



VIRTUAL RESEARCH CENTRE (VRC)

The OECD, PISA and the Impacts on Educational Policy

*by Bernie Froese-Germain
CTF Researcher
September 2010*



**Canadian Teachers' Federation
Fédération canadienne des enseignantes et des enseignants**

Table of Contents

Key external pressures on teachers, teaching and public education	p. 1
The OECD – education indicators & international surveys	p. 6
PISA's impact on education policy	p. 9
PISA criticisms at a glance (issues & concerns)	p. 14
Some observations on PISA's influence in Canada	p. 19
Moving forward	p. 22
Concluding observations.....	p. 24
References / Sources	p. 28
Endnotes.....	p. 31

Key external pressures on teachers, teaching and public education

In his discussion of key pressures driving the push for more external assessment in education – pressures generally shaping the policy environment in which teachers and other educators work – David Robinson (CAUT Associate Executive Director), speaking at the CTF President’s Forum on External Assessment in Ottawa in July 2009, cited the usual suspects: neo-liberal economic globalization (underpinned by the ideology that the market rules); declines in public funding¹; and the new public management.

Among the impacts of these trends and pressures on education has been an emphasis on test-driven accountability and on standardization of teaching and learning in general; a fostering of competition between schools and of commercialization within schools; growing privatization including public-private partnerships (see Education International, 2009) and more subtle forms of privatization such as the privatization of education policy (see for example Thompson, 2009); and more emphasis on “outputs” and less on “inputs”. In the current context a major “output” of schools is the expansion of human capital to enable countries to better compete in the information-based global economy.

Large-scale assessment regularly takes place in most jurisdictions across Canada, a fact not lost on the Fraser Institute and other right wing think tanks such as AIMS (Atlantic Institute for Market Studies) which use the test results as the primary basis for compiling school rankings at both the elementary and secondary level (see Gutstein, 2010). The frequency of external testing at different levels (provincial/territorial, national, international) – coupled with the high visibility accorded by the mainstream media to the results, usually in the form of league tables, and the imperatives of short term political mandates – have all contributed to a focus on improving one’s position within the list of rankings, as well as to a narrow focus on the tested subjects – math, science, reading.

In this era of accountability-by-numbers, the elevated status accorded to large-scale external assessments such as PISA results is symptomatic of a trend towards data-driven policy initiatives in education, and the need for regular sources of outcome data to constantly feed narrow indicators of accountability. Hargreaves and Shirley argue that we’ve been distracted down this “path of technocracy” in which “technocrats value what they measure instead of measuring what they value” (p. 31). They demonstrate how data in education can be misleading, misinterpreted and/or misused, stating that an “*overreliance on data distorts the system and leads it to ignore and marginalize the importance of moral judgment and professional responsibility.*” (p. 31) The culture of standards-based accountability and data-driven school improvement distorts the educational process, and can result in “gaming the system”, leading to “cynical, quick-fix strategies to appease administrative superiors and create the appearances of improvement that would keep politicians and the public at bay.” (p. 40)

To illustrate, they cite the example of an Ontario high school which pre-tested students in advance of the grade 10 literacy test (a graduation requirement) and then focused the efforts of the English department on test preparation for the 20% of students whose marks were just below a pass.

Or the case of a primary school in London (UK) which showed dramatic achievement gains by assigning strong teachers to Year 6 (a key testing point),

drilling those teachers in test preparation procedures, and obliging them to abandon all other areas of the curriculum except the areas that were being tested. Because there was great improvement in Year 6 but none in Year 2 (Key Stage 1) where the weaker

teachers remained, the school was able to register a phenomenal record in demonstrating value-added student progress between the two key stages, and so came to be counted among the most improved schools in the nation. (p. 40)

No analysis of the impact of test-driven accountability to distort educational processes would be complete without mention of the *No Child Left Behind* legislation. Adopted in 2002, NCLB embodies the U.S. obsession with standardized testing, providing numerous examples of how student test performance can be manipulated to show improvements that have little to do with real learning. Described as “George W. Bush’s lasting gift to American public education”, Yakabuski – in an article appropriately entitled “A teachable moment for American schools” – notes that NCLB has resulted in “everyone gaming the system”:

[NCLB] made states eligible for extra federal cash if they could show continuous improvement in student scores on state-administered tests. Unfortunately, the law created a set of perverse incentives for teachers, students and bureaucrats alike. Teachers have increasingly ‘taught to the tests,’ diminishing the emphasis on other worthy material and subjects. Students have ‘learned’ what’s needed to score better on the tests. And states have lowered standards to raise student scores and get their hands on the federal moola. In other words, everyone is gaming the system. The proof is that students in almost every jurisdiction have shown eye-popping improvement on state tests, even though their scores on the federally administered National Assessment of Education Progress tests have been flat since 2002.

Yakabuski also highlights the impact of poverty on student achievement, noting that

income inequality is the elephant in the room of U.S. education policy. It goes almost entirely unmentioned as a causal factor in the low test scores of black and Hispanic students, though the efforts to single-mindedly “lift” math and reading scores are focused squarely on minorities. ...Minorities are no better served than any other group by a system that privileges narrow testing in math and reading to the detriment of literature, art, music, science and geography. *How does a mechanistic emphasis on teaching to the tests inspire them to learn, much less equip them to be productive 21st-century citizens, workers and human beings? If it doesn't, what is public education for anyway?* [emphasis added]

Put another way, are the means and the ends of education out of synch? Sahlberg describes this as “a contradiction between what is measured and what is valued in the system.” (Alberta Teachers’ Association, 2010)

Hargreaves and Shirley report that 85% of educators surveyed in the U.S. agreed that NCLB was not improving schools, and that “shortly before the 2008 U.S. presidential election, the chair of the U.S. House Education and Labor Committee proclaimed that the No Child Left Behind (NCLB) Act had ‘become the most negative brand in America.’” (p. 1). Arguably NCLB has done more to discredit the standardization of education agenda than any other single initiative.

The much anticipated shift in direction in U.S educational policy under Barak Obama’s new administration has so far not been borne out. Karp (2010) observes that, “...instead of a dramatic break with the test, punish, and privatize policies of the Bush era, there’s been so much continuity under Obama that historian Diane Ravitch calls it ‘Bush’s third term in education.’”

A case in point is the \$4.35 billion (U.S.) “Race to the Top” program, a major education initiative rolled out by the Democrats in fall 2009. As the name implies Race to the Top provides federal funding in the form of competitive grants to states implementing various reforms. While the funds are tied to “innocuous-sounding categories” – improve teacher quality and distribution; strengthen standards and assessments; improve data collection; turn around low-performing schools – Karp assures us they are far from being innocuous, containing as he notes

landmines that have become defining features of the administration’s reform plans: linking test scores to teacher evaluation and compensation; rapid expansion of charter schools; development of data systems that facilitate remote control of schools and classrooms; and aggressive intervention for schools with low test scores, including closures, firing of staff, and various forms of state and private takeovers.

The first Race to the Top (RTTT) award winners, announced in April, were Delaware (\$100 million) and Tennessee (\$500 million). Both states agreed to lift caps on charter schools and to base teacher evaluation and compensation on student performance. The concept of merit pay seems to have been given new life by RTTT with proposals popping up with growing frequency.

Race to the Top and its top-down push for market-based education reforms has been tied to the drive for national education standards in the U.S. On the latter, the “common core” state standards initiative as it is known is sponsored by the National Governors Association and the Council of Chief State School Officers, and funded by the Bill & Melinda Gates Foundation. According to Karp,

48 states (excluding Texas and Alaska) have agreed to adopt consultant-written standards in multiple subject areas. States that participate will get points on their RTTT applications (applications that Gates hired consultants to help 25 states write). Once the standards are written, [U.S. Secretary of Education Arne] Duncan will use \$350 million to finance multistate consortia to develop new high-stakes assessments based on the standards.

The degree to which the teaching profession has been and continues to be shaped by these external forces, to the detriment of teachers and teaching, is profound. Pasi Sahlberg, speaking at the Alberta Teachers’ Association Leadership in Educational Accountability conference in April 2008, notes that

higher external expectations through prescribed learning standards and stronger school accountability are the two main drivers of educational change today Competitive pressures, higher productivity, better efficiency and system-wide excellence are also having visible effects on schools and teachers. Schools that compete over students and related resources are shifting their *modus operandi* from moral purpose towards production and efficiency, i.e. measurable outcomes, higher test scores, and better positions in school league tables. ...Efficiency measures have brought standards and testing to the centre of [the] lives of teachers and students in and out of their schools. (p. 3)

Hence the dilemma facing teachers and teaching, a profession “typically driven by ethical motive or intrinsic desire” is that it is caught between two competing forces in schools: “Teachers try to balance their work between the moral purpose of student-centred pedagogy within education as a public good, on one hand, and the drive for higher standards through perceived efficiency of the presentation-recitation mode of instruction and the perspective of education as a private good.” (p. 4)

If student engagement is suffering as a result, it's perhaps not surprising given that teaching to the test and to meet externally imposed accountability targets are not exactly conducive to engaging students daily in their classrooms, undermining both the joy of reading and the joy of teaching reading. Teachers are in many ways caught between a rock and a hard educational place – some might say these external forces are squeezing the human element out of teaching and learning, a serious concern if one believes that fostering caring supportive relationships with students is what lies at the heart of successful teaching.

Some groups such as Aboriginal students are especially vulnerable to these forces. In a recent CTF study exploring the professional experiences and knowledge of Aboriginal teachers in Canadian public schools, Verna St. Denis discusses the impact of market-driven educational reform on Aboriginal teachers' capacity to form meaningful caring relationships with their students and to generally improve the poor quality of education for Aboriginal children:

The [Aboriginal] participants in this study became teachers and remained in the teaching profession because the ethical and moral dimensions of teaching motivated them. But these dimensions can be undermined in a climate of market-driven education policies and practices that are increasingly present in educational systems. (p. 65)

St. Denis notes that her research is consistent with other studies that have found that

the moral and ethical dimensions of teaching need attention, and that teachers' morale is an important factor in effective teaching, one that appears to have been marginalized in an educational climate that defines success in terms of test scores. (p. 65)

As the title of Joel Westheimer's lecture at the 2009 CEA Whitworth Forum strongly suggests – “No child left thinking: Testing, ‘accountability’, and the threat to Canadian democracy” – the accountability stakes are also very high for schools in terms of the implications for teaching critical thinking and citizenship education, and ultimately for democracy. Westheimer (2008) is critical of the general thrust of education reforms in Canada, noting that “in many boards and provinces, ever more narrow curriculum frameworks emphasize preparing students for standardized assessments in math and literacy at the same time that they shortchange the social studies, history, and citizenship education. ...Curricular approaches that spoonfeed students to succeed on narrow academic tests teach students that broader critical thinking is optional.” (p. 7)

Teacher voices are critically important in this debate. Nick Forte, a teacher in London, Ontario, reflects on the changes he's observed to his profession over more than two decades in the classroom. The erosion of teacher professional judgment and autonomy is affecting genuine student learning including the teaching of critical thinking. As Forte puts it, the student's freedom to learn hinges upon the educator's freedom to teach:

I always believed – and still do believe – that education is about learning to think critically, asking questions and using one's imagination: creating a classroom atmosphere where thinking, questioning and imagining is encouraged because this is what permits students the freedom to truly learn. But for students to learn, teachers need the freedom to teach. And increasingly, freedom to teach is in short supply.

Over my 20-plus years of classroom teaching, I have become convinced that the debates about education are now, more than ever, populated by an excess of experts whose purported expertise is inversely proportionate to the amount of time they have

spent inside a classroom. Those of us who have taught in classroom for many years have seen scores of bandwagons come and go – but the vast majority of the time we have not seen these initiatives produce the panacea of results they had professed to obtain. Short-term ideas – often politically-driven, inconsistent and incoherent rather than providing concrete information, realistic ideas or authentic suggestions – are intruding and micromanaging my classroom practice. As a result, I usually find that, more and more, my freedom to use professional judgment and, therefore, my autonomy as an educator has been and is being curtailed. (p. 25)

In a recent CTF teacher survey the detrimental impact of standardized testing on teaching for critical thinking and citizenship education was cited by a number of respondents, including this educator:

I feel that the reason that critical thinking skills and responsible citizenship activities are not being addressed as well as they should be in the classroom is that teachers are still driven by standardized testing scores and covering content instead of focussing on nurturing students to become responsible global citizens.

Policy and other documents from education officialdom are often devoid of terms and expressions such as “caring relationships”, “moral and ethical purposes of education”, “critical thinking”, or the “democratic role of public education”. On the contrary, in a commentary on the U.S. national standards movement, Alfie Kohn notes that reform papers on national standards issued by corporate groups and politicians are filled with terms like “rigorous”, “measurable”, “accountable”, “competitive”, “world-class”, and “raising the bar”. Kohn says that such language

has very real implications for what classrooms will look like and what education is (and isn't) all about. The goal here isn't to nourish children's curiosity, to help them fall in love with reading, to promote both the ability and the disposition to think critically, or to support a democratic society. Rather, a prescription for uniform, specific, rigorous standards is made to order for those whose chief concern is to pump up the American economy and triumph over people who live in other countries.

If you read the FAQs page on the common-core-standards Web site, don't bother looking for words like “exploration,” “intrinsic motivation,” “developmentally appropriate,” or “democracy.” Instead, the very first sentence contains the phrase “success in the global economy,” followed immediately by “America's competitive edge.”

If these standards are more economic than educational in their inspiration, more about winning than learning, devoted more to serving the interests of business than to meeting the needs of kids, then we've merely painted a 21st-century facade on a hoary, dreary model of school-as-employee-training. Anyone who recoils from that vision should strenuously resist a proposal for national standards that embodies it.

The OECD – education indicators & international surveys

A major proponent of this vision of education as provider of human capital for the globalized economy is the Organisation for Economic Co-operation and Development (OECD). In a critique of the OECD education indicators project written a decade ago, Kuehn describes the lens through which this powerful organization views and interprets education:

A central assumption behind the OECD view of education is to consider students as “human capital.” The perspective logically follows from a conception of the role of government as primarily fostering economic growth, with culture and society also seen as subordinate to the economy. When economic objectives become central, then the value of education is measured primarily by its contribution to economic growth (Spring, 1998). The OECD’s education indicator systems are largely focused on elements of education that are seen as developing “human capital” and thus making a contribution to economic growth. (p. 118)

This is as true now as it was a decade ago, and perhaps even more so in the wake of the global economic and financial crisis. The thrust of the 2009 *Education at a Glance* report revolves around the need to invest in education to beat the recession.

On the rationale underlying the process of how the indicators used in the OECD *Education at a Glance* reports are chosen, Kuehn (p. 120) offers this insight, quoting from a 1992 OECD publication entitled “What are international indicators for?” (In *The OECD International Indicators: A Framework for Analysis*):

“The aim is not to create international rankings but rather to provide national policy-makers with a broader understanding of factors influencing the quality of education and permit them to explore a wider range of policy options. In theory, it is desirable to develop a wide range of indicators that accommodates the breadth of existing educational objectives within countries and that allows countries to choose those measures that best match their goals and priorities. In practice, the cost and difficulty of establishing a wide range of indicators adds weight to the argument for concentrating on a limited set of indicators that are acknowledged by all to be significant and to explore how these can be interpreted for divergent national priorities. (Bottani and Walberg, 1992, 9)”

So, while it is possible and even “desirable” for the OECD to develop alternative education indicators, as they see it the cost and challenges make it prohibitive to do so (the current approach is more ‘efficient’). There’s also the fact that by its nature, any attempt to carry out international comparisons of education systems “requires that national systems of gathering and reporting of data be reduced to a common international standard”, resulting in pressure to harmonize indicators (as we’ve seen in the Canadian context – more on this later). Not surprisingly much is sacrificed in the process, not least of which is the “loss of control by communities of their education system” (Kuehn, 1999, pp. 118-119).

Unfortunately, as Kuehn observes – and as we have seen repeatedly in the intervening years since the publication of his article (and indeed since the appearance of the first OECD education indicators report in 1992),

While the aim of the developers of the OECD indicators may not have been to create international rankings, virtually every media report that is based on their indicators charts rankings of countries. Further, those areas on which data is reported take on a

level of importance that is considerably higher than those excluded from study. Areas *not* included reflect the broad range of social and cultural goals that should interest parents, community and country. (p. 120)

Interestingly, the OECD's Programme for International Student Assessment (PISA) was just being launched at the time of Kuehn's article in the late 1990s. PISA has since become the principal source of data on the performance and quality of education systems, as measured by student achievement outcomes, for the OECD indicators reports.

In addition to PISA, a number of new international surveys are being developed and conducted by the OECD², in part due to its experience in developing large-scale international assessments such as PISA:

- PIAAC – Programme for International Assessment of Adult Competencies

According to the OECD website (www.oecd.org/), this international survey of adult skills will take place across OECD and partner countries in 2011, with results published in 2013; it will measure the skills and competencies needed for individuals to participate in society and for economies to prosper, and will help governments better understand how education and training systems can nurture these skills.

- AHELO – International Assessment of Higher Education Learning Outcomes

A publication on the OECD website states that AHELO “will potentially be the largest, most comprehensive assessment of universities yet devised. The aim is to measure various types of learning outcomes and to examine as wide a range as possible of criteria to assess their influence on those outcomes.”

Interestingly, and ironically, considering the widespread use of PISA results for ranking purposes, the OECD publication also emphasizes that

AHELO is not a league table. University league tables, or rankings, fail to do justice to the main task of a university: teaching and learning An unhealthy air of competition surrounds rankings. Universities change educational missions and cut programmes in a desperate effort to climb the ranks and join the top 10, 50 or 100 universities. A slip in the ranks means a loss of potential students Rankings may be harmful not because of what they measure, but because of what they leave out.

Most of this critique could easily be applied to public schools. (The impact of league tables on English schools is described by Mortimore in the section of this paper on PISA criticisms.)

- TALIS – Teaching and Learning International Survey

Results from the first TALIS survey which involved 23 OECD and partner countries in four continents were published in June 2009. The OECD Web site describes TALIS as “the first international survey to focus on the learning environment and the working conditions of teachers in schools and it aims to fill important information gaps in the international comparisons of education systems. TALIS offers an opportunity for teachers and school principals to give their input into education analysis and policy development in some key policy areas.”

TALIS looks at key issues affecting the work of lower secondary school teachers: professional development; teacher beliefs, attitudes and practices; school leadership; and teacher appraisal and feedback.

Just as PISA data have come to figure more prominently in the annual *Education at a Glance* reports, in all likelihood these new surveys now being ramped up will increasingly become part of the data collected and used in the annual Education Indicators reports. TALIS for example was reported on for first time in the 2009 *Education at a Glance* report.

While Canada was not among the participating countries for this first TALIS cycle, it remains to be seen what will happen in future rounds. At the time of writing the OECD is in the process of developing a proposal for a second round of TALIS. It has invited all OECD member and partner countries to participate in a priority-rating exercise intended to guide the content of the next round of TALIS.

EI Senior Research Coordinator Guntars Catlaks argues that the TALIS findings, with regard to teacher appraisal and feedback in particular, could be used to strengthen the argument for crafting systems of performance-based pay for teachers:

An important finding of TALIS is that teachers generally respond positively to appraisal and feedback. They tend to report that it is helpful for their work (improves their teaching skills) and that it increases their job satisfaction and, to a lesser extent, job security. In addition, teachers report that it significantly increases their development as teachers. The report recognises that school-level evaluation can be an important driver of school improvement. However, TALIS strongly argues that there should be an improvement in the reward side of feedback and appraisal. Three quarters of teachers report that they would not receive recognition for the increased quality of their work. Around half of teachers in TALIS countries report that their school principals use effective methods to determine teachers' performance. In addition, three quarters of teachers report that, in their schools, the most effective teachers do not receive more recognition, that their principals do not take steps to alter the monetary rewards of persistently underperforming teachers, or that underperforming teachers are not being dismissed because of poor performance. Thus, TALIS suggests that there should be more emphasis on performance management and paves the way for arguing in favour of performance-based pay.

Another concern – and logical outcome – is the potential linkage between TALIS findings and PISA results (or results from other large-scale standardized tests) in terms of teacher evaluation being based on student performance on PISA. On this point Catlaks notes that,

There could be a potentially “dangerous” connection between TALIS and PISA. The emphasis on individual teachers and their style of teaching, beliefs, cooperative attitudes, and, above all, “effectiveness” can indeed be linked to how these teachers' students perform in PISA, or PISA-like assessments, with dangerous consequences for individual teachers whose students do not perform high enough. In other words, could individual teachers be evaluated according to how well their students perform in PISA-type assessments?

PISA's impact on education policy

PISA is an international standardized test administered to random samples of 15-year-old students in all 30 OECD member countries as well as a growing number of OECD partner countries. It is not designed on the basis of national curricula and programmes but rather, "PISA assesses the extent to which 15-yrs-old students near the end of compulsory education possess the key knowledge and skills [in the areas of science, reading and mathematics] for their full participation in society." (Figazzolo, 2009, p. 3)

To date there have been four cycles of PISA:

- PISA 2000 – 43 countries participated in the assessment; subject focus was reading (primary assessment domain)
- PISA 2003 – 41 countries participated; focus on mathematics
- PISA 2006 – 57 countries involving approx. 400,000 students; focus on science
- PISA 2009 – 65 countries are participating; focus on reading

Key features of PISA according to the OECD (p. 7) include:

- its policy orientation, with design and reporting methods determined by the need of governments to draw policy lessons.
- its contextualization within the system of OECD education indicators, which examine the quality of learning outcomes, the policy levers and contextual factors that shape these outcomes, and the broader private and social returns to investments in education.
- its breadth of geographical coverage and collaborative nature, with more than 60 countries (covering roughly nine-tenths of the world economy) having participated in PISA assessments to date, including all 30 OECD countries.
- its regularity, which will enable countries to monitor their progress in meeting key learning objectives.

These combined features of PISA – policy orientation, contextualization within the OECD education indicators, and the breadth and regularity of testing – make PISA a powerful instrument for shaping education policy among OECD countries and beyond.

In 2007 the OECD conducted its own evaluation of the impact of PISA in member countries in order to:

- evaluate the relevance of PISA in participant countries and economies.
- assess the effectiveness and sustainability of PISA.
- identify the unexpected impacts of PISA.

PISA has no doubt had an influence on education policy, something the OECD views as an important indicator of PISA's effectiveness. Bernard Hugonnier (in his address at the CTF President's Forum in July 2009) highlighted a number of themes emerging from the OECD's evaluation of PISA. Among the themes is that:

- policymakers are considered to be the most significant stakeholder group both in relation to PISA and its results, and in implementing policies in light of PISA.
- local authority officials and school principals are second and third respectively with regards to the implementation of policies in light of PISA.
- PISA is used as a tool for monitoring and evaluating a country's education performance and equity.
- the influence of PISA on policy formation both nationally and locally is increasing over time.
- the influence of PISA seems to be greater at a national level rather than at a local level, and has less impact on school practices and instruction.
- in order for the influence of PISA to be increased at all levels in the system and for PISA to contribute further to sustainable change, the following are required:
 - a better coordinated and strategic approach for the dissemination of PISA results
 - further support for various stakeholder groups in interpreting PISA results and in designing strategies for improvement in light of these results
 - clearer linkage of PISA results to national or federal assessment strategies
 - greater utilization of PISA results by participant countries
- PISA has made an impact on policy in all countries studied, the impact varying from country to country (Canada was one of five case study countries selected for the qualitative data analysis).
- countries that rank relatively high in PISA use the results as a mechanism for evaluating their education system but they do not seem to have introduced many policy initiatives directly as a consequence of PISA.
- on the contrary, in countries that perform relatively poorly a direct policy impact after the publication of the PISA results is identified. The lower a country's ranking, the more awareness is raised at all levels in the system.
- many countries justify their educational reforms through their performance on PISA.
- policymakers, academics and researchers, and the media seem to be the stakeholder groups most aware of PISA.
- the more emphasis policymakers and the media place on the PISA results, the more awareness is raised at all levels in the system.
- PISA is regarded by all participating countries as an extremely important measure of the performance of their school system.

- PISA is accorded a high level of credibility and influence in all jurisdictions.
- the level of integration of PISA results in national assessment systems varies considerably.

Education International, a global federation of teacher unions based in Brussels, set out to explore the extent and the nature of PISA's influence on education policy. In 2008 it carried out a major study of PISA 2006. The study included a survey of EI member organizations as part of its methodology, "analyzing the impact of PISA 2006 on the education policy debate, with a specific focus on how the mass media reported on PISA, how governments drew from its conclusions and how unions reacted to them. In other words, how education is affected by the most important international comparative survey." (Figazzolo, 2009)

The study methodology incorporated the following components:

- classification and analysis of news articles on the international media coverage of PISA 2006 (using media monitoring agency Meltwater News); approx. 12,000 articles were collected between Dec. 2007 and Oct. 2008
- conducting interviews with teacher unions particularly involved in PISA and which represent interesting case studies
- dissemination of a questionnaire to EI affiliates/member teacher unions in the 57 countries participating in PISA 2006 to investigate their level of involvement in PISA, their interface with the media and policymakers on PISA, and their perceptions of the content of PISA and its impact on policy in their respective countries; 32 replies were received from 25 countries, "allowing for a good geographical representation of unions and countries involved." (Figazzolo, p. 11)

The study emphasizes the explicit policy orientation of PISA and more importantly, the powerful impact that PISA (in this case the PISA 2006 cycle) has had in effect on the education policy debate. The report's author, Laura Figazzolo in the EI Research Unit, states that:

What makes PISA different from the other international comparative surveys in education such as the TIMSS and PIRLS studies – conducted by the International Association for the Evaluation of Educational Achievement (IEA) since the late 1950s – is its clear policy orientation. PISA is not only research, for it contains key policy recommendations: "it aims to provide a new basis for policy dialogue and for collaboration in defining and implementing educational goals". This makes it a powerful tool for political influence.

In other words, even if "school and system characteristics cannot provide precise policy prescriptions, they can address educational policies correlated to high performance". PISA provides policy recommendations for governments to improve their country's education system's performance and to enhance its competitiveness. PISA's "implications for policy" are, in fact, led by the principle of increasing school efficiency, an approach to national education based on neo-liberal economics. (pp. 5-6)

The way in which the OECD is able to extend its influence on education policy internationally through the mechanism of PISA is based on the concept of “soft power” or “soft governance” – as Figazzolo explains,

These PISA recommendations are likely to have an extremely significant impact on governments, influencing their policy orientations. In fact, the OECD is able to exert a sort of peer pressure and “soft governance” on national governments, by virtue of its mandate and, in particular, its status as an authoritative, impartial source of evidence. The OECD does not have a legally binding mandate on member countries in the field of education, and cannot use regulations and sanctions; however, it does act through “soft power” instruments, such as recommendations, guidelines, and policy orientations. This specific feature represents a strength in a new policy environment where “soft power” instruments and peer pressure play an increasingly relevant role, particularly given the organization’s capacity to develop international comparisons based on sound data-gathering mechanisms that are globally recognized and accepted. (p. 7)

Figazzolo goes on to demonstrate this impact with a key finding from the study based on an analysis of PISA 2006³:

PISA 2006 has hugely affected the education policy debate, at [a] worldwide as well as national level. On the one side, it has constituted the point of reference for governments in their educational reforms – both centre-left and centre-right governments, but very often with an “economic” approach, i.e. testing pupils, evaluating teachers, with reforms inspired by efficiency or competitiveness criteria. On the other side, PISA has also been employed by unions and opposition parties to blame governments and attack their educational policies.

Policymakers, in government and in opposition, are eager to suggest reforms in education, not necessarily because of PISA findings, but increasingly with reference to the latest PISA results as a source of legitimacy for their actions: “We are going to propose such and such reforms, because, as demonstrated by the PISA scores, our country is lagging behind.” In fact, **half of the [survey] respondents report that their government has launched reforms or initiatives in education as a consequence of or with reference to PISA 2006**: rankings and results have been used to justify and sustain these actions. (p. 12) [emphasis in original]

She cites a number of cases of countries – with governments of varying political stripes – which instituted “efficiency/competitiveness-oriented reforms” in the wake of the release of PISA results.

For example, in response to its poor performance in PISA 2006, Mexico launched in May 2008 a new reform known as the “Alliance for Educational Quality”. It is as Figazzolo notes “a ‘neo-liberal’ reform that has not only used PISA as a point of reference, but has directly involved the OECD in its formulation and implementation It is a very peculiar reform, as it is based, for the first time, on an agreement between the government, the education union (SNTE) and the OECD itself.” (pp. 12-13)

Italy too has used PISA results to “advocate neo-liberal reform of the country’s education system”, proposing to institute draconian measures such as eliminating teaching positions (an estimated 80,000 positions) including specialist teachers, reducing the school week by almost half, and introducing an annual bonus for ‘good teachers’ of 5,000-7,000 Euros, radically altering the country’s education landscape (Figazzolo, p. 16).

In the U.S. and France, PISA data have been used to justify reforms in education based on increasing efficiency and competitiveness. In Germany, Ireland and Australia, PISA results have been used to make the case for more testing and evaluation. Australia for example recently introduced a national testing regime in which students in Years 3, 5, 7 and 9 are required to take national standardized tests in literacy and numeracy. Figazzolo states that, “Notwithstanding the dangers associated with national full cohort testing, however, Australia continues down the path of census, mass standardized testing in contrast to sample testing which could inform policy appropriately and determine the needs of student communities.” (p. 18)

There are also examples of collaboration between governments and unions regarding PISA, such as Ireland’s partnership model of decision-making regarding curriculum and assessment which includes the teacher unions; German teacher unions using PISA findings to advocate against early tracking in schools and for a more socially equitable school system; and unions in the UK using PISA findings to argue against further government expansion of school selection.

In Finland which consistently ranks at the top on PISA, a country with an education system that does not rely on standardized testing, EI’s affiliate OAJ (Opetusalan Ammattijärjestö) has used PISA results in support of its defence of public education, having “managed to focus more public attention on teachers and the issue of teacher education”. Finnish teachers are highly qualified, receive strong support, and enjoy high status within their society; in addition Finnish students are ensured equal access to a high quality education.

Some interesting and revealing findings emerged from EI’s analysis of the media reporting of PISA 2006. For example about 29% of the approx. 12,000 articles published worldwide which they examined reported only on the rankings (i.e. country ranking relative to other countries) without further explanation. Almost 30% of the articles referred to PISA results to advocate for education reforms without analysing the meaning of those results (p. 26). A much smaller percentage, 1-2%, blamed teachers “more or less explicitly” for poor results.

Almost 75% of the articles cited Finland, frequently as a country for others to emulate. In addition to its history, geography and “low intergenerational income elasticity”, teacher quality is high on the list of factors cited by the media as being responsible for Finland’s PISA success:

Keeping the status of teachers high, in terms of both salaries and qualifications, plays a key role Involving teachers in the decision-making process in schools, i.e. how decisions in education are made, represents another fundamental ingredient explaining the Finnish success. Last but not least, many articles refer to the relevance of an early energetic intervention in support of struggling pupils (Figazzolo, p. 29).

What these articles neglected to mention is a successful campaign waged by the Finnish teachers’ union, OAJ, to improve the status of teachers in Finland.

Figazzolo summarizes the PISA media coverage with these observations, stressing the media’s role as a pro-active agent of OECD-driven education reform:

By narrowing down the focus to rankings, they [the media] have often conveyed a very simplistic interpretation of educational issues. In that way, media have often operated as proactive agents of the policy orientation promoted by the OECD and, for that matter, by the more conservative media outlets themselves. Moreover, media worldwide have generally promoted a “copy and paste” approach, based on implanting the features of well-performing school systems into other, poorly performing countries. (p. 1)

PISA criticisms at a glance (issues & concerns)

Given PISA's growing international profile and influence, and in an attempt to address EI's concern "that PISA data should be used more constructively", the Education International Research Institute commissioned an expert paper on alternative models for analysing and representing countries' performance in PISA. The paper's author, Peter Mortimore (2009), summarizes the general criticisms of the program – conceptual, methodological and political – raised in a review of the literature on PISA. Several major issues emerge which are outlined here.

On the issue of cultural differences among participating countries, Mortimore notes that:

For any programme of assessment the challenge of having to treat students from over fifty nationalities in a common manner is daunting. Students from a range of cultural backgrounds may react in different ways to common questions and even to a common formal testing situation. The problem may reduce over time as the Programme [PISA] continues and the expectations of teachers and younger students develop. *Clearly it should not be the intention of test designers to iron out such cultural differences (which may be crucially important to the identities of different peoples). It is vital, however, that the existence of such differences be borne in mind when interpreting PISA outcomes.* [emphasis added]

Related to this are translation issues in terms of some students taking the test in a language other than their native tongue and how this can affect their performance.

Criticisms regarding PISA survey sampling are not new – they revolve around the representativeness of the selected sample and "dangers that some countries will endeavour to increase the proportion of the most able students and decrease the proportion of those deemed to be the least able. But students with learning difficulties and those lacking fluency with the native language also need to be included in a correct proportion." (p. 3)

PISA's use of a cross-sectional survey design (as opposed to a longitudinal survey) has also come under scrutiny. Despite the limitations of such a survey design in that "it is restricted to drawing on correlations between different measures" and hence cannot be used to infer causality, governments and the media have used PISA to imply causality. A longitudinal design would allow for this but would also increase test costs.

Another criticism of PISA is its disregard of national curricula, given PISA's "emphasis on asking questions which can be answered using common sense rather than knowledge of a particular curriculum" (p. 4). In addition, the impact of PISA's narrow focus on math, science and reading on national education systems in terms of a "powerful back-wash" effect cannot be underestimated. Mortimore states that:

Currently, as PISA increasingly is used as the ultimate reference on "quality of education" it shifts the attention of the general public and politicians almost exclusively to the core subjects of mathematics, science and reading. This leaves history, geography, civics, foreign languages, and all the other subjects taught in schools, marginalized

If PISA data are seen merely as a crude estimate of the performance of a nation's educational system and its attempts to achieve equity, the limitations of the curriculum being assessed may not matter. If, however, the data are considered to be the

definitive judgement of national systems, then the results will be damaging. The powerful back-wash will have the effect of permanently dividing the curriculum into a core (literacy, mathematics and science that which is assessed by PISA) and a less important periphery of all other subjects. (p. 4)

In Canada we've seen this trend regarding the narrow range of subjects tested (math, science, literacy/language arts) in provincial testing programs as well as the national testing program (Pan-Canadian Assessment Program).

As we have seen, PISA distorts educational policies as it has “encouraged governments to adopt an overly ‘economic’ approach to education” (p. 4).

Another shortcoming is the lack of sufficient teacher involvement in PISA. While PISA collects information from principals, students and parents via questionnaires, it does not collect information from teachers. On this, Mortimore observes that,

It appears somewhat peculiar, therefore, that those working in classrooms – at the heart of education and who have the most to contribute to the improvement of both quality and equity – are ignored. Of course, some teachers may not wish to get involved; others may not wish to devote precious time to what they see as a marginal activity. Others, however, may welcome the opportunity to participate and, by so doing, may help fill the space which now exists between the school policies – as expressed by the principal – and the students' views and outcomes. (p. 5)

It must be emphasized that the effects of league tables on national school systems and the resultant negative impact on equity in education represent one of the more egregious abuses of PISA results. Mortimore illustrates this negative impact by drawing on the experience in England over two decades:

Although PISA data are presented in great detail, and meticulous attention is paid to the inclusion of standard errors and multiple comparisons of mean performance showing which national differences are statistically significant, discussion in the media is almost exclusively in terms of countries' rank positions. The subsequent league tables so dominate the news that much of the careful work carried out by the PISA Team on the relationship between domains, contextual factors, attitudes and school and system factors is ignored – except by scholars.

The media emphasis on such limited data has encouraged the adoption of league tables in individual countries. According to EI: *“These actions usually lead to the introduction of more measurements, of national testing systems based on the PISA model and methodology, of more scrutinized teaching procedures, and ultimately, to linking teachers' performance and pay to students' test scores”.*

The experience from England, where testing of whole age cohorts using standard achievement tests (Sats) has been made statutory, is that league tables dominate public discussion about schools. Such tables fuel the development of a market economy of school places despite being fundamentally flawed by failing to reflect differences in the intakes to schools.

Many parents, however, have been persuaded by the national government's sponsorship of league tables into thinking that they offer a clear guide to the best schools. Accordingly, they are followed avidly by aspiring parents who wish to do the

best for their children. A further problem is that, where there are more applicants than places, the exercise of choice switches from parents to schools. In England there have been cases of false addresses being given in order to gain access to certain schools, resulting in court proceedings against parents.

Even the price of houses is affected by league tables, with properties in the catchment of a highly ranked school attracting a sizeable premium. High ranked schools are usually swamped with applications. Low ranked schools often struggle to achieve sufficient numbers to make them viable.

League tables also tend to depress equity as teachers are encouraged to focus on those students who may help lift the school's rankings. Those who are likely to be borderline candidates in tests are likely to be given the most help. Students who require special help are less likely to be welcomed as they may use more resources – and do less well in the tests – so pulling down the school's scores.

Furthermore, it is clear from OECD research that parents who are socially, economically or culturally advantaged are more adept at making choices. In a system based on league tables, the most sought after school places are very likely to be commandeered by such parents, leaving the remainder of less desirable ones for those with less advantage.

For the last twenty years, life in English schools has been dominated by high stakes testing. (pp. 5-6)

Mortimore also proposes a series of recommendations to address some of these concerns and criticisms, which are outlined in the final section of this paper.

In the current economic and political climate, PISA appears to be tailor-made to further the OECD's interests globally. According to Mortimore,

OECD and other international bodies have undoubtedly strengthened their influence as discussion about globalisation has increased amongst governments. Although economic policy remains paramount within OECD, its interest in education *“has taken on an increased importance within that mandate, as it has been reframed as central to national economic competitiveness within an economistic human capital framework and linked to an emerging ‘knowledge economy’”*. According to Grek, *“OECD had created a niche as a technically highly competent agency for the development of educational indicators and comparative educational performance measures”*.

Other commentators have argued that because governments' powers over economic trends has weakened – as the latest worldwide financial upheaval has demonstrated – *“the competitive advantage of nations is frequently redefined in terms of the quality of national education and training systems judged according to international standards”*. Thus PISA, supported by the governments of over 50 countries, with its clear technical skills and its flair for presenting complex information in an accessible way, has become one of the world's key arbiters of the quality and the equality of school systems. (p. 7)

The reframing of education as central to strengthening national economic competitiveness by providing the necessary human capital, and the critical role PISA has come to play in moving this agenda forward, is made abundantly clear in a recent OECD study, launched at the World Economic Forum in Davos, Switzerland in January 2010.

This study – entitled *The High Cost of Low Educational Performance: The Long-Run Economic Impact of Improving PISA Outcomes* – claims that modest gains in student achievement as measured by PISA can boost a nation’s economic well-being as measured by gross domestic product (GDP). Andreas Schleicher, head of the indicators and analysis division for the OECD education directorate which produced the study, stated that: “There’s almost a one-to-one match between what people know and how well economies have grown over time”. (Robelen, 2010)

The OECD report attempts to quantify this link between PISA score improvement and increases in national GDP:

A modest goal of having all OECD countries boost their average PISA scores by 25 points over the next 20 years – which is less than the most rapidly improving education system in the OECD, Poland, achieved between 2000 and 2006 alone – implies an aggregate gain of OECD GDP of USD 115 trillion over the lifetime of the generation born in 2010 (as evaluated at the start of reform in terms of real present value of future improvements in GDP). Bringing all countries up to the average performance of Finland, OECD’s best performing education system in PISA, would result in gains in the order of USD 260 trillion. (p. 6)

For its part Canada would apparently benefit to the tune of an additional USD \$3.7 trillion in GDP over an 80-year period, based on an increase of 25 points in average student performance on PISA.

In this recessionary economic climate, the drive for nations to raise their PISA scores, in order to improve their GDP and revive their economies, may well be given further impetus by such studies.

Comments made at a briefing by the report’s co-authors regarding the potential policy implications of the study’s findings are revealing. According to Schleicher (as cited in Robelen),

“You can never translate one system to another, you can’t copy and paste education systems,” Mr. Schleicher said, “but what you can do, what we’re doing here, is you can look at some of the policy levers that emerge from successful education systems and think about how you can situate those kinds of policy levers in a different [system].”

For one, he emphasized that spending more money doesn’t necessarily correlate with improved student outcomes.

“There’s some relationship, but it’s not particularly strong,” he said. “Money is not a predictor of success in systems.”

He added: “When you look at the best-performing systems, what characterizes them is that they get the money to where the challenges are greatest. They get the best teachers into the most challenging schools.”

The other co-author of the report, Eric A. Hanushek, an economist at the Hoover Institution at Stanford University, took this further by estimating “gains in PISA tied to replacing the most ineffective teachers with ‘average’ ones.” (as cited in Robelen):

“So if I replace 2 percent of the teachers with an average teacher, I get ... almost 25 PISA points, and if I replace 6 percent of the teachers, I get 50 PISA points, which is the difference between the [PISA results for the] United States and Canada,” he said. “If I can replace 10 percent of the teachers, ... we can beat Finland.”

Reinforcing the concerns cited earlier by Catlaks, it’s not inconceivable that presumably “ineffective” and “average” teachers would be determined by student test scores, with more effective teachers being enticed into more challenging schools through some type of merit pay scheme, all in order to boost test scores and move one’s ranking ahead of high-performing countries like Finland, or Canada for that matter.

Some observations on PISA's influence in Canada

So how far is PISA's reach, and by extension the OECD's influence and message into Canada? This section will provide a few examples of PISA's influence in the Canadian educational context.

In Canada PISA is administered through a partnership of the Council of Ministers of Education Canada (CMEC), Statistics Canada, and Human Resources and Skills Development Canada (HRSDC). At approximately 22,000 students, Canada's sample size for PISA 2006 was among the largest of participating countries. Approximately 20,000 students in 900 schools in every province were expected to participate in PISA 2009. Oversampling by provinces of the students to be tested allows comparisons to be made not just with other provinces and with Canada as a whole, but also with other countries. The CMEC Web site (www.cmec.ca/) notes that PISA 2009 results "will be reported at the country and provincial levels, with information on both the anglophone and francophone education systems of Manitoba, Ontario, Quebec, New Brunswick, and Nova Scotia." Results of PISA 2009 will be published in December 2010.

As we've seen, governments often interpret PISA results, positive or negative, to make the case for more testing and other similar reforms, and Canada is no exception. On the Canadian media coverage of the PISA 2006 results, Education International (2007) found that:

In Canada (British Columbia), the media relied solely on government sources to contextualize and interpret the PISA results even if there were several other sources they could have used such as unions, or researchers working on education. The result was that the government was able to use the good PISA results to say that the system with regular testing and standardized curriculum was effective and excellent, with these claims being uncontested and unchallenged by other sources. These findings are consistent with others who have looked at the media coverage of education Therefore only presenting one point of view, as [in] the case with Canada where only government sources were used, the media clearly creates a systematic bias in the coverage which almost inevitably will influence the way people think about education. And the media is not only able to select who should have a voice in the debate, they are also able to select how a story should be presented, i.e. they can *frame* the story. This brings us to an essential point with regard to the political use of PISA and what role the media plays in this. (pp. 18-19)

The mainstream media's over-reliance on government sources in the PISA study coverage, combined with its almost exclusive focus on the league tables and how other countries ranked compared to Canada, is a potent force for shaping public knowledge and opinion about our education systems. This creates enormous hurdles for teacher unions and other education organizations to put forth an alternative perspective on PISA results and more importantly, a broader perspective on the goals of public education.

Commenting on the relatively poor PISA performance of Atlantic Canada, educator and researcher Mike Corbett (2004) says that the release of PISA results "is always a subject for hand-wringing and apologetics on the part of Atlantic Canadian educators and bluster and recriminations from groups like the Atlantic Institute for Market Studies. As usual, Atlantic Canada's performance is mediocre by Canadian national standards and somewhere in the middle of the pack internationally". To understand why certain provinces consistently rank high and others rank low in PISA, Corbett says

it's important to look at factors such as provincial family income levels (poverty), percentage of adults who haven't completed high school or post-secondary education, and levels of school funding. Unfortunately such contextual factors rarely get reported on in the media in an attempt to explain the scores.

Another indication of the growing influence of the OECD on education policy making in Canada is the explicit alignment in large-scale external testing programs between the CMEC's Pan-Canadian Assessment Program (PCAP) and the OECD's PISA. PCAP is a cyclical national testing program administered every 3 years in English and French to a representative sample of students from across Canada. It assesses 13-year-old students in the domains of reading, mathematics and science. PCAP will be administered to about 40,000 students from over 1,500 schools across Canada in May 2010.

In his presentation at the CTF President's Forum, Raymond Théberge (Assistant Deputy Minister in the Ontario Ministry of Education and former CMEC Director General) stated that part of the rationale for PCAP (and the redesign of its predecessor, the School Achievement Indicators Program – SAIP) is to complement existing assessments in the provinces and territories and to provide data for thematic reports. Indeed PCAP and PISA “complement each other” in several dimensions: they have the same structure; test scores are presented on the same scale; the same cohort of students was assessed in 2009 through PISA; and the various jurisdictions are able to examine and compare patterns of performance between the PCAP and PISA assessments and validate these results with the results of their own provincial assessments.

Harmonization is also occurring with respect to education indicators in Canada. Coinciding with the release of the OECD's *Education at a Glance – OECD Indicators* report in September 2009, the CMEC issued a communiqué announcing the release of *Education Indicators in Canada: An International Perspective*, a product of the Pan-Canadian Education Indicators Program (PCEIP). PCEIP is an ongoing initiative of the Canadian Education Statistics Council, a partnership between CMEC and Statistics Canada formed in 1998. The release of the first PCEIP indicators report in 1996 was followed by PCEIP reports in 1999, 2003, 2005 and 2007. The traditional PCEIP publication has since evolved into this new report, linked with the release of the OECD's *Education at a Glance*, with the objective of “facilitating the comparison of educational systems in Canada's provinces and territories with those of countries that belong to the OECD.”

The *Education Indicators in Canada* report provides indicators on educational attainment, upper-secondary graduation, tertiary graduation, the academic performance of students, labour market outcomes, the economic benefits of education, expenditures on education, international students, and transitions to the labour market, for Canada, its provinces and territories, and for OECD countries. These indicators, based on the availability of data for the provinces and territories, are aligned or harmonized with the definitions and methodologies of selected indicators in the OECD report.

Such harmonization serves to facilitate the OECD's influence on education policy in Canada, arguably (indirectly) giving the CMEC considerable influence to shape education policy at the national level.

Another example of how the OECD's influence seeps into thinking about our education system is evident in the Conference Board of Canada's *How Canada Performs: A Report Card on Canada*. The section of the report on “Education and Skills” performance, in which Canada receives an “A” grade, focuses on using education to improve Canada's innovation (defined as “the ability to turn

knowledge into new and improved goods and services”), productivity, national economic competitiveness, and to address changing labour market needs and boost individual earnings. It’s a view of education as skills training, not surprising considering the lens through which the Conference Board views education. Most of the data sources used in the report (for the indicators) are drawn from the OECD including *Education at a Glance 2009* and PISA studies.

PISA findings and other OECD data are also used in the Canadian Council on Learning’s Composite Learning Index, a tool praised by the OECD (Mason, 2010).

Moving forward

Through the OECD's extensive reach and powerful influence and reputation as the go-to source for international comparative economic – and increasingly education – data, PISA results have become the data of choice for reporting on the performance of education systems in OECD member and partner countries.

In little more than a decade, the impact that the first three PISA cycles has had on shaping policy making in education to meet narrow economic objectives leads one to conclude that PISA is arguably a form of high-stakes testing. In this case the stakes are high with regard to PISA's observed impact on high-level policy and program decisions made by national governments and governments at other levels, through the auspices of a highly influential organization.

As PISA appears to be here to stay, ignoring it is not an option. Teacher unions need to be able to respond to PISA, even use the reports in a creative positive way. As noted earlier Finland uses PISA results to bolster its case for a strong teaching profession along with the necessary resources to properly support teachers.

EI's paper on alternative approaches for analysing and reporting countries' performance in PISA does provide some direction on this. Describing PISA as "an OECD triumph" (which it undoubtedly is from an OECD perspective) as it "provides governments, academics and voters with rich information about their education systems", Mortimore (2009) proposes that PISA could be "further enhanced" by doing the following (excerpted from p. 11):

- shifting the aims of PISA from a snapshot of national achievements to a more nuanced interpretation of countries' strengths and weaknesses in their development of lifelong learning
- refocusing on how schools and school systems could promote achievement and increase the equity of their educational outcomes
- widening the basis of assessment though inclusion of knowledge from geography and social science domains
- involving teachers in the design and development phases of the PISA cycles
- including information from teachers (through a questionnaire) to enrich the context of the data
- extending the methodology to include a longitudinal element
- reconfiguring to an alphabetical order the minority of tables currently presented in rank positions

While these recommendations for change may be necessary they may not be sufficient to counter the powerful trend towards using PISA results mainly to rank order countries.

Kuehn (2004) imagines the possibilities of an alternative assessment system modelled on the four pillars of education contained in the influential Delors report, *Learning: The Treasure Within* (Report

to UNESCO of the International Commission on Education for the Twenty-first Century), released in 1996. The pillars are: learning to live together, learning to be, learning to do, learning to know. As he notes, the OECD's education indicators project and PISA

are really only interested in the 'learning to do' element, the development of human capital for the economic system. A valid assessment system would not just focus on one element, but would be concerned with all four aspects of learning, some of which do not lend themselves at all to testing as a means of assessment. (pp. 65-66)

Robertson's analysis of the PISA findings provides an interesting example of how teacher unions and other education organizations can make more constructive use of PISA data. Analysing Canadian results from the first PISA cycle (conducted in 2000), she notes that "buried" in the Statistics Canada-produced report are a number of important findings, "predictably overlooked by the media's preoccupation with rankings", the most useful of which "were derived by integrating achievement scores with contextual information". The latter was gathered through a series of student, teacher and principal questionnaires as well as parent telephone interviews. Among the findings was that "Canadian schools have been more successful than most of their international counterparts in mitigating the effects on achievement of students' socioeconomic characteristics." (p. 68)

On this point Saskatchewan was particularly noteworthy for achieving better equity outcomes – measured using different indices of educational equity – than all other provinces and indeed better than any of the 31 other countries studied. Robertson concludes that "in a finding that ought to be making headlines everywhere, this report concludes that in Saskatchewan socioeconomic status exerts no statistically significant effect on student achievement among 15-year olds Something is going on in Saskatchewan that the rest of us should be paying attention to." (p. 71)

This example from Australia illustrates another missed opportunity, a logical path not taken. Australia's poor performance on PISA 2006 equity measures (the gap in its educational outcomes related to socioeconomic background exceeded the OECD average) prompted calls by government for "greater testing accountability regimes", rather than leveraging this finding to advocate for policies addressing the "structural inequities of the Australian education system, the public/private divide, or the underfunding of the public education system" (Figazzolo, 2009, p. 17). Again there are important lessons here for Canada and other countries participating in PISA.

Concluding observations

According to Hargreaves and Shirley (2009), even though it may not seem like it to educators swimming in the fishbowl of standardization, “we are entering an age of post-standardization in education.” (p. 1)

Encouraging signs of this new age include, as noted earlier, the case of Finland which “avoids national standardized tests altogether and reaches high levels of achievement by attracting highly qualified teachers with supportive working conditions, strong degrees of professional trust, and an inspiring mission of inclusion and creativity.” (p. 2) In a refreshing contrast to the incessant call for greater accountability on this side of the pond, Hargreaves and Shirley note that “there is no indigenous Finnish term for *accountability*. Instead, public education is seen as a collective social and professional *responsibility*. In direct opposition to the New Orthodoxy, Finland achieves high performance without constantly taking public measures of it.” (p. 54) [emphasis in original]

The rationale for an approach to accountability based on shared responsibility is well supported by the research. Hargreaves and Shirley state that,

...years of research in school effectiveness still show that most explanations for differences in student outcomes exist *outside* the school. Schools can and do make a significant difference but do not make all or even most of it. They cannot excel alone but need communities and society to work with and alongside them. This means more schools working with communities and more investment in community and family development. We can't expect to raise standards on the cheap. In schools as in business, there is no achievement without investment. High-performing Finland and other Nordic countries benefit from a welfare state that makes strong public investment not just in education, but also in housing, medical care, social services, and community development. While schools contribute a great deal, they are not expected to achieve miracles by themselves. (p. 79)

There is clearly much to be learned about the Finnish approach to education, particularly regarding its treatment of teachers and strong support for public education, accompanied by a strong social safety net.

Some constructive things are also happening in the UK. A government “White Paper” presented to the UK Parliament in June 2009 reviewed trends such as “the abandonment of some system-wide standardized tests and reduced prescription by the government in terms of delivery of literacy and numeracy lessons. Teachers in England and Wales are being given increased autonomy in how they deliver national curriculum. School league tables will be “superseded” by a form of report card which would include a much wider range of indicators than test scores.” (Naylor, 2009) Wales for example has abolished national testing for students up to the age of 14.

Commenting on the White Paper and what these shifting trends might signal for education policy in the UK, Baker (2009) – echoing Hargreaves and Shirley – suggests that we may be starting to rethink educational accountability in terms of an ‘intelligent’ accountability, one that recognizes the need for both high teacher autonomy and high accountability, the latter moving away from standardized testing:

...30 years ago we had a school system that was characterised by high autonomy and low accountability. Since then we have moved to the polar opposite: low autonomy and high accountability. The current shift seems to suggest that this is now becoming as

undesirable as the position 30 years ago. There is growing realisation, I believe, that a more desirable combination is: high autonomy and high accountability. So schools and teachers should be trusted more: trusted to devise their own curriculum, their own teaching styles, and their own forms of assessment. But the price of that freedom is strong accountability. But an accountability that is more intelligent than the rather crude measures we have had until now, which have distorted school behaviours.

There are some encouraging signs in Canada as well – for example the Alberta legislature voted to abolish the Grade 3 provincial test, and Nova Scotia has announced it will eliminate Grade 6 and Grade 9 provincial examinations (Hargreaves & Shirley, p. 2).

While the Canadian media portrayal of teacher and teacher union opposition to testing is all too common, in actual fact teachers are not opposed to testing. They do however have serious concerns about the misuse and overuse of external standardized testing, while supporting the need for broader assessments of student learning that emphasize more than achievement in math, science and literacy as measured by standardized test scores.

A rethinking of educational accountability, with genuine learning for all students as the overriding goal, would dispel the false notion that teachers are opposed to assessment. For example, the Alberta Teachers' Association, in a publication entitled *Real Learning First: The Teaching Profession's View of Student Assessment, Evaluation and Accountability for the 21st Century*, asserts that:

- the primary purpose of student assessment and evaluation is to support student learning, broadly conceived.
- ongoing student evaluation is an integral part of the teaching and learning process; toward this end teachers conduct two main types of evaluation: formative evaluation (assessment *for* learning) and summative evaluation (assessment *of* learning).
- students need timely constructive feedback that supports their learning.
- a variety of evaluation practices are required to determine student achievement, including performance assessments, projects, written work, demonstrations, portfolios, observations, examinations.
- data from these multiple assessments over a period of time are essential to informing teachers' judgments about student growth, development and learning.
- many factors can influence student achievement including individual learning needs, the resources available to support teaching and learning, and the socioeconomic characteristics of the community.
- classroom teachers design student evaluation based on the curriculum that students have been taught – it is unfair and unethical for teachers to evaluate students on material they have not had the opportunity to learn.

- classroom teachers are in the best position to develop evaluation strategies that align with the curriculum and address individual learning needs. (Alberta Teachers' Association, 2009)

Rethinking accountability in education would also need to address the purpose and design of large-scale external standardized testing. In this regard Hargreaves and Shirley make a case for the use of statistically valid random sampling of students for systemwide accountability in education. Despite an emerging consensus that "systemwide accountability ... can be achieved through prudent sampling rather than through a profligate and politically controlling census yet a shrinking number of governments hang on to accountability by census, even though it is subject to widespread abuse. They do it even though it is exorbitantly expensive – diverting scarce resources from teaching and learning needs elsewhere." (pp. 102-103)

Large-scale assessment, conducted through random sampling of students, should be used to assess aspects of the quality of the education system as a whole, such as curriculum effectiveness and how well the system is meeting the needs of particular groups such as Aboriginal students, students with special needs and students from poor families. At the same time a properly designed system could address some of the concerns related to the use of large-scale external testing as a means of sorting students and the ranking and reporting of schools and provinces based on single test results. A number of teacher organizations support the use of random sample-based assessments.

Commenting on the respective roles of external testing and classroom assessment, STF General Secretary Gwen Dueck argues that we must strike the right balance between classroom-based assessment for broad-based student learning conducted by teachers, and large-scale external testing programs for assessing the effectiveness of education systems through a process of random sampling (rather than a sweeping census approach). Drawing on the work of Andy Hargreaves and Pasi Sahlberg, she points out that,

rather than validating the current external accountability milieu we find ourselves in, what we need are new accountability policies and practices that respect the professionalism of teachers and the commitments they bring to the profession. Intelligent accountability, as they term it, builds on mutual accountability, professional responsibility and trust. An accountability framework such as this utilizes a wide variety of data. It combines internal accountability or school-based assessment as some might refer to it, which consists of school processes, self-evaluations, critical reflection and school-community interaction, with levels of external accountability that build on monitoring, sample-based assessment and thematic evaluations appropriate to each individual school and context.

Dueck frames some of the critical challenges ahead, remarking at the CTF President's Forum on External Assessment that,

during this forum we have talked much about how we might validate the practice of 'external assessment' or 'large-scale assessments' and not enough about what we should be teaching our students. What are the critical elements of a holistic, broad based education and how do we measure those elements?

Someone in one of the small group discussions framed it this way: “What knowledge is going to be required for the future our children will experience?” To that, I pose the question: “How do we design a system that shifts our attention from outcomes to that of the inputs, process and context in which we as educators carry out our professional responsibilities and commitments?”

References / Sources

- Alberta Teachers' Association (2010). *Changing Landscapes of the Next Alberta: Shaping a Preferred Future 2009-2029*.
www.teachers.ab.ca/SiteCollectionDocuments/ATA/Features/2009-10/Changing%20Landscapes%20of%20the%20Next%20Alberta.pdf
- Alberta Teachers' Association (2009). *Real Learning First: The Teaching Profession's View of Student Assessment, Evaluation and Accountability for the 21st Century*, Edmonton, AB.
- Baker, Mike (July 5, 2009). "Is it time to trust the teachers?" *BBC News*.
http://news.bbc.co.uk/2/hi/uk_news/education/8133820.stm
- Canadian Teachers' Federation (2010). *Highlights of CTF Teacher Survey – Fostering Responsible Citizenship*.
- Canadian Teachers' Federation (Oct. 2008). "2007-08 Consolidated Provincial/Territorial-Local Government Finances."
Economic & Member Services Notes.
- Catlaks, Guntars (2010). "OECD and relevant policy on quality teaching and measurement and assessment of teachers' performance." Workshop presentation at Education International conference, "Education Beyond the Crisis: Responses to New Challenges", New York, March 2010.
- Conference Board of Canada (Jan. 2010). *How Canada Performs: A Report Card on Canada*.
www.conferenceboard.ca/HCP/Details/education.aspx#what_if
- Corbett, Mike (Dec. 18, 2004). "The learning tower of PISA." Canadian Centre for Policy Alternatives – Nova Scotia Office.
www.policyalternatives.ca/publications/commentary/learning-tower-pisa
- Dueck, Gwen (2009). Panel presentation at the CTF President's Forum on External Assessment, Ottawa, July 14-16, 2009.
- Education International (Sept. 2009). *Public Private Partnerships in Education*. Brussels.
- Education International (2007). *Education International Guide to PISA 2006*.
www.ei-ie.org/annualreport2007/upload/content_trsl_images/435/Education_International_Guide_to_PISA_2006.pdf
- Figazzolo, Laura (2009). *Testing, Ranking, Reforming: Impact of PISA 2006 on the Education Policy Debate*. Brussels: Educational International.
- Figazzolo, Laura (March 2009). "PISA: Is testing dangerous?" *Worlds of Education* (Education International), Issue 29.
- Forte, Nick (Summer 2009). "The student's freedom to learn requires the educator's freedom to teach." *Our Schools/Our Selves*, 18(4), pp. 25-33.
- Froese-Germain, Bernie (2004). *Educational Accountability With a Human Face*. Ottawa: Canadian Teachers' Federation.
- Froese-Germain, Bernie (1999). *Standardized Testing: Undermining Equity in Education*. Ottawa: Canadian Teachers' Federation.
- Gutstein, Donald (May 17, 2010). "War on public schools rages." *Rabble News* (rabble.ca).
- Hargreaves, Andy, & Shirley, Dennis (2009). *The Fourth Way: The Inspiring Future for Educational Change*. Thousand Oaks, CA: Corwin Press.
- Hugonnier, Bernard (2009). "OECD International Comparative Surveys: Impact on Education Policy and Governance." Presentation at the CTF President's Forum on External Assessment, Ottawa, July 14-16, 2009.

- Karp, Stan (Spring 2010). "School reform we can't believe in." *Rethinking Schools Online*, 24(3).
www.rethinkingschools.org/archive/24_03/24_03_NCLBstan.shtml
- Kohn, Alfie (Jan. 14, 2010). "Debunking the case for national standards." *Education Week*, 29(17), Editorial Projects in Education.
www.edweek.org/ew/articles/2010/01/14/17kohn-comm.h29.html?tkn=WXUFzMageZQH55f62lISx04pEy9h1dTAV7g0
- Kuehn, Larry (2004). "Leaning between conspiracy and hegemony: OECD, UNESCO and the tower of PISA." (pp. 57-66) In M. Moll (ed.), *Passing the Test: The False Promises of Standardized Testing*. Ottawa: Canadian Centre for Policy Alternatives.
- Kuehn, Larry (June 1999). "Globalization and the control of teachers' work: The role of the OECD indicators." *Our Schools/Our Selves*, 9(6), pp. 116-129.
- Mason, Gary (Jan. 9, 2010). "Behind the political decision to cut the lifeline: Tories didn't like CCL's message or its independence." *Globe and Mail*.
- Mortimore, Peter (2009). *Alternative Models for Analysing and Representing Countries' Performance in PISA*. Brussels. [Expert paper commissioned by the Education International Research Institute]
www.ei-ie.org/research/en/documentation.php
- Naylor, Charlie. "High autonomy and high accountability: Change in the air for the English and Welsh school systems?" BCTF Research, Sept. 30, 2009.
- Organisation for Economic Co-operation and Development. *PISA – The OECD Programme for International Student Assessment*.
www.oecd.org/dataoecd/51/27/37474503.pdf
- OECD (2010). *The High Cost of Low Educational Performance: The Long-Run Economic Impact of Improving PISA Outcomes*, Paris.
www.oecd.org/dataoecd/11/28/44417824.pdf
- Robelen, Erik W. (Jan. 25, 2010). "Study links rise in test scores to nations' output." *Education Week*, 29(19), Editorial Projects in Education.
www.edweek.org/ew/articles/2010/01/27/19costs.h29.html?tkn=QSWFXrlyTCP7TXJmzzLd2Aehxr63RZJgeJ8l
- Robertson, Heather-jane (2004). "An overlooked success in an under-appreciated province." (pp. 67-71) In M. Moll (ed.), *Passing the Test: The False Promises of Standardized Testing*. Ottawa: Canadian Centre for Policy Alternatives.
- Robinson, David (2009). "The Mismeasure of Education." Presentation at the CTF President's Forum on External Assessment, Ottawa, July 14-16, 2009.
- Sahlberg, Pasi. (2008). *Real Learning First: Accountability in a Knowledge Society*. Paper prepared for the Alberta Teachers' Association conference on *Leadership in Educational Accountability: Sustaining Professional Learning and Innovation in Alberta Schools*, Edmonton, Alberta, Canada, April 18-19, 2008.
www.pasisahlberg.com/downloads/Real%20Learning%20First%20FINAL.pdf
- Sinyolo, Dennis (2009). "International Perspectives." Presentation at the CTF President's Forum on External Assessment, Ottawa, July 14-16, 2009.
- St. Denis, Verna (2010). *A Study of Aboriginal Teachers' Professional Knowledge and Experience in Canadian Schools*. Ottawa: Canadian Teachers' Federation / Canadian Council on Learning.
- Stack, Michelle (2006). "Testing, testing, read all about it: Canadian press coverage of the PISA results." *Canadian Journal of Education*, 29(1), pp. 49-69.
- Théberge, Raymond (2009). "National Implications: CMEC and Beyond." Presentation at the CTF President's Forum on External Assessment, Ottawa, July 14-16, 2009.

Thompson, George (Nov. 6, 2009). "The global privatization of education policy: Lorna Earl conflict of interest is the tip of the iceberg." *The Daily Censored*.

<http://dailycensored.com/2009/11/06/the-global-privatization-of-education-policy-lorna-earl-conflict-of-interest-is-the-tip-of-the-iceberg/>

Westheimer, Joel (Summer 2008). "What kind of citizen? Democratic dialogues in education." *Education Canada*, 48(3), pp. 6-10.

Westheimer, Joel (Spring 2010). "No child left thinking: Democracy at risk in Canada's schools." *Education Canada*, 50(2), pp. 5-8.

http://cea-ace.ca/media/en/Democracy_at_Risk_Spr10.pdf

Yakabuski, Konrad (Sat. April 3, 2010). "A teachable moment for American schools." *Globe and Mail*, p. A13.

www.theglobeandmail.com/news/opinions/a-teachable-moment-for-american-schools/article1520908/

Endnotes

¹ Funding for public education in Canada as a proportion of provincial budgets has decreased over time. The share of consolidated provincial/territorial and local government spending in Canada allocated to elementary and secondary education dropped to its lowest level in two decades, falling from a high of 14.6% in 1992-93 to 11.8% in 2007-08 (Canadian Teachers' Federation, 2008).

² OECD Deputy Director for Education Bernard Hugonnier's keynote presentation at the CTF President's Forum focused on these new international comparative surveys being conducted by the OECD.

³ Dennis Sinyolo, Senior Coordinator in Education International (EI)'s Education and Employment Unit, raised some of these issues in his presentation at the CTF President's Forum on External Assessment, citing the lessons learned from PISA through EI's research.