

## Introduction aux théories des tests : ce sera pour une prochaine fois

Laveault, D. et Grégoire, J. (1997). Introduction aux théories des tests en sciences humaines. Bruxelles : De Boeck Université, 336 p.

Richard Bertrand

Volume 20, numéro 1, 1997

URI : <https://id.erudit.org/iderudit/1091392ar>

DOI : <https://doi.org/10.7202/1091392ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (imprimé)

2368-2000 (numérique)

[Découvrir la revue](#)

Citer ce compte rendu

Bertrand, R. (1997). Compte rendu de [Introduction aux théories des tests : ce sera pour une prochaine fois / Laveault, D. et Grégoire, J. (1997). Introduction aux théories des tests en sciences humaines. Bruxelles : De Boeck Université, 336 p.] *Mesure et évaluation en éducation*, 20(1), 117–120.

<https://doi.org/10.7202/1091392ar>

## **INTRODUCTION AUX THÉORIES DES TESTS : CE SERA POUR UNE PROCHAINE FOIS**

Laveault, D. et Grégoire, J. (1997). Introduction aux théories des tests en sciences humaines. Bruxelles: De Boeck Université, 336 p.

Il est, bien sûr, plusieurs façons d'envisager l'écriture de textes en théories de la mesure de même qu'il existe plusieurs manières de donner un cours sur ledit sujet. Dans le spectre des possibilités, deux pôles : à gauche, « dire peu de choses sur à peu près tout » et à droite, « tout dire sur à peu près rien ». Jusqu'ici, aucun des livres qu'il nous avait été donné de recenser ne se situait *exactement* à l'un ou à l'autre de ces pôles. L'impression nette qui se dégage, en parcourant le volume de Laveault et Grégoire, est que ceux-ci se sont résolument tournés vers le pôle de gauche. S'il faut rendre hommage au courage des auteurs pour la clarté de leur choix, il faut aussi se demander si, au bout du compte, ce n'est pas le lecteur qui en sort pénalisé. En clair : plusieurs des chapitres traités dans ce texte ne devraient pas y figurer, du moins si nous nous fions au titre donné à leur bouquin par les auteurs.

Par exemple, il est très étonnant de voir des notions comme celles abordées aux chapitres 2 et 3 faire partie d'un texte traitant des théories de la mesure, même s'il s'agit d'une *introduction*. Les méthodes statistiques portant sur les comparaisons entre deux moyennes, les comparaisons entre plus de deux moyennes ou les comparaisons multiples auraient en effet avantage à être retranchées du texte, pour plusieurs raisons. D'abord, plusieurs bons textes (par exemple celui de Glass et Hopkins paru en 1996) offrent déjà un excellent traitement de ces méthodes; ensuite, celles-ci apparaissent guère utiles à la compréhension des notions présentées plus loin; enfin, le lecteur un peu attentif découvrira, en feuilletant les pages de ces chapitres, plusieurs erreurs comme appeler  $s_x^2$  la variance des moyennes et

$s^2_{\bar{x}}$ , la variance des scores (p. 42) ou encore comme l'incompatibilité entre la formule de la corrélation (équation 2.22) et celle de la covariance (équation 2.20) (p. 69) et plusieurs formulations gauches dont celle-ci : « En statistiques, nous ne sommes pas seulement intéressés par le calcul des paramètres d'un échantillon. » (p. 41) ou bien « [...] se demander si la différence observée entre deux moyennes est probable » (p. 45). Le présent auteur est convaincu que le volume de Laveault et Grégoire est probablement le seul texte sur les théories de la mesure qui traite de concepts aussi pointus que l'*experimentwise error rate* et le *familywise error rate*, concepts qui, du reste, ne retrouvent aucun écho dans les paragraphes ou les chapitres qui suivent. Les auteurs font, par ailleurs, un traitement très classique des bases de l'analyse de la variance à une dimension, sujet sans réelle utilité pour les modèles de la généralisabilité, tous fondés sur des modèles à plusieurs dimensions. Toutes ces techniques trouvent leur place non pas dans un volume sur les théories de la mesure mais dans un volume spécialisé dans les méthodes statistiques. Il aurait été, par contre, beaucoup plus important de traiter plus à fond les tenants et les aboutissants de la corrélation, concept fondamental pour la compréhension des caractéristiques parfois subtiles du modèle classique. Le traitement plus approfondi de ce concept aurait peut-être même pu empêcher Laveault et Grégoire de commettre une erreur de conception à peine acceptable pour un volume portant sur les théories de la mesure : à la page 143, en effet, les auteurs mentionnent que, la corrélation entre deux tests parallèles étant 0,81, la proportion de variance vraie dans la variance totale est de 0,65 soit «  $(0,81)^2$  ». Or, à l'encadré 2 (p. 139-140), ils venaient de montrer que la fidélité (corrélation entre deux formes parallèles) était égale à la proportion de variance vraie. Ils ont confondu (à deux reprises dans cette même page) deux concepts de base en théorie de la mesure: la fidélité (corrélation entre deux formes parallèles) et la proportion de variance commune entre les deux formes. Enfin, il aurait été facile, à la fin du chapitre 2, de renvoyer le lecteur à des volumes de base portant sur les méthodes statistiques (et il y en a d'excellents!) comme celui de Glass et Hopkins plutôt que de les renvoyer à Siegel et Castellan (1988), un volume très spécialisé portant sur les statistiques non paramétriques, sujet fort pointu et d'une utilité incertaine dans un texte portant sur les théories de la mesure.

Il nous semble que le chapitre 3 n'est pas non plus nécessaire. Plusieurs excellents textes francophones (p. ex. Morissette, 1993) traitent déjà de façon très détaillée de notions de base en mesure et évaluation que sont l'évaluation sommative, l'évaluation formative, les formats d'items, etc. Notons, encore ici, que les auteurs auraient eu avantage à se concentrer sur le peaufinement des concepts propres aux modèles de mesure plutôt que s'acharner à présenter ces notions de base en mesure et évaluation.

Le traitement que font les auteurs des caractéristiques de la théorie de la généralisabilité est assez sobre et comporte peu d'erreurs de conception. Les auteurs s'appuient sur l'excellent texte de Cardinet et Tourneur (1985) pour présenter les concepts essentiels à la compréhension des modèles associés à cette théorie : une décision que nous endossons sans réserve. Deux remarques cependant : premièrement, nous aurions aimé une présentation plus étoffée des tenants et des aboutissants de cette théorie (les auteurs consacrent moins de 8 % du volume à cette théorie) comme le traitement non technique des conditions d'application et des nécessaires mises en garde à la Smith (1978) relatives à la difficulté d'estimation des composantes de variance. Deuxièmement, à la page 186, les auteurs, faisant référence à un plan d'optimisation (tableau 21, p. 185), semblent tomber dans un des pièges que présente la théorie en laissant entendre qu'ils emboîtent la facette de différenciation (les sujets) dans une *nouvelle* facette d'instrumentation (les groupes de juges); or, il est bien connu que cet emboîtement rend impossible l'identification de la variance vraie qui se trouve alors confondue avec la variance d'erreur.

Nous n'avons rien à redire sur le traitement des notions présentées au chapitre 5 sur la validité et au chapitre 7 sur l'interprétation des scores. Peu de textes francophones traitent de ces sujets importants pour les spécialistes de la mesure et de l'évaluation. Nous ne pouvons que féliciter les auteurs de s'y s'être astreints. Tout spécialiste de la mesure et de l'évaluation ne pourra que bénéficier de la lecture de ces deux chapitres. Par contre, la présentation des modèles de réponse aux items (chapitre 8) est nettement trop sommaire: les auteurs, en effet, n'y consacrent qu'un peu plus de 5 % de leur volume. Ce qui étonne compte tenu de la croissance phénoménale qu'ont connu ces modèles au cours des vingt dernières années, autant auprès de ceux qui se spécialisent dans les développements de ces modèles qu'auprès de ceux qui les appliquent. Pour s'en convaincre, on n'a qu'à lire assidûment les revues spécialisées en mesure et évaluation comme *Journal of Educational Measurement*, *Applied Psychological Measurement*, *Educational and Psychological Measurement*, *Psychometrika*, *European Journal of Psychological Assessment* ainsi que plusieurs autres.

Il va de soi que, si le livre n'est pas à la hauteur de nos attentes, nous devons nous demander si c'est parce que le traitement n'est pas assez rigoureux ou alors parce que nos attentes étaient trop élevées. Toutes les petites fautes, coquilles et erreurs de conception nous empêchent de suggérer ce volume aux étudiants du premier cycle universitaire. Nous connaissons cependant plusieurs étudiants des cycles universitaires supérieurs qui

n'auraient aucun mal à distinguer et à corriger la plupart de ces erreurs et donc à tirer profit de la lecture du volume de Laveault et Grégoire. Quant aux chercheurs dont nous sommes et qui attendaient, depuis plusieurs lunes, de s'affranchir des textes anglo-saxons, ce sera pour une prochaine fois.

**Richard Bertrand**  
**Université Laval**

### **Ouvrages cités**

- Cardinet, J. & Tourneur, Y. (1985). Assurer la mesure. Berne : Peter Lang.
- Glass, G.V. & Hopkins, K.D. (1996). Statistical methods in education and psychology (3<sup>e</sup> éd.). Englewood Cliffs, NJ : Prentice Hall.
- Morissette, D. (1993). Les examens de rendement scolaire. Sainte-Foy : Les Presses de l'Université Laval.
- Siegel, S. & Castellan, J.J. (1988). Nonparametric statistics for the behavioral sciences. New York : McGraw-Hill.
- Smith, P. (1978). Sampling errors of variance components in small multifacet generalizability studies, Journal of Educational Statistics, 3 , 319-346.