

**Ordre des conseillers et conseillères d'orientation et des psychoéducateurs et psychoéducatrices du Québec (2003).  
*Normes de pratique du testing en psychologie et en éducation.*  
Montréal : Institut de Recherches Psychologiques**

Carl Bouchard

---

Volume 34, Number 1, 2005

URI: <https://id.erudit.org/iderudit/1097579ar>  
DOI: <https://doi.org/10.7202/1097579ar>

[See table of contents](#)

---

Publisher(s)

Revue de Psychoéducation

ISSN

1713-1782 (print)  
2371-6053 (digital)

[Explore this journal](#)

---

Cite this review

Bouchard, C. (2005). Review of [Ordre des conseillers et conseillères d'orientation et des psychoéducateurs et psychoéducatrices du Québec (2003). *Normes de pratique du testing en psychologie et en éducation*. Montréal : Institut de Recherches Psychologiques]. *Revue de psychoéducation*, 34(1), 167–170.  
<https://doi.org/10.7202/1097579ar>

Les quarante-six pages du livre sont agrémentées d'illustrations dont des dessins d'enfants et laissent passer quelques traits d'humour, qui facilitent la compréhension. Plusieurs termes, expressions, comparaisons ou phrases sont graphiquement mises en évidence pour faciliter la lecture des enfants concernés qui présentent souvent des difficultés dans ce domaine. La deuxième moitié de l'ouvrage encadre des stratégies cognitives et métacognitives reliées aux différentes fonctions exécutives (organisation, planification, activation, inhibition, ajustement, métacognition) propres à soutenir les activités quotidiennes de tout enfant. Inspirée du *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV), on trouve à la page 22 une adaptation des critères diagnostiques du TDA/H destinée aux parents et aux enseignants qui oeuvrent les uns et les autres dans les principaux milieux de vie des enfants.

Le trouble est toujours traité de façon très bipolaire comme pour bien marquer la distinction entre l'inattention et l'hyperactivité alors que le DSM-IV, moins dichotomique, fait état de deux types de prédominance (impulsivité/hyperactivité ou inattention) auquel s'ajoute un tableau qui réunit les deux. Autrement dit, la variable «inattention» serait toujours présente contrairement à l'hyperactivité. Si le livre néglige cette précision, c'est sans doute par souci de simplifier la lecture. Ce choix reste discutable. Une légère précision aurait également pu être apportée quant au taux de prévalence du TDA/H chez l'un et l'autre sexe.

Les 16 références sont distribuées sous trois rubriques : TDA/H chez enfants et ados, TDA/H chez adultes et sites Internet). On fournit aussi les coordonnées de trois ressources d'aide spécialisée.

En somme, un excellent ouvrage du point de vue de l'accessibilité, qui permet à de nombreux parents, enfants et enseignants de démystifier le TDA/H, de mieux reconnaître ses manifestations et de faciliter l'intervention.

**M. Larivée**

- Ordre des conseillers et conseillères d'orientation et des psychoéducateurs et psychoéducatrices du Québec (2003). *Normes de pratique du testing en psychologie et en éducation*. Montréal : Institut de Recherches Psychologiques.

Sous la direction de Georges Sarrazin de l'Université d'Ottawa, l'ordre des conseillers et conseillères d'orientation et des psychoéducateurs et psychoéducatrices du Québec nous offre l'adaptation francophone du « Standards for Educational and Psychological Testing », ouvrage initialement publié par l'American Educational Research Association (AERA), l'American Psychological Association (APA) ainsi que le National Council on Measurement in Education (NCME). Une telle adaptation n'est pas sans complexité. En premier lieu, la traduction du vocabulaire technique en psychométrie trouve de nombreux synonymes dans la langue française. De plus, certaines expressions courantes en anglais subissent une traduction littérale peu orthodoxe sur le plan linguistique. Par exemple, afin de définir la technique de validation par laquelle un nouveau test sera comparé à un critère externe, le terme anglais « concurrent validity » devient en français la « validité concourante », alors qu'il

faudrait écrire « validité concomitante ». En second lieu, il n'existe aucun consensus quant à l'utilisation de ces termes techniques au sein même de la francophonie. Par exemple, pour établir la validité d'un concept mesuré par quelconque instrument, certains auteurs utiliseront le terme de « validité de construit » alors que d'autres préféreront le terme « validité théorique ». Notons cependant que le problème se pose également chez les professionnels anglophones, ce qui complique d'autant l'adaptation des ouvrages d'une langue dans une autre.

Le texte initial visait à établir succinctement des critères pour évaluer la qualité de tests et d'instruments utilisés en psychologie et en éducation. Une première parution en 1954 a été successivement révisée en 1966 et en 1974, pour aboutir finalement à la version augmentée de 1985. En 1991, différents comités issus de l'APA, de l'AERA et du NCME se rencontraient pour discuter des normes et des principes qui devraient guider une autre révision puis pour envisager la constitution des comités. Entre 1993 et 1996, ces différents comités ont soumis leurs recommandations et leurs commentaires quant à la liste des ajouts et des modifications concernant les normes établies en 1985. Finalement, la plus récente version paraissait en 1999. Celle-ci reflète les modifications de certaines lois fédérales américaines concernant l'application des tests aux personnes, y compris des standards visant certaines populations spéciales, et d'autres visant l'encadrement de gestes administratifs basés sur les résultats obtenus avec l'utilisation des tests. En bref, les nouveaux standards ont pour but de mieux protéger les droits des personnes examinées.

L'adaptation francophone de l'ouvrage respecte la structure de la version américaine constituée de trois parties. **La première partie** intitulée « Construction de tests, évaluation et documentation » renferme 6 chapitres touchant respectivement la validité (chapitre 1) ; la fidélité et les erreurs de mesure (chapitre 2) ; le développement et la révision des tests (chapitre 3) ; l'échelonnage, l'étalonnage et la comparabilité des scores (chapitre 4) ; l'administration et la correction des tests et la communication des scores (chapitre 5) ; la documentation afférente aux tests (chapitre 6). **La seconde partie** porte sur l'« Équité en évaluation », elle contient 4 chapitres abordant spécifiquement les thèmes de l'équité en évaluation (chapitre 7) ; les droits et les responsabilités des candidats (chapitre 8) ; l'évaluation des individus d'origines linguistiques diverses (chapitre 9) ; l'évaluation des personnes présentant un handicap (chapitre 10). Enfin, la troisième partie traite des questions touchant l'« Application du testing » et présente les normes relatives aux responsabilités des utilisateurs de tests (chapitre 11) ; du testing et de l'évaluation psychologique (chapitre 12) ; du testing et de l'évaluation en milieu scolaire (chapitre 13) ; du testing d'emploi et d'accréditation (chapitre 14) ; elle traite finalement du testing dans l'évaluation de programmes et des politiques publiques (chapitre 15).

En plus de garder la structure initiale du « Standards for Educational and Psychological Testing », la traduction présente des exposés théoriques au début des chapitres afin de faciliter la compréhension du lecteur en regard des concepts utilisés et qui font l'objet des normes de pratiques. A cet égard, et contrairement aux prétentions du coordonnateur de l'adaptation, la compréhension des concepts utilisés dans l'ouvrage peut certainement représenter un obstacle pour les néophytes. Bien que l'information soit judicieusement présentée, les usagers ne profiteront de l'ouvrage que s'ils disposent déjà d'une formation de base dans les domaines de la

mesure et/ou de la psychométrie mais surtout de l'évaluation. En effet, la compréhension des normes de pratiques dépasse largement la connaissance des concepts et des techniques de validité et de fidélité des tests ; elle requiert également une capacité de réflexion sur la valeur des résultats quantitatifs obtenus (les scores bruts et normalisés) et le jugement évaluatif qui en découle. Comme Laurencelle (1998) le souligne, évaluer c'est plus que mesurer. La mesure exprime, d'une façon ou d'une autre, l'ampleur de l'objet d'étude, mais elle prend toute sa signification dans un contexte bien précis qui permet de l'interpréter et de l'exploiter dans des perspectives sociales, éducatives, économiques et scientifiques. Il est donc nécessaire que cette réflexion sur la valeur de la mesure, et plus particulièrement la valeur des concepts et construits mesurés soit déjà familière au lecteur s'il veut saisir toute la portée des normes de pratique suggérées dans l'ouvrage.

Par ailleurs, compte tenu de la grande diversité culturelle du Québec et de notre proximité avec les Etats-Unis d'où provient la plupart des tests utilisés en psychologie et en éducation, certaines précisions s'imposent concernant le chapitre 9. Il arrive souvent que les utilisateurs de tests se contentent d'une simple traduction pour les appliquer. Tel que mentionné dans l'ouvrage (page 110), le seul processus de traduction ne saurait équivaloir au test d'origine quant aux aspects techniques de la validité et de la fidélité, ni au point de vue du contenu. Autrement dit, la traduction d'un test ne représente que la première d'une longue série qui aboutira à une telle version culturellement adaptée.

A cet effet, Geisinger (1994) propose des lignes directrices en dix étapes basées sur les normes de pratique, lesquelles permettant une validation transculturelle des tests. **La première étape** consiste à traduire et à adapter le test d'origine à la population d'intérêt tout en assurant son caractère d'équivalence. Selon Guillemin, Bombardier et Beaton (1993), cette équivalence se traduit en fait par une série d'équivalences : (1) l'équivalence sémantique, (2) l'équivalence des expressions, (3) l'équivalence expérientielle, (4) l'équivalence conceptuelle. Pour y atteindre, plusieurs techniques de traduction peuvent être utilisées : la traduction traditionnelle par un chercheur bilingue ou un traducteur professionnel, la traduction par un comité d'experts ou encore la rétro-traduction (back-translation) par deux personnes ou plus. Pour plus d'informations à ce sujet, voir Haccoun (1987). **La seconde étape** vise à soumettre cette version préliminaire à un comité d'experts. **La troisième étape** consiste à corriger la version préliminaire compte tenu des commentaires de ce dernier. Par la suite, **en guise de quatrième étape**, une pré-expérimentation est réalisée auprès d'un petit groupe de personnes ciblées par le test en question ( $n = 20$ ) afin d'obtenir de leur part une rétroaction quant à la formulation et à la compréhension de certains items. À la suite des corrections consécutives, **la cinquième étape** consiste à administrer le test à un large échantillon représentatif des individus ciblés. Il s'agit donc une fois de plus de vérifier la clarté des items contenus dans le test afin de s'assurer qu'ils sont énoncés dans un langage accessible et facilement compréhensible par la population visée.

L'adaptation des scores au contexte culturel et le développement des normes constituent les principales tâches de **la sixième étape** proposée par Geisinger (1994). Il s'agit de comparer la distribution des pointages obtenus par la version traduite avec ceux de la version originale du test. Ainsi, la moyenne et l'écart-type permettront de vérifier la variabilité de la mesure et sa pertinence culturelle. Dans l'affirmative, des

normes devront être établies en fonction de la culture cible, incluant la moyenne et son écart-type, les rangs en percentile ainsi que le type de score (Z ou T). **La septième étape**, probablement la plus importante et la plus complexe, nécessite d'énormes ressources financières et temporelles : la vérification empirique de la validité. Pour que la version traduite du test soit validée, elle doit respecter les critères de validité de contenu, de la validité concomitante, de la validité de construit (ce qui prend plusieurs années) et rencontrer les critères de fidélité. Il serait trop long d'exposer ici les procédures techniques appropriées à chacun des critères, mais le lecteur pourrait à cet égard consulter l'ouvrage de Bernier et Pitrulewicz (1997). **La huitième étape** concerne le développement d'un manuel contenant toutes les informations relatives à la validation du test (les coefficients de fidélité et de validité) ainsi que les règles d'administration. Enfin, **les deux dernières étapes** regardent l'entraînement des utilisateurs du test et la collecte de leurs commentaires et suggestions en vue de peaufiner la prochaine version.

Autant de procédures qui mettent en évidence les limites de l'utilisation d'une simple traduction d'items. Il ne s'agit pas ici de décourager les utilisateurs de ces tests mais bien de souligner que l'interprétation des résultats et les décisions dérivées doivent obligatoirement tenir compte du contexte d'évaluation et de toutes autres sources d'informations disponibles susceptibles de nuancer le jugement évaluatif final du professionnel. Ces exigences illustrent bien la rigueur avec laquelle les normes de pratique du testing méritent d'être prises en considération pour éviter des effets néfastes chez les usagers et les intervenants impliqués dans les décisions subséquentes. L'objectif principal de telles normes vise finalement à promouvoir une sage utilisation de tests validés et le respect des droits des personnes évaluées afin d'établir les bases d'une évaluation de qualité. A cet égard, l'adaptation d'une version francophone des normes de pratique représente donc un apport très important à l'encadrement des pratiques du testing par les professionnels d'ici.

### Références

- Bernier, J.-J., et Pitrulewicz, B. (1997). *La psychométrie: traité de mesure appliquée*. Boucherville: Gaëtan Morin.
- Geisinger, K.F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment*, 6, (4), 304-312.
- Guillemain, F., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures : Litterature review and proposed guidelines. *Journal of Clinical Epidemiology*, 46, 1417-1432.
- Haccoun, R., R. (1987). Une nouvelle technique de vérification de l'équivalence de mesures psychologiques traduites. *Revue Québécoise de Psychologie*, 8, 30-39.
- Laurencelle, L. (1998). *Théorie et techniques de la mesure instrumentale*. Sainte-Foy : Presses de l'Université du Québec.