

## IMPACTS DU PIRS EN MILIEU SCOLAIRE

Marielle Simon  
Catherine Turcotte  
Renée Forgette-Giroux  
Université d'Ottawa  
Ottawa, Ontario

**Résumé:** Cet article examine l'impact du Programme d'indicateurs du rendement scolaire (PIRS) sur le milieu éducatif. Ce programme pancanadien évalue le rendement des élèves en mathématiques, en sciences, en lecture, et en écriture. Les données proviennent d'entrevues semi-structurées menées auprès des coordonnateurs ( $n = 20$ ) du programme dans toutes les instances géopolitiques et d'un questionnaire envoyé à travers le Canada à l'intention des dirigeants des conseils scolaires participants ( $n = 147$ ). L'impact du PIRS se décrit selon deux dimensions : positif versus négatif et intrinsèque versus extrinsèque. Les résultats révèlent qu'une certaine faiblesse dans la communication entre les responsables du PIRS et les différents acteurs en éducation joue un rôle important sur le degré d'impact. La discussion offre certaines explications et recommandations à l'endroit des programmes d'évaluation des apprentissages.

**Abstract:** This article examines the impact of the School Achievement Indicators Program (SAIP) on the educational system. SAIP is a Canada-wide program involving the large-scale assessment of student achievement in mathematics, science, reading, and writing. Data were collected using semi-structured interviews conducted with the SAIP's jurisdictional coordinators ( $n = 20$ ) and from surveys sent to participating school boards across Canada ( $n = 147$ ). SAIP's impact is described using two dimensions: intrinsic versus extrinsic and positive versus negative. Results revealed that poor communication between the coordinators and various educational stakeholders plays an important role in the level of impact. Recommendations are made for learning evaluation programs.

Les programmes d'évaluation du rendement scolaire sont des manifestations directes du phénomène d'imputabilité en Amérique du Nord (Sicol, 2002) et partout dans le monde (Firestone, May-

---

Correspondance à l'auteure : Marielle Simon, Faculté d'Éducation, Université d'Ottawa, 145 Jean-Jacques Lussier, C.P. 450, Succ. A., Ottawa, ON K1N 6N5; <msimon@uottawa.ca>

rowetz, & Fairman, 1998; Linn, 2000). Ayant donné lieu à diverses réformes en éducation, ce phénomène découle en partie d'un besoin de responsabilisation des systèmes scolaires envers leurs contribuables (Forgette-Giroux & Simon, 2000). Les programmes d'évaluation à grande échelle affectent de diverses façons les administrateurs, les éducateurs, les élèves, les parents, les futurs employeurs, et la communauté (Coleman, 1997; Earl & Torrance, 2000; Gulek, 2003; Lane & Stone, 2002; Ryan, 2002; Smith, Stevenson, & Li, 1998).

Le Canada possède son propre système d'évaluation systématique du rendement des élèves. Sous la responsabilité du Conseil des Ministres de l'Éducation du Canada, le Programme d'Indicateurs du Rendement Scolaire (PIRS) recueille de façon cyclique depuis 1993 des données sur les acquis des élèves de 13 ans et 16 ans en mathématiques, science, lecture, et écriture à travers le Canada. Un rapport d'évaluation du premier cycle de mise en œuvre (Crocker, 1997) offrait 32 recommandations relatives au PIRS. En 1999, le programme termina son second cycle et fut évalué au printemps 2000 à l'aide d'une étude pancanadienne effectuée auprès des populations touchées particulièrement par le programme d'évaluation en sciences mené en 1999 (Forgette-Giroux & Simon, 2000). Avec ses 30 recommandations, l'évaluation du deuxième cycle du PIRS abordait plusieurs points, y compris son impact sur les pratiques d'enseignement et le milieu scolaire en général.

Cet article découle de cette récente évaluation du PIRS et examine le type d'impact exercé par ce programme sur le milieu scolaire selon la perception des coordonnateurs de l'administration du PIRS et selon celle des responsables des conseils, commissions, et districts scolaires des populations qui ont participé au PIRS en sciences en 1999. La première partie traite du contexte de l'étude et décrit le programme et ses modalités. Elle aborde également les évaluations du PIRS par le Conseil des Ministres de l'Éducation, Canada (CMEC) et la recension des écrits sur les impacts des évaluations à grande échelle en milieu scolaire. Les deuxième et troisième parties présentent la méthodologie et les résultats. L'article se termine par une discussion des résultats, des limites et de la contribution de l'étude.

## CONTEXTE

Cette première section s'attarde plus particulièrement à la nature, au rôle, aux buts, et aux objectifs du PIRS. Elle traite également de l'évaluation ponctuelle et régulière de l'efficacité de ce programme et

de la question de l'impact général des programmes d'évaluation des apprentissages à grande échelle tel que rapporté dans les écrits.

## PIRS

Le PIRS représente une forme particulière de programme d'évaluation à grande échelle du rendement scolaire. En effet, les programmes d'évaluation varient entre eux selon leurs finalités, leurs clientèles visées, leurs modalités de mise en œuvre, et leurs enjeux. Certains programmes, comme la Troisième Enquête Internationale en Mathématiques et en Science (TEIMS) et le Programme International pour le Suivi des Acquis (PISA) s'effectuent à l'échelle internationale, s'adressent aux dirigeants des systèmes d'éducation, s'appuient sur un échantillon d'élèves, et servent à comparer les résultats des populations entre elles comme mesure d'efficacité. D'autres programmes, notamment celui de l'Office de la Qualité et de la Responsabilité en Éducation (OQRE), en Ontario, évaluent tous les élèves d'une année scolaire ciblée afin de situer leur rendement par rapport à des normes établies et de renseigner les écoles et les systèmes scolaires sur leur niveau de performance. Ces programmes diffèrent également en termes d'enjeux puisque certains communiquent les résultats des élèves à titre d'indicateurs uniquement, tandis que d'autres présentent les résultats à des fins de certification. Comparativement, les enjeux associés au PIRS s'avèrent minimes car il agit à titre d'indicateur d'efficacité des systèmes d'éducation provinciaux et territoriaux.

Les programmes d'évaluation à base d'indicateurs fournissent habituellement aux gouvernements, aux conseils, et aux écoles des résultats du rendement des élèves dans le but de prendre des décisions politiques, administratives, et pédagogiques fondées sur des données empiriques (Goertz & Duffy, 2003; McGehee & Griffith, 2001; Schiller & Muller, 2000). À cet égard, le PIRS ne fait pas exception. Dans ce contexte l'indicateur se définit comme une statistique individuelle ou composée reliée à un construit de base en éducation et utile au niveau de l'élaboration de politiques (Shavelson, McDonnell, & Oakes, 1991). Un système d'indicateurs devrait normalement mesurer des composantes distinctes et fournir des informations sur la façon dont ces composantes interagissent pour produire un effet.

Au moment de l'étude, le PIRS visait plus exactement à fournir des renseignements sur le niveau de rendement des élèves de 13 ans et 16 ans en sciences, mathématiques, lecture, et écriture dans l'ensemble du Canada et à comparer les résultats obtenus au fil des

années. Puisque les programmes scolaires diffèrent entre provinces et territoires, les résultats du PIRS indiquent dans quelle mesure les élèves canadiens d'un certain âge atteignent un niveau de rendement similaire. Il comporte trois objectifs : (a) établir des priorités en éducation et améliorer les programmes d'études, (b) assurer une éducation de qualité à tous les jeunes Canadiens et Canadiennes par les discussions et l'échange entre les différents acteurs du monde scolaire, et (c) transmettre des informations sur certains indicateurs au public canadien (cités dans Forgette-Giroux & Simon, 2000).

Puisque le cadre de l'échantillonnage du PIRS ne se prête pas à la diffusion de données et de résultats au plan de l'école et des conseils scolaires, les résultats obtenus paraissent aux niveaux national, provincial, et territorial. Le PIRS tient surtout à renseigner les instances du pays sur le statut pancanadien du rendement scolaire dans le but d'ajuster au besoin leurs priorités en éducation. Toutefois, ces renseignements ne s'obtiennent qu'à partir d'une volonté considérable de la part des instances d'abord, mais aussi des systèmes scolaires, des écoles, des membres du personnel enseignant, et tout particulièrement des élèves, à investir temps, argent, et énergie dans la production de résultats valides. Des évaluations régulières de ce programme sont donc essentielles afin de vérifier son efficacité à rencontrer ses objectifs et à demeurer sensible au milieu scolaire.

### Évaluation du PIRS

Un protocole d'entente avec Développement des ressources humaines Canada (DRHC) mandate le CMEC de vérifier régulièrement et systématiquement dans quelle mesure le PIRS atteint ses objectifs. Au plan de l'impact du PIRS, le rapport Crocker en 1997 faisait état d'une faiblesse dans la communication des résultats au public.

Le second cycle d'évaluation du PIRS, terminé en 1999, fut évalué en 2000. Cette évaluation avait pour but de répondre à trois questions : (a) Dans quelle mesure les objectifs du PIRS ont-ils été atteints au cours du deuxième cycle? (b) Dans quelle mesure le PIRS a-t-il tenu compte des recommandations adoptées par le CMEC suite au Rapport Crocker (1997), et (c) Quels conseils spécifiques peut-on recommander en regard des visées, du fonctionnement, et de l'utilisation future du PIRS? Pour ce faire, deux groupes concernés, c'est-à-dire les principaux coordonnateurs provinciaux et territoriaux et les dirigeants des conseils, commissions, et districts scolaires répartis à travers le Canada, ont été appelés à évaluer diverses composantes dont l'impact du

PIRS en milieu scolaire. Ces groupes avaient également participé au PIRS en science en 1999. Le présent article analyse plus précisément les résultats rapportés par ces deux groupes au sujet du cinquième thème, celui de l'impact du PIRS en milieu scolaire.

#### Impact des évaluations à grande échelle en milieu scolaire

Selon Earl et LeMahieu (1997) l'imputabilité en éducation et les évaluations qui en découlent devraient être vues comme une entreprise humaine qui stimule l'ouverture d'esprit et les discussions entre administrateurs, enseignants, élèves, parents, et membres de la communauté sur les informations disponibles dans le but d'établir des plans d'action et de redressement en éducation. Smith, Stevenson, et Li (1998) ajoutent que ce phénomène provoque des réformes en mobilisant tous les intervenants concernés par l'amélioration du système d'éducation. En effet, les évaluations à grande échelle peuvent améliorer l'enseignement, le rendement, et l'égalité des chances en éducation dans la mesure où elles sont élaborées et implantées de façon appropriée (Glatthorn, 1999; Heubert & Hauser, 1999; Snow & Jones, 2001). Au plan du développement professionnel, par exemple, les formations sur la notation organisées par le programme appuient habituellement l'acquisition d'habiletés d'évaluation chez les enseignants (Gambell & Hunter, 2000). Dans la salle de classe, les enseignants, les élèves, les parents, et la communauté voient l'évaluation comme un reflet des apprentissages acquis ou non par les élèves (Earl & Torrance, 2000). Bref, l'évaluation encourage les enseignants à motiver tous leurs élèves à mieux performer et à aider ceux qui sont à risque à se fixer des buts précis (Coleman, 1997).

Plusieurs autres études se sont attardées à l'impact qu'ont ces programmes d'évaluation sur l'enseignant et sur l'apprentissage en salle de classe (Abrams, Pedulla, & Madaus, 2003; Finbarr & Kelly, 2003; Runté, 1998; Schiller & Muller, 2000; Sloane & Kelly, 2003). Pour Levin (2003), il y a impact lorsque le programme affecte les actions et la réflexion des gens et il peut être positif ou négatif, reconnu ou non, et survenir à court ou à long terme. De plus, selon Darling-Hammond (1994), un impact intrinsèque dans le milieu scolaire agit au plan des valeurs et de la motivation des différents acteurs tandis qu'un impact de nature extrinsèque provoque des changements qui s'introduisent suite à des récompenses ou des sanctions externes aux écoles et aux élèves, ce qui sous-entend que les enseignants n'éprouvent pas nécessairement la volonté de changer leurs pratiques.

Un impact négatif souvent soulevé à l'intérieur des études rapportées concerne l'attitude des enseignants qui se montrent réticents aux évaluations externes puisqu'ils y voient une perte de leur autonomie professionnelle (Abrams, Pedulla, & Madaus, 2003; Schiller & Muller, 2000). Ils disent perdre le contrôle de l'activité d'évaluation dans leur propre salle de classe (Gambell & Hunter, 2000; Runté, 1998). Certains se sentent obligés de préparer les élèves aux fins de l'épreuve plusieurs semaines à l'avance et sont tentés de n'enseigner que la matière évaluée lors du test au lieu de se concentrer sur l'ensemble du curriculum (Abrams, Pedulla, & Maddaus; Earl & Torrance, 2000; Gulek, 2003; McGehee & Griffith, 2001; Reed, 2000). Ce genre d'évaluation risque alors de générer un enseignement centré sur les contenus plutôt que sur l'apprenant (Runté).

D'après Shepard (2000), ces évaluations peuvent provoquer un deuxième impact négatif puisque les tâches proposées dans ce type d'évaluation ne s'inscrivent pas nécessairement dans le nouveau paradigme d'enseignement et l'apprentissage adopté en éducation, c'est-à-dire celui fondé sur la théorie socio-constructiviste. Cet état de fait crée un décalage entre ce qui est vécu en classe et ce qu'on retrouve dans l'évaluation. La perception de l'enseignant en ce qui concerne la pertinence des programmes d'évaluation et son degré d'engagement envers ces derniers jouent pourtant un rôle clé dans l'implantation efficace d'un programme d'évaluation en salle de classe (McGehee & Griffith, 2001). Ainsi, plus les évaluations s'éloignent des pratiques d'enseignement actuelles en salle de classe, plus leur contribution à l'amélioration des systèmes scolaires diminue car les enseignants les ignorent tout simplement (Popham, 1999).

Parfois les programmes d'évaluation mènent à des palmarès de rendement au niveau des écoles et des systèmes d'éducation. Dans ce contexte, des études soulignent l'importance d'interpréter avec prudence les résultats des évaluations à grande échelle car ces palmarès ne tiennent pas nécessairement compte des différences sociales et culturelles retrouvées dans certains secteurs urbains (Darling-Hammond, 1994; Gordon & Reese, 1997; Hoffman, Assaf, & Paris, 2001). En effet, les résultats d'évaluation dépendent souvent de divers facteurs externes à la salle de classe, comme par exemple les caractéristiques de la famille (Caldas & Bankston, 1999; Sicoly, 2002), ce qui empêche de comparer les écoles ou même les classes entre elles. À cet égard, certains proposent des modèles alternatifs où l'examen des résultats se fait en concordance avec les tâches réalisées en classe par les élèves tout au long de l'année scolaire (Buckendahl, Impara, &

Plake, 2002; Rose & Gallup, 2001). Une évaluation continue à l'aide du portfolio, par exemple, permettrait d'avoir une idée du progrès de l'élève du début jusqu'à la fin de l'année et fournirait des données complètes sur les facteurs internes et externes à la salle de classe à considérer lors de l'interprétation des résultats (LeMahieu, Gitomer, & Eresch, 1995).

Les écrits recensés jusqu'ici font état des impacts tantôt positifs ou négatifs engendrés par les programmes d'évaluation. La présente étude s'intéresse plus particulièrement aux questions suivantes en ce qui concerne l'impact. Dans quelle mesure le PIRS, en tant que programme d'évaluation de système, a-t-il un impact en milieu scolaire? De quel type d'impact s'agit-il? Ce dernier est-il fonction d'une simple interprétation des résultats des élèves aux tests à des fins politiques ou administratives ou résulte-t-il de la possibilité accordée aux enseignants et aux différents membres de la communauté de discuter et de réfléchir sur les pratiques d'enseignement en classe et sur le rendement des élèves? L'enquête réalisée auprès de deux des partis concernés offre certaines réponses à ces questions. Les pôles positif versus négatif et intrinsèque versus extrinsèque (Darling-Hammond, 1994) constituent le cadre conceptuel d'interprétation de l'impact du PIRS sur le milieu scolaire.

## MÉTHODOLOGIE

Au printemps 2000, une étude pancanadienne a été réalisée auprès des populations touchées par le programme d'évaluation en sciences mené en 1999, la dernière année de la mise en oeuvre du deuxième cycle du PIRS, dans le but de déterminer son efficacité depuis l'évaluation du premier cycle. Les sections suivantes présentent les participants à l'étude, l'instrumentation, et la procédure suivie.

### Participants

Les participants à l'étude proviennent de deux groupes : (a) les principaux coordonnateurs provinciaux et territoriaux, et (b) les dirigeants des conseils, commissions, et districts scolaires répartis à travers le Canada. Le premier groupe inclut plus spécifiquement le coordonnateur national du CMEC, tous les principaux coordonnateurs des différentes populations canadiennes, des experts impliqués dans le programme, et certains groupes concernés par le PIRS, tels la Fédération canadienne des enseignantes et des enseignants, pour un total

de 27 répondants. Ces derniers comprennent huit participants venant de l'Ouest, sept du Centre, neuf de l'Est, et enfin, trois du Nord du Canada. Ils proviennent des dix provinces et des trois territoires, de milieux francophones et anglophones, minoritaires et majoritaires. Le nombre d'années de participation des coordonnateurs au PIRS varie de un an à dix ans.

Le deuxième groupe représente l'ensemble des 412 conseils, commissions, ou districts scolaires (unités scolaires) des dix provinces et de deux territoires. L'évaluation fut administrée à un échantillon de 31 460 élèves de 13 et 16 ans, tiré de 17 instances canadiennes actuelles y compris les provinces, deux territoires, et les communautés linguistiques. Aux fins de l'évaluation du PIRS, les chercheurs ont sollicité, dans chacune des unités scolaires, une personne ayant joué un rôle actif dans le programme, comme par exemple le directeur général ou d'éducation, le surintendant, le coordonnateur ou l'agent de liaison du PIRS, le conseiller pédagogique ou la direction d'une école. Seul le coordonnateur de la province ou du territoire ne pouvait répondre au questionnaire, étant déjà associé au premier groupe de participants.

### Instrumentation

Deux types d'instruments ont servi à mener l'étude : des entrevues semi-structurées et un questionnaire écrit, distribué par la poste. Les questions des deux instruments se rapportent à cinq thèmes relatifs au PIRS : (a) sa nature et sa place, (b) ses buts et objectifs, (c) son fonctionnement, (d) ses qualités techniques, et enfin, (e) son impact en milieu scolaire.

Le questionnaire comprend au total 13 questions et 73 sous-questions. Trois questions et leurs sous-questions concernent directement l'impact du programme. Elles traitent plus précisément de l'impact du PIRS sur l'établissement des priorités dans les programmes, les programmes d'étude, la perception du public de la qualité de l'éducation, les pratiques d'enseignement, et les pratiques d'évaluation. Les participants devaient répondre en utilisant l'échelle suivante *beaucoup d'impact, plus ou moins d'impact, et pas du tout d'impact*. Des questions ouvertes permettaient également aux répondants de s'exprimer librement sur les aspects de leur choix. Deux experts praticiens, un francophone et un anglophone, ont validé le questionnaire.

Par ailleurs, l'entrevue semi-structurée comporte 26 sous-questions réparties à l'intérieur des cinq thèmes mentionnés. En ce qui a



trait au dernier thème, l'entrevue réfère à l'impact du PIRS sur les programmes d'études, sur la façon de renseigner le public, sur les pratiques d'évaluation, sur les initiatives de formation, et sur les initiatives de recherche.

### Procédure

En juin 2000, les questionnaires furent distribués par la poste aux 412 conseils, districts, ou commission scolaires. Deux rappels téléphoniques furent effectués, un en juillet et un en août, dans le but d'assurer une bonne représentation. Des approches statistiques quantitatives principalement descriptives et une analyse de contenu constituèrent les principales analyses des données.

Les entrevues menées exclusivement par les chercheurs auprès du premier groupe de participants ( $n = 27$ ) se sont déroulées au téléphone ou par mode audioconférence lorsque certains membres devaient être regroupés. Celles-ci furent réalisées après avoir fait parvenir préalablement les questions aux participants et reçu leur consentement écrit. Des notes furent prises durant six entrevues pour lesquelles la permission d'enregistrer n'avait pas été obtenue. Les entrevues d'une durée d'environ 50 à 90 minutes furent transcrites verbatim et analysées par une troisième personne possédant de l'expérience avec le logiciel Nvivo (1998–99). Chaque entrevue fut codée dans sa totalité. Ce codage comportait deux types de catégories : celle faite à partir des questions d'entrevue et celle développée de façon inductive selon les réponses offertes par les participants. Une deuxième analyse fut appliquée afin de synthétiser les diverses réponses en sous thèmes.

### RÉSULTATS

Cette section présente les résultats de l'analyse des données quantitatives et qualitatives recueillies à partir des deux instruments. En premier lieu, les réponses obtenues à l'aide du questionnaire permettent d'examiner l'impact du PIRS sur différents aspects en milieu scolaire. À cet égard, l'étude a obtenu un taux de réponse de 36 %, c'est-à-dire que 147 conseils scolaires sur les 412 sollicités ont accepté d'y répondre. Malgré ce taux de réponse relativement faible, les répondants proviennent autant de petits que de grands conseils, commissions, ou districts scolaires, autant de régions éloignées et rurales que des régions centrales et urbaines, autant des milieux de langue française et anglaise de partout au Canada, ce qui constitue une bonne représentativité des caractéristiques des diverses popula-

tions concernées. L'article ne rapporte aucun résultat de croisement des données selon les diverses sous-populations afin de respecter l'anonymat des groupes de faible taille.

Les résultats montrent que seulement 22 % des dirigeants des unités scolaires voient le programme comme ayant beaucoup d'impact sur l'établissement de priorités dans les programmes scolaires alors que 49 % jugent que le PIRS a plus ou moins d'impact sur ce plan. Des résultats semblables ont été obtenus en ce qui a trait à l'impact direct du PIRS sur les programmes d'études, sur la perception du public de la qualité de l'éducation, et sur les pratiques d'enseignement et d'évaluation en classe. Il semblerait donc qu'environ la moitié des répondants considèrent que le PIRS a plus ou moins d'impact en milieu scolaire de façon générale.

Selon l'analyse des réponses aux questions ouvertes, les participants considèrent que le PIRS jouirait d'un plus grand impact en milieu scolaire s'il fournissait des résultats davantage contextualisés, s'il partageait des stratégies d'enseignement exemplaires, et s'il mettait à la disposition des intervenants en éducation certaines questions actuelles des épreuves. Les répondants se plaignent du manque de rétroaction, désirent une communication plus soutenue de la part du programme, et suggèrent ainsi de rendre les résultats du PIRS plus accessibles au public en général. Les commentaires suivants illustrent cette problématique : « En ce qui nous concerne, nous n'avons pas reçu toutes les informations par rapport aux résultats et dans la pratique, cela n'a rien changé, » « J'insiste à nouveau sur l'importance de mieux faire connaître le PIRS comme outil d'amélioration des apprentissages, » et « We haven't received any feedback. »

En deuxième lieu, les données des entrevues menées auprès des coordonnateurs et des experts du PIRS contribuent à dresser un portrait plus complet de la situation. D'abord, l'impact du PIRS sur les programmes d'études est perçu comme étant plutôt positif selon ce groupe de répondants. Ces derniers rapportent que des enseignants ont intégré dans leur pratique certains éléments du programme d'études portant sur les modèles de résolution de problème en mathématiques ou sur les composantes pratiques en sciences. Plusieurs coordonnateurs mentionnent également la mise en place, suite à l'administration du PIRS en lecture et en écriture, de comités d'enseignants et de chercheurs dans leur milieu respectif afin d'améliorer l'enseignement de la langue. Ils ont aussi constaté l'établissement de clubs de sciences dans certaines écoles grâce au PIRS sciences.

Toutefois, de façon générale, selon les coordonnateurs et experts du PIRS, une exploitation plus approfondie des données servirait davantage à l'amélioration des programmes scolaires. Toujours selon ces répondants, ce programme n'a eu que peu d'impact quant à l'objectif de renseigner le public au sujet de la qualité de l'éducation, faute de diffusion et d'accessibilité aux résultats. Pourtant, à leur avis, les divers intervenants s'intéresseraient aux résultats du PIRS.

La majorité des coordonnateurs estiment que le PIRS a également des effets positifs en ce qui a trait aux pratiques d'évaluation puisque les enseignants adoptent souvent les modèles d'évaluation présentés. Les échelles d'appréciation, par exemple, constituent selon eux des ressources pratiques pour les enseignants. Toutefois, certains coordonnateurs reconnaissent que plusieurs enseignants ne voient dans les résultats qu'un classement de leur école et développent ainsi peu d'attitudes réflexives face à leurs pratiques. Ils remarquent que ces derniers perçoivent le programme comme étant une source de compétition entre les établissements. En effet, au lieu d'aborder le programme comme un indicateur et comme une source de changement, ces enseignants ressentent une certaine anxiété face au programme. De l'avis des coordonnateurs, la normalisation du programme soulèverait des réactions plutôt négatives chez le personnel enseignant. Finalement, malgré l'apport des séances de notation en termes de perfectionnement aux enseignants participants, les coordonnateurs s'entendent pour dire que le PIRS n'a pas un grand impact sur leur développement professionnel et les initiatives de formation.

Ces résultats révèlent des impacts parfois positifs, parfois négatifs du PIRS sur de multiples aspects touchant le milieu scolaire. La discussion suivante approfondit notamment la nature des différents impacts de ce programme.

## DISCUSSION

Cette discussion s'articule autour de la nature de l'impact que peuvent avoir des programmes d'évaluation des apprentissages et des possibilités d'accroître leur impact positif intrinsèque.

### Nature de l'impact

La mise en place de clubs de sciences et de comités locaux regroupant praticiens et chercheurs intéressés à l'enseignement de la langue

constitue, selon Darling-Hammond (1994), un exemple concret d'un impact pédagogique positif *intrinsèque* découlant du programme d'évaluation. Cet impact se construit à l'intérieur de l'école en collaboration avec différents praticiens et les membres de la communauté. Par ailleurs, l'adoption, par les ministères d'éducation, des outils d'enseignement et d'évaluation privilégiés par le programme se veut plutôt un indicateur d'impact positif *extrinsèque*. En d'autres mots, en intégrant dans les programmes d'études provinciaux les modèles du programme d'évaluation, les systèmes d'éducation encouragent fortement leurs enseignants à les adopter progressivement à l'intérieur de leurs pratiques d'enseignement et d'évaluation.

Les impacts négatifs du programme semblent se scinder également en deux selon leur nature intrinsèque ou extrinsèque. Au plan intrinsèque, les résultats obtenus ne constituent pas nécessairement la base d'informations souhaitées sur laquelle les enseignants s'appuient pour améliorer leurs pratiques d'enseignement et d'évaluation. Plusieurs praticiens perçoivent plutôt un tel programme comme étant une source de compétition entre établissements. Pour d'autres, le palmarès a pour effet d'ajuster les pratiques d'enseignement et d'évaluation à des fins politiques plutôt que d'apprentissage, ce qui témoigne d'un impact plutôt extrinsèque.

Finalement, de l'avis des répondants, bon nombre d'enseignants perçoivent ce type de programme comme un test standardisé ayant pour but premier de vérifier leur degré d'adhésion aux attentes des différents ministères de l'éducation. Cette attitude négative provoquée par un manque d'information, par une méconnaissance des objectifs réels, et par la méfiance envers le programme se dit également extrinsèque. Dans ce cas, les enseignants ne s'engagent que de manière superficielle lors de l'administration du programme et n'intègrent dans leur pratique aucun de ses aspects.

Les objectifs du PIRS font référence à un impact indirect et à long terme plutôt que direct et à court terme sur les pratiques. Ainsi, certains responsables de conseils, commissions, et districts scolaires ne sont peut-être pas conscients des effets concrets du programme dans leur milieu scolaire, puisque le changement des pratiques s'opère davantage sur une longue période de temps. De plus, puisqu'il ne sanctionne, ne récompense, et n'a pas de conséquences directes ni sur l'avenir des élèves individuellement, ni sur les ressources attribuées aux établissements, le PIRS ne bouleverse pas les pratiques de façon directe et impérative. Pourtant, il sert à favoriser l'amélioration

des programmes d'études et à assurer une éducation de qualité en valorisant l'échange.

#### Possibilités d'impact positif intrinsèque

Selon Snow et Jones (2001), les programmes d'évaluation à grande échelle peuvent conduire à une amélioration s'ils constituent un stimulus de changement dans le système d'éducation ou s'ils fournissent une base pour modifier le curriculum et l'enseignement, pour favoriser le développement professionnel et pour mieux répartir les ressources. Les résultats de cette étude démontrent, en effet, que le PIRS par exemple, devrait se doter davantage de moyens d'informer et de stimuler les changements accessibles par les différents acteurs du milieu scolaire, en particulier les enseignants et les directions d'école. Des recommandations avaient d'ailleurs été émises à cet effet à l'intérieur du Rapport Crocker, telles « Inciter certaines instances à déployer davantage d'effort à la distribution des rapports et encourager les médias locaux à faire état des résultats avec exactitude et esprit d'analyse » (1997, p. 85). En 2000, Forgette-Giroux et Simon proposaient à leur tour, à titre de recommandation, de publier des fascicules thématiques s'inspirant des catégories principales du questionnaire contextuel du PIRS, tels la motivation des élèves, la gestion de la classe, et les programmes d'études et leur mise en relation avec les résultats du rendement scolaire. Ainsi, le PIRS produirait un impact positif intrinsèque dans les pratiques en informant mieux et en misant sur l'échange et le développement professionnel.

Afin de donner lieu à un impact de ce genre, le PIRS, comme programme d'évaluation, doit trouver des solutions qui amènent les différents partis concernés à discuter des moyens d'améliorer le système sans créer de la tension ou un effet de compétition. Tout en demeurant un indicateur, le PIRS doit développer ses deux premiers objectifs en facilitant sa communication avec les divers intervenants. Ceci s'applique en particulier chez les enseignants qui participent volontiers aux changements initiés en éducation lorsqu'ils perçoivent qu'ils ont un certain contrôle sur le processus décisionnel (Seyfarth, Ysseldyke, & Thurlow, 1998). Des études démontrent que les enseignants et les autres membres de conseils scolaires retirent beaucoup de satisfaction lorsqu'ils s'impliquent à toutes les étapes des programmes d'évaluation à grande échelle, de leur conception jusqu'à la correction des épreuves (Flockton & Crooks, 1997; Hunter & Gambell, 2000; Simon & Forgette-Giroux, 2002).

Dans cette optique, le PIRS peut s'inspirer d'autres programmes d'évaluation qui fournissent aux enseignants un plus grand soutien et davantage d'informations sur les instruments utilisés. L'OQRE en Ontario, par exemple, offre du matériel d'appui à l'enseignant afin de l'aider à préparer les élèves aux tests de troisième et sixième années en mathématiques, lecture, et écriture, au test de mathématiques neuvième année, et au Test Provincial de Compétences Linguistiques (TPCL). Ce matériel, conceptualisé par des enseignants, des conseillers pédagogiques, et des directions d'écoles, permet de préparer les élèves en classe avant l'administration des véritables tests. Dans ce matériel d'appui, l'OQRE préconise toutefois la modélisation de l'enseignement sur le curriculum et non pas uniquement sur le test. En ce qui concerne le TPCL, les enseignants reçoivent des exemples d'annotations de textes d'élèves, un exemple de cahier présenté aux élèves, ainsi qu'un guide de préparation.

La Troisième enquête internationale sur les mathématiques et les sciences (TEIMS-R) révisée et administrée en 1999 propose également un document informatif sur ses tests comprenant des exemples d'items et des échantillons de réponses annotées. Ces initiatives procurent des outils qui permettent aux enseignants de se préparer et leur donnent l'occasion de se familiariser avec de nouveaux contenus et avec des nouvelles méthodes d'enseignement et d'évaluation afin de pouvoir éventuellement se les approprier.

Le Programme international pour le suivi des acquis des élèves (PISA), une autre évaluation à laquelle le Canada participe, évalue la compréhension de l'écrit et la culture mathématique et scientifique. Dans son effort de communiquer les résultats et le fonctionnement du programme, le PISA publie des bulletins faisant état des résultats aux tests et de l'information concernant l'évolution du programme. Ces bulletins fournissent des réponses aux questions importantes posées par différents acteurs du milieu scolaire, telles celles sur la présentation des résultats sous forme d'échelles de compétences. Un fascicule décrit le cadre conceptuel sur lequel se fondent les évaluations du PISA, en précisant les domaines évalués et l'objet des évaluations. Cette publication met à jour les stratégies de mesure des connaissances et des compétences des élèves et présente un modèle d'évaluation dont peuvent s'inspirer les enseignants. Elle dévoile également des informations concernant les caractéristiques des épreuves, l'échelle des résultats, et les contenus évalués.

Bref, en s'inspirant du matériel, des moyens et des outils conçus par d'autres programmes d'évaluation, le PIRS pourrait assurer l'atteinte

de ses objectifs de manière plus efficiente. Tel que démontré dans cet article, le PIRS présente des impacts intrinsèques et extrinsèques, positifs et négatifs, sur le milieu scolaire et les pratiques en classe. Enfin, le PIRS possède un potentiel d'impact positif intrinsèque dans le milieu scolaire qui, jusqu'ici, n'a pas été développé à sa juste mesure.

#### CONTRIBUTIONS ET LIMITES DE L'ÉTUDE

Dans le contexte actuel où des évaluations à grande échelle et des programmes d'indicateurs sont implantés à l'intérieur d'organisations supra-gouvernementales, de gouvernements, d'universités, et d'entreprises (Turgeon & Bernatchez, 2003), cette étude met en lumière les types d'impacts que peut avoir un tel programme en milieu scolaire. Elle dresse un portrait des impacts positifs, négatifs, intrinsèques, et extrinsèques du PIRS selon les principaux coordonnateurs provinciaux et territoriaux ainsi que les dirigeants des unités scolaires réparties à travers les diverses instances canadiennes. Elle enrichit les connaissances établies au sujet de l'impact des évaluations à grande échelle en proposant des suggestions d'amélioration spécifiques et concrètes afin d'amener tous les intervenants du milieu scolaire et le public en général à s'intéresser davantage au PIRS, à s'engager lors de son administration, et à s'inspirer des résultats lors de la mise en place de plans de redressement des systèmes d'éducation. L'étude fait également ressortir le fait qu'une meilleure communication entre les responsables du PIRS et les divers groupes concernés par ce programme, ainsi qu'une plus grande implication des acteurs du milieu scolaire à toutes les étapes de mise en œuvre du programme mèneraient à un impact davantage intrinsèque. Finalement elle fournit des balises d'évaluation de tout programme d'indicateurs ou d'évaluation à grande échelle. Sa méthodologie mixte permet de recueillir des données complémentaires, c'est-à-dire, certaines générales, d'autres davantage spécifiques.

Toutefois l'examen de l'impact du PIRS ne constituait qu'une de cinq composantes de l'évaluation du deuxième cycle du programme et l'instrumentation ne permettait pas de pousser plus loin les analyses. Ceci représente la principale limite de l'étude. Une autre faiblesse se situe au niveau du taux relativement faible des répondants malgré leur représentativité et l'absence de la voix des enseignants à l'endroit de l'utilité des programmes d'évaluation. Le choix de ne faire appel qu'à seulement trois options dans les items du questionnaire écrit a donné lieu à des résultats imprécis. Des recherches subséquentes

auraient avantage à approfondir la nature des impacts et leur degré de présence chez les pratiques d'enseignants en questionnant divers groupes, y compris les enseignants eux-mêmes. Il y a lieu également de comparer ces résultats en fonction des régions géographiques et linguistiques du Canada afin de pouvoir encore mieux concrétiser les actions que le programme d'évaluation pourrait adopter et ainsi d'influencer davantage de manière intrinsèque et positive les pratiques d'enseignement et d'évaluation en salle de classe. Finalement, les chercheurs devraient également s'intéresser à mener des méta-analyses sur la nature de l'impact des divers programmes existants en matière d'évaluation des apprentissages à grande échelle dans le but de produire un portrait cohérent et complet de la situation actuelle. Les quelques études, thèses, recherches, et colloques ayant pour but d'examiner ce thème de plus près se font habituellement de manière ponctuelle et dispersée. Des études synthèses amèneraient les concepteurs à repenser la validité des résultats de ces programmes en termes du niveau d'engagement et de la participation des partis concernés, en particulier les élèves, ou en termes du degré d'harmonisation entre la définition accordée, par les différents intervenants, aux concepts de lecture, d'écriture, de sciences, ou de mathématiques. Elles permettraient ainsi de proposer des recommandations aux plans conceptuels et structurels relatifs à ce type d'initiatives.

## REMERCIEMENTS

Cette étude a été rendue possible grâce à une subvention du Conseil des ministres de l'éducation, Canada.

## RÉFÉRENCES

- Abrams, L.M., Pedulla, J.J., & Madaus, G.F. (2003). Views from the classroom: Teachers' opinions of statewide testing programs. *Theory into Practice, 42*, 18–29.
- Buckendahl, C.W., Impara, J.C., & Plake, B.S. (2002). District accountability without a state assessment: A proposed model. *Educational Measurement: Issues and Practice, 21*, 6–19.
- Caldas, S.J., & Bankston, C. L. III. (1999). Multilevel examination of student, school, and district-level effects on academic achievement. *Journal of Educational Research, 93*, 91–100.



- Coleman, J.S. (1997). Output-driven schools: Principles of design. Dans J.S. Coleman, B. Schneider, S. Plank, K.S. Schiller, R. Shouse, H. Wang, & S.-A. Lee. (Éds.), *Redesigning American education* (pp. 13–38). Boulder, CO : Westview.
- Crocker, R.K. (1997). *Study of the School Achievement Indicator Program*. Toronto : Council of Ministers of Education, Canada (CMEC).
- Darling-Hammond, L. (1994). National standards and assessments: Will they improve education? *American Journal of Education* 102, 478–510.
- Earl, L., & LeMahieu, P. (1997). Rethinking assessment and accountability. Dans A. Hargreaves (Éd.), *Rethinking educational change with heart and mind, 1997 ASCD yearbook*. Alexandria, VA : Association for Supervision and Curriculum Development.
- Earl, L., & Torrance, N. (2000). Embedding accountability and improvement into large-scale assessment: What difference does it make? *Peabody Journal of Education*, 75, 114–141.
- Finbarr, C., & Kelly, A.E. (2003). Issues in high-stakes testing programs. *Theory into Practice*, 42, 2–17.
- Firestone, W., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: The effect of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis*, 20, 95–113.
- Flockton, L., & Crooks, T. (1997). *Reading and speaking assessment results 1996*. Ministry of Education, New Zealand. Dunedin : Educational Assessment Research Unit, University of Otago.
- Forgette-Giroux, R., & Simon, M. (2000). *Évaluation du deuxième cycle du Programme d'indicateurs de rendement scolaire (PIRS)*. Rapport soumis au Conseil des Ministres de l'Éducation, Canada (CMEC).
- Gambell, T., & Hunter, D. (2000, mai). *Professionalism, professional development, and teacher participation in scoring of large-scale assessment*. Communication présentée au 27ième congrès de la Canadian Society for the Study of Education à Edmonton, Alberta.
- Glatthorn, A. (1999). Curriculum alignment revisited. *Journal of Curriculum and Supervision*, 15, 26–34.

- Goertz, M., & Duffy, M. (2003). Mapping the landscape of high-stakes testing and accountability programs. *Theory into Practice, 42*, 4–11.
- Gordon, S.P., & Reese, M. (1997). High stakes testing: Worth the price? *Journal of School Leadership, 7*, 345–368.
- Gulek, C. (2003). Preparing for high-stakes testing. *Theory into Practice, 42*, 51–58.
- Heubert, J.P., & Hauser, R. M. (1999). *High-stakes: Testing for tracking, promotion, and graduation*. Washington, DC : National Academy.
- Hoffman, J.V., Assaf, L.C., & Paris, S.G. (2001). High-stakes testing in reading: Today in Texas, tomorrow? *The Reading Teacher, 54*, 484–492.
- Hunter, D.M., & Gambell, T. J. (2000). Incorporating stakeholders in standard setting: What's at stake? *Canadian Journal of Program Evaluation, 15*, 83–109.
- Lane, S., & Stone, C.A. (2002). Strategies for examining the consequences of assessment and accountability programs. *Educational Measurement: Issues and Practice, 21*, 23–31.
- LeMahieu, P.G., Gitomer, D.H., & Eresh, J.T. (1995). Portfolios in large-scale assessments: Difficult but not impossible. *Educational Measurement: Issues and Practice, 14*, 11–16.
- Levin, B. (2003, août). *Helping research in education to matter more: A brief discussion and proposals for action*. Communication présentée à The University of Manitoba, Winnipeg, MB. Répéré de <[www.cea-ace.ca/media/en/Helping\\_Research\\_in\\_Education\\_Matter\\_BLevin.pdf](http://www.cea-ace.ca/media/en/Helping_Research_in_Education_Matter_BLevin.pdf)>.
- Linn, R.L. (2000). Assessments and accountability. *Educational Researcher, 29*, 4–15.
- McGehee, J.J., & Griffith, L.K. (2001). Large-scale assessments combined with curriculum alignment: Agents of change. *Theory into Practice, 40*, 137–144.
- Popham, W.J. (1999). Where large scale educational assessment is heading and why it shouldn't. *Educational Measurement: Issues and Practice, 18*, 18–26.

- Reed, S. (2000). The too often neglected aspects of state assessment: A state perspective. *State Education Standard, 1*, 12–16.
- Rose, L.C., & Gallup, A.M. (2001). The 33rd annual Phi Delta Kappa/Gallup Poll of the public's attitudes toward the public schools. *Phi Delta Kappan, 83*, 41–58.
- Runté, R. (1998). The impact of centralized examination on teacher professionalism. *Canadian Journal of Education, 23*, 166–181.
- Ryan, K. (2002). Assessment validation in the context of high-stakes assessment. *Educational Measurement: Issues and Practices, 21*, 7–15.
- Schiller, K.S., & Muller, C. (2000). External examination and accountability, educational expectations, and high school graduation. *American Journal of Education 108*, 73–102.
- Seyfarth, A., Ysseldyke, J.E., & Thurlow, M.L. (1998). *An analysis of perceived desirability, feasibility, and actual use of specific criteria for large-scale assessment and accountability systems*. Minneapolis, MN : University of Minnesota, National Center on Educational Outcomes.
- Shavelson, R.J., McDonnell, L., & Oakes, J. (1991). What are educational indicators and indicator systems? *Practical Assessment, Research & Evaluation, 2*. Répéré le 30 août 2005 de <<http://PAREonline.net/getvn.asp?v=2&n=11>>.
- Shepard, L.A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*, 1–14.
- Sicol, F. (2002). What do school-level scores from large-scale assessments really measure? *Educational Measurement: Issues and Practice, 21*, 17–27.
- Simon, M., & Forgette-Giroux, R. (2002). Senior school board officials' perception of a national achievement assessment program. *Education Policy Analysis Archives, 10*. Répéré le 12 février 2004 de <<http://epaa.asu.edu/epaa/v10n46.html>>.
- Sloane, K.C., & Kelly, A.E. (2003). Issues in high-stakes testing programs. *Theory into Practice, 42*, 12–17.

- Smith, M., Stevenson, D., & Li, C. (1998). Voluntary national tests would improve education. *Educational Leadership*, 55, 42–44.
- Snow, C.E., & Jones, J. (2001). Making a silk purse. *Education Week*, 25, 41.
- Turgeon, J., & Bernatchez, J. (2003). Les données secondaires. Dans B. Gauthier (Éd.), *Recherche sociale: De la problématique à la collecte de données*. Sainte-Foy, QC : PUL.

**Marielle Simon** est professeure titulaire en mesure et évaluation à la Faculté d'éducation de l'Université d'Ottawa depuis 1994. Son programme de recherche actuel comprend l'étude des pratiques de notation en salle de classe, l'évaluation formative, et les enquêtes provinciales, nationales, et internationales du rendement scolaire.

**Catherine Turcotte** est étudiante au doctorat en psychopédagogie à la Faculté d'éducation de l'Université d'Ottawa.

**Renée Forgette-Giroux** est professeure titulaire et vice-doyenne aux programmes à la Faculté d'éducation de l'Université d'Ottawa. Ses études portent sur l'évaluation des apprentissages en salle de classe et à grande échelle. Elle s'intéresse particulièrement aux politiques, aux pratiques, aux stratégies, et aux qualités des instruments d'évaluation.