

# Mesure et évaluation en éducation



## Présentation

Jean Cardinet

Volume 26, numéro 1-2, 2003

Généralisabilité

URI : <https://id.erudit.org/iderudit/1088235ar>

DOI : <https://doi.org/10.7202/1088235ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (imprimé)

2368-2000 (numérique)

[Découvrir la revue](#)

Citer ce document

Cardinet, J. (2003). Présentation. *Mesure et évaluation en éducation*, 26(1-2), 1–3.  
<https://doi.org/10.7202/1088235ar>

## Présentation

**Jean Cardinet**

*Anciennement responsable du Service de la recherche de l'IRD, Neuchâtel*

Consacrer tout un numéro spécial de cette revue au thème de la «Généralisabilité» et de ses applications pourra peut-être surprendre. Certains lecteurs pensent certainement que le sujet ne peut pas les intéresser, parce qu'ils le croient trop difficile. D'autres jugent le domaine épuisé, puisque les publications initiales de Cronbach sur cette théorie datent de 1963. En fait, il s'agit dans les deux cas d'idées reçues, et les textes qui suivent devraient pouvoir les redresser.

La théorie de la généralisabilité est certes fondée sur celle de l'analyse de la variance en statistique, mais sa compréhension ne demande la maîtrise que de quelques formules, comme la présentation d'un modèle par Gratien Mokonzi pourra le prouver. Les calculs ne font plus obstacle : les logiciels actuellement disponibles sont faciles à utiliser et à interpréter par quiconque s'intéresse à contrôler, par exemple, la valeur des examens.

Mais cette préoccupation est elle-même aujourd'hui en train de passer au second plan, et c'est là le second objectif de ce numéro : montrer la diversité des domaines d'application possibles de cette théorie de la mesure, qui ne se limite pas à la docimologie.

Daniel Bain, illustre, par un exemple réel, comment les méthodologues peuvent contrôler et enrichir leur expérimentation de séquences didactiques, grâce aux outils qu'il propose. Des instruments fiables sont nécessaires pour déterminer les notions déjà connues sur lesquelles s'appuyer, pour former des groupes d'élèves aux besoins particuliers, pour mesurer les progrès individuels et collectifs, pour contrôler l'atteinte de standards de performance prédéterminés, etc.

Sandra Johnson montre également, sur la base de son expérience personnelle, comment la connaissance de l'importance relative de diverses sources de variation permet de planifier une enquête sur les acquisitions scolaires de façon à réduire le mieux possible la marge d'erreur, pour un coût de prise d'infor-

mation donné. Elle démontre clairement qu'une épreuve n'a pas besoin de différencier les bons et les mauvais élèves, ni les niveaux d'aucune dimension, pour renseigner sur la performance du système scolaire dans son ensemble.

Jean Cardinet modifie la perspective docimologique classique en proposant de certifier les progrès des élèves de façon individualisée. Il faut prouver que ces progrès dépassent l'effet des sources de variation aléatoire connues. Mais il faut choisir ce terme de comparaison de façon rationnellement acceptable, parmi cinq possibilités qui sont présentées et discutées. Il n'est pas besoin pour cela de faire intervenir une quelconque comparaison entre élèves.

Gratien Mokonzi explore de nouvelles méthodes de construction d'épreuves pédagogiques «robustes» en comparant cinq façons différentes de sélectionner les questions que l'on veut utiliser. Il montre que chacune permet de répondre à des besoins spécifiques, c'est-à-dire à des types particuliers de décision à prendre, en améliorant soit la fidélité, soit la précision de ces instruments, sur une échelle appropriée soit à un concours, soit à un examen.

Enfin, Richard Bertrand termine la série d'articles en établissant des ponts entre la théorie de la généralisabilité et la théorie des tests classique, d'une part, et la théorie des réponses aux items, d'autre part. Les non-spécialistes de ces domaines seront heureux de voir apparaître une plate-forme commune derrière les spécificités de chaque approche. Les apports particuliers de la théorie des réponses aux items sont bien mis en lumière. Il ne s'agit pas, en effet, de présenter la théorie de la généralisabilité comme la seule démarche possible pour assurer la mesure, mais d'en montrer simplement l'intérêt et les multiples applications.

À l'heure où les grandes enquêtes internationales suscitent des réactions plus affectives que rationnelles, il est bon de rappeler que le premier souci des responsables qui les commanditent doit être de s'assurer de la valeur des mesures qu'on leur présente. On verra, dans ce numéro, que des méthodes de contrôle existent, sur lesquelles les spécialistes peuvent tomber d'accord.

Inversement, il serait grave de rejeter les méthodes quantitatives sans les connaître, par pur préjugé, surtout si l'on n'a pas voulu faire l'effort de s'informer à leur sujet. Ces méthodes permettent souvent de mieux estimer les deux risques d'erreur que comporte toute alternative et de prendre ainsi des décisions en meilleure connaissance de cause. L'insuffisante fiabilité des examens, par exemple, est scandaleuse, si l'on songe aux conséquences aussi bien individuelles que sociales des erreurs de jugement tolérées jusqu'à maintenant par l'institution scolaire.

Souhaitons que ce numéro d'une revue consacrée à la mesure aussi bien qu'à l'évaluation rappelle à tous ses lecteurs qu'une réponse OUI ou NON, ou un verdict de réussite ou d'échec, sont tout aussi «quantitatifs» qu'un 1 ou un 0. Ce n'est pas parce qu'elles sont numériques que les notes sont à bannir, mais parce qu'elles sont trompeuses. Utilisons donc les chiffres chaque fois qu'ils peuvent rendre l'école plus juste et plus humaine !