

Performance data about schools and school systems can be positive or negative but they are unlikely to be neutral. However, we know from experience that accountability systems and performance assessment often seem more concerned with apportioning blame for school failure than contributing to school improvement. If we truly believe that schools are the centre of change, then we must build our systems from the "inside-out." We must seek ways to provide information systems that are useful to schools and this means involving schools in the design, testing, implementation and monitoring of these systems. Such an approach is expensive and time consuming but any system that does not provide for such involvement is doomed to failure.

As we have concluded elsewhere: "For us, the purpose of self-evaluation is not essentially about public accountability or about informing parental choice. It is about strengthening the quality of learning and teaching and the capacity of the school to support these effectively. If, however, self-evaluation can serve those other purposes too it may do a great service to public and parental understanding of what good schools are really all about."⁷

Endnotes

1. M. Barlow and H.J. Robertson, *Class Warfare: The Assault on Canada's Schools* (Toronto: Key Porter Books, 1994); S.B. Lawton, *Busting Bureaucracy to Reclaim our Schools* (Montreal: IRPP, 1995).

2. P. Coleman, *Learning About Schools: What Parents Need To Know and How They Can Find Out* (IRPP, 1994) p. 43.

3. See J. MacBeath, B. Boyd, J. Rand and S. Bell, *Schools Speak for Themselves: Towards a Framework for Self-evaluation* (London: The National Union of Teachers, 1996).

4. A. Bacon, "The Teachers' Perspective on Accountability," *Canadian Journal of Education*, Vol. 20 (1995), pp. 85-91, 85.

5. See J.E. Ysseldyke and M.E. Thurlow, *Self-study Guide to the Development of Educational Outcomes and Indicators* (Minneapolis, MN: National Center on Educational Outcomes, 1993).

6. L. Stoll and D. Fink, *Changing Our Schools: Linking School Effectiveness and School Improvement* (Buckingham UK: Open University Press, 1996), pp. 166-67.

7. W.J. Smith, L. Moos and J. MacBeath, "School Self-assessment: Quality in The Eye of the Stakeholder," Paper presented at the annual conference of the American Educational Research Association, San Diego, CA (April 1998) p. 16.

William J. Smith is the Director of the Office of Research on Educational Policy (OREP) and an Assistant Professor in the Department of Educational Studies at McGill University (wj-smith@cel.lan.mcgill.ca).

by J. Douglas Willms and Elizabeth A. Sloat

SCHOOLING OUTCOMES FOR YOUTH AT RISK

De nombreux jeunes sont exposés à divers problèmes — décrochage, délinquance, problèmes de santé mentale et physique. Dans le passé, c'est l'État qui assurait les programmes d'intervention destinés à aider les jeunes à faire face à ces difficultés, mais à mesure que s'accroît la demande de tels services et que diminuent les ressources publiques, des pressions s'exercent en vue d'amener les collectivités locales à fournir le plus tôt possible les services et le soutien nécessaires. Et les bailleurs de fonds veulent savoir si les programmes qu'on leur demande de financer produisent effectivement les résultats escomptés. Les auteurs passent en revue les démarches en vue de mettre au point un modèle pratique et relativement peu coûteux permettant d'évaluer les programmes communautaires destinés à ces jeunes.

At-risk youth in our schools

An increasing number of youth must cope with a variety of negative life experiences while growing up, such as racial and ethnic prejudice, severe learning and behaviour problems, inadequate parenting, family vio-

lence and poverty. Consequently, these youth are “at risk” in many ways: they are more likely to drop out of school, face prolonged unemployment, participate in delinquent activities and experience mental and physical health problems. The growing number of youth deemed at risk poses significant challenges for educators and the school system in general. An Ontario study found that in a sample of 3,000 children, over five percent were classified as having a conduct disorder.¹ In the US, roughly 40 percent of students at the elementary and secondary school levels are considered to be at-risk of educational failure.² Not only is the prevalence of children with learning and behavioural disorders high, but the children who suffer from these conditions have, in addition to antisocial behaviours, associated impairments such as poor school performance and difficulties in getting along with peers, teachers and parents.³ Thus, these children present serious challenges to teachers and other students.

The important question for administrators and policy makers at the school, district and provincial levels is whether their response to this problem is meeting the needs of students with behaviour and learning problems, while ensuring that all students receive a quality education. There is considerable evidence that the way schools respond makes a difference, particularly for students from less advantaged backgrounds. For example, a recent study of the literacy skills of Canadian youth⁴ indicated that youth in Quebec and the Prairie provinces fared better in their literacy skills than youth in other provinces. An intriguing finding was that youth whose parents had high socioeconomic status fared relatively well in all provinces; whereas youth from less advantaged backgrounds varied considerably in their literacy skills among the 10 provinces. The success of Quebec and the Prairie provinces was largely attributable to the higher scores of youth with less advantaged backgrounds. These findings raise questions about whether schools, communities, school districts and provincial boards of education can improve schooling outcomes by improving their programs for youth at risk.

The chief response by the schools — retaining at-risk children for one or more grades — does not work.⁵ Pull-out programs, whereby students at risk are removed from the regular classroom for remedial instruction, have been largely abandoned because of their damaging “labeling” effects,⁶ and because the skills learned in remedial instruction often fail to bolster regular classroom learning.⁷ However, recent evaluations suggest that in-class remedial efforts are also ineffective.⁸ Moreover, because even one child with a conduct disorder can negatively affect the learning climate of a classroom or school, many parents seek safe havens for their children through private schooling, charter schools, enrichment programs or French immersion classes. Yet these solutions segregate advantaged from disadvantaged students and worsen the chances of success for children at risk.⁹ Consequently, many educational leaders have

called for the *restructuring* of schools in fundamental ways. They would like to see smaller learning communities, where students are not tracked or streamed. The restructured schools would offer a core academic program, achieve greater parental involvement, give teachers and students greater autonomy, and evaluate students and schools in more authentic ways.¹⁰

A number of school-based models have been proposed for restructuring schools in the US. Three prominent models, *Accelerated Schools*,¹¹ *Success for All Schools*¹² and the *School Development Program*,¹³ have been adopted by schools throughout the US, and to some extent worldwide. All three models emphasize prevention over remediation, a highly contextualized curriculum with strong components in reading and language, parental participation and site-based governance structures.¹⁴ These schools provide a more supportive environment for disadvantaged children. In Canada, a number of schools have embarked on restructuring efforts, and generally the philosophy of “middle schools” (which serve children in grades 6 to 8) is consistent with the restructuring movement. However, there is no widely accepted Canadian model for restructuring schools or for supporting children at risk.

Whether intervention programs are provided by the schools or by other social and judicial agencies, most of them receive at least some of their funding from the provincial or federal governments. The escalating demand for services, however, combined with a decline in government resources, means that local communities must provide the needed services. Consequently, private and corporate donors, community agencies and municipal governments are increasingly being called upon to help meet the needs of vulnerable youth. This shift has also increased the call for accountability. Schools and community agencies are being asked to demonstrate the effectiveness of their programs in clearly measurable terms. Sponsors want to be assured that the programs they support truly achieve their purported benefits and long-term outcomes.

This presents a major challenge because programs serving youth at risk are extremely difficult to evaluate. In many cases, organizations have not achieved consensus on what outcomes are most important, and often they lack the resources and expertise necessary to develop an assessment framework, or to reliably measure student outcomes. Usually there are strong pressures from within organizations to conduct *formative* evaluations, with an emphasis on improving service delivery. Thus most evaluations have focussed on the demographic characteristics of the population served, the type and frequency of services received by differing types of clients, and the costs of service delivery. Yet funding agencies and private donors tend to want *summative* evaluation: they want to know whether these programs truly alter the life course of the youth served. Are the programs effective for youth with different types of problems? What elements of these programs are most successful in affecting youth outcomes? Can the

effects be realized in different kinds of settings?

This paper sets out a model for the evaluation of school- and community-based programs serving youth at risk, and provides a framework for its implementation. The development of a practical and inexpensive evaluation plan that can be used in diverse settings constitutes the first phase of a two-year study funded by the Toronto Sick Children's Hospital Foundation. The next phase of the research entails testing and improving the model by applying it to an evaluation of Partners for Youth, a community-based program located in Fredericton, New Brunswick. The evaluation will assess the program's effects on youth over a three-year period, the duration for which youth are normally enrolled in the Partners program. Our model differs from many of the approaches described in the evaluation literature in that it attempts to identify and measure student outcomes. The next section of the paper provides a brief description of Partners for Youth. It is followed by sections describing our schema for measuring student outcomes and some aspects of the design of our evaluation. We hope that the presentation of our model at this early stage in our research will contribute to the dialogue among educators about the goals of education for disadvantaged youth and whether we are adequately meeting these goals.

Partners for Youth as a case study

Partners for Youth, located in Fredericton, New Brunswick, is a community program that aims to improve the educational and health outcomes of youth at risk. The organization coordinates the efforts of several partners: schools, businesses, the two local universities, a number of community agencies and the provincial government. It delivers a school-based program based on the principles of adventure-based counseling, which integrates skill-building activities with regular group counselling. (It also operates a special foster care program for children who are wards of the province.) The program provides regular group activities and counselling in the school, and on weekends and during a two-week summer program, a series of challenging, adventurous activities, such as rock climbing and winter camping. These experiences offer children an opportunity to experience success, sometimes for the first time, and to build trust, set challenging goals and work cooperatively in a group. The program is staffed mainly by volunteer teachers who work closely with other school staff and support agencies. Three criteria for the success of Partners for Youth are that it have positive, measurable effects on children's education and health outcomes, that it improves the learning climate of the schools, and that its program be transferable to other

schools in Canada.

The theoretical framework underlying Partners for Youth is based on a "cumulative risk model."¹⁵ The model is derived from longitudinal research which suggests the effects of exposure to multiple risk factors are cumulative and associated with diverse disorders. Many children follow a pattern whereby risk factors

associated with poverty, biological handicaps, poor temperament and cognitive deficits prior to entering school lead to school misconduct, combined with low achievement and learning disabilities in the early grades. By the middle school grades, these

Funding agencies and private donors want to know whether intervention programs truly alter the life course of the youth served.

children typically display a low commitment to educational activities, disaffection toward school, poor social bonding and poor peer relations. These characteristics continue into the secondary school years, which place these children at a very high risk of mental and physical illness, and juvenile delinquency.¹⁶ Many of the youth served by Partners for Youth have dealt with multiple risk factors, and a number have followed this pattern from their pre-school years. Others have simply encountered a difficult transition period, or have decided themselves that they could benefit from this type of program. Nevertheless, like many youth at this age, they too tend to be disaffected from school and have low self esteem. The program therefore attempts to foster a stronger commitment toward school, improve self esteem and develop strong social bonds and positive peer relations.

A schema for measuring student outcomes

There are 17 outcomes in our evaluation model that have been identified as the most important outcomes for measuring the success of programs serving at-risk youth. These outcomes, displayed in Figure 1, have been categorized into three strands: personal development, social behaviour and school engagement. The schema was established by conducting a detailed review of the literature pertaining to at-risk youth, analysing the goals of Partners for Youth and other comparable programs serving at-risk youth, and consulting with stakeholders who work with at-risk youth, including program staff, volunteers and social workers.

The outcomes included in the first strand, *personal development*, include measures pertaining to a child's personal qualities, characteristics and general sense of place, belonging and well being. Improving personal factors like self esteem and locus of control are common to most programs serving at-risk youth. Many programs strive to improve the relationships the youth have with their parents and their peers. Our model also includes two outcomes particular to the circumstances of children who are wards of the provincial government

Figure 1.
A Schema for the Measurement of Student Outcomes

Evaluation Outcomes for Programs Serving At-Risk Youth

Personal Development pertains to an individual's personal qualities, characteristics, and general sense of place, belonging, and well being.

Self Esteem is an individual's appraisal of their self-worth, personal dignity, and self confidence.

Locus of Control and Personal Responsibility describes one's sense of personal control over their fate, and how this is reflected in goal-setting, planning, and problem-solving.

Sense of Security and Well Being is the degree to which youth feel secure as opposed to feeling vulnerable and in danger.

Family Relations refers to the extent to which youth strive to resolve conflict, develop positive relationships, and share in family responsibilities.

Social Support pertains to youths' perceptions about their level and availability of social support from friends, family, and other sources.

Knowledge of Birth Family and Level of Contact refers to the type and degree of contact children living outside of their natural family have with their birth parents, and the background knowledge they have about their parents, siblings, and relatives.

Type and Frequency of Care Placement pertains to the frequency that youth move in and out of care (often referred to as "foster care drift"), and the kind of care facilities experienced.

Social Behaviour encompasses participants' social context and the interpersonal nature of their peer relationships.

Relationships With Other Children considers a youth's peer network and its degree of influence on various attitudes, values, beliefs, and behaviours.

Prosocial and Antisocial Behaviours considers a range of both positive and negative behaviours that demonstrate either a responsible orientation to others or a violation of the rights of others and of societal norms.

Use of Free Time measures out-of-school behaviours to determine what activities are engaged in outside of school and how youth use their free time.

Use of Tobacco, Alcohol, and Illegal Drugs measures the extent to which youth participate in substance use and substance abuse.

Ability to Work Well With Others relates to factors associated with engaging positively in teamwork and working to aid group members in achieving positive results.

School Engagement consists of outcomes associated with personal, social and academic engagement and achievement in schooling.

Academic Achievement refers to the growth in academic achievement based on standardized test scores and teachers' grades.

Acceptable School Conduct covers a range of variables associated with the type and degree of critical incidents occurring at school, a youth's level of attendance at school, and the frequency of school placement and relocations for purposes other than regular promotion.

Engagement in Academic Activities considers the extent to which youth commit to and value their education, and their level of academic aspirations.

Engagement in School Life considers the degree to which youth become involved in the social fabric of their school through interactions with others.

Sense of Belonging at School determines the degree to which youth feel accepted or alienated by others within their school community.

the longer term.

The third outcome strand, *school engagement*, includes outcomes associated with personal, social and academic engagement. Some programs serving at-risk youth primarily aim to provide a "safe haven" in which youth can develop personally and socially, but school success is not part of their mission. Other programs, including Partners for Youth, strive to help youth understand the goals of formal schooling, and increase their participation. A growing literature suggests that youth who are more engaged in school activities, including extracurricular activities, are more likely to succeed academically, and less likely to drop out during secondary school. Academic achievement is only one element of this strand. It also includes a range of variables associated with school conduct (critical incidents like cheating on tests, truancy and detention, for example), the extent to which youth engage in the social life of the school, and the degree to which they feel accepted or alienated by their peers and teachers.

in foster care situations: a child's knowledge of and level of contact with his or her birth family, and the type and frequency of care placement a child has experienced.

The outcomes in the second strand, *social behaviour*, encompass youths' social context and the interpersonal nature of their peer relationships. There are two outcomes of particular note in this strand: the monitoring of prosocial and antisocial behaviours, and the ability to work well with others. The monitoring of prosocial and antisocial behaviour is particularly important because youth with emotional and conduct disorders are vulnerable to experiencing mental and physical health problems as adults. The ability to work well with others is also important because the capacity to work collaboratively has important implications for school success in the short-term, and for workforce success in

Implementation and design

The information relevant to these outcomes was gathered this year as a baseline, and will be collected on an annual basis. Data collection entails administering three questionnaires: one to the youth themselves, a second to each youth's teacher, and the third to the person designated as the person most knowledgeable (PMK) about that youth, in most cases the birth parent or official guardian (for children who are wards of the province, the PMK may be the foster parent or a social worker). The youth questionnaire asks the children to provide information pertaining to all but two of the outcomes — academic achievement and school conduct. The PMK is asked about the child's family relations, relationships with other children, prosocial and antisocial behaviours, the use of free time, the ability to work

with others, academic achievement, school conduct, engagement in academic activities and engagement in school life. The child's teacher is asked for information about the child's relationships with other children, prosocial and antisocial behaviour, ability to work with others and academic achievement.

Five of the outcomes — relationships with other children, prosocial and antisocial behaviour, ability to work well with others, engagement in academic activities and sense of belonging at school — are included in all three of the questionnaires. This strengthens the reliability and validity of measurement, and allows for comparisons among respondents. An important component of the questionnaires is the behaviour checklist, which is comprised of a standard set of questions designed to gauge prosocial and antisocial behaviour. It consists of 46 questions related to conduct disorder, hyperactivity, emotional disorder, indirect aggression, physical aggression, inattention, prosocial behaviour, difficult behaviours and risk-taking behaviours. The checklist is useful for determining a child's perception of his or her general behaviour and engagement in risk-taking behaviours, and for determining how these perceptions compare with those of the PMK and the teacher's. It also enables youth to be tracked along a number of significant at-risk indicators.

The primary goals of research design in evaluation studies are to identify and isolate the effects of a program on client outcomes, and to discern the extent to which these effects are generalizable to other clients and settings. The chief problem in isolating program effects is that clients are likely to change even without intervention. This is particularly problematic in the evaluation of youth services, because many youth participate in delinquent activities during a brief period in their early teens, but these problems disappear during the latter stages of secondary school. Also, other events can occur during the course of an intervention which can change behaviour, such that it is difficult to discern whether it was the program or the event that caused the desired effect. The preferred approach for controlling for these threats to the internal validity of an evaluation is to include a control group. But this is usually infeasible because it is unethical to deny treatment to youth in need of services. Also, collecting data for a control group is expensive, and for this reason alone, impractical in most non-research settings.

There are two alternatives to design that employ a control group, and both will be used in this evaluation. One is to collect longitudinal data for the treated individuals, and determine their growth trajectories for each outcome measure. In essence, each child serves as his or her own control. This strategy improves the reliability of measurement, but it does not account for the possibility that the children could have made similar changes without any intervention. The other strategy is to match the treated individuals on key characteristics (such as risk factors) to others who are not receiving the treatment. Generally this is difficult because one does

not have a large enough pool of potential control subjects to match on more than one or two key variables.

This study will match the individuals in the Partners for Youth program with comparable children who are being studied in Canada's National Longitudinal Study of Children and Youth (NLSCY). The NLSCY is following over 20,000 children, roughly 4000 of which are in the age range covered by Partners. Thus, it is feasible to identify a large group of comparable children. The essence of our evaluation design, then, is to compare the growth trajectories of the children in the Partners program with the growth trajectories of a large set of children followed in the NLSCY who have comparable characteristics. This matching technique, combined with the longitudinal analysis of each child's growth trajectory for each outcome measure, will help discern the overall effects of the program for each youth.

Endnotes

1. D.R. Offord, "Prevention of Behavioral and Emotional Disorders in Children," *Journal of Child Psychology and Psychiatry*, Vol. 28 (1987), pp. 9-19.
2. H. Levin, "Increasing Human Capital in At-Risk Populations," *Policy Options*, Vol. 18, no. 6 (1997), pp. 39-42.
3. D.R. Offord and B.G. Waters, "Socialization and Its Failure," in M.D. Levine, W.B. Carey, A.C. Crocker and R. T. Gross (eds.), *Developmental-Behavioral Pediatrics* (Philadelphia: Saunders, 1983), pp. 650-82.
4. J.D. Willms, *Literacy Skills of Canadian Youth*, Report no. 89-552-MPE, no. 1 (Ottawa: Statistics Canada, 1997).
5. G.B. Jackson, "The Research Evidence on the Effects of Grade Retention," *Review of Educational Research*, Vol. 45 (1975), pp. 613-45; L. Shepard and M. Smith, *Boulder Valley Kindergarten Study: Retention Practices and Prevention Effects* (Boulder, CO: University of Colorado Laboratory of Educational Research, 1985).
6. E.G. Cohen, *Designing Groupwork: Strategies for Heterogeneous Classrooms* (New York: Teachers College Press, 1994).
7. P. Johnston, R. Allington and P. Afflerbach, "The Congruence of Classroom and Remedial Instruction," *Elementary School Journal*, Vol. 85 (1985), pp. 465-77.
8. F.X. Archambault, "Instructional Setting and Other Design Features of Compensatory Education Programs," in R.E. Slavin, N.L. Karweit, N.A. Madden (eds.), *Effective Programs for Students at Risk* (Needham Heights, MA: Allyn & Bacon, 1989); N.A. Madden and R.E. Slavin, "Effective Pull-out Programs for Students at Risk," in R.E. Slavin, N.L. Karweit, N.A. Madden (eds.), *Effective Programs for Students at Risk* (Needham Heights, MA: Allyn & Bacon, 1989).
9. A.F. McPherson and J.D. Willms, "Equalisation and Improvement: Some Effects of Comprehensive Reorganization in Scotland," *Sociology*, Vol. 21, no. 4 (1987), pp. 509-39; J.D. Willms, "Social Class Segregation and its Relationship to Pupils' Examination Results in Scotland," *American Sociological Review*, Vol. 51 (1986), pp. 224-41.
10. M. Fullan, *Successful School Improvement: The Implementation Perspective and Beyond* (Buckingham: Open University Press, 1992); Sui-Chu Ho and J.D. Willms, "The Effects

of Parental Involvement on Eighth Grade Achievement," *Sociology of Education*, Vol. 69 (1996), pp. 126-41; V.E. Lee and J.B. Smith, "Effects of School Restructuring on the Achievement and Engagement of Middle-Grade Students," *Sociology of Education*, Vol. 66 (1993), pp. 164-87; J.D. Willms, *Monitoring School Performance: A Non-technical Guide for Educational Administrators* (Lewes: Falmer Press, 1992).

11. H.M. Levin, "Accelerated Schools for Disadvantaged Students," *Educational Leadership*, Vol. 44, no. 6 (1987), pp. 19-21.

12 N.A.Madden, R.E. Slavin, N.L. Karweit and B.J. Livermon, "Effective Pullout Programs for Students at Risk," in R.E. Slavin, N.L. Karweit, N.A. Madden (eds.), *Effective Programs for Students at Risk* (Needham Heights, MA: Allyn & Bacon, 1989).

13. J.P. Comer, *School Power* (New York: Free Press, 1980); J.P. Comer, "Educating Poor Minority Children," *Scientific American*, Vol. 259, no. 5 (1988), pp. 42-48.

14. J.A. King, "Meeting Educational Needs of At-Risk Students: A Cost Analysis of Three Models," *Educational Evaluation and Policy Analysis*, Vol. 16, no. 1 (1994), pp. 1-19.

15. J.D. Coie and M.R. Jacobs, "The Role of Social Context in the Prevention of Conduct Disorder," *Development and Psychopathology*, Vol. 5 (1993), pp. 263-75; H. Yoshikawa, "Prevention as Cumulative Protection: Effects of Early Family Support and Education on Chronic Delinquency and its Risks," *Psychological Bulletin*, Vol. 115 (1994), pp. 28-54.

16. J.D. Hawkins and D.M. Lishner, "Schooling and Delinquency," in E. H. Johnson (eds.), *Handbook on Crime and Delinquency Prevention* (New York: Greenwood, 1987), pp. 179-221.

17. S.W. Raudenbush and W.S. Chan, *Growth Curve Analysis in Accelerated Longitudinal Designs with Application to the National Youth Survey* (East Lansing, MI: Michigan State University, 1991).

J. Douglas Willms is Professor and UNB/CIAR Research Chair, Atlantic Centre for Policy Research in Education, University of New Brunswick. **Elizabeth Sloat** is Research Associate, Atlantic Centre for Policy Research in Education, University of New Brunswick.



J. Douglas Willms



Elizabeth A. Sloat

by Kelly Bedard

STREAMING AND INEQUALITY: WHO WINS?

La répartition des étudiants en classes homogènes est un élément important de la politique d'éducation. L'âge des étudiants au moment d'être répartis en classes spécialisées et leur ventilation parmi les différentes classes ont une influence sur leurs résultats. Compte tenu de la variété des structures éducationnelles qui existent et de l'influence qu'exerce l'éducation sur l'accès des étudiants au marché du travail, il importe d'examiner les institutions auxquelles les enfants sont exposés avant d'être assez vieux pour faire leurs propres choix en matière d'éducation et de carrière. L'auteure trace un portrait de ce que nous savons au sujet de l'impact de la répartition en classes homogènes sur les résultats des étudiants.

Streaming is an important element of educational policy. Both the age at which students are streamed into specialized programs and the proportion of students placed in each stream influence student outcomes. Given the variety of educational structures that exist, and the influence that education has on labour market outcomes, it is important to examine the institutions that children are exposed to before they are old enough to make their own academic and career choices.

Most child-oriented institutions sort children into