتمي إليه، وتقديرات الثبات للمقاييس الفرعية (العوامل) الأربع، باستعمال ثلاث طرق: طريقة معامل ألفا، وطريقة معامل ثبات المفهوم، وأخيرا طريقة أوميجا الموزونة، بهدف المقارنة بين نتائجها، وتوضيح كيفية حسابها.

وتجدر الإشارة إلى أنه لحساب معامل ألفا لكرونباخ بدلالة (بوجود) قيم التشبعات بدلا من استعمال القيم الخام، استعملنا المعادلة التالية المشتقة من المعادلة الأصلية رقم (٣) التي سبق استعراضها وتحليلها.

$$Alphd(\alpha) = \frac{k}{k-1}\left(1-\frac{k}{k+\left(\sum\lambda_i\right)^2-\sum\lambda_i^2}\right) \quad (١٨)$$

و لا تنطوي هذه المعادلة على رموز جديدة، فهي رموز مألوفة، ولقد سبق تعريفها. وتتطلب فقط معرفة قيم التشبع على عامل معين (λ_i)، وعدد الفقرات (k)(٥٠).

ولتوضيح طريقة تقدير معاملات الثبات الثلاث، سنركز على كيفية حساب ثبات عامل أو بعد المعرفة الوقائعية الذي انطوى على مؤشرين أو فقرتين.

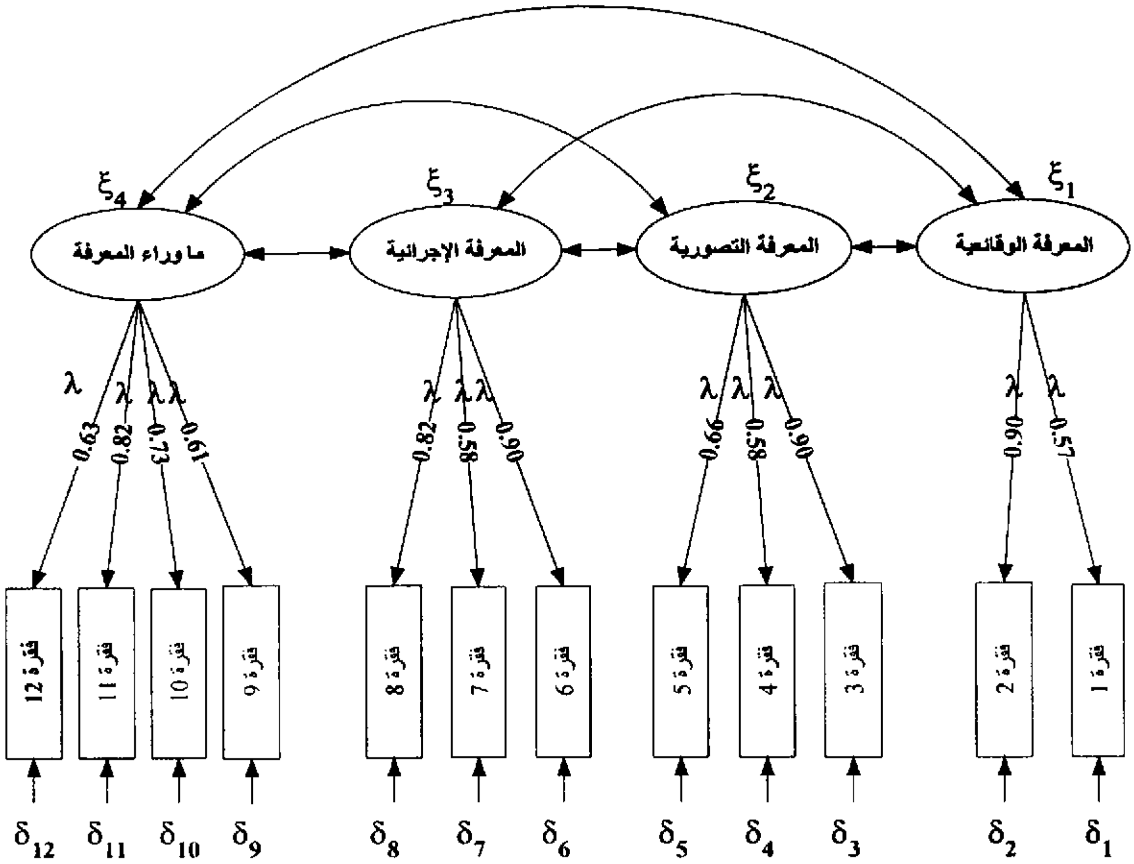
ولحساب معامل ألفا لبعد المعرفة الوقائعية بمعرفة عدد الفقرات (فقرتان)، وقيم التشبع (يوجد تشبعان: ٠.٥٧١ ؛ ٠.٩٠٤)، نعوض في المعادلة رقم (١٨):

$$Alpha(\alpha) = \frac{2}{2-1}\left(1-\frac{2}{2+(1.48)^2-1.14}\right) = 0.68$$

ولحساب ثبات المفهوم (CR) لبعد المعرفة الوقائعية نحتاج فقط إلى معرفة قيم التشبع(تشبعان: ٠.٥٧١ ؛ ٠.٩٠٤)؛ وبالتعويض في حدود المعادلة رقم (١٦)، فإن ثبات المفهوم هو:

$$CR = \frac{(1.48)^2}{(1.48)^2 + 0.86} = 0.72$$

ولحساب معامل أوميجا الموزونة (Ω_w)، نحتاج فقط إلى معرفة قيم التشبع (يوجد تشبعان: ٠.٥٧١ ؛ ٠.٩٠٤)؛ وبالتعويض في حدود المعادلة رقم (١٧)، فإن معامل ثبات أوميجا الموزونة يساوي:



الشكل رقم (١) تدل الأشكال البيضوية التي يرمز لها بزيتا ξ على المفهوم أو المتغير الكامن الذي يراد قياسه. ويفترض النموذج وجود أربع أبعاد أو عوامل (مفاهيم) كامنة : المعرفة الواقعية، المعرفة التصورية، المعرفة الإجرائية، وأخيرا ما وراء المعرفة. كل مفهوم أو بعد كامن من الأبعاد الأربعة قيس بعدد من الفقرات أو المؤشرات التي عبر عنها في النموذج بالمستطيلات. وتدل الأسهم التي تنطلق من المتغيرات الكامنة (الأشكال البيضوية) إلى المتغيرات المقاسة أو الملاحظة (الفقرات أو المؤشرات) على القاسم المشترك، أو العلاقة المشتركة بين المفهوم الذي يراد قياسه وبين الفقرة. وتدل رموز لمبدا λ على التشبعات. فمثلا يدل $\lambda_{12,4}$ (٠,٦٣) على تشبع الفقرة رقم (١٢) على البعد أو العامل رقم (٤) أو معامل الارتباط بين البعد والفقرة الذي يساوي ٠,٦٣ ؛ ويفضل تربيع التشبع لكي يدل على العلاقة المشتركة (التباين المشترك) بين البعد والمؤشر أو الفقرة. فتربيع ٠,٦٣ الذي يساوي ٠,٤ تقريبا يدل على أن البعد أو المتغير الكامن أو مفهوم ما وراء المعرفة يفسر ٤٠٪ من تباين الفقرة، أو يحدد الفقرة (يؤثر فيها) بنسبة أربعين في المئة. أما باقي التباين الذي تحتوي عليه الفقرة (٦٠٪) فيدل على التباين الخاص بالفقرة الذي لا تشترك به مع البعد، ويدل أيضا على تباين الخطأ العشوائي. والبواقي هذه تشير إليها الأسهم المتجهة نحو المستطيلات أي الفقرات، ويرمز لها غالبا بدلتا δ وتدعى عادة بالخطأ في التحليل العاملي أي التباين المتبقي سواء أكان تباينا خاصا بالفقرة، أم تباينا راجعا للخطأ العشوائي.

الجدول رقم (٢) يظهر الجدول تشبعات الفقرات على العوامل باستعمال التحليل العاملي التوكيدي، كما يظهر معاملات ألفـــا، ومعاملات ثبات المفهوم، ومعاملات أوميجا الموزونة.

المؤشرات المقاسة أو الفقرات	العوامل أو الأبعاد الكامنة: المعرفة الوقائعية(م و)	المعرفة التصورية(م ت)	المعرفة الإجرائية(م ج)	ما وراء المعرفة(م م)
فقرة (م و)١	٠.٥٧١			
فقرة (م و)٢	٠.٩٠٤			
فقرة (م ت)٣		٠.٩٠٦		
فقرة (م ت)٤		٠.٥٨٤		
فقرة(م ت) ٥		٠.٦٦٣		
فقرة (م ج)٦			٠.٨٢٤	
فقرة (م ج)٧			٠.٧٧٤	
فقرة(م ج) ٨			٠.٨٢٠	
فقرة (م م)٩				٠.٦١٧
فقرة (م م)١٠				٠.٧٣٠
فقرة (م م)١١				٠.٨٢٢
فقرة(م م) ١٢				٠.٦٣٠
معامل ألفا α	٠.٦٨	٠.٧٥٤	٠.٨٤٨	٠.٧٩٠
معامل ثبات المفهوم CR	٠.٧١٧	٠.٧٩٦	٠.٨٤٨	٠.٧٩٦
معامل أوميجا الموزونة Ω_w	٠.٨٣٢	٠.٨٥٥	٠.٨٥	٠.٨١٩

$$\Omega_w = \left[\frac{0.571^2}{1-0.571^2}+\frac{0.904^2}{1-0.904^2}\right] \Big/ \left[1+\frac{0.571^2}{1-0.571^2}+\frac{0.904^2}{1-0.904^2}\right] = 0.83$$

وباتباع نفس الطريقة، يمكن حساب معاملات الثبات لبقية العوامل أو المقاييس الفرعية. وعند مقارنة قيم المعاملات الثلاث، نجد أن معامل أوميجا الموزونة تتميز بارتفاع قيمها مقارنة بمعاملات ألفا التي يسفر استعمالها عن الحد الأدنى لمعامل الثبات الحقيقي. أما معاملات ثبات المفهوم فتتأرجح بينهما مدا وجزرا، إذ تحتل أحيانا موقعا وسطا بينهما، وقد تقترب من معامل ألفا أحيانا، أو من معامل أوميجا أحيانا أخرى.

متضمنات ختامية وتوصيات

إن المعالجة التحليلية السابقة لافتراضات معامل ألفا، ولبنيتها المنطقية والرياضية، ولحدودها، وتذبذب دقتها في تقدير الثبات استهدفت تبصير القارئ بمواطن قوة وقصور معامل ألفا حتى يتسنى له توظيفه توظيفا سليما، وقراءته قراءة دقيقة. وفي هذا السياق، لقد تبين من التحليل السابق أن دقة معامل ألفا تتأثر كثيرا بمدى تحقق افتراضاته في بيانات القياس. وافتراضات معامل ألفا هي ذاتها افتراضات نموذج القياس الذي يدعى بنموذج "طاو" المترادف في الأساس. ومعنى ذلك، أن معامل ألفا يتطلب أن تكون الدرجات الحقيقية متساوية في تباينها، وإن اختلفت في متوسطاتها لكونها تتفاوت فيما بينها بقيمة ثابتة؛ وأن تقيس كل فقرة ذات البعد الذي تقيسه الفقرات الأخرى، أي أن يتسم الاختبار بالتجانس أو بأحادية البعد. فعند توفر هذه الخصائص في درجات القياس، فإن استعمال معامل ألفا يؤدي إلى تقدير دقيق لمعامل الثبات.

غير أن هذه الافتراضات، وإن كانت أقل صرامة من افتراضات النموذج المتوازي، وافتراضات نموذج "طاو" المترادف، إلا أنها تبقى مع ذلك صعبة المنال، إذ من النادر أن تتوفر هذه المواصفات في واقع بيانات القياس. وما دام الأمر كذلك، تتأثر دقة معامل ألفا في تقدير الثبات بحيث تسفر فقط عن تقدير الحد الأدنى لمعامل الثبات الحقيقي، الذي قد يبتعد أو يقترب من القيمة الحقيقية لمعامل الثبات.

أما في حالة ارتباط أخطاء القياس، وهو أمر وارد جدا، فإن استعمال معامل ألفا قد لا يؤدي فحسب إلى تقدير تقليصي للثبات الحقيقي، بل يؤدي – في الغالب – إلى تقدير تضخيمي للثبات.

من متضمنات ما تقدم، أنه يمكن أن نتوصل إلى تقدير أدق للثبات من تقدير معامل ألفا، إن توفرت لنا أداة تقدير قائمة على افتراضات، أكثر واقعية، وأقل تقييدا. ولعل الأدوات التي تطرقنا إليها، ومنها معامل ثيتا، ومعامل أوميجا، ومعامل ثبات المفهوم، ومعامل أوميجا الموزونة، التي تحقق تقديرا أدق للثبات، وتنسجم أكثر من معامل ألفا مع نموذج القياس التقاربي الذي يعد أكثر تحررا وواقعية في افتراضاته، تمثل نماذج من الأدوات التي تنسجم مع هذا التوجه.

ومن استنتاجات عملية تحليل البنية المنطقية والإحصائية لمعامل ألفا، أنه لما كانت نتائج المعامل تتأثر بمقامه (تباين الدرجة الكلية للمقياس)، أكثر مما

تتأثر ببسطه (مجموع تباين فقرات المقياس)، فإن ارتفاع التجانس الداخلي للمقياس، وبالتالي ارتفاع الثبات، رهن بارتفاع تباين الدرجة الكلية للمقياس. وارتفاع الدرجات الكلية للمقياس رهن بمدى تباين العينة بحيث يزداد الثبات بازدياد تباين العينة، أو تناقص تجانسها. ومعنى ذلك، أن قيم الثبات لذات المقياس قد تختلف باختلاف العينات، إذ يستتبع اختلاف العينة اختلاف في مدى تباين درجات المقياس ككل. ينبثق من هذا الاستنتاج أمر هام وهو أن الثبات ليس صفة أو خاصية في المقياس تلازم المقياس وجودا وعدما، بغض النظر عن اختلاف العينات، أو تباين مواقف القياس. بل هو إمكان لا يتحقق نهائيا، أو إجراءات أو عمليات تتعلق بتقدير اتساق الإجابات على مثيرات تتخذ في الغالب شكل فقرات لمقياس معين. وبالتالي، ينبغي التحفظ على تعابير مثل "ثبات المقياس" أو "قياس الثبات"، واستبدالها بطريقة تعبير أكثر دقة وانسجاما مع منطق الثبات، كأن نقول "ثبات درجات المقياس" بدلا من "ثبات المقياس"، و"تقدير الثبات" بدلا من "قياس الثبات" لأن الثبات ليس صفة مستقرة موجودة داخل المقياس تقاس على مستوى العينة، وانما يستدل عليه لا على مستوى مجموعة الفقرات التي تشكل مقياسا معينا، وإنما على مستوى مجال افتراضي أوسع من الفقرات (تمثل مجتمع الفقرات)، بحيث أنه يفترض في فقرات المقياس أن تشكل عينة ممثلة من مجتمع الفقرات الافتراضي.

ومن متضمنات تحليل بنية معامل ألفا، أنه لا ينبغي أن يتسرع الباحث في الحكم على قيمة معامل ألفا المرتفعة بأنها تدل على ارتفاع الاتساق الداخلي للاختبار (ارتفاع الثبات). لا بد أن نتعامل مع نتائج معامل ألفا التي تزودنا بها حزمة SPSS بروية، وألا يكتفي الباحث بتوظيف القيمة المطلقة لمعامل ألفا سواء أكانت مرتفعة أومعتدلة. فمثلا، عندما يكون معامل ألفا مرتفعا، لا بد أن نتدبر هذه النتيجة، أهي بالفعل ناجمة عن ارتفاع الاتساق الداخلي للمقياس، أي عن مستوى كاف للارتباطات بين الفقرات . وبالتالي لا بد أن يعاين الباحث مصفوفة الارتباطات بين الفقرات، أو ارتباط كل فقرة بالدرجة الكلية لباقي الفقرات، أو مربع معامل الارتباط بين كل فقرة وباقي الفقرات الأخرى والذي يدل على التباين المشترك (المساحة المشتركة) بين الفقرة وباقي الفقرات. فإذا كشفت هذه المعاينة عن وجود مستويات كافية ومقبولة من الارتباطات، دل ذلك على أن معامل ألفا المرتفع يدل بالفعل على تجانس داخلي مرتفع لفقرات المقياس. أما إذا أظهرت المعاينة أن الارتباطات بين الفقرات، أو بين كل فقرة وباقي الفقرات، ضعيفة رغم ارتفاع معامل ألفا، دل ذلك على التأثير القوي لطول الاختبار على ارتفاع معامل ألفا. ومغزى هذا المفارقة الأخيرة بين ارتفاع معامل ألفا وانخفاض الارتباطات البينية للفقرات، أن ارتفاع معامل ألفا لا يدل بالضرورة على ارتفاع التجانس الداخلي لفقرات المقياس.

وفي ضوء ما تقدم، يمكن تقديم التوصيات التالية:

١ـ أن تراعى الافتراضات التي يتطلبها معامل ألفا في بيانات القياس، ذلك أن دقة معامل ألفا تعتمد إلى حد كبير على توفر هذه الافتراضات في بيانات القياس.

٢ـ أحيانا، قد يلجأ الباحث إلى استعمال طريقة سبيرمن - براون للتجزئة النصفية إلى جانب استعمال معامل ألفا لكرونباخ لتقدير ثبات المقياس ذاته، من باب التنويع في أدلة الثبات، وليسر الحصول عليها عند استخدام الحزم الإحصائية. وإذا كان ثمة ما يبرر تقدير الثبات عن طريق الاتساق الداخلي للمقياس، فيفضل استعمال معامل ألفا لكرونباخ دون استعمال طريقة التجزئة النصفية لسبيرمن - براون لسببين رئيسيين: **أولهما** لأن قيمة معامل الثبات قد تختلف باختلاف طريقة التجزئة النصفية في حين أن معامل ألفا يمثل متوسط قيم طرق التجزئة الممكنة للمقياس عند تقدير الثبات، ولهذا فهو أكثر دقة واستقرارا، وأقل تذبذبا من طريقة التجزئة النصفية. **وثانيهما**، لأن طريقة التجزئة النصفية لسبيرمن - براون تقوم على مسلمات أو افتراضات نموذج التوازي التام الذي يقتضي أن تكون الدرجات الحقيقية لأجزاء الاختبار متساوية، وأن يكون تباين خطأ هذه الأجزاء متساويا أيضا؛ أما معامل ألفا فيقوم على افتراضات نموذج قياس آخر أقل تشددا وأكثر واقعية في افتراضاته وهو نموذج "طاو" المترادف في الأساس الذي يسمح بتباين الخطأ، وباختلاف مستويات الدرجات الحقيقية أو متوسطاتها.

٣ـ ينبغي العزوف عن استعمال الصيغة المبسطة لمعامل كيودر- رتشاردسن المعروفة برقم (٢١) KR-21 لأنها تقوم على افتراض تساوي مستوى الصعوبة لكافة فقرات المقياس. ولما كان هذا لا يتحقق في واقع القياس، فإن معامل الاتساق أو الثبات الناجم عن استعمال هذه المعادلة يتسم بعدم الدقة مقارنة بمعامل الاتساق الناتج عن استعمال المعادلة العشرين لكيودر - رتشاردسن (KR-20) أو استعمال معامل ألفا.

٤ـ عند تقدير الثبات عن طريق الاتساق، يفضل الاكتفاء باستعمال معامل ألفا لكرونباخ ، دون اللجوء إلى معادلة كيودر- رتشاردسن (KR-20)، سواء أكانت فقرات المقياس ثنائية الدرجة Dichotomously scored items ، أو متصلة تنطوي على نطاق أوسع من الدرجات؛ وذلك لسببين: **أولا** لأن معادلة كيودر- رتشاردسن رقم ٢٠ (KR-20) تعتبر حالة خاصة من المعادلة الأعم التي تتمثل في معامل ألفا. **وثانيا** لأن معادلة كيودر- رتشاردسن قامت على افتراضات نموذج التوازي التام الذي يتعذر تحققه في الواقع: لا بد أن يتوفر شرط تساوي الدرجات الحقيقية وأيضا تساوي تباين الخطأ العشوائي بين فقرات المقياس؛ في حين أن معامل ألفا فيكتفي بتوفر افتراضات نموذج "طاو" المترادف في الأساس، ذلك لأن افتراضات هذا النموذج أكثر واقعية من افتراضات

نموذج التوازي التام، إذ لا يشترط وجوب تساوي تباين الخطأ، ويسمح بإمكان اختلاف متوسطات الدرجات الحقيقية للفقرات.

٥ـ لما كانت قيمة معامل ألفا أكثر تأثرا بتباين الدرجة الكلية للاختبار بمقام المعادلة، وأقل تأثرا بمجموع تباين الفقرات بالبسط ؛ ونظرا لكون تباين الدرجة الكلية للاختبار يرتفع عند ازدياد تباين العينة في الصفة المقاسة، وينخفض عند ازدياد تجانسها، فينبغي أن يستخدم الباحث عينة متباينة من الأفراد في الظاهرة المقاسة، بدلا من استعمال العينة المتجانسة.

٦ــ إن تصور الثبات باعتباره صفة جوهرية للاختبار، أو خاصية لصيقة بالمقياس، تلازمه وجودا وعدما بغض النظر عن مدى تجانس العينة أو تباينها، ينبغي أن يتغير لصالح اعتبار الثبات عملية تتعلق بدرجات المقياس، لا المقياس ذاته. فالمقياس الذي أظهر مستوى ثبات مرتفع عند تطبيقه على عينة متباينة، قد يظهر مستوى ثبات منخفض عند تطبيقه على عينة أخرى متجانسة. ولذلك، ينبغي على الباحث ألا يكتفي بالأدلة الموجودة في الدراسات السابقة عن ثبات درجات الأدوات التي وظفها في بحثه، بل يسعى بنفسه إلى تقويم ثبات درجات أدواته باستخدام عينة بحثه.

٧ـ إن معامل ألفا المرتفع لا يدل بالضرورة على ارتفاع اتساق فقرات المقياس، وبالتالي على ارتفاع معامل الثبات. إن ارتفاع معامل ألفا قد ينتج أساسا عن طول المقياس وليس بسبب اتساق فقراته. ولكي يطمئن الباحث على أن فقرات مقياسه تتوفر على مستوى كاف من الاتساق ، فينبغي أن يلقي نظرة فاحصة على معاملات الارتباط بين الفقرات، بحيث ينبغي ألا يقل متوسط معاملات الارتباط عن (٠.٣).

٨ـ ينبغي على الباحث أن يتدبر أيضا نتيجة معامل ألفا المرتفعة، وألا يتسرع في تأويلها على أنها تدل على اتساق داخلي مرتفع لارتفاع معامل ألفا، ولارتفاع العلاقات الارتباطية بين الفقرات، إلا بعد التحقق من أن ارتفاع الارتباطات بين الفقرات ، وبالتالي ارتفاع معامل ألفا، ليس وليد تشابه دلالات فقرات المقياس، وإن اختلفت شكلا وصياغة. إن الفقرات التي تكرر بعضها بعضا. والتشابه الصارخ بين كثير من الفقرات في المعنى رغم الاختلاف الظاهري بينها، يؤدي إلى تضخيم مصطنع ومزيف لمعامل ألفا، بحيث أننا لا نعدم أن نجد معامل ألفا يصل إلى ٠.٩٧ أو ٠.٩٨ بل ٠.٩٩ في بعض البحوث. ينبغي أن لا يعزب عن ذهن الباحث، أن الارتفاع الشديد في معامل الثبات بحيث يكاد يقارب الواحد الصحيح، يخل بفلسفة التعدد ذاته لفقرات المقياس. فالحكمة من تعدد الفقرات في قياس المفهوم، أنه لا يمكن الامساك بحقيقة هذا المفهوم باستعمال فقرة واحدة، ولكن يمكن الإحاطة به باستعمال عدد من الفقرات. غير أنه يشترط في كل فقرة من هذه الفقرات أن تشترك مع الفقرات الأخرى في تحديد المفهوم، لكن إذا اكتفت الفقرة بالمشاركة بنفس المساهمة أو الإثراء الذي قدمته

الفقرات الأخرى، ولم تضف شيئا خاصا في إثراء المفهوم أو تحديده تنفرد به الفقرة، فتعد هذه الفقرة تكرارا لما قدمته الفقرات الأخرى، واستنساخا لها. وبالتالي نحصل على مقياس قد ازداد طوله أو فقراته بدون أي إثراء جديد، أو تحديد جوانب جديدة للمفهوم، وأيضا نحصل على فقرات شديدة الترابط بسبب تشابهها التام في الدلالة والمعنى رغم اختلافها الظاهري. ولذلك يتضخم معامل ألفا نتيجة هذه الزيادة في عدد الفقرات، ونتيجة ارتفاع ارتباطها لتشابهها.

٩- من الضروري القيام بدراسات تركز أساسا على معاملات الاتساق الأخرى سواء أكانت تلك التي تطرقنا إليها باختصار، أو التي لم نتطرق إليها، ومقارنتها بمعامل ألفا من زاوية الافتراضات التي تقوم عليها، ونموذج القياس الذي يناسبها، والعوامل التي تحدد دقتها، والعوامل التي تؤدي إلى تحيزها.

المراجع

أولاً: المراجع العربية

جابر، جابر عبد الحميد و كاظم، أحمد خيري. *مناهج البحث في التربية وعلم النفس*. القاهرة: دار النهضة العربية، (١٩٩٠).

سيد، علي أحمد و سالم، أحمد محمد. *التقويم في المنظومة التربوية*، الرياض: مكتبة الرشد، (٢٠٠٤).

صبري، ماهر إسماعيل والرفاعي، محب محمود كامل. *التقويم التربوي: أسسه وإجراءاته*. الرياض: مكتبة الرشد، (٢٠٠١).

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

عبد الخالق، أحمد محمد. *قياس الشخصية*. الكويت: مطبوعات جامعة الكويت، (١٩٩٦).

ثانياً: المراجع الأجنبية

Allen, M. J. & Yen, W. M. (1979). Introduction to Measurement Theory. Monterey, CA: Brooks/Cole.

Anastasi, A, & Urbina, S. (1997). Psychological Testing. New Jersey: Prentice Hall.

Anderson, L., Krathwohl, D., Airasian, P., Cruikshank, K., Pintrich, P., and Raths, J. (2001). A taxonomy of learning , teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives. N. Y.: Longman.

Armor, D (1974) Theta reliability and factor scaling. In Costner, H. (Ed.) Sociological Methodology, pp. 17-50. San Francisco: Jossey-Bass.

Armor, D (1974) Theta reliability and factor scaling. In Costner, H. (Ed.) Sociological Methodology, pp. 17-50. San Francisco: Jossey-Bass.

Armor, D (1974) Theta reliability and factor scaling. In Costner, H. (Ed.) Sociological Methodology, pp. 17-50. San Francisco: Jossey-Bass.

Bacon, D.R., Sauer, P. L. & Young, M. (1995) Composite reliability in structural equation modeling. Educational and Psychological Measurement. Vol. 55, pp. 394-406.

Bacon, D.R., Sauer, P. L. & Young, M. (1995) Composite reliability in structural equation modeling. Educational and Psychological Measurement. Vol. 55, pp. 394-406

Brown, W. (1910). Some experimental results in the correlation of mental abilities. British Journal Of Psychology, Vol. 3, pp. 269-322.

Brunner, M., & Heinz-Martin Sub (2005) Analysing the reliability of multidimensional measures: An example from intelligence research. Educational and Psychological

Carmines, E. & Zeller, R. (1979).Reliability and validity assessment. Baverly Hills: Sage.

Carmines, E. & Zeller, R. (1979).Reliability and validity assessment. Baverly Hills: Sage

Carmines, E. **& Zeller, R.** (1979).Reliability and validity assessment. Baverly Hills: Sage.

Caruso, J. C. (2000). Reliability generalization of the NEO personality scales. Educational and Psychological Measurement. Vol. 60; pp, 236-254.

Clark, L.A. & Watson, D. (1995) Constructing validity: basic issues in scale development. Psychological Assessment, Vol. 7, pp. 309-319.

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

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

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

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

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

Cronbach L. J., (1947). Test "reliability": Its meaning and determination. Psychometrica, Vol. 16, pp 297-334.

Cronbach L. J.; Schonemann, P. & McKie, D. (1965). Alpha coefficient for stratified-parallel tests. Educational and Psychological Measurement. Vol. 25; pp. 291-312.

Cronbach, L. J. (2004) My current thoughts on coefficient alpha and successor procedures. Cse report 642. National Center for research on evaluation. University of California, Los Angeles.

Cronbach, L. J. (1951). Coefficient Alpha and the internal structure of tests. Psychometrica, Vol. 16, pp. 297-334.

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

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

Feldt, L. S. & Qualls, A. L. (1996). Bias in coefficient alpha arising from heterogeneity of test content. Applied Measurement in Education, Vol. 9; pp. 277-286.

Feldt, **L. S. & Qualls, A. L**. (1996). Bias in coefficient alpha arising from heterogeneity of test content. Applied Measurement in Education, Vol. 9; pp. 277-286.

Fornell, C. & Larker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research, Vol. 18, pp. 39-50.

Ghiselli, E. E,; Campbell, J. P. & Zedeck, S. (1981) Measurement Theory for the behavioural sciences. San Francisco: W. H. Freeman.

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. Vol. 66, No 6.

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. Vol. 66, No 6.

Graham, J.M. (2006) Congeneric and (essentially) Tau-Equivalent Estimates of Score Reliability: What they are and how to use them.

Lord, F. M. & Novick, M. R. (1967). Statistical theories of mental test scores. Reading, MA: Addison-Wesley

Lord, F. M. & Novick, M. R. (1967). Statistical theories of mental test scores. Reading, MA: Addison-Wesley

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

Measurement. Vol. 65, pp. 227-240. Brunner, M., & Heinz-Martin Sub (2005) Analysing the reliability of multidimensional measures: An example from intelligence research. Educational and Psychological Measurement. Vol. 65, pp. 227-240.

Miller, M. B. (1995). Coefficient alpha: A basic introduction from the perspective of classical test theory and structural equation modelling. Structural Equation Modeling. Vol. 2. pp 255-273.

N.C.M.E & A.P.A & A.E.R.A., (1985). Standards for educational and psychological testing. Washington, DC: American Psychological Association, Inc.

Netemeyer, R. (2001) Commentary: Can a reliability coefficient be too high?. Journal of Consumer Psychology, Vol. 10, pp. 55-69.

Netemeyer, R. (2001) Commentary: Can a reliability coefficient be too high?. Journal of Consumer Psychology, Vol. 10, pp. 55-69.

Novick, M. R. & Lewis, C. (1967). Coefficient alpha and reliability of composite measurements. Psychometrica, Vol. 32, pp 1-13.

Novick, M. R. & Lewis, C. (1967). Coefficient alpha and reliability of composite measurements. Psychometrica, Vol. 32. pp 1-13.

Novick, M. R. & Lewis, C. (1967). Coefficient alpha and reliability of composite measurements. Psychometrica, Vol. 32, pp 1-13.

Nunnally, J. C. & Bernstein, I. H. (1994). Psychometric Theory (3rd ed.). New York: McGraw-Hill.

Educational and Psychological Measurement. Vol. 66, pp. 930-944.

Green, B. G. & Hershberger, S. L. (2000) Correlated errors in true score models and their effect on coefficient alpha. Structural Equation Modeling, Vol. 7, N0 2, pp 251-270

Green, S. B., & Hershberger, S. L. (2000). Correlated errors in true score models and their effect on coefficient alpha. Structural Equation Modeling, Vol. 7, pp 251-270.

Green, S. B., & Hershberger, S. L. (2000). Correlated errors in true score models and their effect on coefficient alpha. Structural Equation Modeling, Vol. 7, pp 251-270.

Green, V. & Carmines, E. G. (1980) Assessing the reliability of linear composites. In Schuessler. K. F. (Eds.) Sociological Methodology . San Francisco: Jossey-Bass.

Green, V. & Carmines, E. G. (1980) Assessing the reliability of linear composites. In Schuessler, K. F. (Eds.) Sociological Methodology . San Francisco: Jossey-Bass.

Green,S. B., Lissitz, R. W., & Mulaik, S. A. (1977). Limitation of coefficient alpha as an index of test unidimensionality. Educational and Psychological Measurement. Vol. 37; pp, 827-838.

Hattie, J., (1985). Methodology review: Assessing unidimensionality of tests and items. Applied Psychological Measurement, Vol. 9; pp. 139-164.

Heise, D. R. & Bohrnstedt, G. W. (1970). Validity, invalidity, and reliability. . In Borgatta, E. F. and Bohrnstedt, G. W. (Eds.) Sociological Methodology pp. 104-129. San Francisco: Jossey-Bass.

Kamaroff, E. (1997). Effect of simultaneous violations of essential tau-equivalence and correlated errors on coefficient alpha. Applied Psychological Measurement, Vol. 21, pp337-348.

Kuder, G. F. & Richardson, M. W, (1937) The theory of the estimation of test reliability. Psychometrica, Vol. 2, pp. 151-160.

Multivariate Behavioral Research, Vol. 32, pp. 329-353.

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

Raykov, T. (1998).Coefficient alpha and composite reliability with interrelated nonhomogeneous items. Applied Psychological Measurement, Vol. 22, pp 375-385.

Raykov, T. (2001a).Bias of Cronbach's coefficient alpha for fixed congeneric measures with correlated errors, Applied Psychological Measurement, 25, pp 69-76.

Raykov, T. (2001b). Examination of congeneric scale reliability using covariance structure models with nonlinear constraints. British Journal of Mathematical and Statistical Psychology, Vol. 32, pp. 213-221.

Raykov, T. (2004a). Estimation of maximal reliability: A covariance structure modeling approach. British Journal of Mathematical and Statistical Psychology, Vol. 57, pp. 21-27.

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

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

Reinhardt, B. (1996) Factors affecting coefficient alpha: A mini Monte Carlo study. In Thompson, B. (Ed.) Advances in social science methodology (Vol. 4; pp 3-20). Greenwich, CT: JAI Press.

Robinson, J. P. ; Shaver, P. R. & Wrightsman, L. S. (1991) Criteria for scale selection and evaluation. In Robinson, J. P. ;Shaver, P. R. & Wrightsman, L. S. Measures of personality and social psychological attitudes (pp. 1-15). San Diego, CA: Academic.

Nunnally, J. C. (1967). Psychometric Theory . New York: McGraw-Hill.

Nunnally, J. C. (1978). Psychometric Theory (2nd ed.). New York: McGraw-Hill.

Onwuegbuzie, A. J., & Daniel, L. G. (2002). A fram-ework for reporting and interpreting internal consistency reliability estimates. Measurement and evaluation in counselling and development. Vol, 35: pp. 89-103.

Pedhazur, E. J. & Schmelkin, L. P. (1991) Measur-ement, Design, and Analysis: An integrated approach. New Jersey: Lawrence Erlbaum Associates.

Pedhazur, E. J. & Schmelkin, L. P. (1991). Measure-ment, Design, and analysis: an integrated approach. Hillsdale, NJ: Lawrence Erlbaum.

Pedhazur, E. J. & Schmelkin, L. P. (1991). Measure-ment, Design, and analysis: an integrated approach. Hillsdale, NJ: Lawrence Erlbaum.

Raju,N. S. (1982). On test homogeneity and maximum KR-20. Educational and Psycholo-gical Measurement. Vol. 42: pp. 145-152.

Raykov, T. (1997a). Estimation of composite relia-bility for congeneric measures. Applied Psychological Measurement, Vol. 21, pp 173-184.

Raykov, T. (1997a). Estimation of composite reliability for congeneric measures. Applied Psychological Measurement, 21, pp 173-184.

Raykov, T. (1997a). Estimation of composite reliability for congeneric measures. Applied Psychological Measurement, Vol. 21, pp 173-184.

Raykov, T. (1997a). Estimation of composite reliability for congeneric measures. Applied Psychological Measurement, Vol. 21, pp 173-184.

Raykov, T. (1997a). Estimation of composite reliability for congeneric measures. Applied Psychological Measurement, 21, pp 173-184.

Raykov, T. (1997b). Scale reliability, Cronbach's coefficient alpha, and violation of essential tau-equivalence with fixed congeneric components.

Streiner, D. L. (2003) Starting at the beginning: An introduction to Coefficient Alpha and internal consistency. Journal of Personality Assessment, Vol. 80, pp. 99-103.

Terwilliger, J. S., & Lele, K. (1979). Some relationships among internal consistency, reproducibility, and homogeneity. Journal of Educational Measurement. Vol. 16; pp. 101-108.

Thompson, B. (1994). Guidelines for authors. Educational and Psychological Measurement. Vol. 54; pp, 837-847.

Thompson,B. & Vacha-Haase, T. (2000). Psychometrics is datametrics: the test is not reliable. Educational and Psychological Measurement. Vol. 60; pp, 174-195.

Zimmerman, D. W. (1972). Test reliability and the Kuder-Richardson formula: Derivation from probability theory, Educational and Psychological Measurement. Vol. 32. Vol. 939-954.

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. Vol. 53, Vol. 33-49.

Rozenboom, W. W. (1966). Foundations of the theory of prediction. Homewood, IL: Dorsey.

Sawilowsky, S. S. (2000). Reliability: Rejoinder to Thompson and Vacha-haase. Educational and Psychological Measurement. Vol. 60; pp, 196-200.

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

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

Streiner, D. L. (2003) Starting at the beginning: An introduction to Coefficient Alpha and internal consistency. Journal of Personality Assessment, Vol. 80, pp. 99-103.

Streiner, D. L. (2003) Starting at the beginning: An introduction to Coefficient Alpha and internal consistency. Journal of Personality Assessment, Vol. 80, pp. 99-103.

Streiner, D. L. (2003) Starting at the beginning: An introduction to Coefficient Alpha and internal consistency. Journal of Personality Assessment, Vol. 80, pp. 99-103.

The Logical Structure of Cronbach's Coefficient Alpha, and its Precision in Estimating Reliability under Measurement Models Assumptions.

M'hamed Tighezza

Professor, Dept. of Psychology, College of Education, King Saud University, Riyadh, Saudi Arabia

(Received 27/4/1428H; accepted for publication 29/10/1428H)

Abstract The study examined three main research questions: 1- What is the logical structure of Cronbach's Coefficient Alpha and its implications? 2- Under what measurement models and cicumstances the use of Coefficient Alpha produce either accurate or biased assessment of reliability? 3- What are the alternative Consistency Coefficients that provide the researcher with more accurate estimation of reliabilty in the absence of assumptions required by Coefficient Alpha?

With respect to the first question, it was demonstrated that the total variance (compared with the sum of item variances) and the test length are the most determinant components of Coefficient Alpha. These findings imply that the sample used shoud be heteregenous and that Coefficient Alpha is not a pure indicator of internal consistency.

To deal with the second question, four measurement models were examined: 1- The Parallel Model, 2- The Tau-equivalent Model, 3- The Essentially Tau-equivalent Model, and 4- The Congeneric Model.

Coefficient alpha is based on the Essentially Tau-equivalent Model assumptions concerning test item data. It provides precise assessment of reliability under the three first aforementioned measurement models. However, coefficient alpha tends to underestimates reliability under violation of the assumptions of essential tau-equivalence (being usually the case) of test item scores, and tends to overestimate reliability under violation of the assumption of uncorrelated error scores. When test data satisfy the least restrictive and more realistic model: the congeneric model, alpha provides lower bound estimate of reliability.

Regarding the third question, Some alternative assessment of reliability in the presence of the congeneric models, and when some restrictive assumptions of the essential tau-equivalence model are not present in the items data, are examined. Therefore, Theta coefficient, Omega coefficient, Construct Reliability coefficient, and Weighted Omega coefficient are succinctly described. For didactic purposes, three coefficients: Alpha, Construct Reliability, and weighted Omega, were computed using data emanating from a research example.

The study concluded with some research implications and recommendations