Revue de psychoéducation

Adieu coefficient alpha de Cronbach! J'ai trouvé plus fidèle que toi…

Goodbye Cronbach's alpha! I found more reliable than you…

Sébastien Béland and Denis Cousineau

Volume 47, Number 2, 2018

URI: https://id.erudit.org/iderudit/1054068ar
DOI: https://doi.org/10.7202/1054068ar

See table of contents

Publisher(s)

Revue de Psychoéducation

ISSN

1713-1782 (print)
2371-6053 (digital)

Cite this article

Béland, S. & Cousineau, D. (2018). Adieu coefficient alpha de Cronbach! J'ai trouvé plus fidèle que toi…. *Revue de psychoéducation*, *47*(2), 449–460. https://doi.org/10.7202/1054068ar

Article abstract

Many researchers in psychometrics argue that Cronbach's alpha (1951) is not an adequate measure of reliability. It is with the idea to improve data analysis practices that this short article will present the theoretical foundations of reliability, summarize the main limits of Cronbach's alpha, and concisely present two promising alternatives: great lower bound and omega. Open softwares are mentioned to show how to obtain an estimate of these coefficients.

Revue de psychoéducation
Volume 47, numéro 2, 2018, 449-460

Mesure et évaluation

Adieu coefficient alpha de Cronbach! J'ai trouvé plus fidèle que toi…

Goodbye Cronbach's alpha! I found more reliable than you...

S. Béland[1]
D. Cousineau[2]

[1] Université de Montréal, Canada
[2] Université d'Ottawa, Canada

Résumé

De nombreux chercheurs en psychométrie soutiennent que le coefficient alpha de Cronbach (1951) n'est pas une mesure adéquate de la fidélité (reliability) *d'un test. C'est dans l'optique d'améliorer les pratiques d'analyse de données que ce court article va présenter les fondements théoriques de la fidélité, synthétiser les principales limites de l'alpha en plus de présenter de façon concise deux alternatives très prometteuses: le* great lower bound *et l'omega. Des logiciels gratuits et en téléchargement libre sont mentionnés pour obtenir un estimé de ces coefficients.*

Mots-clés : **Fidélité; coefficient alpha; coefficient omega; coefficient glb; logiciel en téléchargement libre**

Abstract

Many researchers in psychometrics argue that Cronbach's alpha (1951) is not an adequate measure of reliability. It is with the idea to improve data analysis practices that this short article will present the theoretical foundations of reliability, summarize the main limits of Cronbach's alpha, and concisely present two promising alternatives: great lower bound and omega. Open softwares are mentioned to show how to obtain an estimate of these coefficients.

Keywords: **Reliability; alpha coefficient; omega coefficient; glb coefficient; open software**

Correspondance :

Sébastien Béland

Université de Montréal
Pavillon Marie-Victorin
90, avenue Vincent d'Indy
Montréal (Québec) H2V 2S9

sebastien.beland@umontreal.ca

Introduction

La fidélité (*reliability*) est un concept fondamental en mesure, en plus d'être une composante importante de la validité. Souvent associée à l'idée de consistance (Buckingham, McCall, Otis, Rugg, Trabue et Courtis, 1921; Cronbach, 1951), de stabilité (Friedenberg, 1995) et de précision (Lord et Novick, 1968), sa définition est restée relativement inchangée au cours des dernières décennies.

Il existe traditionnellement plusieurs façons d'exprimer le degré de fidélité d'un test à l'aide d'un seul coefficient. En voici trois exemples. D'abord, il est possible de faire la passation de versions équivalentes d'un même test (*parallel-form reliability*). Une corrélation élevée entre les deux versions du test sera considérée comme une évidence de fidélité. Une autre stratégie consiste à repasser le même test à deux moments différents (*test-retest reliability*). Encore une fois, une corrélation élevée entre les scores obtenus à ces passations permet d'être optimiste à l'égard de la fidélité d'un test. Enfin, il est possible d'estimer la fidélité en analysant les scores obtenus lors d'une seule administration (*internal consistency reliability*) comme le fait, par exemple, le coefficient alpha (noté α) de Cronbach.

Afin de mieux comprendre le concept de fidélité, voici un court exemple didactique supposant l'existence de tests équivalents A et B, qui présentent une bonne fidélité et qui ont été dispensés à un groupe de répondants. Ces derniers peuvent avoir obtenu un score qui peut différer à ces deux tests. Par contre, un individu qui obtient un score plus élevé au test A qu'au test B, ne doit pas voir sa *position* changer au sein du groupe évalué. Ainsi, en présence de tests fidèles, les individus qui ont des scores faibles au test A doivent être les mêmes qui ont obtenu des scores faibles au test B et *vice versa* (Friedenberg, 1995).

Le concept de fidélité bénéficie d'un regain d'intérêt depuis quelques années (Bentler, 2017; Laveault, 2012; McNeish, 2017; Revelle et Zinbarg, 2009; Sijtsma, 2009; Trizano-Hermosilla et Alvarado, 2016) grâce aux nombreuses critiques formulées à l'égard du coefficient α de Cronbach (1951). Nous allons, dans ce court article, tenter d'expliquer pourquoi ce populaire coefficient doit être abandonné au profit d'alternatives plus prometteuses. Ansi, nous allons centrer notre argumentaire sur deux coefficients alternatifs: le Great lower bound (*glb*) et les indices de type omega (ω). Nous espérons que les auteurs qui envisagent publier dans la *Revue de psychoéducation* adopteront rapidement ces coefficients.

Fondements théoriques de la fidélité[1]

Il faut retourner au modèle du score vrai (*true score model*) pour comprendre les fondements de la fidélité. Celui-ci prend la forme suivante :

$$Y = V + E \qquad (1)$$

1. La conception de la fidélité qui prévaut dans les modèles de la théorie de la réponse aux items ne sera pas traitée dans cet article.

où le score observé *Y* d'un individu à un test est composé de deux éléments théoriques et non observables (Bertrand et Blais, 2004) : un score vrai *V* et une erreur de mesure *E*. L'élément *E* est fondamental dans la présente réflexion puisqu'il représente l'imprécision dans la prise de mesure. Celui-ci peut être symbolisé par la différence entre le score observé et le score vrai, soit $E=Y-V$. De plus, mentionnons que *E* est de nature aléatoire, ce qui veut dire que l'erreur n'est pas systématique ou volontaire dans le modèle du score vrai. Ainsi, dans une situation hypothétique où nous faisons passer un grand nombre de fois le même test au même individu, l'élément *V* est une constante propre à un individu donné (on utilise parfois le terme *trait latent* dans les écrits en psychométrie). Les erreurs de mesure *E*, de leur côté, sont des « accidental deviations [that] are different in every individual case […] and occur quite impartially in every direction according to the known laws of probability » (Spearman, 1904; cité par Revelle (en préparation)).

Pour clarifier l'idée des multiples passations d'un test à un même individu, il est pertinent de s'inspirer de l'exemple classique de Lazarsfeld (1959). Imaginons que nous demandions à un illustre inconnu, appelons-le Monsieur Larivée, son degré d'accord à l'aide d'une échelle de mesure allant de « 0 » (pas du tout d'accord) à « 50 » (entièrement d'accord) sur la gratuité du café sur le campus de l'université. Ajoutons que ce dernier se dit ambivalent sur le sujet. Imaginons, de plus, que nous avons accès à une machine inédite qui a la capacité d'effacer instantanément la mémoire à court terme lorsque celle-ci est placée sur la tête d'un individu. Après avoir posé une première fois la question sur la gratuité du café à M. Larivée et où il nous a présenté son score observé *Y*, nous lui « lavons la mémoire » en utilisant la machine à effacer la mémoire avant de lui reposer la question. Cette procédure est répétée un grand nombre de fois. Nous observons que notre cobaye est parfois plutôt en accord avec la gratuité scolaire, alors qu'à d'autres moments, il est plutôt en désaccord. Puisque la question a été posée plusieurs fois et que les erreurs sont de nature aléatoire, la moyenne des scores représentant son degré d'accord sur la gratuité du café est son score vrai *V*.

La fidélité d'un test ne peut évidemment pas se calculer à partir du score observé *Y* d'un seul individu : il faut avoir en main les scores observés d'un ensemble d'individus. Cela permet d'envisager la variance de *Y*, qui est la somme de la variance vraie *V* et de la variance d'erreur *E* : $Var(Y)=Var(V)+Var(E)$. Cette formulation permet d'aboutir à la conception théorique de la fidélité, qui est le rapport de la variation dans le score vrai, notée $Var(V)$, relativement à la variation existant dans le score observé, notée $Var(Y)$. Cette idée prend mathématiquement ces différentes formes :

$$\rho^2 = \frac{Var(V)}{Var(Y)} = \frac{Var(V)}{Var(V)+Var(E)} = 1 - \frac{Var(E)}{Var(Y)} \quad (2)$$

où ρ^2 (rho carré) est nécessairement borné entre 0 et 1. Ainsi, plus la valeur de *Var*(*V*) se rapproche de la valeur de *Var*(*Y*), plus le coefficient ρ^2 prend une valeur proche de 1 et plus le test est considéré comme étant fidèle. Pour illustrer cette idée, nous reprenons l'exemple de Guilford et Fruchter (1973, p. 398). Les scores hypothétiques de dix individus sont présentés au tableau 1. Le lecteur doit garder en tête que cette illustration est uniquement didactique, car le modèle du score vrai suppose une situation où il y a un très grand nombre d'individus qui sont mesurés.

Tableau 1. Exemple de 10 scores observés visant à expliciter les éléments de l'équation 1 et de l'équation 2

Score vrai (*V*)	Erreur de mesure (*E*)	Score observé (*Y*=*V*+*E*)
45	0	45
25	+10	35
15	+2	17
5	-2	3
20	-4	16
35	-2	33
30	-4	26
25	-2	23
25	+2	27
25	0	25
Somme_V : 250	Somme_E : 0	Somme_Y : 250
Moyenne_V : 25	Moyenne_E : 0	Moyenne_Y : 25
Var(*V*) : 116,7	*Var*(*E*) : 16,9	*Var*(*Y*) : 133,6

Puisque les erreurs de mesure sont de nature aléatoire, nous observons que leur somme est égale à zéro. Et tel que stipulé un peu plus haut, la somme de *Var*(*V*) et de *Var*(*E*) est égale à *Var*(*Y*). De plus, la figure 1 permet d'exprimer de façon visuelle l'idée de la fidélité.

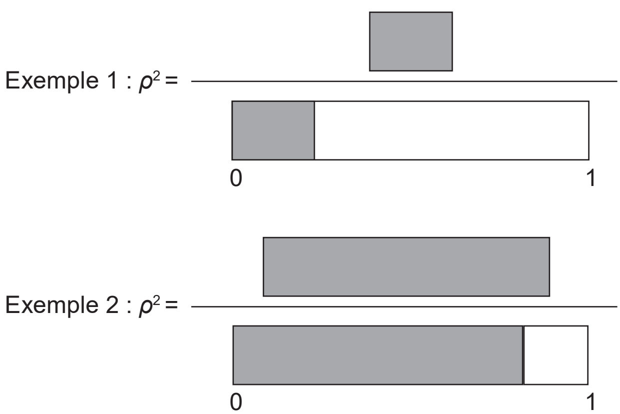

Figure 1. Représentation visuelle de la fidélité. *Var*(*V*) est en gris et *Var*(*E*) est en blanc. *Var*(*Y*) équivaut à 100 % de la variance observée. Autant dans l'exemple 1 que dans l'exemple 2, le dénominateur, qui prend une valeur entre 0 et 1, représente toute la variance des scores observés. Pour cette raison, ρ^2 ne peut pas être inférieur à zéro ou supérieur à 1. Nous observons que dans l'exemple 1, la fidélité est limitée, car la section en gris qui représente *Var*(*V*) est beaucoup plus petite que la section en blanc qui représente *Var*(*E*). Dans l'exemple 2, c'est plutôt l'inverse qui se produit : la section en gris est beaucoup plus imposante, ce qui témoigne d'une fidélité élevée.

Pour conclure, nous allons utiliser les informations présentées au tableau 1 et à l'équation 2, pour calculer ρ^2 :

$$\frac{116{,}7}{133{,}6} = 0{,}87.$$

ce qui témoigne d'une bonne fidélité.

Adieu coefficient α de Cronbach!

Bien que l'α ait été présenté en 1951 par Lee J. Cronbach, on retrouve un coefficient similaire dans les travaux de Guttman, en 1945. L'α s'inspire de la dernière forme de l'équation 2 pour estimer ρ^2, soit :

$$\rho^2 = 1 - \frac{Var(E)}{Var(Y)}. \quad (3)$$

Pour le α de Cronbach[2], $Var(E)$ est la somme de la variance à chacun des items et $Var(Y)$ est la somme de la matrice de variance-covariance pour tous les items du test.

Ce coefficient est probablement le plus connu et le plus utilisé de tous les coefficients psychométriques. Dunn, Baguley et Brunsden (2013) nous informent d'ailleurs que l'α a été repéré dans plus de 17 600 publications depuis sa création. Malheureusement, celui-ci présente de nombreuses limites d'application qui en font une mauvaise méthode pour estimer la fidélité. En voici quelques exemples. D'abord, l'α de Cronbach repose sur le postulat que la variance doit être identique pour tous les items du test. Or, il peut y avoir, par exemple, des effets plafond ou plancher qui réduisent les variances de certains items. Dans les études empiriques, nous observons plutôt que certains items peuvent être fortement liés à un construit, alors que d'autres le sont très peu (McNeish, 2017). Une violation de ce postulat engendre une sous-estimation de la fidélité réelle (Miller, 1995). Ensuite, les données analysées à l'aide de ce coefficient présentent parfois un résultat étonnant : α peut prendre une valeur négative (Martineau, 1982). Or, la fidélité doit théoriquement être bornée entre 0 et 1 selon la définition présentée à l'équation 2 et exprimée à la figure 1. Un autre problème inhérent au coefficient α est lié à l'existence d'un biais d'estimation lorsque les réponses aux items ne sont pas de nature continue et/ou ne suivent pas une distribution normale (Gadermann, Guhn et Zumbo, 2012; McNeish, 2017). Il faut aussi mentionner le fait que le coefficient α n'est pas bien équipé pour analyser des données multidimensionnelles. Or, un grand nombre de chercheurs accumulent des preuves de validité sur des tests comportant plusieurs dimensions. À ce titre, l'α n'est ne concorde pas avec les pratiques de nombreux chercheurs en psychoéducation.

2. Veuillez consulter l'annexe pour obtenir plus d'information.

Nous allons, pour conclure cette section, reprendre le constat glacial de Sijtsma (2009): « the only reason to report *α* is that top journals tend to accept articles that use statistical methods that have been around for a long time such as *α* » (p. 118).

À la rencontre de deux coefficients plus fidèles : *glb* et *ω*

Il existe plusieurs coefficients de fidélité qui sont de meilleures alternatives à l'*α* de Cronbach. Nous allons, dans ce qui suit, nous concentrer sur deux cas qui semblent particulièrement prometteurs.

Une première alternative à l'*α* est le coefficient Great lower bound (*glb*) de Jackson et Agunwamba (1977). Les éléments nécessaires au calcul de ce coefficient s'inspirent de l'équation 1, mais sous la forme suivante :

$$Cov(Y) = Cov(V) + Cov(E) \tag{4}$$

où *Cov*(*Y*) est la matrice de covariance inter-item des scores observés qui se décompose en une matrice de covariance inter-item des scores vrais Cov(V) et une matrice de covariance inter-item des erreurs *Cov*(*E*). À l'instar du coefficient *α* de Cronbach, *glb* s'inspire de la forme de la dernière équivalence de l'équation 2 afin d'estimer ρ^2.

Green et Yang (2009), McNeish (2017) et Sijtsma (2009) ont récemment montré que ce coefficient performait généralement mieux que *α*. Néanmoins, le lecteur doit savoir qu'il faut utiliser *glb* avec prudence lorsque l'ensemble de données est limité. Comme Shapiro et ten Berge (2000) l'écrivaient « it is well known that the *glb*, based on small samples (even a sample of one thousand subjects is not generally enough) may severely overestimate the population value, and statistical treatment of the bias has been badly missing » (p. 413).

Une deuxième alternative d'intérêt est la série de coefficients omega (noté *ω*) et popularisée par McDonald (1985; 1999). Cette approche de la fidélité est basée sur le modèle d'analyse factorielle à un facteur commun de Spearman, qui est une mise à l'échelle du modèle de score vrai présenté à l'équation 1.

Selon McDonald, le modèle d'analyse factorielle à un facteur commun peut être exprimé de la façon suivante :

$$Y = C + U \tag{5}$$

où *Y* est le score total au test, *C* est la part de *Y* qui est exprimée par le facteur commun de l'analyse factorielle, *U* est la part de *Y* qui est exprimée par un facteur unique à un item et ne se distingue pas de l'erreur de mesure. Postulant que *C* et *U* sont indépendants, nous obtenons :

$$Var(Y) = Var(C) + Var(U) \tag{6}$$

ce qui nous ramène à la deuxième équivalence de l'équation 2 :

$$\rho^2 = \frac{Var(V)}{Var(Y)} = \frac{Var(V)}{Var(V)+Var(E)} = \frac{Var(C)}{Var(C)+Var(U)}. \quad (7)$$

où *Var(V)*=*Var*(*C*) et *Var*(*E*)=*Var*(*U*).

Les coefficients *ω* sont particulièrement intéressants, car ils sont très flexibles tout en pouvant s'adapter à des tests multidimensionnels. Par contre, il faut s'assurer que le modèle d'analyse factorielle utilisé présente une bonne adéquation avec les données analysées, sans quoi *ω* ne peut pas être utilisé. Magré cela, Dunn, Baguley et Brunsden (2013) sont sans équivoque: « given the consensus in the psychometric literature that *α* is rarely appropriate and given the good performance of omega when the assumptions of *α* are not met, it is recommended that psychologists change to the routine reporting of *ω* in place of *α* » (p. 409).

Le tableau suivant synthétise les principales forces et faiblesses des coefficients ω et glb.

Tableau 2. Principales forces et faiblesses des coefficients *ω* et *glb*

	Force	Faiblesse
ω	L'hypothèse théorique (et irréaliste) d'équivalence de la variance pour tous les items du test n'est pas requise pour calculer *ω*. De plus, ce coefficient peut être utilisé avec différents modèles d'analyse factorielle à un facteur commun (exploratoire et/ou confirmatoire).	Le modèle d'analyse factorielle utilisé pour calculer *ω* doit présenter une bonne adéquation (*fit*) aux données.
glb	*glb* évalue l'erreur de mesure de façon beaucoup plus efficace que le coefficient *α*.	Ce coefficient surestime la fidélité lorsque l'échantillon analysé est de petite taille.

Enfin, nous allons laisser McNeish (2017), qui a comparé plusieurs coefficients de fidélité dont *α*, *glb* et *ω*, clore cette section: « any of the alternatives is preferable to continued use of Cronbach's *α* » (p. 12). Adieu coefficient *α* de Cronbach!

Où trouver *glb* et *ω*?

À notre connaissance, les coefficients *glb* et *ω* ne sont pas encore disponibles dans SPSS, mais ils peuvent être calculés gratuitement à l'aide des logiciels suivants. D'abord, une librairie R nommée psych (Revelle, 2017) permet de calculer les coefficients *glb* (voir les fonctions glb.algebraic et glb.fa) et *ω* (voir les fonctions omega et omegah). Le lecteur trouvera des informations supplémentaires dans le manuel d'utilisateur de cette librairie en suivant ce lien : https://cran.r-project.org/web/packages/psych/psych.pdf

Une autre librairie R, MBESS (Kelley, 2017), permet aussi de calculer *ω* en plus de fournir les intervalles de confiance de ce coefficient. Malheureusement, cette librairie n'inclut pas le coefficient *glb* au moment d'écrire ces lignes. Voici le lien vers le manuel d'utilisateur : https://cran.r-project.org/web/packages/MBESS/MBESS.pdf

glb et *ω* sont aussi intégrés dans le logiciel en téléchargement libre avec interface visuelle JASP (https://jasp-stats.org/). Après avoir chargé l'ensemble de données à analyser, le lecteur doit d'abord ouvrir l'onglet « Descriptives » et sélectionner « Reliability Analysis ». Ensuite, il doit cliquer sur les cases des coefficients *glb* et *ω*, qui se trouvent dans la section « Scale statistics ».

Quelques mots à l'attention des futurs auteurs de la Revue de psycho-éducation

glb et *ω* doivent être interprétés de la même façon : plus la valeur du coefficient approche de 1, plus le test est considéré comme fidèle (voir l'équation 2 et la figure 1). Nous suggérons de systématiquement rapporter (i) la valeur du coefficient et (ii) le nombre d'items analysés. En ce qui concerne *ω*, il est aussi pertinent d'intégrer des informations sur l'adéquation (*fit*) des données au modèle d'analyse factorielle utilisé (voir Kline (2011) pour plus de détails sur le sujet). Et lorsque cela est possible, nous encourageons fortement les chercheurs à rapporter les intervalles de confiance. À titre de rappel, la librairie MBESS permet de calculer de telles intervalles pour le coefficient *ω*.

Références

Bentler, P. M. (2017). Specificity-enhanced reliability coefficients. *Psychological Methods, 2*2, 527-540 doi: 10.1037/met0000092

Bertrand, R. et Blais, J.-G. (2004). *Modèle de mesure : l'apport de la théorie de la réponse aux items*. Québec, QC : Presses de l'Université du Québec.

Buckingham, B. R., McCall, W. A., Otis, A.S., Rugg, H. O., Trabue, M. R. et Courtis, S. A. (1921). Report of the standardization committee. *Journal of Educational Research, 4*, 78-80.

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

Dunn, T. J., Baguley, T., et Brunsden, V. (2013). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. British *Journal of Psychology, 105*, 399-412. doi: 10.1111/bjop.12046

Friedenberg, L. (1995). *Psychological testing: Design, analysis, and use*. Boston, MA: Allyn and Bacon.

Gadermann, A. M., Guhn, M. et Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation, 17*, 1–13.

Green, S.A. et Yang, Y. (2009a). Commentary on coefficient alpha: a cautionary tale. *Psychometrika.* doi: 10.1007/s11336-008-9098-4.

Guilford, J. P. et Fruchter, B. (1973). *Fundamental statistics in psychology and education* (5e édition). New York: McGraww-Hill.

Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika, 10*, 255– 282.

Jackson, P. H. et Agunwamba, C. C. (1977). Lower bounds for the reliability of the total score on a test composed of non-homogeneous items: I: Algebraic lower bounds. *Psychometrika, 42*, 567–578. doi: 10.1007/BF02295979

Kelley, K. (2017). MBESS: The MBESS R Package. R package version 4.4.2. https://CRAN.R-project.org/package=MBESS

Kelley, K. et Pornprasertmanit, S. (2016). Confidence intervals for population reliability coefficients: Evaluation of methods, recommendations, and software for composite measures. *Psychological methods, 21*, 69-92, doi: 0.1037/a0040086

Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling* (3e éd.). New York, NY: Guilford Press.

Laveault, D. (2012). Soixante ans de bons et mauvais usages du alpha de Cronbach. *Mesure et évaluation en éducation 35*, 1–7. doi: 10.7202/1024716ar

Lazarsfeld, P. F. (1959). Latent structure analysis. Dans S. Koch: *Psychology: a study on a science,* 3, New York: McGraw-Hill.

Lord, F. M., et Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley

Martineau, G. (1982). Exploration des valeurs possibles du coefficient α de Cronbach. *Revue des sciences de l'Éducation, 8*, 135-143.

McDonald, R. P. (1985). *Factor analysis and related methods*. Mahwah, NJ: Lawrence Erlbaum.

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

McNeish, D. (2017, May 29). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*. Advance online publication. doi: 10.1037/met0000144

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

Revelle, W. (2017) psych: Procedures for Personality and Psychological Research., Evanston, IL: Northwestern University Press. https://CRAN.R-project.org/package=psych Version = 1.7.8.

Revelle, W. (en préparation). *An introduction to psychometric theory with applications in R.* URL: https://www.personality-project.org/r/book/

Revelle, W., et Zinbarg, R. E. (2009). Coefficients Alpha, Beta, Omega, and the glb: Comments on Sijtsma. *Psychometrika, 74*, 145–154. doi: 10.1007/s11336-008-9102-z

Shapiro, A. et ten Berge, J. M. (2000). The asymptotic bias of minimum trace factor analysis, with applications to the greatest lower bound to reliability. *Psychometrika, 65*, 413–425.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of cronbach's alpha. Psychometrika, 74, 107–120. doi: 10.1007/s11336-008-9101-0

Tavakol, M. et Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of Medical Education, 2*, 53-55.

Trizano-Hermosilla, I. et Alvarado, J. M. (2016). Best alternatives to cronbach's alpha reliability in realistic conditions: Congeneric and Asymmetrical Measurements. Frontiers in psychology, 26, 769. doi: 10.3389/fpsyg.2016.00769

Appendice : comparer *α*, *ω* et *glb* en utilisant la simulation informatique

Trois coefficients de fidélité ont été discutés dans ce court article. D'abord, l'*α* de Cronbach prend la forme suivante:

$$\alpha = \frac{J}{J-1}\left(1 - \frac{\sum_{j=1}^{J} \mathrm{Var}(Item_j)}{Var(Total)}\right) \quad (8)$$

où J est le nombre d'items à un test, $Var(Item_j)$ est la variance de l'item j (j valant entre 1 et J) et $Var(Total)$ est la somme de toutes les variances et covariances au test. Ensuite, le coefficient *glb* s'écrit mathématiquement :

$$glb = 1 - \frac{tr(Cov(E))}{Var(Total)} \quad (9)$$

où $tr(Cov(E))$ est la trace[3] de la matrice de covariances inter-item (voir l'équation 4) et $Var(Total)$ est la somme de toutes les variances et covariances au test. Enfin, le coefficient ω est équivalent, dans sa forme la plus simple, à :

$$\omega = \frac{\left(\sum_{j=1}^{J} \lambda_j\right)^2}{\left(\sum_{j=1}^{J} \lambda_j\right)^2 + \sum_{j=1}^{J} \psi_j^2} \quad (10)$$

où λ_j est un coefficient de saturation (*factor loading*) de l'item j et ψ_j^2 est la variance unique de l'item j.

Pour comparer ces trois coefficients de fidélité, et voir leur capacité à chiffrer correctement la fidélité, une stratégie intéressante consiste à utiliser la simulation informatique. L'idée derrière cette approche est de sélectionner une population théorique avec certaines caractéristiques connues, générer des échantillons aléatoires à partir de cette population et utiliser des estimateurs sur ces échantillons. Ici, les trois estimateurs considérés sont l'*α*, le *ω* et le *glb*. Si un estimateur est efficace, il permettra de retrouver, en moyenne, la valeur vraie choisie pour la population. Sinon, on dira qu'il est biaisé, en générant une valeur soit trop petite, soit trop grande.

Dans la présente étude, nous avons simulé un seul scénario. Des études plus exhaustives se trouvent dans Revelle et Zinbarg (2009), Sijtsma (2009) et Trizano-Hermosilla et Alvarado (2016). Ce scénario est idéal en ce sens qu'il respecte tous les postulats sous-jacents aux coefficients de fidélités examinés dans cet article. Nous supposons que les participants sont testés 10 fois (10 items) et qu'en moyenne, ils ont une habilité égale à 5 sur une échelle de 0 à 10. L'écart type de leurs réponses est égal à 2. Par ailleurs, la covariance est constante et égale à l'unité. Avec ces informations, la fidélité vraie dans la population est égale à 0,925. Pour chaque simulation, nous avons généré 250 participants, soit un nombre semblable à ce que l'on trouve dans plusieurs études empiriques en

3. La trace est la somme des éléments de la diagonale d'une matrice.

psychoéducation. Finalement, ce scénario est répété 1000 fois pour étudier le comportement de *α*, *ω* et *glb*.

La figure 2 montre la distribution des trois coefficients de fidélité, alors que la ligne grise est la valeur vraie de 0,925 dans la population. Comme on peut le voir, l'α de Cronbach sous-estime presque toujours la vraie fidélité. Seulement un petit pourcentage des simulations de ce coefficient est surévalué. Ceci montre un biais important du α de Cronbach. Pour les deux autres coefficients de fidélité, on ne voit aucun biais important.

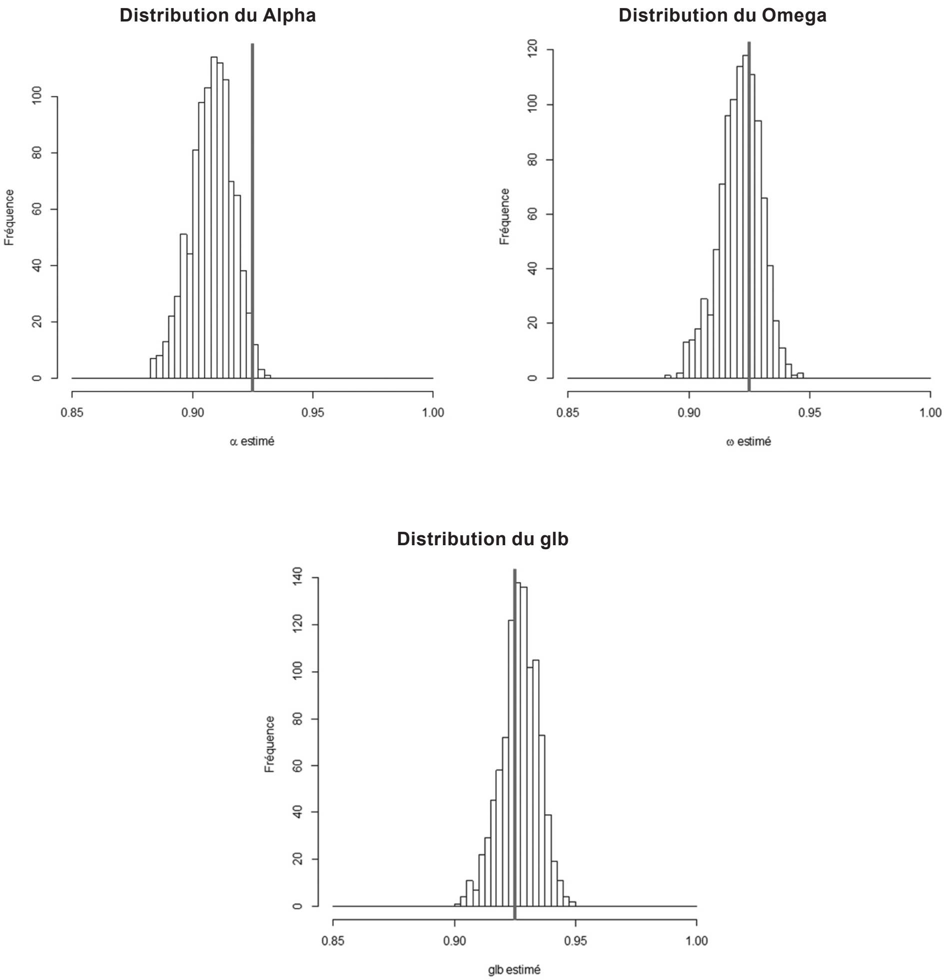

Figure 2. Distribution des coefficients *α* (panneau du haut, à gauche), *ω* (panneau du haut, à droite) et *glb* (panneau du bas). La ligne grise indique la fidélité vraie dans la population simulée (0,925).

En moyenne, l'*α* de Cronbach obtient une fidélité de 0,908, ce qui est une sous-estimation de la valeur pour la population. Cette sous-estimation est d'ailleurs notable étant donné que les échantillons simulés sont grands (250 participants). Elle est aussi significativement non nulle ($t(999) = 1.97$, $p < .05$). Pour de plus petits échantillons, le biais est encore plus prononcé.

En ce qui concerne les deux autres coefficients, ils sont en moyenne très proches de la fidélité vraie. Pour les distinguer, il faudrait simuler un nombre considérablement plus grand d'échantillons que les 1000 présentement générés.

Dans ce scénario très limité, les postulats sous-jacents au calcul de la fidélité sont respectés. Il n'y a donc aucune raison particulière pour que le coefficient *α* sous-estime systématiquement la fidélité de la population. Comme il existe au moins deux alternatives moins biaisées (ou sans aucun biais, notre simulation de taille limitée ne permet pas de trancher), nous avons ici un autre argument pour cesser d'utiliser l'*α* de Cronbach.