

# Le coefficient de fiabilité et la multidimensionalité des mesures en systèmes d'information

Par  
Mohamed El Louadi  
Université de Tunis  
Institut Supérieur de Gestion  
41, Rue de la Liberté -- Cite Bouchoucha  
Le Bardo, Tunis 2000  
[mohamed.ellouadi@isg.mu.tn](mailto:mohamed.ellouadi@isg.mu.tn)

Novembre 2001

# Le coefficient de fiabilité et la multidimensionalité des mesures en systèmes d'information

## Résumé

Pour qu'une mesure soit considérée acceptable dans les recherches en sciences de gestion en général et en systèmes d'information en particulier, il faut au moins qu'elle satisfasse le test de fiabilité. Ce test, également appelé le test de fidélité, utilise souvent le coefficient alpha de Cronbach (1951). Sans s'étendre sur les questions de validité des construits, cet article s'appuie sur l'exemple de la mesure de la satisfaction de l'utilisateur avec un système d'information pour montrer que l'application du coefficient alpha pour déterminer la fiabilité d'une mesure se fait souvent d'une manière désinvolte et au détriment de l'esprit de la théorie du test. Nous concluons que le concept même du test de fiabilité est encore mal compris par les chercheurs en systèmes d'information.

## Introduction

La fiabilité d'une mesure se calcule différemment selon que la mesure est unidimensionnelle ou multidimensionnelle. Dans le cas où la mesure est unidimensionnelle, sa fiabilité est déterminée au niveau de toute la mesure. Dans le cas où la mesure est multidimensionnelle, sa fiabilité est déterminée en fonction de celles des dimensions qui la constituent. Ces dimensions sont soit déterminées a priori par la théorie soit révélées suite à la collecte des données.

Dans la littérature sur la recherche en systèmes d'information, les auteurs appliquent indifféremment le même traitement à leurs mesures qu'elles soient unidimensionnelles ou multidimensionnelles. Il s'agit alors de savoir si ce traitement est justifiable sur le plan de la rigueur scientifique. Cela est d'autant plus pertinent que la discipline des systèmes d'information est une discipline relativement jeune. La question posée est donc de savoir si le traitement de la fiabilité d'une mesure est indifférent selon que la mesure est unidimensionnelle ou multidimensionnelle.

La question est justifiée par le sentiment de l'existence d'un problème dans la littérature en systèmes d'information où cet aspect est quelque peu négligé, où la fiabilité est traitée d'une manière quasi unique et où des valeurs de coefficients de fiabilité indûment élevées sont publiées.

La revue de la littérature sur la recherche en systèmes d'information effectuée par Straub (1989) avait déjà révélé que 62% des 117 études empiriques recensées ne contenaient aucune évaluation psychométrique des mesures utilisées. Des études publiées dans des revues scientifiques aussi renommées que le *MIS Quarterly*, *Communications of the ACM* et *Information & Management*, 83% ne comportaient aucune évaluation de la fiabilité des mesures utilisées. La tendance ne se limite d'ailleurs pas uniquement à la fiabilité des mesures mais également à la validité des construits (El Louadi, 1995). Plusieurs autres chercheurs en systèmes d'information se sont inquiétés de la situation qui semble prévaloir dans le domaine de la recherche et surtout dans le courant du développement des mesures (Goodhue, 1988; Hamel, 1988; Kim, 1989; Klenke, 1992; Melone, 1990).

Nous nous proposons, dans cet article, de démontrer sur la base d'un exemple empirique que la fiabilité d'une mesure multidimensionnelle diffère sensiblement selon qu'elle est calculée au niveau de toute la mesure, comme si elle était unidimensionnelle, ou en fonction de la fiabilité de ses dimensions. L'objectif est de démontrer que ne serait-ce qu'au niveau du calcul, le traitement correct de la fiabilité d'une mesure multidimensionnelle n'est pas plus ardu que celui de la fiabilité d'une mesure unidimensionnelle. Ce faisant, nous démontrons que même si la fiabilité ainsi obtenue est inférieure, elle garantit une plus grande rigueur scientifique.

### *Le problème de la fiabilité des mesures*

Lorsqu'un concept ne se prête pas aisément à la mesure directe, il n'y a d'autre recours que de le mesurer indirectement, c'est-à-dire par le biais de ses manifestations externes. Le chercheur soucieux de mesurer ce concept choisit de saisir ces manifestations en utilisant autant d'indicateurs que possible. Les concepts multidimensionnels ne peuvent être mesurés qu'à l'aide de plusieurs indicateurs chacun tentant de capter une manifestation différente du concept. C'est ainsi que souvent, les chercheurs (Fornell, 1992) se trouvent face à un choix: poser plusieurs questions liées à un même concept (maximisant la fiabilité) ou poser des questions liées à plusieurs concepts (maximiser l'information).

En guise d'exemple, le revenu d'un ménage est non seulement un concept mesurable directement, mais il l'est à l'aide d'un seul indicateur. La surface d'un plan géométrique quoique mesurable directement, requiert deux indicateurs, la longueur et la largeur. D'autres concepts sont non seulement de nature multidimensionnelle mais aussi non directement mesurable; la personnalité est un exemple de concept multidimensionnel qui incorpore des traits de personnalité divers tels que l'autosuffisance, l'introversiion et la tendance à la domination (Flanagan, 1935).

Les chercheurs en sciences de gestion et en systèmes d'information ont souvent à composer avec des concepts qui sont (i) non directement mesurables et (ii) multidimensionnels. Par ailleurs, et en ce qui concerne les mesures en général, il est rare que la valeur observée (score) reflète la valeur exacte de ce qui est mesuré. Ce problème devient d'autant plus aigu lorsque les concepts sont multidimensionnels.

La différence entre le score réel et le score observé d'une mesure constitue une erreur due à plusieurs facteurs dont Sellitz, Lawrence, Wrightsman, et Cook (1976, pages 164-168) avaient identifié sept:

1. Des différences réelles entre les individus chez qui on mesure un concept ou un trait de caractère tel que la propension à révéler leurs sentiments par exemple;
2. Des différences dues à des facteurs circonstanciels tels que la fatigue ou l'humeur de la personne qui subit le test au moment où elle la subit;
3. Des différences dues aux conditions dans lesquelles le test est administré, est-ce que le sujet a répondu au test alors qu'il était chez lui ou dans un endroit où il se sentait moins à l'aise, etc.;
4. Des différences dues à la méthode choisie pour administrer le test: entrevue directe avec des contacts visuels fréquents, enquête par téléphone ou questionnaire anonyme, etc.;
5. Des différences dues aux indicateurs choisis pour former la mesure dans le test; il suffirait qu'un seul indicateur soit modifié pour que toute la mesure en soit affectée;
6. Des différences dues à la qualité de l'instrument, incluant sa clarté et les différentes interprétations qu'un sujet pourrait avoir des composantes du test;
7. Des différences dues aux facteurs mécaniques et aux erreurs machinales que les sujets commettent en répondant au test telles que toujours cocher la case "neutre" dans un questionnaire.

#### *Définition de la fiabilité d'une mesure*

Si l'on prend pour acquis que le score réel d'une mesure n'est ni directement observable ni directement mesurable, la relation statistique entre le score réel ( $X_R$ ) et le score observé ( $X_O$ ) est donnée par l'équation suivante:

$$X_O = X_R + \varepsilon$$

où  $\varepsilon$  est l'erreur de mesure qui est due à un ou plusieurs des facteurs identifiés par Sellitz et al. (1976).

En psychométrie, la qualité d'une mesure s'établit par le rapport entre la variance du score observé et la variance du score réel:

$$\text{Fiabilité} = \frac{\text{Variance}(X_o)}{\text{Variance}(X_R)}$$

Plus ce rapport se rapproche de l'unité et plus la mesure est considérée fiable (Guilford, 1954). Ce rapport reflète également le degré de cohérence interne de la mesure. En d'autres termes, il reflète le degré d'interdépendance entre les échelles de la mesure. Or il apparaît que ce rapport est d'autant moins mesurable que la valeur réelle du score. On ne peut par conséquent que l'estimer à partir des propriétés de la mesure. Ces propriétés ont été identifiées au début du siècle (voir Symonds, 1928).

Plusieurs estimateurs de ce rapport ont été suggérés par les psychométriciens. Churchill et Peter (1984), par exemple, en ont répertorié six parmi les plus populaires dont le coefficient de Spearman-Brown, celui de Cronbach et celui de Kuder-Richardson. D'entre tous ces estimateurs, le coefficient de fiabilité  $\alpha$  de Cronbach (1951) est jugé le plus important (Nunnally, 1967, page 96) et le plus utilisé non seulement en recherche sur les systèmes d'information mais ailleurs (Peter, 1979).

Le coefficient  $\alpha$  de Cronbach (1951) représente une interprétation du rapport  $X_o / X_R$  qui n'est pas nécessairement équivalente à celle offerte par Kuder et Richardson (1937) ou par Spearman et Brown (Brown, 1910; Spearman, 1910; Pedhazur et Schmelkin, 1991) par exemple. Le sens donné à ce coefficient est que sa racine carrée représente la corrélation entre  $X_o$  et  $X_R$  (Nunnally, 1967; pages 191-196).

### Le coefficient $\alpha$ de Cronbach

#### *Le coefficient $\alpha$ et ses composantes*

Mathématiquement, le coefficient  $\alpha$  est donné par la formule suivante:

$$\alpha = \frac{k \times r_{ij}}{1 + (k-1) \times r_{ij}}$$

où  $k$  est le nombre d'échelles dans la mesure et  $r_{ij}$  la moyenne des corrélations entre toutes les paires d'échelles ou l'inter-corrélation moyenne. Théoriquement, la valeur de  $\alpha$  ne peut dépasser l'unité et doit se situer entre -1 et 1. Elle peut cependant être négative ou largement supérieure à 1 compte tenu de la présence de  $r_{ij}$  qui, rappelons-le, est un coefficient de corrélation (Cronbach et Hartmann, 1954)<sup>1</sup>. Par ailleurs, une fiabilité trop élevée peut parfois être aussi psychométriquement indésirable qu'une fiabilité trop faible (Churchill et Peter, 1984; Peterson, 1994).

Cette formule peut être représentée graphiquement pour illustrer l'effet conjugué de ses deux composantes  $k$  et  $r_{ij}$ . Cette représentation est telle que montrée dans la figure 1.

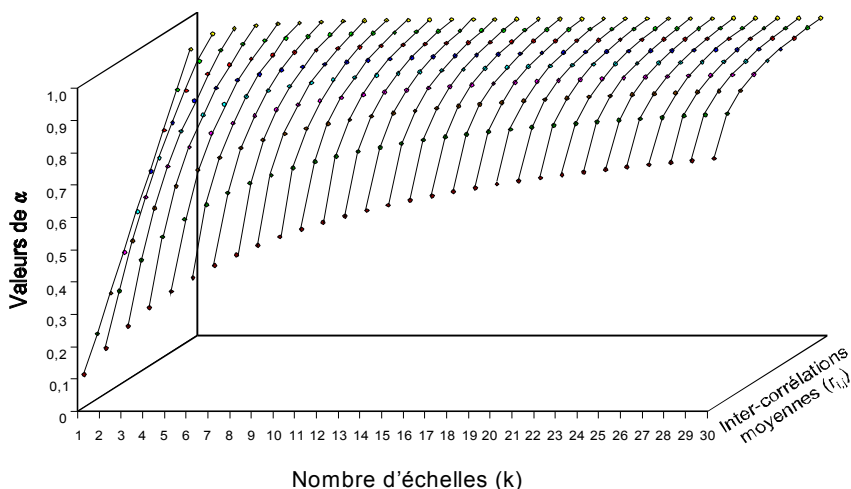
La figure confirme la croyance selon laquelle la valeur absolue du coefficient  $\alpha$  augmente avec le nombre  $k$  d'échelles composant la mesure<sup>2</sup>. Cette augmentation de  $\alpha$  est d'autant plus accentuée que l'inter-corrélation moyenne se rapproche de l'unité. Tout ensemble d'échelles peut assumer

<sup>1</sup> En fait, si  $r_{ij}$  est égal à -1 et  $k$  à 3,  $\alpha$  est égal à 3!

<sup>2</sup> Le terme "échelle" utilisé dans cet article se réfère à un ensemble d'indicateurs représentant chacun une question dans un test. Il est pris comme un synonyme du terme "dimension" et correspond à l'équivalent anglais "scale" tel que défini par Peter (1979, page 7). A l'exception de ce qui a trait aux corrélations, tout ce qui s'applique aux échelles, s'applique également à des indicateurs.

une valeur  $\alpha$  très proche de l'unité pourvu qu'il soit très élevé et que l'inter-corrélation moyenne entre les échelles ne soit ni négative ni nulle.

Notons que rares sont les chercheurs qui publient la valeur de l'inter-corrélation moyenne, la tendance est surtout de souligner les valeurs plus ou moins élevées du coefficient et le nombre d'échelles composant la mesure.



**Figure 1.** Représentation graphique de l'évolution de  $\alpha$  en fonction du nombre des échelles ( $k$ ) et de l'inter-corrélation moyenne des échelles ( $r_{i,j}$ )

Les valeurs acceptables du coefficient  $\alpha$  sont plus subjectives que scientifiques. Certains chercheurs se font forts d'invoquer les valeurs "acceptables" qui, selon Nunnally (1967, page 226), se situant entre 0,50 et 0,60 pour les recherches préliminaires et entre 0,80 et 0,90 pour les recherches de base; les valeurs supérieures à 0,90 sont de rigueur surtout dans le contexte des recherches appliquées<sup>3</sup>.

Par ailleurs, Van de Ven et Ferry (1988) offrent un exemple de valeurs espérées pour  $\alpha$  sous la forme d'un tableau que nous reproduisons ci-dessous.

Dans le tableau 1, Van de Ven et Ferry (page 79) montrent que contrairement à  $\alpha$ ,  $r_{i,j}$  est insensible au nombre d'échelles dans une mesure. Dans l'exemple de Van de Ven et Ferry, la valeur moyenne de  $r_{i,j}$  décroît à mesure que la mesure devient plus complexe, c'est-à-dire à mesure que le nombre de ses dimensions augmente. Par conséquent, les valeurs espérées de  $\alpha$  deviennent de moins en moins élevées et ce, bien que la valeur de  $k$  ait été fixée à 3 (voir le tableau 1).

**Tableau 1.** Intervalles espérés de la valeur de  $\alpha$  d'après Van de Ven et Ferry (1988, p. 79).

	Nombre de dimensions dans le construit		
	1	2	3 ou plus
1. Nombre d'échelles dans la mesure ( $k$ )	3	3	3
2. Inter-corrélation moyenne ( $r_{i,j}$ )	0,50-0,65	0,30-0,45	0,10-0,25
3. Valeurs espérées de $\alpha$	0,70-0,90	0,55-0,70	0,35-0,55

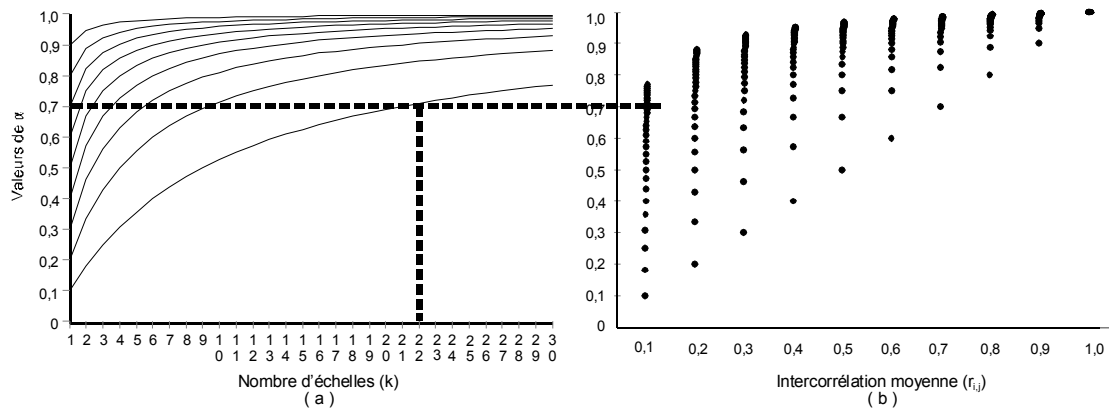
<sup>3</sup> Incidemment, ces valeurs changèrent dans l'édition parue en 1978: 0,70 pour les recherches préliminaires, 0,80 pour les recherches de base et 0,90-0,95 pour les recherches appliquées (Nunnally, 1978, pages 245-246).

Ainsi,  $\alpha$  est fonction non seulement du nombre d'échelles dans la mesure mais aussi de la nature du construit mesuré. Si le construit est simple (ou unidimensionnel), certaines valeurs s'appliquent. Si le construit est complexe (ou multidimensionnel), d'autres valeurs sont indiquées.

Donc, les chercheurs se fient davantage à la valeur numérique du coefficient qu'ils obtiennent qu'à la nature du construit ou à la nature de la relation qui existe entre ses composantes et publient essentiellement des données concernant leurs mesures qui ont davantage trait à la quantité (le nombre d'échelles) qu'à la qualité (la cohérence interne ou multidimensionalité).

Par ailleurs, et si l'on examine la figure 1 de deux angles différents, tel que cela est proposé dans la figure 2, plusieurs conclusions s'imposent:

- Le nombre d'échelles dans la mesure n'a pas toujours la même influence sur les valeurs de  $\alpha$  (Figure 2a) puisque celle-ci dépend également de l'inter-corrélation moyenne entre ces échelles;
- A chaque valeur de  $r_{ij}$  correspondent plusieurs valeurs possibles de  $\alpha$  dépendant du nombre d'échelles présentes dans la mesure (Figure 2b);
- L'influence du nombre d'échelles est plus marquée dans le cas des valeurs faibles de  $r_{ij}$ ; par contre, si une mesure est composée de suffisamment d'échelles, elle peut atteindre un  $\alpha$  de 0,70 même si l'inter-corrélation moyenne n'est égale qu'à 0,1 (Figures 2a et 2b);
- Une mesure dont les échelles ont une inter-corrélation moyenne de 0,1 peut atteindre des valeurs de  $\alpha$  supérieures à 0,75 si le nombre de ces échelles est supérieur à 27;
- Une mesure dont les échelles ont une inter-corrélation moyenne de 0,5 est assurée d'avoir un coefficient  $\alpha$  supérieur à 0,50 et ce, quelque soit le nombre de ces échelles pourvu qu'il soit supérieur à 1.



**Figure 2.** Les valeurs de  $\alpha$  en fonction (a) du nombre d'échelles ( $k$ ) et (b) de l'inter-corrélation moyenne ( $r_{ij}$ )

Ces implications sont importantes parce qu'elles montrent le peu d'informations qu'un coefficient  $\alpha$  révèle sur la qualité psychométrique d'une mesure. Qu'une mesure satisfasse les critères de Nunnally n'est pas suffisant pour qu'elle soit adoptée comme une mesure fiable, et surtout pas comme une mesure homogène. Les points c et d, tels qu'énoncés ci-dessus, sont particulièrement importants vu la gravité de leurs implications: c'est l'inter-corrélation moyenne qui est dépendante du construit et le nombre d'échelles qui est une caractéristique de la mesure.

En résumé, un  $\alpha$  de valeur relativement élevée indique seulement qu'il existe très peu de variance dans la mesure qui soit spécifique aux échelles individuelles. Donc,  $\alpha$  n'est informatif que lorsqu'il existe un facteur commun unique dans la mesure (Cotton, Campbell et Malone, 1957). S'il existe un facteur commun unique dans la mesure, c'est-à-dire lorsque la mesure est unidimensionnelle, alors  $\alpha$  est une mesure de fiabilité de ce facteur.

## *L'homogénéité et la cohérence interne des mesures*

Une mesure est homogène lorsqu'elle n'est composée que d'une seule dimension. L'homogénéité d'une mesure peut être indiquée par l'inter-corrélation moyenne quoiqu'un ensemble d'échelles peut être relativement inter-corrélé et multidimensionnel (Green, Lissitz et Mulaik, 1977). Le coefficient  $\alpha$ , quand il est considéré dans sa totalité, n'a jamais été un indicateur de l'homogénéité d'une mesure (Cronbach, 1951; Green et al., 1977) et encore moins un indicateur de l'unidimensionalité d'un construit. Plus problématique encore est le fait que certains chercheurs en systèmes d'information (Subramanian et Nilakanta, 1994) continuent à confondre les deux concepts de la cohérence interne et celui de l'homogénéité et définissent l'un à l'aide de l'autre:

“Internal Consistency. The degree of homogeneity or cohesiveness between items that purport to measure the same construct.” (Subramanian et Nilakanta, 1994, page 15).

Pour Cortina (1993), la distinction entre cohérence interne et homogénéité est loin d'être subtile:

“[Alpha] is a function of the extent to which items in a test have high commonalities and thus low uniqueness. It is also a function of interrelatedness, although one must remember that this does not imply unidimensionality or homogeneity” (page 100).

Le terme “cohérence interne” décrit l'interdépendance élevée qui existe entre les échelles d'une mesure alors que le terme “homogénéité” décrit une situation où un ensemble d'échelles mesurent toutes la même dimension d'un même construit. Si le construit est unidimensionnel (homogène), alors ces échelles mesurent le construit (Crano et Brewer, 1973).

En outre, l'homogénéité d'une mesure implique automatiquement sa cohérence interne mais la cohérence interne d'une mesure, si élevée soit-elle, n'implique pas nécessairement son homogénéité (Crane et Brewer, 1973).

Il existe plusieurs méthodes pour déterminer si une mesure est hétérogène. La première se présente sous la forme de l'inter-corrélation moyenne ( $r_{ij}$ ). Si, par exemple, l'inter-corrélation moyenne des échelles est égale à 0,52, cela veut dire que les échelles mesurent 27% (le carré de 0,52) de *quelque chose* qui leur est commun. Restent 73% qui mesurent des aspects divers qui ont peu à voir avec ce *quelque chose*. Mais, rappelons-le, l'inter-corrélation moyenne, même relativement élevée, n'est pas une condition suffisante pour conclure à l'unidimensionalité.

La procédure de l'analyse de fiabilité, telle que définie dans le programme statistique SPSS<sub>x</sub> (Norusis, 1988) par exemple, restitue quelques indices utiles quant à la possibilité qu'un ensemble d'échelles soit hétérogène. Cette procédure doit être la première analyse entreprise par un chercheur afin d'épurer ses mesures (Churchill, 1979). Si une proportion relativement importante des échelles se porte candidate pour l'élimination afin d'augmenter la fiabilité de la mesure, cette mesure est susceptible d'être hétérogène. La mesure doit être épurée de ces échelles. Les échelles ainsi épurées doivent alors être re-évaluées en aparté pour explorer la possibilité qu'elles puissent former une autre dimension de la mesure. C'est la méthode qui semble être la plus populaire dans le domaine des systèmes d'information. Rares toutefois sont les chercheurs qui essayent d'incorporer les échelles éliminées dans une nouvelle dimension de la mesure.

Dans le cas où aucune échelle ne se porte candidate pour élimination, il n'y a pas plus d'assurance que la mesure soit unidimensionnelle. Cortina (1993) offre une statistique qui mesure le degré de précision du coefficient  $\alpha$  obtenu. Cette statistique,  $\tau$ , se calcule en fonction du nombre  $k$  d'échelles présentes dans la mesure et de l'écart-type  $ET_r$  correspondant à l'inter-corrélation moyenne  $r_{ij}$ :

$$\tau = \frac{ET_r}{\sqrt{\frac{k}{2} \times (k-1) - 1}}$$

La statistique  $\tau$  est plus indicative que déterministe; plus sa valeur est élevée, plus l'hétérogénéité de la mesure est indiquée; en d'autres termes, plus  $\tau$  se rapproche de 0, plus la mesure a de chances d'être unidimensionnelle (ou homogène).

Une deuxième méthode, plus populaire, est l'utilisation de l'analyse factorielle (Pedhazur et Schmelkin, 1991; Thurstone, 1947) pour détecter plusieurs facteurs au sein d'un ensemble d'échelles à condition que certaines conditions soient réalisées quant au ratio du nombre d'échelles par rapport au nombre d'observations (Hair, Anderson, Tatham et Grablovsky, 1979).

Ce ratio est souvent ignoré. Wan et Wah (1990), par exemple, ont pu publier les résultats d'une analyse factorielle sur 64 observations en utilisant 33 échelles, un ratio de 1,94/1 bien inférieur à celui recommandé par Kerlinger (10/1; 1973) ou même à celui proposé par Hair et al. (5/1; 1979) et dégager huit facteurs.

Dans le cas où la mesure est unidimensionnelle, l'analyse factorielle avec composantes principales doit révéler un premier facteur dont la contribution à la variance totale de la mesure est au moins six fois plus importante que celle du deuxième facteur (Bynner, 1988).

Les deux méthodes du degré de précision et de l'analyse factorielle ne sont malheureusement que des manipulations de variances et de moyennes. Quoique Nunnally (1967) considère que l'analyse factorielle constitue l'essence même de la mesure des construits (page 101) et bien que Kerlinger (1973) et Pedhazur et Schmelkin (1991) lui attribuent une utilité particulière dans le processus de validation des construits, il est important de garder à l'esprit le fait qu'en l'absence d'une théorie préalable sur la multidimensionalité d'un construit, un chercheur peut l'utiliser pour dériver le nombre et le type de facteurs qu'il désire. Souvent, cette procédure restitue des facteurs purement imaginaires inter-reliés uniquement par la magie des nombres. Rares sont les chercheurs qui suivent le conseil de Harris (1967) qui recommande l'utilisation de plusieurs techniques d'extraction de facteurs et non pas uniquement les composantes principales afin de neutraliser "l'effet de méthode". Seule la structure factorielle qui est commune à toutes les extractions sera jugée stable et robuste puisque indépendante de la méthode d'extraction (Tabachnick et Fidel, 1989). Par contre, et dans le cas où le chercheur est guidé par la théorie en ce qui concerne la nature des facteurs à dégager, l'analyse factorielle confirmatoire est plus indiquée (Pedhazur et Schmelkin, 1991; Renterberg et Gustafsson, 1992).

#### *Le coefficient de fiabilité d'une mesure multidimensionnelle*

Si plusieurs dimensions sont présentes dans une mesure, soit par déduction (appuyée par les données en utilisant une des techniques résumées ci-dessus) ou par induction (supportée par la théorie), l'application de la formule du coefficient  $\alpha$  n'est plus valable. Le coefficient mesurant le degré de cohérence interne d'une mesure ne peut plus être appliqué dans le cas d'une mesure qui n'est pas homogène.

Dans le cas d'une mesure hétérogène, la formule donnée par Nunnally (1967; pp. 226-235) est plus appropriée:

$$\alpha = \left| 1 - \frac{\sum \sigma_i^2 - \sum \alpha_i \sigma_i^2}{\sigma^2} \right|$$

où  $\alpha$  est le nouveau coefficient, que nous appellerons le "coefficient de Nunnally" dans le reste de cet article, appliqué dans le cas d'une mesure hétérogène (multidimensionnelle),  $\sigma_i$  est l'écart-type associé au facteur  $i$ ,  $\sigma$  est l'écart-type de la combinaison linéaire des facteurs composant la mesure et  $\alpha_i$  le coefficient de Cronbach correspondant au facteur  $i$ .

Il va de soi que si une mesure est décomposée en plusieurs facteurs, chaque facteur aura un nombre d'échelles strictement inférieur à celui de la totalité de la mesure. Par ailleurs, et par essence même, les facteurs auront une homogénéité supérieure à celle de la totalité de la mesure.



Est-ce que ces facteurs ont une cohérence interne supérieure à celle de la mesure dont ils font partie? En d'autres termes, quels seront les coefficients  $\alpha$  des facteurs comparativement au coefficient  $\alpha$  de la totalité de la mesure?

Une ébauche de réponse a été proposée par Churchill et Peter (1984) qui ont testé l'hypothèse que le coefficient  $\alpha$  de Nunnally d'une mesure hétérogène est supérieur au coefficient  $\alpha$  de Cronbach. Le test de cette hypothèse a été négatif. Il est néanmoins regrettable que Churchill et Peter aient testé leur hypothèse non pas sur la base des construits mais sur la base de paramètres  $k$ ,  $r_{ij}$  et  $\alpha_i$  sans distinction aucune entre les différents construits sous l'étude.

### Un exemple: la satisfaction de l'utilisateur

La satisfaction de l'utilisateur (SU) avec un système informatique est considérée comme l'un des construits les plus importants parce qu'il constitue la variable dépendante dans nombre de cadres théoriques et de recherches en systèmes d'information (Bailey et Pearson, 1983; Galletta et Lederer, 1989; Hamilton et Chervany, 1981; Ives, Olson et Baroudi, 1983; Powers et Dickson, 1973).

Plusieurs instruments ont été développés pour mesurer le construit dont celui de Gallagher (1974), Larcker et Lessig (1980), Bailey et Pearson (1983), Ivarson (1987) et Doll et Torkzadeh (1988). De ces instruments, celui initialement développé par Pearson (1977; Pearson et Bailey, 1979, Bailey et Pearson, 1983) est le plus répandu et le plus utilisé dans la recherche dans ce domaine.

L'instrument original était composé de 39 échelles composées chacune de quatre indicateurs pour un total de 156 indicateurs exprimés sous forme de questions. Le questionnaire a été repris par Ives et al. (1983) afin d'affiner ses propriétés psychométriques.

Le résultat de l'effort de Ives et al. (1983) fut une version abrégée de l'instrument qui ne contenait plus que 13 échelles avec deux questions chacune (voir le tableau 2). Il n'empêche que les échelles originales ont été utilisées sous différentes formes par plusieurs autres chercheurs. La version à 13 échelles, que nous appellerons SU13, fut reprise par Baroudi et Orlikowski (1988) et soumise à d'autres tests statistiques et psychométriques. Galletta et Lederer (1989) utilisèrent cette même version dans leur analyse sur la fiabilité test-retest de SU13.

**Tableau 2.** La version à 13 échelles de Ives et al. (1983)<sup>1</sup>.

1. Relations avec le personnel informatique
2. Réponses aux demandes de modification du système
14. Formation informatique dispensée à l'utilisateur
15. Compréhension du système par l'utilisateur
16. Participation de l'utilisateur au développement du système
18. Attitude du personnel informatique
19. Fiabilité de l'information produite par le système
25. Pertinence de l'information produite par le système
28. Véracité de l'information produite par le système
29. Précision de l'information produite par le système
30. Communication avec le personnel informatique
32. Temps requis pour développer de nouvelles applications
36. Exhaustivité de l'information produite par le système

<sup>1</sup>. La numérotation des échelles est celle utilisée par Ives et al. (1983). Nous l'avons gardée pour faciliter le repérage et la comparaison des échelles entre les différentes versions du questionnaire.

Les différentes versions de SU avaient atteint des niveaux de fiabilité élevés: 0,97 pour la version de Ives et al. (1983,  $k=39$ ,  $n=200$ ), 0,91 pour la version de Raymond (1985; 1987,  $k=20$ ,  $n=464$ ), 0,97 pour la version de Tait et Vessey (1988,  $n=42$ ) et 0,89 pour la version de Baroudi et Orlikowski (1988,  $k=13$ ,  $n=358$ ). Voir le tableau 3.

**Tableau 3.** Quelques versions de SU13 et leurs coefficients de fiabilité  $\alpha$ .

	Bailey et Pearson (1983)	Ives et al. (1983)	Treacy (1985)	Raymond (1985, 1987)	Barki et Huff (1985, 1990)	Tait and Vessey (1988)	Baroudi et Orlikowski (1988)	Montazemi (1988)	Baronas et Louis (1988)	Wan et Wah (1990)	Kappelman et McLean (1991)
Nombre d'échelles (k)	39	39	10/25 <sup>2</sup>	20	26	11	13	35	9	26	17
Fiabilité totale de la mesure ( $\alpha$ )	nr	0,97	nr	0,91	0,96	0,97	0,89	0,94	nr	0,96	0,90
Intervalle des coefficients de fiabilité des échelles individuelles	.75-.. <sup>1</sup>	.82-.97	.36-.75	.81-.96	nr	.78-.98	.84-.94	.88-.98	nr	nr	nr
Taille de l'échantillon (n)	29	200	297	464	42	42	358	77	92	64	146
Nombre de facteurs (s'il y a lieu)	na	5	3	4	1	na	3	1	nr	8	nr
Fiabilité des facteurs publiée (Oui/Non)	Non	Non	Oui	Non	Non	Non	Oui	Non	Non	Non	Non

<sup>1</sup>. Seuls les coefficients de fiabilité minimum et moyens ont été publiés. Le coefficient de fiabilité moyen était égal à 0,93 (Bailey et Pearson, 1983, page 535). Des 39 échelles, 32 avaient une fiabilité supérieure à 0,90.

<sup>2</sup>. Dix questions avaient été adressées à 68 agents intermédiaires (Treacy, 1985) et 25 questions avaient été adressées à 142 décideurs. Il n'y eut que 87 personnes qui répondirent à la totalité du questionnaire composé de 21 questions.

na: non applicable ou non effectué

nr: non reporté

Presque chacune des versions de SU a également été assujettie à une analyse factorielle. Les deux tentatives de Ives et al. (1983) ont généré cinq et trois facteurs avec la version des 39 échelles, Raymond (1985; 1987) a dégagé quatre facteurs mais seulement à partir des 14 échelles des 20 qu'il avait utilisées et Baroudi et Orlikowski (1988) ont dégagé les mêmes trois facteurs préalablement obtenus par Ives et al. (1983) et qui sont donnés dans le tableau 4.

Nous noterons qu'à l'exception de l'échelle no. 2 (Réponses aux demandes...), les mêmes échelles sont stables et se retrouvent dans les deux structures. Ces échelles sont identifiées dans la dernière partie du tableau 4. Notons par ailleurs que certaines échelles se retrouvent groupées dans les mêmes facteurs et ce, quelque soit le nombre d'échelles de départ, le nombre de facteurs obtenus et le chercheur effectuant l'analyse ou l'échantillon.

En fait, ce n'est pas tant la nature ni le nombre de facteurs qui importent dans le cadre de la fiabilité des mesures, mais bien l'existence de ces facteurs. C'est surtout lorsqu'une mesure manifeste cette propriété d'une manière cohérente et répétitive que nous sommes forcés de sérieusement considérer l'hypothèse de son hétérogénéité.

**Tableau 4.** Les deux structures factorielles obtenues à partir des 13 échelles de SU13.

La structure factorielle obtenue par Ives et al. (1983)

Facteur 1	Facteur 2	Facteur 3
1. Relations avec le personnel ...	14. Formation ...	19. Fiabilité de l'information
2. Réponses aux demandes ...	15. Compréhension ...	25. Pertinence de l'information
18. Attitude du personnel informatique	16. Participation ...	28. Véracité de l'information
30. Communication avec le personnel		29. Précision de l'information
32. Temps requis pour développer ...		36. Exhaustivité de l'information

La structure factorielle obtenue par Baroudi et Orlikowski (1988)

Facteur 1	Facteur 2	Facteur 3
1. Relations avec le personnel ...	2. Réponses aux demandes ...	19. Fiabilité de l'information
2. Réponses aux demandes ...	14. Formation ...	25. Pertinence de l'information
18. Attitude du personnel informatique	15. Compréhension ...	28. Véracité de l'information
30. Communication avec le personnel	16. Participation ...	29. Précision de l'information
32. Temps requis pour développer ...		36. Exhaustivité de l'information

Les facteurs stables

Facteur 1	Facteur 2	Facteur 3
1. Relations avec le personnel ...	14. Formation ...	19. Fiabilité de l'information
18. Attitude du personnel informatique	15. Compréhension ...	25. Pertinence de l'information
30. Communication avec le personnel	16. Participation ...	28. Véracité de l'information
		29. Précision de l'information

Le construit de la satisfaction de l'utilisateur a donc constamment été identifié comme un construit hétérogène (multidimensionnel) à travers les différentes versions de la mesure SU. Kappelman et McLean (1991), par exemple, ne purent dégager une structure factorielle unidimensionnelle (page 342) d'une version abrégée composée de 17 échelles.

Il n'y a donc pas de justification, au sens strictement psychométrique du terme, à calculer la fiabilité d'une mesure manifestement multidimensionnelle en utilisant la formule de Cronbach.

Ainsi, les coefficients publiés par Ives et al. (1983), par exemple, ne reflètent en rien la fiabilité ou la cohérence interne de la mesure toute entière puisque celle-ci s'est révélée hétérogène. Le coefficient de 0,96 de Wan et Wah (1990) (voir le tableau 3) n'est pas plus valide que leur structure factorielle. Pourtant, certains chercheurs, notamment Hawk et Raju (1991), pensent précisément le contraire:

"use of Cronbach's alpha as an indicator of the overall reliability of the UIS instrument is not a serious problem" (p.1168).

Ces auteurs semblent donc considérer que ne pas utiliser la formule appropriée lorsque la mesure se révèle multidimensionnelle est sans conséquences sérieuses.

Que la fiabilité d'une mesure dans sa totalité soit supérieure à celle de n'importe quel des facteurs qui la composent est probable lorsque l'analyse statistique de la fiabilité suggère un ensemble d'échelles sans amélioration possible.

Par contre, il est concevable qu'au moins un facteur ait une fiabilité supérieure à la totalité de la mesure lorsque plusieurs échelles sont éliminées pour former un facteur séparé puisque la mesure est ainsi considérée "épurée" (Churchill, 1979). Mais souvent, l'on se retrouve dans des situations moins claires.

Dans la section qui suit, nous tentons d'illustrer les différences qui peuvent se présenter avec des données réelles. Pour ce faire, nous avons utilisé les données de El Louadi, Galletta et Sampler (1998) concernant le degré de satisfaction de l'utilisateur avec un système d'information.

#### *Un exemple empirique*

Les données concernent 68 utilisateurs de systèmes tels que décrits par El Louadi et al. (1998). La satisfaction de l'utilisateur avait été mesurée avec la version à 13 échelles de SU13 telle que dégagée par Ives et al. (1983) et validée par Baroudi et Orlikowski (1988).

Nous avons commencé par calculer le coefficient de fiabilité des 13 échelles. Le coefficient  $\alpha$  obtenu était égal à 0,93 avec un degré de précision (Cortina, 1993) de 0,017 et une inter-corrélation de 0,52.

Quoique le degré de précision était assez proche de 0 pour présumer de l'homogénéité de la mesure et que l'analyse de fiabilité ne présentait aucune amélioration possible, le fait que l'inter-corrélation moyenne des 13 échelles n'était égale qu'à 0,52, nous incita à entreprendre une analyse factorielle afin de tester l'unidimensionalité de la mesure. Pour ce faire, nous avons suivi les recommandations de Harris (1967) et Tabachnick et Fidel (1989) en utilisant la rotation varimax et plusieurs méthodes d'extraction. Les résultats de cette analyse sont reproduits dans le tableau 5. Ces résultats montrent que c'est à un nombre de facteurs égal à 2 que la même structure émerge dans toutes les méthodes d'extraction utilisées. Cette structure factorielle n'est donc pas un artifice de la méthode d'extraction.

**Tableau 5.** Analyses factorielles sur SU13 en utilisant plusieurs méthodes d'extraction.

Échelles	PC		Alpha		GLS		Image		ML		PAF	
	1	2	1	2	1	2	1	2	1	2	1	2
1. Relation avec le personnel ...	<b>0,87</b>	0,13	<b>0,84</b>	0,14	<b>0,86</b>	0,15	<b>0,84</b>	0,15	<b>0,86</b>	0,14	<b>0,85</b>	0,14
30. Communication avec le...	<b>0,87</b>	0,26	<b>0,86</b>	0,26	<b>0,89</b>	0,28	<b>0,86</b>	0,28	<b>0,88</b>	0,27	<b>0,86</b>	0,26
16. Participation ...	<b>0,85</b>	0,19	<b>0,82</b>	0,20	<b>0,80</b>	0,22	<b>0,79</b>	0,22	<b>0,80</b>	0,21	<b>0,82</b>	0,20
15. Compréhension ...	<b>0,85</b>	0,16	<b>0,82</b>	0,18	<b>0,83</b>	0,15	<b>0,79</b>	0,17	<b>0,79</b>	0,17	<b>0,81</b>	0,17
18. Attitude du personnel...	<b>0,83</b>	0,23	<b>0,80</b>	0,23	<b>0,84</b>	0,24	<b>0,82</b>	0,24	<b>0,84</b>	0,23	<b>0,80</b>	0,23
32. Temps requis pour développer	<b>0,71</b>	0,30	<b>0,67</b>	0,31	<b>0,65</b>	0,29	<b>0,64</b>	0,29	<b>0,64</b>	0,30	<b>0,68</b>	0,45
14. Formation ...	<b>0,70</b>	0,46	<b>0,69</b>	0,45	<b>0,67</b>	0,46	<b>0,66</b>	0,46	<b>0,66</b>	0,45	<b>0,66</b>	0,31
2. Réponses aux demandes ...	<b>0,59</b>	0,46	<b>0,58</b>	0,43	<b>0,54</b>	0,46	<b>0,55</b>	0,46	<b>0,55</b>	0,44	<b>0,57</b>	0,44
29. Précision de l'information	0,15	<b>0,93</b>	0,16	<b>0,94</b>	0,15	<b>0,93</b>	0,17	<b>0,89</b>	0,16	<b>0,92</b>	0,16	<b>0,93</b>
28. Véracité de l'information	0,15	<b>0,92</b>	0,16	<b>0,91</b>	0,15	<b>0,95</b>	0,17	<b>0,89</b>	0,17	<b>0,92</b>	0,16	<b>0,91</b>
36. Exhaustivité de l'information	0,21	<b>0,84</b>	0,25	<b>0,77</b>	0,23	<b>0,77</b>	0,23	<b>0,76</b>	0,23	<b>0,78</b>	0,24	<b>0,77</b>
19. Fiabilité de l'information	0,39	<b>0,80</b>	0,40	<b>0,76</b>	0,38	<b>0,79</b>	0,39	<b>0,76</b>	0,39	<b>0,77</b>	0,40	<b>0,76</b>
25. Pertinence de l'information	0,51	<b>0,59</b>	0,50	<b>0,54</b>	0,51	<b>0,54</b>	0,50	<b>0,54</b>	0,49	<b>0,55</b>	0,50	<b>0,54</b>
<b>Valeurs propres</b>	<b>7,6</b>	<b>2,0</b>	<b>7,3</b>	<b>1,8</b>	<b>7,3</b>	<b>1,9</b>	<b>7,1</b>	<b>1,7</b>	<b>7,1</b>	<b>1,9</b>	<b>7,3</b>	<b>1,8</b>
<b>% Variance</b>	<b>58,3</b>	<b>15,5</b>	<b>55,9</b>	<b>13,9</b>	<b>56,2</b>	<b>14,8</b>	<b>54,6</b>	<b>12,8</b>	<b>55,0</b>	<b>14,6</b>	<b>56,0</b>	<b>13,8</b>
<b>% de la variance totale expliquée</b>	<b>58,3</b>	<b>73,8</b>	<b>55,9</b>	<b>60,8</b>	<b>56,2</b>	<b>71,0</b>	<b>54,6</b>	<b>67,4</b>	<b>55,0</b>	<b>69,6</b>	<b>56,0</b>	<b>69,8</b>

**Légende:**

PC: Principal Components  
 Alpha: Alpha Factoring  
 GLS: Generalized Least Squares  
 Image: Image Factoring  
 ML: Maximum Likelihood  
 PAF: Principal Axis Factoring

Nous avons ensuite calculé le coefficient de fiabilité de chacun de ces deux facteurs. Les deux coefficients obtenus étaient égaux à 0,92 et 0,91 respectivement avec des inter-corrélations moyennes égales à 0,65 et 0,67 ( $\tau_1=0,016$  et  $\tau_2=0,039$ ). Basés sur ces résultats, nous avons entrepris de calculer le coefficient de fiabilité de toute la mesure en fonction de ceux des deux facteurs en utilisant la formule de Nunnally (1967). Le résultat obtenu était:  $\alpha=0,81$  ( $\sigma_1^2=1,287$ ,  $\sigma_2^2=0,707$ , et  $\sigma^2=0,854$ ). Voir le tableau 6.

**Tableau 6.** Différences entre le coefficient de Cronbach et celui de Nunnally.

	Toute la mesure	Premier facteur	Deuxième facteur
Nombre d'échelles (k)	13	8	5
Variance ( $\sigma^2$ )	0,854	1,287	0,707
Inter-corrélation moyenne ( $r_{ij}$ )	0,52	0,65	0,67
$\sigma_r$	0,0221	0,0068	0,0143
Précision ( $\tau$ )	0,017	0,016	0,039
$\alpha$ de Cronbach	<b>0,93</b>	<b>0,92</b>	<b>0,91</b>
Coefficient de Nunnally	-	<b>0,81</b>	

Ainsi, quand une mesure est multidimensionnelle, le coefficient  $\alpha$  de Nunnally de toute la mesure peut être sensiblement inférieur au coefficient  $\alpha$  de Cronbach de toute la mesure ( $\alpha=0,81 < \alpha=0,93$ ).

Par contre, les coefficients de fiabilité respectifs des facteurs (0,92 et 0,91) sont très comparables mais légèrement inférieurs à celui de toute la mesure.

### Discussion

Au moins dans le cas de la mesure multidimensionnelle de la satisfaction de l'utilisateur avec un système d'information, le coefficient de fiabilité le plus élevé est celui qui est calculé au niveau de toute la mesure comme si elle était unidimensionnelle. Par contre, le coefficient de fiabilité le plus faible est celui qui est calculé en tenant compte de ses dimensions. Mais quoique plus faible en valeur ( $\alpha=0,81$ ), ce coefficient est le plus scientifiquement correct. En fait, malgré sa valeur plus élevée, le coefficient obtenu à partir de toute la mesure ( $\alpha=0,93$ ) n'a aucune valeur informative quant à la fiabilité de la mesure étant donné qu'elle n'est pas unidimensionnelle (homogène).

Une des raisons probables pour lesquelles les chercheurs continuent à préférer le traitement qui leur procure le coefficient le plus élevé est qu'une mesure est composée d'un plus grand nombre d'échelles et garantit donc une valeur de fiabilité plus élevée.

Quoique ne s'étant basé que sur un seul exemple, notre analyse indique bien que, souvent, les chercheurs sous-estiment l'importance de l'évaluation psychométrique de leurs mesures. Quand ils s'en préoccupent, ils se fient davantage à des formules et à des recettes telles que celle relative à la relation entre le nombre des échelles d'une mesure et la fiabilité de cette mesure et ce, même s'il n'est pas dit qu'augmenter le nombre des échelles améliore systématiquement la fiabilité d'une mesure.

Cette recette est poursuivie aux dépens de l'autre critère affectant la fiabilité d'une mesure et qui est son homogénéité.

Dans ce contexte, le chercheur soucieux de rigueur et de conformité avec les canons d'une science guidée gagnerait à n'utiliser la formule de Cronbach que dans le cas d'une mesure unidimensionnelle, c'est-à-dire lorsque le chercheur est convaincu que le construit est

unidimensionnel et que cette unidimensionalité est présente dans les données qu'il a recueillies. Il est dérisoire d'appliquer le coefficient de Cronbach lorsque le construit mesuré est, au départ, multidimensionnel. Le coefficient  $\alpha$  de la mesure de ce construit sera manifestement plus élevé que le coefficient  $\alpha$  de Nunnally qui est plus approprié. L'exemple que nous avons présenté l'illustre bien. Un chercheur désireux de faire valoir ses mesures a toujours le recours de les épurer des échelles inutiles ou de les fractionner en autant de facteurs que la théorie de ce construit l'indique. Dans ce cas, le coefficient  $\alpha$  de Cronbach s'applique sur les facteurs et le coefficient  $\alpha$  de Nunnally s'applique sur la combinaison linéaire de ces facteurs, c'est-à-dire sur la totalité de la mesure. Certes le risque demeure que le coefficient de la combinaison linéaire soit inférieur en valeur au coefficient  $\alpha$  appliqué à toute la mesure, comme cela a été démontré dans notre exemple, mais il sera plus correct.

### Conclusion

L'utilisation de mesures fiables est un pré-requis à la conduite et à la poursuite d'une recherche scientifique rigoureuse. Dans une discipline aussi jeune que les systèmes d'information, il importe que les outils (mesures) utilisés soient testés le plus rigoureusement possible. Le coût additionnel de la rigueur que nous préconisons est minime au regard de la perte de crédibilité scientifique dont Keen (1980) avait pourtant parlé il y a plus de vingt ans. Contrairement à d'autres disciplines qui jouissent d'une tradition de recherche souvent séculaire, de définitions consensuelles des construits et des concepts, d'une panoplie importante de revues et de conférences, d'un héritage en termes de théorèmes et d'axiomes unanimement acceptés, la discipline des systèmes d'information est naissante, a des spécificités plus pragmatiques que théoriques et emploie des cadres de recherche et des variables souvent empruntés à d'autres disciplines dites "disciplines de référence" (Keen, 1980).

Il importe désormais que chercheurs, évaluateurs et responsables de revues scientifiques soient plus vigilants car l'inflation de la valeur des coefficients ne sert pas l'avancement de la science. Il importe qu'une discipline n'ayant pas encore une tradition théorique cumulative aspire à un minimum de rigueur. Si certains, comme Hawk et Raju, pensent que l'utilisation d'un coefficient plutôt qu'un autre n'est vraiment pas un problème, que coûte-t-il vraiment d'utiliser le bon?

Cet article a trois limitations. La première est qu'il s'est basé sur une seule mesure et sur une seule collecte de données pour illustrer son propos. La deuxième est qu'il s'est contenté de traiter de la question de la fiabilité aux dépens d'autres aspects aussi importants, tels que la validité des construits. Le domaine des systèmes d'information souffre également de ces problèmes tout aussi, sinon plus importants. La troisième c'est qu'il ne traite pas des autres outils qui, tels que l'analyse factorielle confirmatoire et les équations structurelles apportent, elles aussi, leurs lots de promesses qui pourraient être utilisées par les chercheurs dans leur effort de validation de leurs mesures.

## Références

- Bailey, J.E. et Pearson, S.W., Development of a Tool for Measuring and Analyzing Computer User Satisfaction, *Management Science*, Vol.29, N°5, Mai 1983, 530-545.
- Barki, H. et Huff, S.L, Change, Attitude to Change, and Decision Support System Success, *Information & Management*, Vol.9, 1985, 261-268.
- Barki, H. et Huff, S.L., Implementing Decision Support Systems: Correlates of User Satisfaction and System Usage, *Infor*, Vol.28, 1990, 89-101.
- Baronas, A-M. K. et Louis, M.R., Restoring a Sense of Control During Implementation: How User Involvement Leads to System Acceptance, *MIS Quarterly*, Vol.12, No.1, Mars 1988, 111-124.
- Baroudi, J.J. et Orlikowski, W.J., A Short-form Measure of User Information Satisfaction: A Psychometric Evaluation and Notes on Use, *Journal of Management Information Systems*, Vol.4, 1988, 44-59.
- Brown, W., Some Experimental Results in the Correlation of Mental Abilities, *British Journal of Psychology*, Vol.3, 1910, 296-322.
- Bynner, J., Factor Analysis and the Construct Indicator Relationship, *Human Relations*, Vol.41, No.5, 1988, 389-405.
- Churchill, G.A., A Paradigm for Developing better Measures of Marketing Constructs, *Journal of Marketing Research*, Vol.16, No.1, 1979, 64-73.
- Churchill, G.A. et Peter, J.P., Research Design Effects on the Reliability of Rating Scales: A Meta-Analysis, *Journal of Marketing Research*, Vol.21, Novembre 1984, 360-375.
- Cortina, J.M., What is Coefficient Alpha? An Examination of Theory and Applications, *Journal of Applied Psychology*, Vol.78, No.1, Février 1993, 98-104.
- Cotton, J.W., Campbell, D.J., et Malone, R.D., The Relationship between Factorial Composition of Test Items and Measures of Test Reliability, *Psychometrika*, Vol. 22, 1957, 347-358.
- Crano, W.D. et Brewer, M.B., Principles of Research in Social Psychology, McGraw-Hill, New York, 1973.
- Cronbach, L.J., Coefficient Alpha and Internal Structure of Tests, *Psychometrika*, Vol.16, 1951, 297-334.
- Cronbach, L.J. et Hartmann, W., A Note on Negative Reliabilities, *Educational and Psychological Measurement*, Vol.14, 1954.
- Doll, W.J. et Torkzadeh, G., The Measurement of End-user Computing Satisfaction, *MIS Quarterly*, Vol.12, 1988, 259-274.
- El Louadi, M., On the Appropriate Assessment of Reliability and Construct Validity in IS Research: The Case of the User Information Satisfaction Instrument, *Proceedings of the Academy of Business Administration*, Reno, Nevada, 1995, 570-578.



- El Louadi, M., Galletta, D.F. et Sampler, J. An empirical validation of a contingency model for information requirements determination, *The Data Base for Advances in Information Systems*, Vol.29, N°3, 1998, 31-51.
- Fornell, C. A National Customer Satisfaction Boremeter: The Swedish Expérience, *Journal of Marketing*, Vol. 56, No.1, 1992, 6-12.
- Flanagan, J.C., *Factor Analysis in the Study of Personalty*, Stanford University Press, Stanford, 1935.
- Gallagher, C.A., Perceptions of the Value of a Management Information System, *Academy of Management Journal*, Vol.17, No.1, 1974, 46-55.
- Galletta, D.F. et Lederer, A.L., Some Cautions on the Measurement of User Information Satisfaction, *Decision Sciences*, Vol.20, 1989, 419-438.
- Goodhue, D., I/S Attitudes: Toward Theoretical and Definitional Clarity, *The Data Base for Advances in Information Systems*, 1988, 6-15.
- Green, S.B., Lissitz, R.W. et Mulaik, S.A., Limitations of Coefficient Alpha as an Index of Test Unidimensionality, *Educational and Psychological Measurement*, Vol.37, No.4, Hiver 1977, 827-838.
- Guilford, J.P., *Psychometric Methods*, 2ème édition, McGraw-Hill, New York, 1954.
- Hair, J.F., Anderson, R.E., Tatham, R.L., et Grablowsky, B.J., *Multivariate Data Analysis*, Petroleum Publishing Co., Tulsa, Oklahoma, 1979.
- Hamel, K., Construct Measurement in Management Information Systems Research: The Case of User Satisfaction and User Involvement, *Proceedings of the 26<sup>th</sup> ORSA/TIIMS Annual Conference*, Denver, Colorado, Octobre 1988, 24-26.
- Hamilton, S. et Chervany, N.L., Evaluating Information Systems Effectiveness, *MIS Quarterly*, Vol.5, 1981, 79-86.
- Harris, C.W., On Factors and Factor Scores, *Psychometrika*, Vol.32, 1967, 363-379.
- Hawk, S.R. et Raju, N.S., Test-Retest Reliability of User Information Satisfaction: A Comment on Galletta and Lederer's Paper, *Decision Sciences*, Vol.22, No.5, 1991, 1165-1170.
- Iivari, J., User Information Satisfaction (UIS) Reconsidered: An Information System as the Antecedent to UIS, *Proceedings of the 8<sup>th</sup> International Conference (ICIS)*, Pittsburgh, Pennsylvanie, 6-9 Décembre 1987, 57-73.
- Ives, B., Olson, M.H., et Baroudi, J.J., The Measurement of User-Information Satisfaction, *Communications of the ACM*, 1983, Vol.26, 785-793.
- Kappelman, L.A. et McLean, E., The Respective Roles of User Participation and User Involvement in Information System Success, *Proceedings of the 12<sup>th</sup> International Conference International Conference on Information Systems (ICIS)*, 16-18 Décembre 1991, New York, 339-349.
- Keen, P.G.W., MIS Research: Référence Disciplines and a Cumulative Tradition, *Proceedings of the First International Conference on Information Systems (ICIS)*, Philadelphia, PA, 8 au 10 décembre, 1980, 9-18.

- Kerlinger, F.N., *Foundations of Behavioral Research*, 2ème édition, Holt, Rinehart, & Winston, New York, 1973.
- Kim, K.K., User Satisfaction: A Synthesis of three Different Perspectives, *Journal of Information Systems*, 1989, 1-12.
- Klenke, K., Construct Measurement in Management Information Systems: A Review and Critique of User Satisfaction and User Involvement Instruments, *Infor*, Vol. 30, No.4, Novembre 1992, 325-348.
- Kuder, G.F. et Richardson, M.W., The Theory of the Estimation of Test Reliability, *Psychometrika*, Vol.2, Septembre 1937, 151-160.
- Larcker, D.F. et Lessig, V.P., Perceived Usefulness of Information: A Psychometric Examination, *Decision Sciences*, Vol.11, 1980, 121-134.
- Montazemi, A.R., Factors Affecting Information Satisfaction in the Context of the Small Business Environment, *MIS Quarterly*, Vol.12, No.2, Juin 1988, 239-256.
- Melone, N.P., A Theoretical Assessment of the User-Satisfaction Construct in Information Systems Research, *Management Science*, Vol.36, 1990, 76-91.
- Norusis, M.J., *SPSS-X Advanced Statistics Guide*, 2ème édition, SPSS Inc., 1988.
- Nunnally, J.C., *Psychometric Theory*, 1ère édition, McGraw-Hill, New York, 1967.
- Nunnally, J.C., *Psychometric Theory*, 2ème édition, McGraw-Hill, New York, 1978.
- Pearson, S. W., Measurement of Computer User Satisfaction, *These de Doctorat*, Arizona State University, Temple, 1977.
- Pearson, S. W. et Bailey, J.E., Measurement of Computer User Satisfaction, *Installation Management Review*, Vol.8, 1979, 1-8.
- Pedhazur, E. J. et Schmelkin, L. P. *Measurement, Design, and Analysis: An Integrated Approach*. Hillsdale, Lawrence Erlbaum Associates, Publishers, New Jersey, 1991.
- Peter, J.P., Reliability: A Review of Psychometric Basics and Recent Marketing Research, *Journal of Marketing Research*, Vol.16, Février 1979, 6-17.
- Peterson, R.A., A Meta-Analysis of Cronbach's Coefficient Alpha, *Journal of Consumer Research*, Vol.21, Septembre 1994, 381-391.
- Powers, R.F. et Dickson, G.W., MIS Project Management Myths, Opinions, and Reality, *California Management Review*, Vol.15, 1973, 147-156.
- Raymond, L., Organizational Characteristics and MIS Success in the Context of Small Business, *MIS Quarterly*, Vol.9, 1985, 37-52.
- Raymond, L., La Satisfaction de l'Utilisateur d'un Système d'Information dans une PME: Une Étude Méthodologique, *P.M.O. Revue de Gestion des Petites et Moyennes Organisations*, Vol.2, 1986, 39-70.

- Raymond, L., Validating and Applying User Satisfaction as a Measure of MIS Success in Small Organizations, *Information & Management*, Vol.12, 1987, 173-179.
- Renterberg, S.E. et Gustafsson, J.E., Confirmatory Factor Analysis and Reliability: Testing Measurement Model Assumptions, *Educational and Psychological Measurement*, Vol.52, 1992, 795-811.
- Sellitz, C., Lawrence, S., Wrightsman, S, et Cook, S.W., *Research Method in Social Relations*, Holt, Rinehart & Sons, 1958.
- Spearman, C., Correlation Calculated with Faulty Data, *British Journal of Psychology*, Vol.3, 1910, 271-295.
- Straub, D.W., Validating Instruments in MIS Research, *MIS Quarterly*, Vol.13, No.2, Juin 1989, 147-169.
- Subramanian, A. et Nilakanta, S., "Measurement -- A Blueprint for Theory Building in MIS," *Information & Management*, Vol.26, 1994, 13-20.
- Symonds, P.M., Factors Influencing Test Reliability, *Journal of Educational Psychology*, Vol.19, Février 1928, 73-87.
- Tabachnick, B.G. et Fidell, L.S., *Using Multivariate Statistics*, 2ème édition, Harper Collins Publishers, Inc., New York, 1989.
- Tait, P. et Vessey, I., The Effect of User Involvement on System Success: A Contingency Approach, *MIS Quarterly*, Vol.12, 1988, 91-107
- Thurstone, L.L., *Multiple-Factor Analysis*, University of Chicago Press, Chicago, 1947.
- Treacy, M.E., An Empirical Examination of a Causal Model of User Information Satisfaction, *Proceedings of the 6<sup>th</sup> International Conference on Information Systems (ICIS)*, Indianapolis, Décembre 1985.
- Van de Ven, A.H. et Ferry, D.L., *Measuring and Assessing Organizations*, Wiley & Sons, New York, 1988.
- Wan, T.B. et Wah, L.T., Validation of a User Satisfaction Instrument for Office Automation Success, *Information & Management*, Vol.18, No.4, Avril 1990, 203-208.