

Ottawa Nov 01

Understanding SAIP: assessment as political text

Jane Gaskell

University of British Columbia

Gilles Fournier's paper is clear, factual and points out where SAIP is planning to go.

How to respond? I am a researcher who no longer works with large scale data sets, though I did it as a graduate student. My background is in sociology and education and I have moved, like many educational researchers, to the study of policy and of educational process. In these remarks, I will explore the social and political location of SAIP, as befits a sociologist, and then ask how it can help in the project of public understanding and discussion of education.

A bit of biography will situate my remarks. I was a graduate student in the U.S. just after the Coleman report (Coleman, 1966) on educational opportunity was released. Coleman's research was funded by Congress to "make a report concerning the lack of availability of equal educational opportunities for individuals by reason of race, color, religion or national origin in the public educational institutions at all levels in the United States.(p.iii)" As the mandate suggests, policy makers assumed it would document the inferiority of the schools black students attended. The report's unexpected conclusions about the relative lack of impact that schools had on achievement, and about the importance of integration for black students fascinated policy makers, researchers and the

public. Discussion of the report formed my graduate education and my ideas about how research and public debate could inform each other . Discussion of the Coleman Report produced theoretical work reconceptualizing educational inequality, policy debates about bussing and school effectiveness, methodological advances in quantitative analysis and academic rivalries that rivetted not only academics, but the Atlantic Monthly and the New York Times (Jencks, 1972; Levine & Bane, 1975; Mosteller & Moynihan, 1972). Hanushek (1997) calls it “the largest and probably still most influential study of education to date” (p.301). I want to have that kind of debate in Canada, to liven up my graduate seminars and focus researchers, the media and policy makers across the country on how to improve schooling. I have no doubt that large scale, comparative data are important, but I want the discussions of them to reach a high standard.

Because large scale research is so expensive, social and political processes influence the data that are collected, the way they are analyzed, the meaning that is drawn and the uses to which it is all put. The Coleman report arose out of concerns about civil rights; SAIP comes out of concerns about standards and global competitiveness. In the late 80’s and early 90’s, Canadian politicians were worrying about the quality of Canada’s educational system in order to boost the economy. SAIP arrived at the same time as the Prosperity Initiative, which was concerned with skills in the global economy. In that initiative, the federal government, while “accepting” provincial jurisdiction in education (did it have any choice?) said it “believes it has a role to play in promoting excellence and supporting provincial efforts to improve the acquisition of knowledge and skills. These are essential elements of our future prosperity and economic security.”(Government of Canada, 1991:

i) Since the federal government has no direct jurisdiction over education, it worked with the provinces. Since it was concerned about skill levels, it mandated an assessment.

SAIP was, then, part of a federal agenda to improve and bring some coherence to education across the country. It was part of a broader strategy which provided federal funding for new educational surveys, conferences, consortia, interventions and workshops. As a result, we were blessed with, among other things, the School Leavers Survey, the National Graduates survey, IALS, the longitudinal survey of children, CEA's study of exemplary schools, the "Stay in School" initiative and TIMSS. All of these succeeded in raising the profile of educational research and engendering public discussion, mostly in relationship to "raising standards."

SAIP marked a significant entry by CMEC into research. Its design, like the design of the exemplary schools study (Gaskell, 1996), reflects the difficult federal/provincial politics of education in Canada (Manzer, 1994; Nagy & Lupart, 1994). SAIP exists, but barely, ringed around with hedging and defensiveness. It has not become a bold and provocative statement which generates debate; rather, it is a careful, very limited set of indicators to be discussed only at the provincial level. Fournier's paper reflects this mood, looking at "the advantages that large-scale assessments can provide," rather than looking also at the disadvantages, limitations and trade-offs they entail, and the controversies they provoke.

SAIP's design was shaped by the low level of agreement among provinces and their concern about comparing provincial approaches to educational policy. What started as a broad concern for enrolment flows and outcomes across the curriculum became a limited set of achievement indicators, available only at the provincial level. What started as a move towards national goals and standards became "expectations" for students, with no tie to curriculum content or policy. What started as research on schools cannot be analyzed at the level of the school or the school board. The results are not analyzed and discussed by CMEC, severely limiting its ability to be, as it claims to be, "a forum in which jurisdictions can share information and best practices in order to enhance the quality of education."

SAIP, then, can be read as a political sign in Canada's federal system. A defence of SAIP may signal threats to its existence, perhaps in favour of other surveys, perhaps in favour of even less pan-Canadian effort now that the provinces are doing more themselves.

Fournier's paper proposes that SAIP is important because it makes provincial governments accountable for the success of their students. Accountability is, politically, a strong ground for the defense. It was key to the federal Learning Initiatives program, it was one of two overarching themes identified by CMEC in its joint declaration in 1993 (the other was life long learning) and it was reaffirmed at the First National Consultation on Education in 1994. But can SAIP achieve this goal?

Because SAIP measures only achievement on standardized tests, and only in reading, writing, math and science, it is clearly only a very partial kind of accountability. Test scores in core subjects are only a small part of what teachers are trying to achieve, so they are unlikely to use these scores as a measure of their own performance. They are also a small part of what parents/voters want from schools. This is not to say that reading and writing and science are unimportant, but they are not enough to hire or fire a minister of education on, especially when the variation is so small and the reasons for it are so unclear. If scores across the board plummeted in one province, there would be an outcry. But this is unlikely, and if it happened, we would question the test and the way it was scored. Scores shouldn't change that fast, as New York city administrators realized when CTB/McGraw Hill made mistakes with its tests (Gewirtz, 2001). Long term changes among specific subpopulations might best show the result of bad educational practice and policy, but even long term changes can be attributed to many causes, and aren't likely to be noticed till the politician responsible is out of office.

Moreover, it is not clear that SAIP is part of a political process that uses the results to make ministers accountable. Does the public know the results of SAIP? Is it part of the political debate at election time? Will it change the way people vote? It would be interesting to carry out research on what the public knows and thinks about SAIP, but my guess is that the results are not salient for people, and don't affect their votes. Part of the reason that SAIP rankings don't seem to generate a lot of excitement and news coverage (unlike provincial tests) is that they are not linked to schools and districts. The public, if it wants to know, wants to know about the schools its children attend, and is less concerned

about the provincial average. As a result, the argument that SAIP holds ministers accountable is not a strong one.

The purposes of this kind of testing should be sought elsewhere. I find the argument that SAIP teaches experts across the country about large scale assessment convincing and important. We need that kind of expertise. Helping small jurisdictions that don't do it themselves, enhancing cost effectiveness by working together, encouraging the collaboration of provinces among themselves and with the federal government, and learning from the process of carrying out SAIP are all phenomena that promise to improve research and policy in education, if indirectly. SAIP becomes a springboard for Canada-wide discussions, and who knows what kinds of issues and insights might arise while testing experts from across the country are together, discussing what they do.

The most important aim of SAIP, and other research like it, however, should surely be to improve schooling. CMEC's agenda is to "ensure that our provincial and territorial systems of education are among the most innovative and flexible in the world." While it is clear that too much emphasis on standardized achievement tests can hurt the quality of education (McNiel, 2000), we are not there yet in Canada, and if we get there, it won't be SAIP that takes us. This kind of testing, when it is combined with the collection of data on students and teachers as is now happening, (even though it has taken too long to get there) allows research on what kinds of educational experiences are related to achievement, and this can help with the discussion of improving schools. Shades of the Coleman report. It allows us to explore linkages and connections over time, examining

how student characteristics relate to learning, and asking how the way teaching is done and school systems are managed affect test scores. Comparing policies across 50 US states has provided a context for research that leads to a lively discussion of the role of graduation standards and teacher qualifications. The differences across Canadian provinces would also be useful to compare and discuss.

However, we must be absolutely clear that causes are not determined by correlations alone. In the Coleman report, the number of vacuum cleaners in a home was highly correlated with reading scores, but buying vacuum cleaners won't raise test scores. Preliminary analysis of the Coleman report also suggested that "schools don't make a difference," but further analysis has shown how schools do matter. Causal statements about what happens in schools, unlike statements about correlations among variables, are always tentative and they need to draw on more than quantitative research. As Goldthorpe (2000) puts it, cause cannot be assumed from the "robust dependence" of one variable on another, or even from the "consequential manipulation" associated with different treatments. Research needs to focus also on the "generative process" that makes connections in the world of human actors. Quantitative research can find regularities that are not apparent to the casual observer, though she might suspect. Quantitative researchers can speculate about why and under what range of conditions the regularities they find will persist. But then research must look for the social processes that underlie and produce the regular connection of one "variable" with another.

The variables in education are not things, but social processes. Introducing a computer into a classroom will change the ways students and teachers interact, and the knowledge they work with. Interactions in classrooms and with written documents turn lower social class background into lower reading scores. It happens more in some schools than others. Lowering class size only changes learning if the teachers starts teaching differently with fewer students. Only understanding these social processes can help teachers think about their practice, and help administrators with the policies that will encourage good practice. As society is always changing, finding generative process is an endless search with no single definitive answer for all children, at all times, in all schools. Research must connect with a continuing professional conversation, for it will not replace that conversation.

SAIP, and other assessment programs, then, should help schools become what is being described these days as “learning organizations” (Senge, 1990). They should foster individual and school and system-wide responsibility for continuous improvement. To do this, they need to provide much more information and disseminate it better. SAIP scores are a text that is interpreted and fed into the educational and political process (Stone, 1997). At the moment, it is a shame to say that the Fraser Institute is the most provocative interpreter of assessment results, and it is not very informative or helpful. As researchers, we can’t be like Werner von Braun and make it go up, without caring where it comes down. We need to bring together the findings of SAIP with other things we know about schools and engage the public in a rivetting and sophisticated discussion, one more time. The conclusion to Fournier’s paper suggests this might well move discussion

towards an equity agenda, for the achievement of aboriginal youth and gender differences are more striking than patterns of provincial achievement across Canada. But wherever it leads, we should use the research to create a public and an academic debate instead of being afraid of its political repercussions.

Bibliography

Coleman, J. S. (1966). Equality of Educational Opportunity. Washington, D.C.: National Centre for Educational Statistics,.

Gaskell, J. (1996). One Big Study. The Local and the National in the Exemplary Schools Project. Alberta Journal of Educational Research, 42(2), 195-209.

Gewirtz, C. (2001). Test firm, New York City officials say scores were 'overstated'. Education Week, 20(41), 14.

Goldthorpe, J. H. (2000). Causation, Statistics and Sociology, On Sociology: Numbers, Narratives and the Integration of Research and Theory (pp. 137-160). Oxford: Oxford University Press.

Government of Canada. (1991). Learning Well..Living Well: Issues for Discussion. Ottawa: Prosperity Initiative.

Hanushek, E. (1997). Outcome, Incentives, and Beliefs: Reflections on Analysis of the Economics of Schools. Educational Evaluation and Policy Analysis, 19(4), 301-308.

Jencks, C. (1972). Inequality: A Reassessment of the Effect of Family and Schooling in America. New York: Basic Books.

- Levine, D. M., & Bane, M. J. (Eds.). (1975). The Inequality Controversy: Schooling and Distributive Justice. New York: Basic Books.
- Manzer, R. (1994). Public Schools and Political Ideas: Canadian Public Educational Policy in Historical Perspective. Toronto: University of Toronto Press.
- McNiel, L. M. (2000). Contradictions of School Reform: Educational Costs of Standardized Testing. New York: Routledge.
- Mosteller, F., & Moynihan, D. (Eds.). (1972). On Equality of Educational Opportunity: Papers deriving from the Harvard University Faculty Seminar on the Coleman Report. New York: Vintage Books.
- Nagy, P., & Lupart, J. (Eds.). (1994). Is There a National Role in Education? Ottawa: Canadian Education Association and Canadian Society for the Study of Education.
- Senge, P. M. (1990). The Fifth Discipline: The Art and Practice of the Learning Organization. New York: Doubleday.
- Stone, D. (1997). Policy Paradox: The Art of Political Decision Making. New York: Norton.