

## UTILISER LE COEFFICIENT OMEGA DE MCDONALD À LA PLACE DE L'ALPHA DE CRONBACH

SÉBASTIEN BÉLAND *Université de Montréal*

DENIS COUSINEAU *Université d'Ottawa*

NATHALIE LOYE *Université de Montréal*

**RÉSUMÉ.** Des chercheurs discutaient récemment de l'importance de rester à jour au sujet des plus récentes avancées en méthodes quantitatives. À ce titre, de nombreux auteurs ont exposé leur souhait de voir les chercheurs abandonner le populaire coefficient alpha de Cronbach. C'est dans une optique de diffusion et de vulgarisation que ce court article a comme objectif de présenter l'alternative qui semble la plus prometteuse pour mesurer la fidélité d'un test, le coefficient omega de McDonald, qui est basée sur l'analyse factorielle à un facteur commun.

### USING THE MCDONALD'S OMEGA COEFFICIENT INSTEAD OF CRONBACH'S ALPHA

**ABSTRACT.** Researchers have recently discussed the importance of being up to date about the latest advances in quantitative methods. Many authors have expressed their desire to see researchers abandon the popular Cronbach's alpha. It is with a perspective of diffusion and popularization that this short article aims to present the most promising alternative to measure the reliability of a test, the McDonald's omega coefficient, which is based on common factor analysis.

### INTRODUCTION

**A**lors que le concept de validité peine depuis près d'un siècle à être défini de manière consensuelle, celui de fidélité a une définition à la fois stable et relativement facile à comprendre. Apparentée à la consistance, à la stabilité ou à la reproductibilité, la fidélité se vérifie en s'assurant que le résultat que l'on obtient en collectant des données avec un instrument de mesure resterait le même si la mesure était reprise avec le même instrument ou avec un instrument alternatif comparable (Crocker et Algina, 1986).

Peut-être à cause de cette apparente simplicité, la fidélité finit souvent par se confondre avec le coefficient  $\alpha$  (alpha) de Cronbach (1951). Pourtant, *chiffrer* la fidélité n'est pas si simple, car le résultat obtenu doit fournir une indication fiable sur la précision de l'outil de mesure (celui-ci peut être, par exemple, une échelle pour mesurer la motivation scolaire ou un test pour mesurer la compréhension d'un dialogue en langue espagnole). Or de nombreux auteurs (Laveault, 2012 ; Revelle et Zinbarg, 2009 ; Sijtsma, 2009 ; Trizano-Hermosilla et Alvarado, 2016) ont fait valoir des critiques contre l'usage de ce coefficient pour estimer la fidélité et suggèrent de l'abandonner. Le texte qui suit présente quelques-uns de leurs arguments pour convaincre le lecteur qu'il est temps de tourner la page et de passer au coefficient omega. Il survole aussi la façon de le calculer.

### *La nature de la fidélité*

Pour comprendre l'approche classique de la fidélité, il faut retourner au modèle du *score vrai* (*true score model*). Celui-ci est basé sur l'idée que chaque individu est doté d'une capacité stable à réaliser une tâche, mais que cette capacité ne s'exprime pas toujours de manière parfaite dans une situation de mesure. Mathématiquement, ce modèle s'écrit :

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

où  $Y$  est le score observé d'un individu qui se décompose en un score vrai  $V$  et une erreur de mesure aléatoire  $E$ . Dans ce modèle,  $V$  est une valeur fixe pour un individu donné, alors que  $E$  symbolise l'imprécision de l'outil de mesure et fluctue à chaque prise de mesure. Conséquemment, le score observé  $Y$  varie lui aussi de façon aléatoire.

Reposant sur l'hypothèse théorique qu'un même test puisse être passé de multiples fois à un groupe d'étudiants, et postulant que l'erreur ne dépend pas du score vrai, l'équation 1 implique que la variance observée<sup>1</sup> est la somme de la variance vraie et de la variance d'erreur :  $Var(Y) = Var(V) + Var(E)$ . Cette formulation permet de définir la fidélité (notée  $\rho^2$  – rho carré – par convention) comme le rapport de la variance du score vrai  $Var(V)$  sur la variance du score observé  $Var(Y)$  :

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

où  $\rho^2$  prend nécessairement une valeur entre 0 et 1 puisque  $Var(V)$  est positif et inférieur à  $Var(Y)$ . Ainsi, plus la valeur de  $Var(V)$  sera près de la valeur de  $Var(Y)$ , plus le coefficient  $\rho^2$  prendra une valeur proche de 1 et plus le test sera considéré comme étant fidèle. Pour le dire autrement, un test fidèle révèle le score vrai, entaché de peu d'erreurs. Ainsi, parmi deux instruments de mesure d'une habileté donnée, on préférera évidemment le plus fidèle.

Il est malheureusement impossible d'avoir accès à la valeur exacte de  $Var(V)$ , car comme le disait Borsboom (2008) : l'intelligence générale, l'habileté spatiale, l'attitude et l'auto-efficacité [i. e., le  $V$  contenu dans l'équation 1] ne sont pas dans la tête des personnes. Quand on ouvre la tête d'une personne, nous trouvons une espèce de gelée grise, pas des variables psychologiques (p. 42, traduction libre). Pour contourner cette difficulté, les psychométriciens ont développé différentes approches pour *estimer* la valeur de  $\rho^2$ , dont le fameux coefficient  $\alpha$  (alpha) de Cronbach (1951).

### Grandeur et décadence du coefficient $\alpha$ de Cronbach

L' $\alpha$  de Cronbach (1951) est probablement le coefficient le plus connu et le plus utilisé en psychométrie. Dunn, Baguley et Brunson (2013) rapportaient d'ailleurs que celui-ci avait été mentionné dans plus de 17 600 publications depuis le début des années 1950.

Cronbach, qui puise sa réflexion dans la troisième forme de  $\rho^2$  présentée à l'équation 2, argumente d'une part en disant que la façon dont les réponses aux items fluctuent d'une personne à l'autre doit être un estimé de la variance d'erreur  $Var(E)$  ; d'autre part que l'ensemble des variabilités dans les réponses des individus, les variances et les covariances,<sup>2</sup> doit être un estimé de la variance du score observé,  $Var(Y)$ . Rassemblant ces deux intuitions, Cronbach a proposé le coefficient  $\alpha$  pour estimer la fidélité  $\rho^2$ :

$$\alpha = \frac{J}{J-1} \left( 1 - \frac{\sum_{j=1}^J Var(Item_j)}{Var(Total)} \right) \quad (3)$$

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.

Le Tableau 1 présente une matrice de variances-covariances entre quatre items qui permettra au lecteur de mieux comprendre comment calculer ce coefficient.

TABLEAU 1. *Matrice de variances-covariances entre quatre items*

Item	1	2	3	4
1	,799	,365	,413	,313
2	,365	1,275	,572	,446
3	,413	,572	1,452	,477
4	,313	,446	,477	1,191

D'abord, nous savons que  $J = 4$  puisqu'il y a quatre items analysés. L'élément  $Var(Item_j)$  est la somme des éléments qui se trouvent dans la diagonale d'une matrice de variances-covariances ( $,799 + 1,275 + 1,452 + 1,191 = 4,717$ ) alors que  $Var(Total)$  est la somme des 16 nombres contenus dans cette même matrice ( $,799 + ,365 + ,413 + ,313 + ,365 + 1,275 + \dots + ,477 + 1,191 = 9,889$ ). Il ne reste plus qu'à remplacer les éléments dans l'équation 3 :

$$\alpha = \frac{4}{4-1} \left( 1 - \frac{4,717}{9,889} \right) = ,697$$

En ce qui concerne l'interprétation de  $\alpha$ , Tavakol et Dennick (2011) indiquent qu'une valeur minimale de ,7 est nécessaire pour obtenir un  $\alpha$  satisfaisant.

Malheureusement, il est connu que ce coefficient présente de nombreux problèmes dont trois sont décrits ici (Laveault, 2012 ; Revelle et Zinbarg, 2009 ; Sijtsma, 2009 ; Trizano-Hermosilla et Alvarado, 2016). Le premier est que l' $\alpha$  de Cronbach repose sur une hypothèse difficile à démontrer qui est que  $Var(V)$  doit être identique pour tous les items du test. Cette hypothèse théorique stipule que chaque item mesure le même construit, avec la même précision, mais avec différents niveaux d'erreur (Graham, 2006). Celle-ci est irréaliste, presque systématiquement violée et a comme effet d'engendrer une sous-estimation de la fidélité réelle (Miller, 1995). Conséquemment, le coefficient  $\alpha$  de Cronbach ne doit jamais être utilisé avec un test contenant différents formats d'item, comme c'est pourtant le cas dans un grand nombre de questionnaires ou de tests. Le deuxième est qu'il arrive que le coefficient  $\alpha$  prenne une valeur négative si certaines covariances sont aussi négatives (Martineau, 1982). Or, la fidélité doit être bornée entre 0 et 1 selon la définition présentée à l'équation 2. Enfin, plusieurs chercheurs justifient le retrait d'un item d'un test en calculant un «  $\alpha$  si l'item est enlevé » ( *$\alpha$  if item deleted*). Malheureusement, cette pratique est à proscrire, car elle est dépendante de l'échantillon analysé ; ce qui peut faire artificiellement augmenter ou diminuer la valeur du coefficient  $\alpha$  (Dunn et coll., 2013).

Sijtsma (2009) déclarait : la seule raison pour rapporter le coefficient alpha est que les journaux scientifiques de premier plan tendent à accepter les articles qui utilisent des méthodes statistiques présentes depuis longtemps (p. 118, traduction libre). Dans l'esprit de ses propos, nous considérons qu'il est aujourd'hui temps de mettre de côté ce coefficient pour adopter des alternatives plus efficaces. La suite de cet article présente celle qui semble la plus prometteuse, le coefficient omega de McDonald.

UN MEILLEUR COEFFICIENT DE FIDÉLITÉ : LE  $\omega$  DE MCDONALD

Le coefficient  $\omega$  (omega) permet d'estimer  $\rho^2$  à partir d'un modèle d'analyse factorielle à un facteur commun (*common factor analysis model*). Ce modèle a comme objectif général de regrouper les items d'un test qui mesurent un même facteur commun (ou général) auquel peuvent être ajoutés d'autres facteurs.

Selon McDonald (1999), le modèle d'analyse factorielle à un facteur commun appliqué à des items homogènes est une extension du modèle de score vrai présenté à l'équation 1. Celui-ci peut être exprimé de la façon simplifiée suivante :

$$Y = \mu + C + U \tag{4}$$

où  $Y$  est le score total au test,<sup>3</sup>  $C$  est la part de  $Y$  qui est exprimée par le facteur commun,  $U$  est la part de  $Y$  qui est exprimée par une compétence unique à cet item et ne se distingue pas de l'erreur de mesure. Enfin,  $\mu$  est une constante sans variance qui est, en pratique, triviale. Postulant que le facteur commun est indépendant des facteurs uniques, on déduit que :

$$Var(Y) = Var(C) + Var(U) \tag{5}$$

où par analogie avec l'équation 2,  $Var(V) = Var(C)$  et  $Var(E) = Var(U)$ . Enfin, McDonald (1999) définit la fidélité comme étant égale à

$$\rho^2 = \frac{Var(V)}{Var(Y)} = \frac{Var(V)}{Var(V) + Var(E)} = \frac{Var(C)}{Var(C) + Var(U)} \tag{6}$$

comme précédemment. Évidemment, les éléments  $C$  et  $U$  ne sont pas directement observables. Cependant, l'analyse factorielle permet de chiffrer la contribution du facteur commun  $C$  aux scores observables en calculant des *coefficients de saturation*  $\lambda_j$  (*factor loading*). La mise au carré de  $\lambda_j$  permet ensuite d'obtenir la proportion de variance d'un item dû au facteur commun (*communalities*). La différence  $1 - \lambda_j^2$  est alors la proportion de la variance due à l'erreur de mesure et aux facteurs uniques, notée  $\psi_j^2$  (*uniqueness*).

Selon McDonald (1985, 1999), la somme des coefficients de saturation, une fois mise au carré, fournit une estimation de  $Var(C)$  et donc de  $Var(V)$ , alors que la somme des coefficients des facteurs uniques donne une estimation de  $Var(U)$ , et donc de  $Var(E)$ . Tout cela mène à :

$$\omega = \frac{\left(\sum_{j=1}^J \lambda_j\right)^2}{\left(\sum_{j=1}^J \lambda_j\right)^2 + \sum_{j=1}^J \psi_j^2} \tag{7}$$

où  $\lambda_j$  est le coefficient de saturation (*factor loadings*) de l'item  $j$  et  $\psi_j^2$  est la variance unique de l'item  $j$ . Ainsi, seulement deux ingrédients sont nécessaires pour calculer  $\omega$ . Le lecteur curieux pourra consulter l'Annexe B pour obtenir plus d'information.

Le Tableau 2 présente des données permettant au lecteur de comprendre comment calculer ce coefficient.

TABLEAU 2. Exemple de données avec un facteur général basées sur une matrice de corrélations

Item	$\lambda_j$	$h_j^2 = \lambda_j^2$	$\psi_j^2 = 1 - h_j^2$
1	,57	,33	,67
2	,64	,41	,59
3	,66	,43	,57
4	,56	,31	,69
Somme	2,43	-	2,52

La deuxième colonne du Tableau 2 présente tous les coefficients de saturation du facteur commun. Ainsi, plus un coefficient est élevé, plus cet item est révélateur de ce facteur commun. Dans ce cas-ci, l'item 3 présente par exemple une meilleure saturation que l'item 4 au facteur commun. L'élément  $h_j^2$ , de son côté, est la communauté (*communalities*) de l'item, soit la variance de chaque item qui est expliquée par le facteur commun. Finalement, la dernière colonne présente la variance unique à l'item,  $\psi_j^2 = 1 - h_j^2$ , c'est à dire la proportion de la variance qui n'est pas attribuable au facteur commun. Ainsi,

$$\omega = \frac{(2,43)^2}{(2,43)^2 + 2,52} = ,70$$

s'interprète de la même façon que  $\alpha$ : plus la valeur s'approche de 1, plus le test est considéré comme étant fidèle. Ainsi, une valeur égale ou supérieure à ,70 témoigne d'une fidélité satisfaisante.

L'exemple précédent postule qu'il n'y a qu'un seul facteur mesuré par le test. Pourtant, les chercheurs analysent souvent des données de questionnaires mesurant plusieurs facteurs (ou dimensions, au nombre de  $D$ ) spécifiques en plus d'un facteur général  $G$  sous-tendant tous les items. L'avantage de  $\omega$  est qu'il se généralise à un nombre quelconque de facteurs : le coefficient prend alors la forme suivante pour considérer l'ajout des coefficients de saturation des facteurs supplémentaires :

$$\omega = \frac{[(\sum_{j=1}^J \lambda_{Gj})^2 + \sum_{d=1}^D (\sum_{j=1}^J \lambda_{dj})^2]}{[(\sum_{j=1}^J \lambda_{Gj})^2 + \sum_{d=1}^D (\sum_{j=1}^J \lambda_{dj})^2] + \sum_{j=1}^J \psi_j^2} \quad (8)$$

Nous allons illustrer le calcul à l'aide des données présentées au Tableau 3.

TABLEAU 3. Exemple de résultats d'une analyse factorielle contenant deux facteurs en plus d'un facteur commun G

Item	$\lambda_{dj}$	$\lambda_{1j}$	$\lambda_{2j}$	$h_j^2$	$\psi_j^2$
1	,70	,33	-	,60	,40
2	,63	,30	-	,49	,51
3	,54	,25	-	,35	,65
4	,49	-	,50	,49	,51
5	,41	-	,45	,37	,63
6	,27	-	,27	,15	,85
Somme	3,04	,88	1,22	-	3,55

Puisque  $\omega$  repose sur le modèle d'analyse factorielle à un facteur commun, nous observons des coefficients de saturation à tous les items pour un facteur général. Ensuite, le premier facteur est révélé par les items 1 à 3 et le deuxième facteur, par les items 4 à 6. Mentionnons aussi que  $h_j^2 = \lambda_{Gj}^2 + \lambda_{1j}^2 + \lambda_{2j}^2$  et  $\psi_j^2 = 1 - h_j^2$ . Enfin, selon les données présentées au tableau 3, nous obtenons :

$$\omega = \frac{[(3,04)^2 + (0,88)^2 + (1,22)^2]}{[(3,04)^2 + (0,88)^2 + (1,22)^2] + 3,55} = ,76$$

Il existe également une variante du coefficient  $\omega$ , le coefficient  $\omega_{\text{Hiérarchique}}$ , qui ne sera pas rapporté dans cet article. Nous encourageons le lecteur intéressé à consulter McDonald (1985, 1999) et Widhiarso et Ravand (2014). Pour conclure, nous ajoutons que Kelley et Pornprasertmanit (2016) recommandaient de calculer les intervalles de confiance du coefficient  $\omega$  à l'aide de la méthode par bootstrap.

### Illustration

Le Programme international de recherche en lecture scolaire (PIRLS) est un test qui évalue le rendement en lecture des élèves de 4<sup>e</sup> année (Conseil des ministres du Canada, 2012). Nous allons utiliser une partie des réponses de 4 762 élèves canadiens à ce test en guise d'illustration. Les 15 items sélectionnés portent sur la compréhension d'un court texte intitulé « *Enemy Pie* ». Chacun des items de cette section est dichotomisé en une bonne ou une mauvaise réponse.

Le Tableau 4 présente la moyenne (M) et l'écart-type (ÉT) de chacun des 15 items. Nous observons que Q15 et Q5 sont respectivement les items les moins bien réussis, alors que Q10 et Q13 sont les items les mieux réussis par les élèves. Les écart-types sont bornés entre ,35 et ,50.

TABLEAU 4. Statistiques descriptives de 15 items du PIRLS 2011

Item	M	ÉT
Q1	,83	,37
Q2	,82	,38
Q3	,79	,40
Q4	,50	,50
Q5	,43	,50
Q6	,80	,40
Q7	,51	,50
Q8	,78	,41
Q9	,81	,39
Q10	,86	,35
Q11	,76	,43
Q12	,67	,47
Q13	,86	,35
Q14	,61	,49
Q15	,37	,49

TABLEAU 5.  $\omega$  si l'item est enlevé

Item	$\omega$
Q1	,828
Q2	,830
Q3	,827
Q4	,830
Q5	,824
Q6	,823
Q7	,824
Q8	,822
Q9	,828
Q10	,818
Q11	,823
Q12	,820
Q13	,825
Q14	,823
Q15	,825



Un coefficient  $\omega$  de ,83 (Tableau 4) a été calculé pour les items liés au test « *Enemy Pie* », ce qui représente une bonne fidélité. De plus, nous présentons une analyse «  $\omega$  si l'item est enlevé » au Tableau 5. Les résultats montrent que la suppression d'un item n'a pas d'impact important la valeur du coefficient  $\omega$ .

### CONCLUSION : POURQUOI UTILISER LE COEFFICIENT $\omega$ PLUTÔT QUE $\alpha$ ?

Nous proposons plusieurs arguments pour répondre à cette question. Tout d'abord, le coefficient  $\omega$  offre des conditions d'utilisation plus souples que  $\alpha$ . Le fait qu'il accepte que les données soient multifactorielles (ou multidimensionnelles) correspond bien à la réalité des données collectées en sciences de l'éducation. Deuxièmement,  $\omega$  offre une estimation de la fidélité qui est cohérente avec celle de la population lors d'analyses du type «  $\omega$  si l'item est enlevé ». Le troisième argument en est un de cohérence entre la définition de la fidélité et le résultat de l'estimation. Ainsi,  $\omega$  présente le clair avantage de ne jamais être un nombre négatif, contrairement au coefficient  $\alpha$ . Enfin, Revelle et Zinbarg (2009) nous offrent un argument de plus. Ils ont comparé 13 estimateurs de la fidélité et leur conclusion est sans équivoque : c'est  $\omega$  qui présente les estimations les plus précises de  $\rho^2$  dans des simulations où le  $\rho^2$  véridique est connu.

Une difficulté d'utilisation est toutefois qu'il faut préalablement réaliser une analyse factorielle à un facteur commun (confirmatoire ou exploratoire) pour pouvoir estimer la valeur de  $\omega$ . Cette difficulté est toutefois balayée si on utilise une procédure qui automatise le calcul dans R (voir l'Annexe A) ou le logiciel gratuit avec interface JASP (<https://jasp-stats.org/>). Dans le cas de SPSS, il est uniquement possible de calculer  $\omega$  sur des tests contenant un seul facteur général. Évidemment, tout ce qui est discuté dans cet article est valable uniquement si le modèle d'analyse factorielle présente une bonne adéquation (*fit*) aux données.

Nous allons, pour conclure, reprendre les paroles de Dunn et coll. (2013) : étant donné le consensus dans les écrits en psychométrie que l'alpha est rarement approprié et considérant la bonne performance de l'omega lorsque les postulats d'utilisation de l'alpha ne sont pas respectés, il est recommandé que les psychologues [tout comme les gens en éducation] changent leur façon de faire en rapportant l'omega plutôt que l'alpha (p. 409, traduction libre).

### NOTES

1. La variance est une mesure de dispersion qui consiste *grosso modo* en la moyenne de toutes les différences mises au carré entre une donnée et la moyenne de son échantillon.
2. La covariance est *grosso modo*, la moyenne des écarts conjoints des répondants à deux items. La covariance est négative lorsque la corrélation entre ces deux items est négative.
3. Nous maintenons cette notation pour éviter toute confusion chez le lecteur.

## RÉFÉRENCES

- Borsboom, D. (2008). Latent variable theory. *Measurement*, 6(1-2), 25-53.
- Conseil des ministres du Canada. (2012). PIRLS 2011 : Le contexte au Canada [Rapport]. Repéré à [https://cmec.ca/Publications/Lists/Publications/Attachments/294/PIRLS\\_2011\\_FR.pdf](https://cmec.ca/Publications/Lists/Publications/Attachments/294/PIRLS_2011_FR.pdf)
- Crocker, L. et Algina, J. (1986). *Introduction to classical and modern test theory*. New York, NY : CBS College Publishing.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 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(3), 399-412. doi:10.1111/bjop.12046
- Graham J. M. (2006). Congeneric and (essentially) tau-equivalent estimates of score reliability what they are and how to use them. *Educational and Psychological Measurement*, 66(6), 930-944. doi: 10.1177/0013164406288165
- Kelley, K. (2017). MBESS: The MBESS R Package. R package version 4.4.2 [logiciel]. Repéré à <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(1), 69-92, doi: 0.1037/a0040086
- Laveault, D. (2012). Soixante ans de bons et mauvais usages du alpha de Cronbach. *Mesure et évaluation en éducation*, 35(2), 1-7. doi: 10.7202/1024716ar
- Martineau, G. (1982). Exploration des valeurs possibles du coefficient  $\alpha$  de Cronbach. *Revue des sciences de l'éducation*, 8(1), 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.
- Miller, M. B. (1995). Coefficient alpha: A basic introduction from the perspectives of classical test theory and structural equation modeling. *Structural Equation Modeling*, 2(3), 255-273.
- Revelle, W. (2017). psych: Procedures for personality and psychological research [logiciel]. Repéré à <https://CRAN.R-project.org/package=psych Version = 1.7.8>.
- Revelle, W. et Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. *Psychometrika*, 74(1), 145-154. doi:10.1007/s11336-008-9102-z
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, 74(1), 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(7), 769. doi: 10.3389/fpsyg.2016.00769
- Widhiarso, W. et Ravand, H. (2014). Reliability coefficient for multidimensional measures. *Review of Psychology*, 21(2), 111-121.

## ANNEXE A. CALCULER LE COEFFICIENT $\omega$ À L'AIDE DE LOGICIELS EN TÉLÉCHARGEMENT GRATUIT

La librairie `psych`, version 1.7.5 et suivante (Revelle, 2017) permet de calculer le coefficient  $\omega$  en seulement quelques lignes de syntaxe. Une fois que l'utilisateur aura téléchargé certaines librairies, il doit les charger dans sa console R pour débiter la séance d'analyse :

```
require(psych)
require(foreign)
require(GPARotation)
```

L'étape suivante consiste à charger les données en mémoire à l'aide du code suivant :

```
fichier <- choose.files()
X <- read.spss( fichier, use.value.labels = FALSE)
```

Si le fichier ne contient que les items et rien d'autre, vous êtes prêts :

```
omcoef <- omega( X )
omcoef$omega.tot
```

Sinon, il faut identifier les colonnes pertinentes, par exemple, la colonne du début du questionnaire et de la fin du questionnaire. Ainsi, si les colonnes sont nommées CE01 à CE36 dans cet ordre :

```
start_col <- match("CE01",names(X))
end_col <- match("CE36",names(X))
Y <- data.frame(X[start_col : end_col])
omcoef <- omega( Y )
omcoef$omega.tot
```

Il est important de noter que la librairie `MBESS` (Kelley, 2017) permet aussi de calculer  $\omega$ . De plus, celle-ci offre la possibilité de calculer les intervalles de confiance de ce coefficient. Enfin,  $\omega$  a été intégré à `JASP` (<https://jasp-stats.org/>), un logiciel en téléchargement libre qui a l'avantage de présenter une interface. Pour cela, l'analyste doit ouvrir l'onglet « Descriptives » et sélectionner « Reliability Analysis ». Ensuite, il doit sélectionner la case intitulée  $\omega$ , qui se trouve dans la section « Scale statistics ».

## ANNEXE B. QUELQUES DÉTAILS TECHNIQUES POUR LE LECTEUR CURIEUX

Rappelons d'abord que le modèle de score vrai (*true score model*) s'écrit :

$$Y = V + E \quad (9, \text{ voir aussi l'équation 1})$$

Le modèle d'analyse factorielle à un facteur commun est une mise à l'échelle du modèle de score vrai et il prend la forme suivante :

$$Y_j = \mu_j + \lambda_j F + E_j \quad (10)$$

où  $Y_j$  est la réponse de l'étudiant à l'item  $j$ ,  $\mu_j$  est une constante permettant à chaque item d'avoir une difficulté qui lui est propre,  $\lambda_j$  est un coefficient de saturation qui s'apparente à la discrimination dans le modèle à deux paramètres de la Théorie de la réponse aux items,  $F$  est un score factoriel et  $E_j$  l'erreur de mesure aléatoire. Pour rendre le lien entre les deux équations précédentes plus explicite, McDonald (1999) mentionne que l'équation 10 peut aussi être écrite de la façon simplifiée suivante :

$$Y = \mu + V + E \quad (11)$$

Selon cet auteur (1999), le score total au test (nous maintenons encore une fois la notation  $Y$  pour faciliter la démonstration) est équivalent à :

$$Y = \sum_{j=1}^J \mu_j + \left( \sum_{j=1}^J \lambda_j \right) F + \sum_{j=1}^J E_j \quad (12)$$

qui peut être simplifié de la façon suivante :

$$Y = \mu + C + U \quad (13, \text{ voir aussi l'équation 4})$$

où  $Y$  est le score total au test,  $C$  est la part de  $Y$  qui est exprimée par le facteur commun,  $U$  est la part de  $Y$  qui est exprimée par une compétence unique à cet item et ne se distingue pas de l'erreur de mesure. Ici,  $\mu$  est une constante. La variance du score total, elle, se décompose de la façon suivante :

$$\text{Var}(Y) = \text{Var}\left(\sum_{j=1}^J \mu_j\right) + \text{Var}\left[\left(\sum_{j=1}^J \lambda_j\right)F\right] + \text{Var}\left(\sum_{j=1}^J E_j\right) \quad (14)$$

qui se réduit à

$$\text{Var}(Y) = \text{Var}\left[\left(\sum_{j=1}^J \lambda_j\right)F\right] + \text{Var}\left(\sum_{j=1}^J E_j\right) \quad (15)$$

car l'élément  $\sum_{j=1}^J \mu_j$  est trivial. Après quelques manipulations algébriques, nous obtenons finalement

$$\text{Var}(Y) = \left(\sum_{j=1}^J \lambda_j\right)^2 + \sum_{j=1}^J \psi_j^2 \quad (16)$$

où  $\text{Var}(C) = \left(\sum_{j=1}^J \lambda_j\right)^2$  et  $\text{Var}(U) = \sum_{j=1}^J \psi_j^2$  sachant que la variance unique à l'item est  $\psi_j^2 = 1 - h_j^2$ . Tout cela nous ramène à l'équation de la fidélité présentée un peu plus haut, soit :

$$\rho^2 = \frac{Var(V)}{Var(Y)} = \frac{Var(V)}{Var(V) + Var(E)} = \frac{Var(C)}{Var(C) + Var(U)} \quad (17, \text{ voir aussi l'équation 6})$$

Finalement, le coefficient  $\omega$  prend la forme suivante pour des données contenant un seul facteur général :

$$\omega = \frac{\left(\sum_{j=1}^J \lambda_j\right)^2}{\left(\sum_{j=1}^J \lambda_j\right)^2 + \left(\sum_{j=1}^J \psi_j^2\right)} \quad (18, \text{ voir aussi l'équation 7})$$

et la forme suivante dans une situation où il y a un ou plusieurs facteurs secondaires qui s'ajoutent au facteur général :

$$\omega = \frac{\left[\left(\sum_{j=1}^J \lambda_{Gj}\right)^2 + \sum_{d=1}^D \left(\sum_{j=1}^J \lambda_{dj}\right)^2\right]}{\left[\left(\sum_{j=1}^J \lambda_{Gj}\right)^2 + \sum_{d=1}^D \left(\sum_{j=1}^J \lambda_{dj}\right)^2\right] + \sum_{j=1}^J \psi_j^2} \quad (19, \text{ voir aussi l'équation 8})$$

SÉBASTIEN BÉLAND est professeur adjoint au département d'administration et fondements de l'éducation, à l'Université de Montréal. Ses intérêts de recherche suivent deux axes : la mesure dans le domaine de l'éducation et l'évaluation des apprentissages dans les programmes d'études en art, au niveau postsecondaires. Il est aussi membre du Groupe de recherche interuniversitaire sur l'évaluation et la mesure en éducation à l'aide des technologies de l'information et de la communication (GRIÉMÉtic). [sebastien.beland@umontreal.ca](mailto:sebastien.beland@umontreal.ca)

DENIS COUSINEAU est professeur de l'Université d'Ottawa à l'école de psychologie et fondateur de la revue *The Quantitative Methods for Psychology* ([www.tqmp.org](http://www.tqmp.org)). Ses recherches portent sur les modèles de temps de réponse, l'attention visuelle et la recherche visuelle. Il travaille aussi en mathématique sur l'estimation de paramètres pour les distributions de Weibull et de Pareto. [Denis.Cousineau@uottawa.ca](mailto:Denis.Cousineau@uottawa.ca)

NATHALIE LOYE est professeure agrégée au département d'administration et fondements de l'éducation, à l'Université de Montréal. Ses intérêts de recherche portent sur les modèles de mesure appliqués à des données issues de tests ou de questionnaires dans les domaines de l'éducation et de la santé. Elle s'intéresse particulièrement à la validité des données et des instruments. Nathalie est directrice du Groupe de recherche interuniversitaire sur l'évaluation et la mesure en éducation à l'aide des technologies de l'information et de la communication (GRIÉMÉtic). [nathalie.loye@umontreal.ca](mailto:nathalie.loye@umontreal.ca)

SÉBASTIEN BÉLAND is adjunct professor in the Département d'administration et fondements de l'éducation at the Université de Montréal. His research interest focuses on two main topics: measurement in the field of education and postsecondary learning assessment in arts. He is also a member of the Groupe de recherche interuniversitaire sur l'évaluation et la mesure en éducation à l'aide des technologies de l'information et de la communication (GRIÉMÉtic). [sebastien.beland@umontreal.ca](mailto:sebastien.beland@umontreal.ca)

DENIS COUSINEAU is professor at the School of Psychology, University of Ottawa, and founder of *The Quantitative Methods for Psychology* ([www.tqmp.org](http://www.tqmp.org)). His areas of research include response time models, attention, and visual search. He also worked in mathematics on parameter estimation for the Weibull and the Pareto distributions. [Denis.Cousineau@uottawa.ca](mailto:Denis.Cousineau@uottawa.ca)

NATHALIE LOYE is associate professor in the Département d'administration et fondements de l'éducation at the Université de Montréal. Her research interests include measurement models applied to data from questionnaires or tests in the fields of education and health. She is particularly interested in the validity of data and instruments. Nathalie is the director of the Groupe de recherche interuniversitaire sur l'évaluation et la mesure en éducation à l'aide des technologies de l'information et de la communication (GRIÉMÉtic). [nathalie.loye@umontreal.ca](mailto:nathalie.loye@umontreal.ca)