

## UN MOT DE LA RÉDACTRICE INVITÉE

En 2002, le Conseil de recherche en sciences humaines du Canada (CRSH) en collaboration avec le Conseil des statistiques canadiennes en éducation (CSCE) mettait sur pied un nouveau programme de subventions dans le cadre du lancement de l'Initiative de la nouvelle économie (INÉ). Le CSCE représente également un partenariat entre le Conseil des ministres de l'Éducation, Canada (CMEC) et Statistique Canada. Ces initiatives avaient comme but d'encourager les chercheurs à mener des analyses secondaires sur les bases de données recueillies aux plans régionaux, nationaux, et internationaux.

Les chercheurs ont effectivement accès à nombreuses bases de données à grande échelle régies par divers organismes. Le CSCE par exemple gère le Programme d'indicateurs pancanadiens de l'éducation (PIPCE) et le Programme pancanadien de recherche en éducation (PPRE). Par ailleurs, Statistique Canada compte une trentaine d'études actives dont l'Enquête auprès des jeunes en transition (EJET) et l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ). Ces bases de données s'avèrent populaires auprès des chercheurs académiques. Les provinces et les territoires administrent également leurs propres programmes d'évaluation du rendement des élèves en plus du Programme d'indicateurs du rendement scolaire (PIRS) du CMEC en lecture, écriture, mathématiques, et sciences. Ce dernier fut remplacé récemment par le Programme pancanadien en éducation (PPCE). Le programme international pour le suivi des acquis (PISA) sous l'égide de l'Organisation pour la coopération et le développement économique (OCDE), le Programme international du rendement en lecture scolaire (PIRLS), et la Troisième enquête internationale sur l'enseignement des mathématiques et des sciences (TEIEMS) monopolisent annuellement l'attention et les ressources de plusieurs des provinces et d'un grand nombre de pays.

Tous ces programmes s'appuient sur des cadres théoriques, méthodologiques, et analytiques qui contribuent à la mise sur pied de bases de données brutes de grande envergure. Ces dernières sont d'abord traitées en fonction des intentions de départ déterminées par les agences qui en sont responsables. Toutefois, elles font parfois l'objet d'analyses secondaires par des experts-chercheurs en vue de déterminer, entre autres, l'impact de variables contextuelles ciblées sur le rendement des élèves.

Malgré la richesse des données recueillies depuis les 20 dernières années, leur exploitation aux fins d'études approfondies des facteurs déterminants du rendement des élèves et plus particulièrement ceux qui sont à risque d'échec n'est qu'au stade embryonnaire. En effet, dans son rapport, Crocker (2002) écrit :

Les élèves du Canada se classent actuellement aux premiers rangs des évaluations internationales pour ce qui est des disciplines principales. On n'en constate pas moins des différences substantielles et persistantes d'une instance à l'autre, entre groupes linguistiques et selon les sexes ou la situation socioéconomique. On n'a jamais utilisé les résultats des évaluations à grande échelle pour élaborer des politiques de nature à combler ces écarts.  
(p. 23)

D'où la mise sur pied de programmes conjoints de subventions comme le CRSH–CSCE, à l'appui d'analyses secondaires de bases de données brutes afin de nourrir les politiques en matière d'éducation à travers le Canada.

Dans le cadre du programme CRSH–CSCE, une rencontre annuelle durant la période 2002 à 2005 réunit chercheurs, praticiens, dirigeants, et politiciens en vue de partager les intentions, les travaux réalisés, les préoccupations, les résultats anticipés, et les résultats obtenus d'une vingtaine de projets. Lors de ces rencontres, il devient évident que les analyses secondaires *post hoc* des données issues des programmes d'évaluation du rendement des élèves donnent des résultats plutôt décevants et font face à des défis de taille. Les chercheurs se rendent compte d'un besoin d'échanger davantage sur la question et de collaborer avec les concepteurs des enquêtes tout au long de l'élaboration, de la mise en œuvre, et de la dissémination des résultats des programmes d'évaluation afin de produire des résultats utiles et pertinents. Les réunions organisées par le CRSH–CSCE ne suffisent pas à débattre à fond les préoccupations des chercheurs.

À titre de suivi à ces réunions, l'équipe de l'unité de recherche en mesure et évaluation des apprentissages (MÉA) de la Faculté d'éducation de l'Université d'Ottawa dont je suis en partie responsable organise, en juin 2006, un atelier bilingue et international intitulé « Vers une conception partagée des enquêtes sur le rendement scolaire en lecture et en écriture ». L'atelier regroupe une vingtaine de participants choisis soigneusement selon des critères précis : réputation académique,

degré d'expertise, expérience reliée au thème de l'atelier, position d'influence, représentativité, équilibre, et variété. Il offre un forum d'échanges de connaissances théoriques, de pratiques actuelles, et de politiques pertinentes en matière des programmes d'évaluation à grande échelle de la lecture et de l'écriture dans le but de les rendre davantage utiles, attrayants, et valides aux yeux de tous les partis concernés.

Pour ce faire, l'atelier s'articule autour des trois questions suivantes. (a) Les intervenants dans les écoles, les concepteurs des programmes d'évaluation à grande échelle, les experts en lecture et en écriture, et les chercheurs en mesure et évaluation se partagent-ils tous la même définition des concepts de lecture et écriture? (b) Dans quelle mesure les programmes d'évaluation de la lecture et de l'écriture tiennent-ils compte de ces définitions, d'où la notion de validité? (c) Comment rendre ces évaluations attrayantes, accessibles, et utiles aux yeux des divers partis concernés et en particulier auprès des intervenants du milieu scolaire? Le numéro spécial de la revue met donc en vedette des articles directement reliés aux conférences prononcées lors de l'atelier. Les articles se regroupent de manière non exclusive selon trois thèmes : (a) conception de la lecture et de l'écriture, (b) mesure de ces concepts, et (c) utilisation des résultats.

Les articles de Hansen et de Klinger et Luce-Kapler ainsi que le compte rendu de Peterson du livre rédigé par l'International Association for the Evaluation of Educational Achievement réfèrent au premier thème. Hansen décrit la conception qu'ont les enseignants et les chercheurs des approches pédagogiques de la lecture et de l'écriture. Son exposé met en relief le potentiel des évaluations à grande échelle de capter les diverses dimensions d'apprentissage de la lecture et de l'écriture en salle de classe. Par ailleurs Klinger et Luce-Kapler décrivent le point de vue des élèves en ce qui concerne l'évaluation de la lecture et de l'écriture dans un contexte à enjeux importants. Peterson critique la conception qu'offrent les responsables du PIRLS en matière de la lecture. Ces études soulèvent la divergence dans les conceptions de la lecture et de l'écriture des divers partis concernés et leur impact sur la validité des programmes d'évaluation.

En lien avec le deuxième thème se trouvent les articles d'Ercikan et Barclay-McKeown et de Simon, Roberts, Tierney, et Forgette-Giroux et la note de Blais sur la recherche et les méthodes, quoique l'article de Simon et al. s'inscrive également dans le troisième thème. Ercikan et Barclay-McKeown discutent de la pertinence limitée des programmes

d'évaluation de la lecture et de l'écriture par rapport à l'amélioration de l'apprentissage en salle de classe en raison de problèmes reliés à la mesure. Parallèlement, en relatant leurs expériences liées à l'application d'analyses secondaires des données de trois programmes d'évaluation du rendement, Simon, Roberts, Tierney, et Forgette-Giroux rapportent les défis conceptuels, pratiques, statistiques, et techniques reliés à l'analyse de données provenant des populations linguistiques en milieu minoritaire. Enfin, Blais traite des limites de l'échantillonnage et de la validité des conclusions qui s'y rattachent.

En ce qui concerne l'utilisation des résultats des programmes d'évaluation à des fins théoriques, politiques, et pratiques, les articles de Savard, Sévigny, et Beaudoin et de Bouchamma, Lapointe, et Richard rapportent les résultats d'analyses secondaires menées sur les données issues du programme pancanadien d'évaluation du rendement en écriture. Sur le même thème, IsaBelle rend compte d'un livre qui s'inspire des analyses secondaires en vue de guider les politiques et les pratiques scolaires en matière de la lecture et de l'écriture. Bref le numéro spécial de la Revue permet de tirer diverses leçons en ce qui a trait aux programmes d'évaluation du rendement scolaire. Celles-ci font l'objet du commentaire de la rédactrice invitée à la fin de ce numéro.

Marielle Simon  
Rédactrice invitée  
Ottawa, décembre 2007

## RÉFÉRENCE

- Crocker, R.K. (2002). *Résultats d'apprentissage : Analyse critique du domaine au Canada*. Rapport présenté au Conseil des statistiques canadiennes de l'éducation. Ottawa, ON : CSCÉ. Répéré décembre 2006 de <[http://www.cmec.ca/stats/pcera/LearningOutcomes\\_StateoftheField\\_RCrocker2002f.pdf](http://www.cmec.ca/stats/pcera/LearningOutcomes_StateoftheField_RCrocker2002f.pdf)>.

## GUEST EDITOR'S REMARKS

In 2002, the Social Sciences and Humanities Research Council of Canada (SSHRC), in co-operation with the Canadian Education Statistics Council (CESC), established a new grant program under the Initiative on the New Economy (INE). The CESC also represents a partnership between the Council of Ministers of Education, Canada (CMEC) and Statistics Canada. One of the purposes of these initiatives is to encourage researchers to conduct secondary analyses using regional, national, and international databases.

In fact, researchers have access to numerous large-scale databases managed by various agencies. The CESC, for example, manages the Pan-Canadian Education Indicators Program (PCEIP) and the Pan-Canadian Education Research Program (PERP). Statistics Canada has some 30 active studies, including the Youth in Transition Survey (YITS) and the National Longitudinal Survey of Children and Youth (NLSCY). These databases have proven popular with academic researchers. The provinces and territories also administer student achievement assessment programs, in addition to the CMEC's School Achievement Indicators Program (SAIP) in reading, writing, mathematics, and science, which has recently been replaced by the Pan-Canadian Assessment Program (PCAP). The Programme for International Student Assessment (PISA) of the Organisation for Economic Co-operation and Development (OECD), the Progress in International Reading Literacy Study (PIRLS), and the Third International Mathematics and Science Study (TIMSS) annually monopolize the attention and resources of many of these provinces and a large number of countries.

All those assessment programs are based on theoretical, methodological, and analytical frameworks that contribute to the establishment of large-scale databases of raw data. The data are first processed based on the initial intentions of the governing agencies. However, at times, they are subjected to secondary analyses by experts and researchers to determine, among other things, the impact of target contextual variables on student achievement.

Despite the wealth of data collected in the past 20 years, their use for the purposes of in-depth studies of the factors that determine student achievement, particularly for those at risk of failure, is only at an early stage. In fact, in his report, Crocker (2002) writes:

Canadian students are now in the top ranks on international assessments in core subject areas. Nevertheless, significant and persistent differences are found across the jurisdictions, between official language groups and by gender and socio-economic circumstances. The results of large-scale assessments have not been used to develop policies that would reduce these disparities. (p. 22)

Hence the creation of joint grant programs, such as the SSHRC–CESC program, to support secondary analyses of databases of raw data for the development of education policy across Canada.

Within the SSHRC–CESC program, an annual meeting has been held from 2002 to 2005 where researchers, practitioners, administrators, and politicians shared intentions, projects carried out, concerns, expected results, and the results obtained from some 20 projects. At the meetings, it became clear that *post hoc* secondary analyses of data from student achievement assessment programs provide somewhat disappointing results and face major challenges. Researchers realize there is a need for more discussion on the issue and for co-operation with survey designers throughout the development and implementation of assessment programs, and the dissemination of their results, to produce useful, relevant information. The meetings organized through the SSHRC–CESC were insufficient for a full debate of researchers' concerns.

As a follow-up to the meetings, the Measurement, Evaluation and Assessment (MEA) Research Unit team in the Faculty of Education at the University of Ottawa, which I am partly responsible for, organized an international bilingual workshop entitled "Toward a Shared Conception of Large-scale Assessments of Reading and Writing" in June 2006. The workshop brought together approximately 20 participants carefully selected based on specific criteria, such as academic reputation, level of expertise, experience related to the workshop theme, position of influence, representativeness, balance, and variety. It provided a forum for the exchange of theoretical knowledge, current practices, and relevant policies on large-scale reading and writing assessment programs aimed at making them more useful, attractive, and valid to all the parties involved.

To achieve its goal, the workshop covered the following three issues. (a) Do school practitioners, large-scale assessment program designers, reading and writing experts, and assessment and evaluation research-

ers all share the same definition of reading and writing concepts? (b) To what extent do reading and writing assessment programs take those definitions into consideration, hence the notion of validity? (c) How can assessments be made more attractive, accessible, and useful to the various stakeholders, particularly to practitioners in school communities? The special issue of the *Journal* thus features articles directly related to the workshop presentations. The articles cover three themes, although not exclusively: (a) reading and writing concepts, (b) assessment of those concepts, and (c) use of results.

The articles by Hansen and Klinger and Luce-Kapler and the review by Peterson of the *International Association for the Evaluation of Educational Achievement* book deal with the first theme. Hansen describes the concept teachers and researchers have of reading and writing instructional approaches. Her article highlights the potential of large-scale assessments for capturing the various dimensions of learning to read and write in the classroom. Klinger and Luce-Kapler study the assessment of reading and writing in a high-stakes context from a student viewpoint. Peterson critiques the PIRLS test designers' construct of reading. These studies point out divergences in reading and writing concepts among the various stakeholders and their impact on the validity of assessment programs.

The articles by Ercikan and Barclay-McKeown and Simon, Roberts, Tierney, and Forgette-Giroux and the research and practice note by Blais deal with the second theme, although the article by Simon et al. also relates to the third theme. Ercikan and Barclay-McKeown discuss the limited relevance of reading and writing assessment programs in improving classroom learning because of measurement problems. Similarly, relating their experience with the application of secondary analyses of data from three achievement assessment programs, Simon et al. report on the conceptual, practical, statistical, and technical challenges involved in the analysis of data from minority language populations. Finally, Blais argues the limitations of sampling and the validity of related findings.

On the theme of the use of the results of assessment programs for theoretical, political, and practical purposes, the articles by Savard, Sévigny, and Beaudoin and Bouchamma, Lapointe, and Richard report the results of secondary analyses of data from the Pan-Canadian Assessment Program for reading and writing. On the same theme, IsaBelle reviews a book that uses secondary analyses to guide policies and school practices in reading and writing. The special issue of

the Journal offers a variety of lessons learned in school achievement assessment programs with comments by the guest editor at the end of the issue.

Marielle Simon  
Guest Editor  
Ottawa, December 2007

**REFERENCE**

Crocker, R.K. (2002). *Learning outcomes: A critical review of the state of the field in Canada*. Report submitted to Canadian Education Statistics Council. Ottawa, ON: CESC. Retrieved December 2006 from <[http://www.cmec.ca/stats/pcera/LearningOutcomes\\_StateoftheField\\_RCrocker2002.pdf](http://www.cmec.ca/stats/pcera/LearningOutcomes_StateoftheField_RCrocker2002.pdf)>.