

L'USAGE ABUSIF ET INCONSIDÉRÉ D'ALPHA CRONBACH ET SES PALLIATIFS

Ali TADJINE¹

The wrong use of alpha'Cronbach

Abstract

The use of measuring instruments in any scientific research is one of the most critical methodological steps. This step is of such importance that it determines the efficiency and effectiveness of the research as a whole. It must guarantee their validity and loyalty, not subjecting them to the requirements of psychometric is simply exposing the results of research to be null. Validity and reliability are two procedures in testing, measuring instrument design. Cronbach's coefficient alpha is probably the most frequently reported measure of reliability and most certainly the most abused one. Concerning reliability, Cronbach alpha weaknesses were examined, and some recommended alternatives method for assessing reliability such as stratified alpha, and composite reliability (construct reliability or Rho) were considered.

Key words: *Validity, Reliability, Cronbach's alpha, Stratified alpha, Composite reliability.*

L'usage des instruments de mesure dans toute recherche scientifique constitue une des étapes méthodologiques les plus critiques. Cette étape est d'une importance telle qu'elle conditionne l'efficacité et l'efficacité de la recherche dans sa globalité. Elle doit garantir la validité et la fidélité des instruments de mesure ; ne pas les soumettre aux exigences de la psychométrie c'est tout simplement exposer les résultats de la recherche à être frappé de nullité. Confrontés à cette exigence d'expertise particulière, les chercheurs surtout en sciences sociales se trouvent souvent démunis en termes de compétences en psychométrie et combler ce déficit par le recours à un travail de mimétisme des travaux académiques exécutés sans réelle maîtrise, ce qui conduit à l'émergence de procédures que les chercheurs suivent malgré qu'elles recèlent des erreurs qui deviennent communes et surtout acceptées sans se rendre compte de l'impact négatif et préjudiciable. Cette réalité est perceptible au quotidien dans notre métier d'enseignants et d'encadreurs de thèses. Il est utile de préciser que les chercheurs en sciences de la nature sont épargnés de cette problématique du fait qu'ils n'élaborent pas eux-mêmes leurs instruments de mesure, ils n'ont que l'obligation de vérifier l'homologation, le calibrage et l'étalonnage, ce qui n'exige pas le recours à la psychométrie. Dans le cadre de cet

¹ Professeur de psychologie des organisations et gestion des ressources humaines – Université Abdelhamid Ibn Badis, Mostaganem – Algérie, Courriel : alitadjine@yahoo.fr.

article, nous nous intéressons à certaines procédures en usage par les chercheurs concernant la vérification des caractéristiques psychométriques des instruments de mesure que ce soient tests, questionnaires ou tout autre instrument.

Sans trop s'attarder sur la définition de la validité que je reprends de façon superficielle de la littérature spécialisée en la matière à savoir « être capable de mesurer ce qu'il est censé mesurer », la typologie préconisée par l'Association américaine de psychologie faisant référence en la matière présente quatre possibilités à savoir la validité du contenu, la validité du critère prédictif, la validité du critère concomitant et la validité du construit. Il est à préciser que cette typologie qui date des années soixante du siècle dernier a mis fin à une anarchie de classification de la validité que nous pouvons trouver dans les anciens livres de psychométrie, mais qui malheureusement certains plus récents reprennent, ce qui continue à fourvoyer et perturber de nombreux chercheurs. Pour ce qui est de la fidélité, il y a lieu de préciser que de nombreux ouvrages spécialisés la définissent suivant une approche réductrice à savoir la capacité de l'instrument de mesure d'avoir les mêmes résultats en cas de répétition (test –retest), alors qu'en réalité elle ne se limite pas à ce seul aspect, mais essaye d'investiguer la part des erreurs aléatoires sur la performance de l'instrument de mesure et peut être également être calculé par les méthodes basées sur la consistance interne et les méthodes mixtes dites des formes parallèles. Dans le cadre de cet article, je m'intéresse à l'usage de la méthode basée sur la consistance interne et plus spécifiquement au recours excessif et en progression croissante de la technique d'alpha de Cronbach. La méthode basée sur la consistance interne suppose la réalisation de conditions que le chercheur doit obligatoirement prendre en considération.

Les conditions d'utilisations du coefficient d'alpha de Cronbach peuvent être résumées comme suit :

Le coefficient alpha est une mesure de la cohérence interne. Cette cohérence interne s'exprime par l'émergence d'une consistance interne, c'est-à-dire l'existence de degrés d'inter corrélation entre les items élevés donnant une structure homogène de l'instrument de mesure. La cohérence interne supposée est une condition nécessaire, mais non suffisante de l'homogénéité des items d'une échelle, cette dernière impliquant l'unidimensionnalité de l'instrument en question (Green *et al.*, 1977). Lorsque les items ne peuvent être considérés comme tau-équivalents (expression statistique appropriée), la fidélité de cohérence interne ne peut être estimée avec exactitude et représente plutôt la borne inférieure de la valeur réelle de fidélité ; ce qui n'est pas suffisant pour s'approprier l'instrument de mesure et l'utiliser de façon appropriée. Pour que le coefficient alpha tende vers sa valeur maximale possible, les items doivent être tau-équivalents, c'est-à-dire être en fortes corrélations et ne différer entre eux que par une constante (Cortina, 1993).

La question qui nous interpelle est : est-ce que vraiment on s'assure en procédant à l'usage d'alpha de Cronbach que les conditions d'utilisation sont respectées ? La réponse est non. La lecture et l'exploitation de nombre de recherches scientifiques et de thèses utilisant ce type de technique laissent entrevoir un manque flagrant de précaution et de respect de la condition d'utilisation. Parallèlement à cet

aspect de méconnaissance dans l'utilisation et de manque de maîtrise flagrant que de nombreux auteurs ont dénoncé et ont démontré les erreurs et la méconnaissance d'usage à l'instar de Cortina 1993, Schmitt 1996, Cronbach et Shavelson (2004) que nous reprendrons ultérieurement. Dans le cadre de cet article, nous nous intéressons au fait que alpha de Cronbach est excessivement utilisé, situation qui nous interpelle : Pourquoi ce recours presque exclusif à cet usage ? Thomas J. Dunn, Thom Baguley & Vivienne Brunsden (2013) ont recensé que Alpha de Cronbach a été utilisé dans plus de 17600 publications. Le nombre d'utilisations inappropriées du coefficient s'est vite accru, au point où plusieurs auteurs (Cortina, Green, Lissitz, & Mulaik, 1977 ; Schmitt, 1996) et Cronbach lui-même (2004) ont cru nécessaire de rappeler et de préciser les limites et les conditions de son utilisation appropriée.

Apparemment, la réponse réside dans sa relative facilité et simplicité d'utilisation surtout que de nombreux logiciels le permettent. Cet usage excessif et parfois non efficient se trouve être encore plus accentué par le fait que de nombreux chercheurs se fourvoient dans l'application des prescriptions de la psychométrie. Nous mentionnons à titre indicatif le recours à l'analyse de la consistance interne comme indicateur de validité.

(Validity of internal consistency) qui s'intéresse aux corrélations entre les items entre eux, et aux corrélations entre les axes et l'instrument dans sa globalité. Cette pratique n'est en fait que la reproduction de l'application d'alpha de Cronbach. En effet, le chercheur en procédant de la sorte sans se rendre compte, arrive au même résultat en suivant un chemin différent. Cette propension à l'usage de la consistance interne de l'instrument comme indicateur de validité est la conséquence de l'hégémonie de l'idée que l'approche quantitative est plus déterminante alors que ce n'est pas le cas. Nombreux sont les chercheurs qui soumettent à l'appréciation chiffrée aussi bien la fidélité que la validité alors qu'en réalité, la validité n'est pas appréhendable à l'instar de la fidélité par le recours à l'approche quantitative. La validité ne peut être exclusivement réduite à une dimension statistique, mais va au-delà pour s'assurer la cohérence théorique. Assertion et prescription que de nombreux chercheurs enfreignent et s'engagent dans le calcul de la consistance interne comme indicateur de validité par le recours à des calculs qui dans le fond ne sont qu'une duplication du coefficient d'alpha de Cronbach. Outre cet aspect préjudiciable de cette pratique, il est également utile de signaler, que lors de cet usage, une erreur fatale est communément admise à savoir le recours à l'exploitation de la signification statistique, qui conduit à l'élagage d'items parfois structurants simplement parce que le coefficient de corrélation n'atteint pas le sacro sein seuil de 0.70 d'usage. En se comportant de la sorte, l'instrument de mesure risque de se trouver amputer de parties essentielles, alors que ces parties (items) présentent une validité théorique évidente, mais la manipulation statistique, en procédant au calcul de la consistance interne et en comparant les coefficients de corrélation, décrète l'absence de signification. Le chercheur non averti, en se fiant aveuglément à la signification statistique, risque de mettre en œuvre un instrument de mesure ne présentant pas les caractéristiques psychométriques nécessaires du fait de la non-

maitrise d'utilisation des enseignements de la psychométrie ; ce qui constitue une entrave de taille et une problématique qui prend de l'ampleur.

De ce qui précède, il devient évident que s'intéresser à la méthodologie et la valorisation et dissémination des résultats de la recherche ne prend du sens que si l'ensemble des étapes de production de ces résultats sont soumis à la rigueur scientifique et présentent les conditions de validité. Comme il vient d'être spécifié, la valeur des instruments de mesure doit faire l'objet d'attention particulière, et dans le cadre de ce souci particulier, nous continuons de focaliser l'attention sur les limites d'utilisation du coefficient d'Alpha de Cronbach, qui a été élaboré en 1951 par J. Cronbach, ce coefficient « alpha » a d'abord été conçu, comme son nom l'indique, comme la première d'une série de mesures de calculs des différentes propriétés des scores (Cronbach & Shavelson, 2004). Cronbach en sa qualité de promoteur de la logique de calcul de la fidélité de l'instrument de mesure basée sur la consistance interne, a proposé plusieurs équations pour y parvenir dont le plus connu est le coefficient d'Alpha de Cronbach qui est fortement influencé par :

- la moyenne des corrélations ou la moyenne des covariances des items,

- par la variance des résultats qui constituent le résultat global de l'instrument de mesure,
- par sa longueur,
- par le nombre de ses items,
- par son homogénéité sans être unidimensionnel ou multidimensionnel.

Ses différentes caractéristiques devraient faire l'objet de précautions particulières par le chercheur sinon la logique mathématique qui soutient son fonctionnement perd toute pertinence et opérationnalité, ce qui devient source d'erreurs et d'égarements.

L'équation d'alpha de Cronbach exprime le seuil de variance de la valeur exacte (variance systémique non aléatoire) par rapport à la variance de l'instrument dans sa globalité. L'une des équations d'alpha de Cronbach se base sur la somme des valeurs de covariance entre les items de l'instrument par rapport à la valeur totale de l'instrument. Par variance totale de l'instrument, il faut comprendre l'ensemble des valeurs de covariance entre les items de l'instrument auquel il faut ajouter l'ensemble des valeurs de variance des items de l'instrument. La force ou faiblesse d'alpha de Cronbach est fonction du niveau des corrélations entre les items et par la même occasion la covariance entre les items. Cette caractéristique laisse supposer que ce coefficient reflète les corrélations et les covariances entre les items de l'instrument, ce qui constitue la consistance interne des items ou la moyenne des corrélations entre les items. La situation idoine est que les niveaux des corrélations entre les items soient élevés et du coup permet l'existence de la consistance interne et par conséquent de la fidélité de l'instrument de mesure. Cette situation espérée se trouve être mise en difficulté par l'existence d'éléments qui conduisent à la difficulté d'évaluer la fidélité à savoir la longueur de l'instrument, l'existence de très faibles corrélations entre les items. La pratique nous révèle à titre d'exemple qu'un instrument de mesure de 5 items avec un niveau de corrélation de 0.1 peut augmenter et atteindre un niveau d'alpha de Cronbach à 0.36 en augmentant simplement le

nombre des items. La longueur de l'instrument de mesure en influençant alpha de Cronbach peut également donner l'impression de consistance interne de l'instrument malgré que les axes le composant n'ont aucune relation, un faisceau d'axes disparates imbriqués dans un instrument de mesure peut atteindre un coefficient d'alpha de Cronbach élevé donnant l'impression que l'ensemble des items fonctionnent dans le cadre d'une approche unidimensionnelle alors qu'en réalité le niveau d'alpha de Cronbach n'est que le résultat de manipulations mathématique (Cortina 1993). Parallèlement à cette réalité que nombre de chercheurs ne prennent pas en considération, il y est également utile de signaler une autre possibilité d'avoir un alpha de Cronbach élevé et conclure de façon erronée à l'existence de la qualité de la fidélité, alors que ce niveau élevé n'est en aucun cas l'expression de la consistance interne, mais un indicateur de l'existence d'une redondance des items (*item over-redundancy*). Une duplication des mêmes items exprimée différemment. Les items ne sont pas indépendants les uns des autres, mais une suite de reformulations des mêmes idées, qui soumise à l'analyse statistique décrète l'existence d'un alpha de Cronbach élevé qu'il faut surtout pas considéré comme indicateur de fidélité de l'instrument (Streiner, 2003 ; Netemeyer, 2001). De ce qui précède, Il ressort clairement que le niveau élevé du coefficient d'alpha de Cronbach n'est en aucun cas le reflet de l'existence de la consistance interne de l'instrument, c'est-à-dire que la qualité psychométrique de fidélité n'est pas forcément prouvée en s'y fiant.

Le chercheur en utilisant le coefficient d'alpha de Cronbach et en supposant qu'il prend les précautions d'usage et maîtrise la technique conformément aux enseignements de la psychométrie, une question s'impose : Quel est le niveau acceptable de ce coefficient ? De nombreux chercheurs s'y sont intéressés et nous pouvons synthétiser dans le cadre de cet article en survol et de façon superficielle les contributions de certains d'eux à l'instar de Nunnally, 1967, 1978, Bernstein 1994, Streiner, 2003.

En se référant aux prescriptions de certains spécialistes en la matière à l'instar de Nunnally (1977, 1978), il apparaît que le seuil minimal de fidélité en général et du coefficient d'alpha de Cronbach en particulier était déterminé à 0.70 puis augmentait à 0.80 pour atteindre 0.90 si l'instrument est à usage de diagnostic dans un processus clinique. Cette exigence de seuil élevé du coefficient doit être prise avec précaution et ne pas le prendre comme une garantie indéfectible et indubitable de fidélité. Comme il vient d'être spécifié, il se peut qu'un niveau élevé d'alpha de Cronbach ne soit que le résultat de la longueur de l'instrument ou encore de la redondance des items (Raju, 1982, Cortina, Feld, 1993, Hattie, 1985, Steiner, 2003, et autres). En prenant en considération cette réalité, il devient évident pour tout chercheur de ne pas se fier sans précaution au niveau élevé du coefficient d'alpha de Cronbach ; son évaluation et la prise de décision quant à sa véritable valeur sont tributaires de la structure de l'instrument et du respect de la méthodologie de sa conception.

Parallèlement à ces quelques remarques concernant l'utilisation du coefficient d'alpha de Cronbach, il est également utile de préciser que des variantes peuvent être utilisées, ce qui permet aux chercheurs d'éviter les biais de ses utilisations abusives.

– Le coefficient d'Alpha stratifié : Il est connu que l'hypothèse de base pour l'utilisation du coefficient d'alpha de Cronbach soit que l'instrument soit unidimensionnel, c'est-à-dire structuré autour d'un seul facteur ou une seule dimension. Cette condition d'homogénéité de l'instrument se trouve être difficilement assurée sur le terrain, ce qui réduit le champ d'application d'alpha de Cronbach de façon efficiente. La parade peut être par le recours au coefficient d'alpha stratifié. En acceptant un coefficient d'alpha de Cronbach même élevé comme indicateur de consistance interne et par conséquent de fidélité alors que l'instrument n'est pas homogène conduit le chercheur à faire preuve de naïveté et d'affirmer des conséquences fallacieuses. Cronbach lui-même suivi par Shonemann, Mc Kie 1965 et après Schmidt et Hunter (1996) préconisent dans le cas d'évaluation de fidélité d'un instrument de recherche non unidimensionnel l'utilisation du coefficient d'alpha stratifié beaucoup plus adapté en précisant la condition que chaque axe de l'instrument de mesure ne s'intéresse qu'à une seule dimension.

Outre le coefficient d'alpha stratifié, on peut également utiliser la méthode de fidélité composite ou la fidélité du construit, parfois présentée par le symbole c_r (composite reliability) ou par la lettre grecque Rho ou encore Omega. En utilisant la méthode (c_r) pour évaluer la consistance interne de l'instrument, la pratique laisse entrevoir un avantage pour cette méthode par rapport au coefficient d'alpha de Cronbach. En considérant un instrument (questionnaire) composé de trois axes : la connaissance cognitive, la connaissance instrumentale et la connaissance méta connaissance. Prenant à titre d'exemple le premier axe et considérant qu'il est constitué de trois items. Les saturations des trois items sur le facteur sont : 0.90, 0.58 et 0.66. Le coefficient de fidélité composite (c_r) est de 0.80 alors que le coefficient d'alpha de Cronbach est de 0.75 Tighessa (2017). Cet exemple exprime la supériorité du (c_r) par rapport au coefficient d'alpha de Cronbach. Cette supériorité est encore plus flagrante en cas d'augmentation des saturations des items.

Cette caractéristique a permis à la méthode de la fidélité composite d'être l'objet d'engouement surtout en sciences sociales. Engouement qu'il faut modérer, car cette méthode présente certains revers à l'instar d'existence de saturation(s) négative(s) qui abaissent le coefficient (c_r), alors que traditionnellement le coefficient de fidélité considère les valeurs absolues des saturations. Il est également utile de signaler le fait que tout ajout de nouveau item influence à la baisse le coefficient (c_r).

Il est clair que ces remarques et bien d'autres ont fait l'objet de nombreuses publications dans les revues spécialisées, mais l'analyse de nombre de thèses et de rapports de recherches laisse entrevoir un manque flagrant de maîtrise ainsi que des interprétations erronées des indicateurs statistiques comme s'ils avaient des interprétations intrinsèques ; alors qu'elles n'ont de sens qu'en s'inscrivant dans une approche contingente et éclairée.

En guise de conclusion, il est d'importance vitale de préciser que l'utilisation du coefficient d'alpha de Cronbach qui est, comme il a été déjà dit, sans doute l'une des mesures les plus répandues de la fidélité ; c'est aussi l'une de celles qui — du fait de sa popularité et de sa facilité d'utilisation — s'avère être l'une des moins bien utilisées. Il faut retenir que ce coefficient ne couvre qu'une faible proposition des besoins de calculs de la fidélité et que son utilisation doit respecter des conditions bien particulières et doit être limitée aux cas bien précis à savoir l'existence de la cohérence interne. Condition d'homogénéité et de cohérence interne nécessaire, mais non suffisante. Tout concepteur ou utilisateur d'instrument de mesure doit veiller au respect des conditions psychométriques et à l'application stricte des conditions d'applications des différents indicateurs à l'instar du coefficient d'alpha de Cronbach.

RÉFÉRENCES BIBLIOGRAPHIQUES

1. AERA, & NCME. (2000). (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education). *Standards for educational and psychological testing*. Washington, DC : American Educational Research Association.
2. Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78(1), 98-104.
3. Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 12, 1-16.
4. Cronbach, L. J., & Shavelson, R. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement*, 64, 391-418.
5. Green, S. B., Lissitz, R. W., & Mulaik, s. a. (1977). Limitations of coefficient alpha as an index of test unidimensionality. *Educational and Psychological Measurement*, 37, 827-838.
6. Nunnally, J. C. (1967). *Psychometric Theory*. New York : McGraw-Hill.
7. Nunnally, J. C. (1978). *Psychometric Theory (2nd ed.)*. New York : McGraw-Hill.
8. Nunnally, J. C. & Bernstein, I. H. (1994). *Psychometric Theory (3rd ed.)*. New York : McGraw-Hill.
9. Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, 8(4), 350-353.
10. Raju, N. S. (1982). On test homogeneity and maximum KR-20. *Educational and Psychological Measurement*, 42, 145-152.
11. Streiner, D. L. (2003). Starting at the beginning: An introduction to Coefficient Alpha and internal consistency. *Journal of Personality Assessment*, 80, 99-103.
12. Schmidt, F. L. & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenario. *Psychological Methods*, 1, 199-223.

13. Tighessa M'hamed. (2017) (Nouvelles orientations dans l'évaluation de la validité et la fidélité des instruments de mesure — analyse d'évaluation théorique et pratique. *Revue des sciences psychologiques et d'éducation en langue arabe*.