THE COMPLETE STORY: A POPULATION-BASED PERSPECTIVE ON SCHOOL PERFORMANCE AND EDUCATIONAL TESTING

Noralou P. Roos, Marni Brownell, Anne Guevremont, Randy Fransoo, Ben Levin, Leonard MacWilliam, & Leslie L. Roos

All children born in Manitoba in 1984 were tracked for 18 years to assess their grade-12 performance on a provincial examination according to a student's socio-economic status. The proportion of youths in families receiving social assistance judged to have passed their language arts exam dropped from 80 per cent to 12 per cent, depending on whether one counts only those in the cohort who took the test on time in 2002 or all youths born in 1984 who should have taken the test in 2002. Getting better data on performance and doing something about the discrepancies should become a Canadian priority.

Key words: educational opportunity, exam performance, socio-economic status, testing, longitudinal studies

Tous les enfants nés au Manitoba en 1984 ont été suivis sur une période de 18 ans en vue d'évaluer leur rendement en 12e année lors d'un examen provincial, tenant compte de -leur statut socioéconomique. La proportion de jeunes issus des familles recevant de l'aide sociale et considérés comme ayant réussi leur examen au plan des compétences linguistiques passe de 80 % à 12 %, selon que l'on compte seulement ceux qui, dans la cohorte, ont subi l'examen à temps en 2002 ou tous les jeunes nés en 1984 qui auraient dû subir l'examen. En matière d'égalité des chances, la performance du système scolaire actuel au Canada laisse à désirer.

Mots clés : possibilités éducatives, résultats d'examen, statut socioéconomique, analyse longitudinale.

Reports comparing educational test results across different schools within a region (Most, 2003), across regions (Bussière et al., 2001; Canadian Education Statistics Council, 2003; Willms,1997; Wirt, Choy, Provasnik, Rooney, & Tobin, 2003) and across countries (Organization for Economic Co-peration and Development [OECD], 2001; Martin et al., 2000; Mullis et al., 2000) are now common. However, testing tells only part of the story. Exam results illustrate the performance of individuals taking the tests, only children enrolled in schools in one particular grade are assessed, those children who have made it far enough in the system to be examined. What is not known is how many students missed a test because they were not in school on test day, because they had fallen one or more years behind their cohort, or because they had dropped out. Raudenbush and Kim (2002) suggest that the most important shortcoming of cross-national data comparisons is bias in how populations are defined, for example "by the meaning of 'grade' in the presence of retention or selective dropout" (p. 292). Programme for International Student Assessment (PISA, 2000) dealt with the failure-to-progress issue in a relatively sophisticated manner, focusing on exam achievement of 15 year olds (actually those who were enrolled in school and 15 years and 3 months to 16 years and 2 months at the start of the assessment period), regardless of their standing in a school system (OECD, 2001, p. 231). However, only 88 per cent of the Canadian schools originally sampled agreed to participate (56 per cent of American sampled schools agreed), and students in the sampled schools in both countries showed a 15 per cent non-participation rate (OECD, 2001, p. 235). The results were weighted to account for non-responders within schools and for non-participating schools, but these weightings did not account for the likelihood that the non-assessed population had lower test scores.

Educational performance is repeatedly shown to be strongly associated with the socio-economic characteristics of a child's neighborhood (Ainsworth, 2002) and with the relative levels of education, income, and occupation of the family in which a child is raised. Indeed, socio-economic status remains the single most powerful predictor of educational and other life outcomes (Gorard, Fitz, & Taylor, 2001; Ma & Klinger, 2000).

The same factors that predict poor performance also predict dropout rates, absenteeism, and grade retention (Audas & Willms, 2001; Fields & Smith, 1998; Haveman & Wolfe, 1994). The question arises: does typical reporting of education test results underestimate the relationship between socio-economic status and educational outcomes? Schools themselves have little information on what happens to children who do not register: have they enrolled in another school, left the region, or dropped out?

Population-based data such as those in Manitoba, and similar resources available elsewhere in Canada such as British Columbia (Chamberlayne et al., 1998), are invaluable for giving a more accurate assessment of educational attainment across at-risk groups. The Manitoba data available for this project provided an essentially complete registration and follow-up of a population of 1.1 million individuals from the early 1970s (Roos et al., 1996). We have used these data to follow all children born in Manitoba in 1984, who remained in the province for 18 years (80%), to assess their educational achievement according to key socio-economic risk factors: membership in a family receiving social assistance, affluence of neighborhood of residence, and maternal age at child's birth.

METHOD

The Population Health Research Data Repository held at the Manitoba Centre for Health Policy has been built over the last 25 years with the cooperation of the Ministries of Health, Education, Citizenship and Youth, and Family Services and Housing. This database includes information on all children born in the province. We took birth information and mother's age at the time of the birth of the child from universal hospital discharge abstracts. A population research registry permits identifying whether an individual has left the province or, if an individual remains, specifying place of residence at various points in time (Roos & Nicol, 1999). The database also includes information on all students enrolled in schools in Manitoba and information from the provincial standard tests taken in grade 12. These tests contribute 30 per cent to students' final course mark. Individuals pass the language arts test by scoring 50 per cent or more on any of the following exams:

English, French as a second language (students in the French immersion program), and French as a primary language. Individuals pass the mathematics test by scoring 50 per cent or better on the pre-calculus exam, the consumer math exam, or the applied math exam. Because the type of math exam taken is also related to socio-economic status (24 per cent of grade-12 students in the highest income neighborhoods take the pre-calculus test compared to seven per cent of those in the lowest income neighborhoods), the measure "per cent passing the math test" underestimates socio-economic differences in achievement.

Of those children born in Manitoba in 1984, 80 per cent were resident in the province in June 2002. The 20 per cent of the cohort who left the province between 1984 and 2002 were slightly more likely to have had older mothers and were somewhat more likely to have been residents of the poorest neighborhoods in the census year, 1986 (24 per cent versus 20 per cent). Because socio-economic status in neighborhood of residence may be more meaningfully defined in urban areas, we restricted our analyses to cohort members resident in Winnipeg in 2002. About eight per cent (approximately 500 youths) of our cohort have Treaty First Nations (Aboriginal) status. These individuals tend to have poorer economic status and educational outcomes (Canadian Educational Statistics Council, 2003; Indian Affairs and Northern Development, 2000). To facilitate generalization to other jurisdictions, Treaty First Nations youth living in Winnipeg (N = 282) were excluded from this analysis. Our method for excluding First Nations youth misses approximately half of such individuals; hence, perhaps 4 per cent of the analyzed cohort (250 of the 5894 youths) were First Nations. A few cohort members (69) were advanced a year and graduated before 2002, the year they would have been expected to take the exam; these individuals and those who had died are not included in the analysis. Students enrolled in private and parochial schools in Winnipeg were included. After all exclusions, 5894 youths born in Manitoba in 1984 were available for these analyses.

If these individuals had proceeded through school on the usual course with others born that year, they would have been writing the grade-12 standard tests in January or June 2002, and graduating in June 2002. We assessed achievement not only by whether the student passed

the test, but also by whether the student was absent or did not complete the test, had dropped back one or more years (so in 2001/02 was enrolled in grades 9, 10 or 11), or had dropped out of school entirely. To clarify the presentation, 2002 is used subsequently in the text when exam/graduation dates are mentioned.

Record linkage between the research registry and the educational enrollment file enabled us to identify those from the cohort who were in the province but not enrolled in school (Black & Roos, 2005). Quality of the linkage was high. Just 2.8 per cent of the students enrolled in school in 2002 could not be linked to the registry as of the end of December 2001. Of this small group, a certain percentage would not be expected to be in the registry. These included foreign students, Canadian students moving to Manitoba whose health coverage had not yet been transferred from their home province, and immigrant students not yet eligible for Manitoba coverage. Hence, we assumed that cohort members identified by the registry as resident in the province, but not found in the enrollment file, were accurately identified as withdrawn from school. A very small number of youths from higher income families leaving the province for boarding school were inappropriately categorized as having dropped out of school, lowering the measured success rates for high-SES students.

Although some parents may deliberately hold their children back a year at school entry, an analysis of children in kindergarten in 1998/99 through 2002/03 showed that this occurred for only 2.05% of Winnipeg children. However, children born in December (and, to a lesser extent, November) are considerably more likely than those born earlier in the year, to be in Grade 11 or lower at the "age appropriate" time. Much of this discrepancy is likely due to cut-off dates for school entry. Some children born in November and December would start school with their birth cohort and some would start the following year. Analyses eliminating those born in December were compared with those including all 12 months; small (1-2%) differences in frequencies were noted; if anything they accentuate the gradients across socio-economic groups in achievement.

To measure socio-economic status, public-use data from the 2001 Canadian Census were aggregated at the geographic unit of the

dissemination area; a dissemination area had a population of 400 to 700 persons. The areas were ranked from highest to lowest income and then grouped into five population quintiles; each quintile contained approximately 20 per cent of the urban population.[1] Each individual was assigned to a dissemination area according to his or her residential postal code in 2001, and each dissemination area was categorized as belonging to one of five income quintiles based on average household income for area residents. A similar technique classified areas and children in 1986. The ordering of neighbourhood income was quite stable, with correlations around .85 over five year census intervals. Manitoba studies using self-reported household income showed results on health status measures quite similar to those using neighbourhood income (Mustard, Derksen, Berthelot & Wolfson, 1999). Individuals were also categorized according to whether their family received social assistance in the recent past, defined as having received one or more months of income support over the two-year period, 2000-2001.

RESULTS

Population- versus Grade-based Performance

The left-hand part of Table 1 highlights those individuals from the 1984 cohort writing the grade-12 exams in 2002. Overall, 89 per cent of those taking the language arts test (3,368 individuals) passed, while 76 per cent of those writing one of the mathematics tests (2,526 individuals) passed. However, these results report only the performance of those born in 1984, reaching grade 12 by 2002, and completing the exam. The right-hand half of Table 1 tells the complete story. This part of the table includes all cohort members living in Manitoba in 2002, not just those writing the exam on time. This more accurate picture identifies the 14 per cent of the cohort who had dropped the course before writing the language arts test, were absent on the day of the test, or did not complete it, the 13 per cent of the cohort at least one grade behind, and the 8 per cent of the cohort who had dropped out of school completely, leaving just 57 per cent who passed. When the mathematics test results are analyzed in the same way, a similar picture emerges; only 43 per cent passed.

Table 1: Grade 12 Standard Test Performance

	Cohort Students Writing Test*		Entire 18 Year Cohort**	
Language Arts Test	**N**	**Percent**	**N**	**Percent**
Passed	3,368	89 %	3,368	57%
Failed	415	11%	415	7%
Absent/Dropped			850	14%
Enrolled Grades 9-11			792	13%
Not Enrolled			469	8%
Number of Individuals	3,783		5,894	
Math Test				
Passed	2,526	76%	2,526	43%
Failed	808	24%	808	14%
Absent/Dropped			1,299	22%
Enrolled Grades 9-11			792	13%
Not Enrolled			469	8%
Number of Individuals	3,334		5,894	

*Includes all those born in Manitoba 1984, resident in Winnipeg in June 2002 who wrote standard test in 2002.

**Includes all those born in Manitoba in 1984 and resident in Winnipeg in June 2002 regardless of whether they wrote the test or not.

Performance by Socio-economic Indicators

The full impact of this more accurate perspective emerges from examining performance by our three socio-economic indicators: family receipt of social assistance, affluence of neighborhood of residence, and age of the mother at her child's birth. Approximately 9 per cent of the 2002-resident cohort (538/5894) were members of families receiving social assistance, but only 2.2 per cent of the test writers were in such families. The proportion of youths in families receiving social assistance

judged to have passed their language arts exam drops from 80 per cent to 12 per cent, depending on whether we counted only those taking the test on time or all children born in 1984 who should have taken the test (Table 2). The mathematics test shows essentially the same pattern (Table 2), with the pass rate dropping from 76 per cent to 10 per cent.

Table 3 assesses performance according to the relative affluence of the neighborhood of residence in 2002, the year during which the exam would have been written had individuals been in the age-appropriate grade. Among test writers, the language arts pass rate ranges from 85 per cent to 92 per cent, while the math pass rate ranges from 70 per cent to 79 per cent. For the entire cohort, though, pass rates for the poorest neighbourhoods are less than half those in the wealthiest (33% versus 72% for language arts, 23% versus 52% for math). Not only did the pass rates fall, but the gradient between higher and lower income areas gets considerably steeper using this truer perspective.

Taking relative affluence of neighborhood of residence when the child was two years old (in 1986) for describing performance produces similar results[2]. Performance by maternal age at the time the child was born also differs dramatically depending on whether only those writing or all those who should have written are considered (Table 4). Among test writers, 78 per cent of the children of teenage mothers passed the language arts test compared with 90 per cent of the offspring of older mothers. For the entire cohort, however, fully 25 per cent of children born in 1984 to teenage mothers had dropped out of school, 33 per cent had been retained at least one grade, and 13 per cent were either absent on the day of the test or had dropped the course. Only 22 per cent of this cohort passed the test, compared to 64 per cent of those youth whose mothers were 25 years or older at the time of the child's birth. Similar, slightly stronger results are obtained by using maternal age at the time of the birth of a mother's first child.

Table 2: Standard Test Performance and Social Assistance

	Cohort Students Writing Test Received Social Assistance				Entire 18 Year Cohort Received Social Assistance			
	Yes*		No**		Yes*		No**	
	N	%	N	%	N	%	N	%
Language Arts Test								
Passed	67	80%	3,301	89%	67	12%	3301	62%
Failed	17	20%	398	11%	17	3%	398	7%
Absent/ Dropped					94	18%	756	14%
Enrolled Grades 9-11					215	40%	577	11%
Not Enrolled					145	27%	324	6%
Number of Individuals	84		3,699		538		5,356	
X^2=7.55 p=0.0006					X^2=801.8 p<0.0001			
Math Test								
Passed	55	76%	2,471	76%	55	10%	2,471	46%
Failed	17	24%	791	24%	17	3%	791	15%
Absent/Dropped					106	20%	1,193	22%
Enrolled Grades 9-11					215	40%	577	11%
Not Enrolled					145	27%	324	6%
Number of Individuals	72		3,262		538		5,356	
X^2=0.016 p=0.006					X^2=775.1 p<0.0001			

*Includes children in families receiving social assistance from the province for one or more months over the period 2000 – 2001.
**Includes everyone else in the cohort as defined in Table 1.

The Pearson chi-square statistic (reported for each of the 4 components in the above table) collectively compares the observed cell frequencies with the expected frequencies under the assumption of no association between the row and column variables in the table. When the p-value is less than 0.01 we conclude that there is a statistically significant association.

Table 3: Standard Test Performance 2002 by Students' Neighborhood of Residence

	Cohort Students Writing Test Neighborhood of Residence					**Entire 18 Year Cohort Neighborhood of Residence**				
	Q1 Poor -est	**Q2**	**Q3**	**Q4**	**Q5 Wealth -iest**	**Q1 Poor -est**	**Q2**	**Q3**	**Q4**	**Q5 Wealth -iest**
Language Arts Test (percent)										
Passed	87%	86%	85%	89%	92%	33%	45%	54%	62%	72%
Failed	13%	14%	15%	11%	8%	5%	8%	9%	8%	6%
Absent/ Dropped						15%	16%	15%	15%	12%
Enrolled Grades 9-11						27%	20%	14%	10%	7%
Not Enrolled						20%	11%	9%	5%	3%
Number of Indivi-duals	343	380	631	1,048	1,381	909	719	1,003	1,504	1,759
X^2=27.8 p<0.0001						X^2=689.4 p<0.0001				
Math Test (per cent)										
Passed	70%	71%	73%	79%	77%	23%	32%	40%	51%	52%
Failed	30%	29%	27%	21%	23%	10%	13%	15%	14%	16%
Absent/ Dropped						20%	23%	23%	21%	23%
Enrolled Grades 9-11						27%	20%	14%	10%	7%
Not Enrolled						20%	11%	9%	5%	3%
Number of Indivi-duals	298	329	550	966	1,191	909	719	1,003	1,504	1,759
X^2=18.7 p=0.0009						X^2=655.4 p<0.0001				

The Pearson chi-square statistic (reported above) collectively compares the observed cell frequencies with the expected frequencies. When the p-value is less than 0.01 we conclude that there is an association between the row and the column variables in the table.

Table 4: Standard Test Performance by Age of Mother at Birth of the Cohort Child

	Cohort Students Writing Test Mother's Age in 1984			Entire 18 Year Cohort Mother's Age in 1984		
	<=19	**20-24**	**25+**	**<=19**	**20-24**	**25+**
Language Arts Test (per cent)						
Passed	78%	85%	90%	22%	43%	64%
Failed	22%	15%	10%	6%	7%	7%
Absent/Dropped				13%	18%	14%
Enrolled Grades 9-11				33%	19%	10%
Not Enrolled				25%	13%	5%
Number of Individuals	87	655	3,041	308	1,297	4,289
X^2=23.4 p<0.0001				X^2=503.1 p<0.0001		
Math Test (per cent)						
Passed	70%	74%	76%	15%	33%	48%
Failed	30%	26%	24%	7%	12%	15%
Absent/Dropped				20%	24%	22%
Enrolled Grades 9-11				33%	19%	10%
Not Enrolled				25%	13%	5%
Number of Individuals	67	572	2,695	308	1,297	4,289
X^2=2.85 p=0.2410				X^2=482.5 p<0.0001		

The Pearson chi-square statistic (reported above) collectively compares the observed cell frequencies with the expected frequencies. When the p-value is less than 0.01 we conclude that there is an association between the row and the column variables in the table.

Taking a Closer Look

The above data reflect movement of the 1984 cohort through the educational system but somewhat misrepresent how cross-sectional differences would typically be reported across advantaged and disadvantaged groups. In any given year, those taking an examination

will include students who have been retained one or more years. These retained individuals have a higher probability of failure; hence the differences typically observed in any given year across the socio-economic groups will be somewhat larger than is seen in our one-year perspective based on the birth cohort. For example, excluding Treaty First Nations youth but including all those taking the language arts test in 2002 (not just those born in 1984) changes the range of differences. The language arts pass rate range for students in social assistance families, versus those not, is 73 to 87 per cent passing (compared with 80 to 89 per cent in Table 2), and 80 to 91 per cent for those resident in the poorest versus the wealthiest neighborhoods (compared with 87 to 92 per cent in Table 3).

We have assumed that those not participating in testing would have performed more poorly than those taking the exam. Is this fair? Our best measure of educational achievement for high school students is "do they graduate?" Our cohort of 5894 youths, starting with the year they entered grade 9, was tracked for up to six years to determine what happened to the test-takers and the non-takers. Almost all those who passed the exam went on to graduate (90%). Those not participating in testing were much less likely to graduate within two years of the test than those who wrote and failed. Seventy-six per cent of those failing the exam graduated within two years compared to 32 per cent of those absent on the test day, 19 per cent of those held back one grade or more, and only three per cent of those not enrolled.

Multivariate Analyses of Factors Related to Achievement

We used logistic regression to examine the simultaneous impact of the various factors on a student's probability of passing the language arts test. These results are presented in Table 5. Among test-takers and for the entire cohort, males are approximately half as likely to pass the test as females, and students whose mother was very young (17 years or less) at their birth were much less likely to pass than students whose mother had been 25 years or older. However, test-takers differ from the entire cohort in most other respects. Among test takers, neighbourhood of residence did not appear to make much difference, but for the entire cohort,

Table 5: The Odds of Passing Language Arts Test

	Cohort Students Writing Test* (N=3783)		Entire 18 Year Cohort** (N=5894)	
	Odds Ratio	**Confidence Limits**	**Odds Ratio**	**Confidence Limits**
Sex				
Male/Female	.49	(.40-.61)	.54	(.49-.61)
Mother's Age in 1984				
< 17 vs 25 years +	.24	(.10-.59)	.24	(.13-.43)
18 – 19	.60	(.31-1.19)	.29	(.21-.41)
20 – 24	.69	(.54-.89)	.57	(.50-.66)
Received Income Assistance				
Yes/No	.60	(.34-1.06)	.15	(.12-.20)
Neighborhood of Residence				
Q1 (Poorest) vs. Q5 (Wealthiest)	.65	(.44-.95)	.33	(.27-.40)
Q2	.56	(.40-.81)	.45	(.37-.54)
Q3	.52	(.38-.70)	.54	(.45-.63)
Q4	.69	(.52-.91)	.67	(.57-.78)

*The area under the ROC (receiver operating characteristic) curve was 0.64 in this logistic regression. Model calibration showed a good fit between estimated and actual values; the Hosmer Lemeshow goodness of fit test was insignificant (p=.41 X^2 = 6.13 and 6 d.f)

**The area under the ROC (receiver operating characteristic) curve was 0.71 in this logistic regression. Model calibration showed a good fit between estimated and actual values; the Hosmer Lemeshow goodness of fit test was insignificant (p=.50 X^2 = 6.35 and 7 d.f)

differences were striking; the poorest neighbourhood students were only one third as likely to pass their language arts test as the wealthiest-neighbourhood students. Youths in families receiving social assistance face even harsher challenges (although again this is not apparent from the test-takers). Focusing on the entire cohort, for youths in families receiving assistance the odds of passing the language arts test at the age appropriate time were 85 per cent lower than the odds for youths in

families not receiving assistance, even after controlling for the other factors (sex, mean household income in neighborhood of residence, and mother's age at the birth of the child.)

The differences in achievement documented here could not have been predicted from common biomedical measures used to evaluate children at birth. Low birth weight babies are at risk for a number of developmental, cognitive, and health problems, and babies with low Apgar[3] scores are at increased risk of later health and developmental problems. However, both the percentage of children with normal birth weight and the percentage with normal Apgar scores varied little across the different socio-economic groups we review here (Brownell et al., 2004).

DISCUSSION

This article shows the clear advantage of approaching educational achievement from the population perspective. The Manitoba Population Health Data Repository provides great potential for other such investigations, both those driven by researchers and those suggested by provincial ministries. Routinely collected administrative data facilitate measurement of several types of variables, those relating to individuals, families, neighborhoods, and schools. Information on health, including measures of mental health, over the life course, can be tracked. Poor child health will influence educational achievement (Alderman, Behrman, Lavy, & Menon, 2001), while parental disability has been shown to be a significant negative predictor of a child's educational attainment (Haveman & Wolfe, 1994). Variables pertaining to marital status, family size, and family structure can be assessed at any point in time using the registry.

Information on place of residence over time can highlight potentially disruptive effects of change of residence, can be used to estimate socioeconomic mobility, and to assess the effects of length of exposure to particular kinds of neighborhoods. Both compositional and contextual neighborhood information are available (Macintyre & Ellaway, 2003). The Canadian census, carried out every five years, provides a wealth of data on each dissemination area, which can be aggregated upward to larger areas (Martens, Frohlich, Carriere, Derksen & Brownell, 2002).

Extensive contextual information on recreational programs, crime rates, and housing characteristics has been gathered for urban neighborhoods. Finally, various types of information on schools are being incorporated into the database.

The longitudinal data (up to 34 years worth) on individuals and families are particularly rich. Haveman and Wolfe (1994) have suggested studies of siblings can help deal with the potential bias by controlling common family background characteristics such as parental aspirations or motivation. For example, sibling analyses tend to lead to smaller estimates of the effects of teen fertility (Ribar, 1999). Designs combining siblings and neighbors are especially appropriate because "parallel analyses for children growing up in the same neighborhood but not in the same family can indicate how much of what is important in the shared environments of siblings lies outside the immediate family" (Duncan & Raudenbush, 2001, p. 363). With grade-12 examination results currently available for seven years (from years 1995-2002, with 2000 missing), performance data for large numbers of both siblings and non-siblings can be analyzed.

ASSESSING PERFORMANCE

We introduce these data to educational groups with the following: "None of what you're about to see is surprising, yet you'll be surprised," followed by "The role of research is to attach numbers to the obvious, to make it undeniable." The Manitoba data emphasize the remarkably different levels of educational achievement across socio-economic groups. They also confirm that assessments of educational achievement not capturing the entire population of children who should be taking the test will overestimate the performance of groups at risk for poor outcomes and provide distorted, inaccurate comparisons of school performance. The 95 per cent participation requirement of the No Child Left Behind testing program (U.S. Dept. of Education, 2006) is a step in reducing selectivity bias but does not solve the problem. By ignoring those not at their age appropriate grade level and those who have dropped out, SES differences in performance are markedly underestimated.

Samples in international comparative studies often have trouble reaching low achievement schools; even the schools responding miss those children (mostly of lower socio-economic status) who have left school, failed to be promoted, failed to complete the exam, or are absent on the day of the test. Panel surveys are also likely to under-represent the gradient in outcomes. They typically draw few families from high poverty urban neighborhoods (Shonkoff & Phillips, 2000), and children from such families who respond initially are more likely to be lost to follow up. Finally, individuals also tend to under-report risk factors and overstate achievement; Statistics Canada's Survey of Labour and Income Dynamics (SLID) markedly understates both welfare dependence in British Columbia and overstates high-school completion by welfare recipients (Warburton & Warburton, 2004).

An education reform strategy that relies mainly on cross-sectional testing to identify and label problem schools will grossly underestimate the performance advantage provided by higher socio-economic status. In younger grades, although students have not yet withdrawn from school, many are likely to have already fallen behind one or more grades for their age. Indeed, improper exemptions, absences, and retentions have been hypothesized to account for the increased pass rate on exams in some U.S. states (Heubert, 2002). In Massachusetts, for example, improved pass rates among tenth graders in 2001 followed increased ninth-grade retention in previous years. Efforts to keep adolescents of lower socio-economic status in school, as well as efforts to improve their performance within the school system, are needed.

There is also a good news story in these data which must be highlighted. For those children who stay in school and are not retained, the differences in achievement across socio-economic groups are modest. Programs that keep children from disadvantaged backgrounds in school and progressing through with their peers are critically important.

Manitoba represents a good jurisdiction in which to study this issue. Canadian children score better on international comparisons than Americans, and Manitobans score at or above the Canadian average (Willms, 1997). Winnipeg results are particularly compelling given that educational funding per child is relatively equal across areas (Task Force on Educational Funding, 2001) with some extra funding going to schools

having more students from low income families (Manitoba Education and Youth, 2003). In many U.S. communities, per pupil spending differs markedly according to a child's place of residence with much more being spent on more affluent children. Thus, per pupil public school spending in Massachusetts recently varied from $5758 to $6503 in less wealthy communities such as Worcester and Lowell to $9116 and $9356 in affluent Weston and Carlisle (Most, 2003). U.S. schools with a larger proportion of poor children are more likely to have less qualified teachers, to be overcrowded, and to use buildings in less than adequate repair (Evans & Kantrowitz, 2002). Thus, our picture of the educational experience of Winnipeg youth from disadvantaged backgrounds may well be optimistic for disadvantaged youth in the settings where per student funding is not at least similar across jurisdictions.

CONCLUSION

It is well known that many students from lower SES situations do not do well in school. Our research demonstrates that "what we think we know," at least what we derive from conventionally reported grade-12 standard test results, strongly underestimates how poorly lower SES students perform. These findings challenge testing policies for monitoring school performance and highlight the importance of changing the trajectories of disadvantaged children. Early childhood development programs (Shonkoff & Phillips, 2000), year-round educational opportunities in disadvantaged areas (Fischer, Hout, Jankowski, Swidler & Voss, 1996), qualified teachers, and increasing the educational attainment of the next generation's parents (Haveman & Wolfe, 1994) are likely candidates for successful interventions. If we are to move from a welfare state to a social investment state (Giddens, 1998, 2000, 2001), then social investments must be made differently. Education as currently functioning in Canada is underperforming as an equalizer of opportunity.

ACKNOWLEDGEMENTS

This research was funded by the Canadian Population Health Initiative (CPHI), a program of the Canadian Institute for Health Information (CIHI), by the RBC Financial Group Child Health Award, by a Canada Research Chair held by N. Roos, and a CIHR New Investigator Award held by Brownell. We gratefully

acknowledge those individuals who helped make this research possible. These include from the Department of Education, Citizenship and Youth, John Van Walleghem, Richard Perrault, Carol Crerar, Jean Britton, Ken Clark and Shirley McLellan; from the Ministry of Family Services and Housing, Harvey Stevens, Grant Doak, Gudrun Fritz and Jan Forster; from the Ministry of Health, Louis Barre.

NOTES

[1] Because the quintiles were developed using total urban population, each Winnipeg quintile will not necessarily contain 20 per cent of the high school population.

[2] These analyses, as well as those showing performance according to family upward or downward mobility, are available from the authors.

[3] Apgar scores measure the physiological well-being of newborn babies. A score is given for each of five vital signs: appearance, pulse, reflex, muscle tone, and breathing pattern. Very low scores are associated with poor neurologic outcomes (Stanley, 1994), whereas borderline scores are associated with decreased visual attentiveness in the first year of life (Lewis, Bartels, Campbell, & Goldberg, 1967).

REFERENCES

Ainsworth, J. W. (2002). Why does it take a village? The mediation of neighborhood effects on educational achievement. *Social Forces,* 81(1), 117-52.

Alderman, H., Behrman, J. R., Lavy, V., & Menon, R. (2001). Child health and school enrollment: A longitudinal analysis. *Journal of Human Resources,* 36(1), 185-205.

Audas, R., & Willms, J. D. (2001). Engagement and dropping out of school: A life-course perspective. Working Paper of the Applied Research Branch of Strategic Policy. Ottawa, ON: Ministry of Supply and Services Canada.

Black, C., & Roos, L. L. (2005). Linking and combining data to develop statistics for understanding the population's health. In D. J. Friedman, E. L. Hunter, & R. G. Parrish (Eds.), *Health statistics in the 21st century: Implications for health policy and practice* (pp. 214-240). Oxford, UK: Oxford University Press.

Brownell, M., Roos, N., Fransoo, R., Guevrèmont, A., MacWilliam, L., Derksen, S., Dik, N., Bogdanovic, B., & Sirski, M. (2004). *How do educational*

outcomes vary with socioeconomic status? Key findings from the Manitoba Child Health Atlas 2004. Winnipeg, MB. Manitoba Centre for Health Policy.

Bussière, P., Cartwright, F., Crocker, R., Ma, X., Oderkirk, J., & Zhang, Y. (2001). *Measuring up: The performance of Canada's youth in reading, mathematics and science: OECD PISA study: First results for Canadians aged 15.* Ottawa, ON: Ministry of Supply and Services Canada.

Canadian Education Statistics Council. (2003). *Education indicators in Canada: Report of the Pan-Canadian education indicators program 2003*. Toronto, ON: Canadian Education Statistics Council.

Chamberlayne R., Green, B., Barer, M. L., Hertzman, C., Lawrence, W. J., & Sheps, S. B. (1998). Creating a population-based linked health database: A new resource for health services research. *Canadian Journal of Public Health,* 89(4), 270-273.

Duncan, G. J., & Raudenbush, S. (2001). Getting context right in quantitative studies of child development. In A. Thornton (Ed.), *The well-being of children and families: Research and data needs* (pp. 356-83). Ann Arbor, MI: The University of Michigan Press.

Evans, G. W., & Kantrowitz, E. (2002). Socioeconomic status and health: The potential role of environmental risk exposure. *Annual Review of Public Health,* 23, 303-31.

Fields, J. M., & Smith, K. (1998). *Poverty, family structure, and child well-being: Indicators from the SIPP.* Population Division Working Paper No. 23. Washington, DC: U.S. Bureau of the Census.

Fischer, C. S., Hout, M., Jankowski, M., Lucas, S., Swidler, A., & Voss, K. (1996). *Inequality by design: Cracking the Bell Curve myth*. Princeton, NJ: Princeton University Press.

Giddens, A. (1998). *The third way: The renewal of social democracy*. Cambridge, MA: Polity Press.

Giddens, A. (2000). *The third way and its critics.* Cambridge, MA: Polity Press.

Giddens, A. (2001). (Ed.). *The global third way debate.* Cambridge, MA: Polity Press.

Gorard, S., Fitz, J., & Taylor, C. (2001). School choice impacts: What do we know? *Educational Researcher,* 30(7), 18-23.

Haveman, R., & Wolfe, B. (1994). *Succeeding generations: On the effects of investments in children.* New York: Russell Sage Foundation.

Heubert, J. P. (2002). Disability, race, and high-stakes testing of students. In Losen & Orfield (Eds.), *Racial inequity in special education* (pp. 137-165). Cambridge, MA: Harvard Education Press.

Indian Affairs and Northern Development. (2000). *Comparison of social conditions, 1991 and 1996: Registered Indians, Registered Indians living on reserve and the total population of Canada.* Ottawa, ON: Ministry of Supply and Services Canada.

Lewis, M., Bartels, B., Campbell, H., & Goldberg, S. (1967). Individual differences in attention. *American Journal of Diseases of Children,* 113(4), 451-65.

Ma, X., & Klinger, D. A. (2000). Hierarchical linear modelling of student and school effects on academic achievement. *Canadian Journal of Education,* 25(1), 41-55.

Macintyre, S., & Ellaway, A. (2003). Neighborhoods and health: An overview. In Kawachi & Berkman (Eds.), *Neighborhoods and health* (pp. 20-42). New York: Oxford University Press.

Manitoba Education and Youth. Funding of schools 2003/2004 school year: Winnipeg, (2003). Retrieved January 19, 2006, from *http://www.edu.gov.mb.ca/ks4/finance/schfund/2003-2004/2003-04_Funding_Book.pdf*

Martens, P. J., Frohlich, N., Carriere, K. C., Derksen, S., & Brownell, M. (2002). Embedding child health within a framework of regional health: Population health status and sociodemographic indicators. *Canadian Journal of Public Health,* 93, Suppl 2.; S15-20.

Martin, M. O., Mullis, I. V. S., Gonzalez, E. J., Gregory, K. D., Smith, T. A., Chrostowski, S. J., Garden, R. A., & O'Connor, K. M. (2000). *TIMMS 1999 international science report*. Boston, MA: Boston College.

Most, D. (2003). What are we doing to our kids? *Boston Magazine*, 41(10), 112-119.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Gregory, K. D., Garden, R A., O'Connor, K. M., Chrostowski, S. J., & Smith, T. S. (2000). *TIMMS 1999 international mathematics report.* Boston, MA: Boston College.

Mustard, C. A., Derksen, S., Berthelot, J. A., & Wolfson, M. (1999). Assessing ecologic proxies for household income: A comparison of household and

neighbourhood level income measures in the study of population health status. *Health and Place,* 5(2), 157-71.

Organisation for Economic Co-operation and Development. (2001). *Knowledge and skills for life: First results from PISA 2000.* Paris: Organisation for Economic Co-operation and Development. Retrieved January 18, 2006, from http://www.pisa.oecd.org/dataoecd/44/53/33691596.pdf

Raudenbush, S. W., & Kim, J. S. (2002). Statistical issues in analysis of international comparisons of educational achievement. In Porter & Gamoran (Eds.), *Methodological advances in cross national surveys of educational achievement* (pp. 267-294). Washington, DC: National Academy Press.

Ribar, D. (1999). The socioeconomic consequences of young women's childbearing: Reconciling disparate evidence. *Journal of Population Economics,* 12(4), 547-65.

Roos, L L., & Nicol, P J. (1999). A research registry: Uses, development, and accuracy. Journal of Clinical Epidemiology, 52(1), 39-47.

Roos N. P., Black, C. D., Frohlich, N., DeCoster, C., Cohen, M. M., Tataryn, D. J., Mustard, C. A., Roos, L. L., Toll, F., Carriere, K. C., Burchill, C., MacWilliam, L., & Bogdanovic, B. (1996). Population health and health care use: An information system for policy makers. *Milbank Quarterly*, 74(1), 3-31.

Shonkoff, J., & Phillips, D. (Eds.) (2000). *From neurons to neighborhoods – The science of early childhood development.* Washington, DC: National Academy Press.

Stanley, F. J. (1994). Cerebral Palsy trends: Implications for perinatal care. *Acta Obstetetricia et Gynecologica Scandanavica,* 73, (1), 5-9.

Task Force on Educational Funding. (2001). *Rethinking education funding: Challenges and opportunities report of the Association of Manitoba Municipalities, 2001*. Retrieved September 8, 2006, from http://www.mast.mb.ca/Communications/Publications/AMM_Rpt.htm# Expenses per Student

U.S. Department of Education. (2006). *No child left behind*. Retrieved September 8, 2006, from http://www.ed.gov/nclb

Warburton, R. N., & Warburton, W. P. (2004). Canada needs better data for evidence-based policy: Inconsistencies between administrative and

survey data on welfare dependence and education. *Canadian Public Policy*, 30(3), 241-55.

Willms, J. D. (1997). *Literacy skills of Canadian youth* (Report prepared for Statistics Canada). Fredericton, NB: Atlantic Centre for Policy Research in Education, University of New Brunswick.

Wirt, J., Choy, S., Provasnik, S., Rooney, P., & Tobin, R. (2003). *The condition of education*. Washington DC: National Center for Educational Statistics (NCES 2003067).