Government of Canada Gouvernement du Canada

cmec
Council of Ministers of Education, Canada
Conseil des ministres de l'Éducation (Canada)

Learning from PISA and Other Large-Scale Assessment Projects

Proceedings of the 2018 Pan-Canadian Research Conference
Council of Ministers of Education, Canada
October 3–4, 2018
Metro Toronto Convention Centre
Toronto, Ontario

Table of Contents

Foreword

It is with great pleasure and pride that I write this foreword to the *Proceedings of the 2018 Pan-Canadian Research Conference: Learning from PISA and Other Large-Scale Assessment Projects*. In collaboration with Employment and Social Development Canada (ESDC), the Council of Ministers of Education, Canada (CMEC), organized the conference, which was held on October 3–4 in Toronto, Ontario, to highlight the use of Canadian data in studies related to the Programme for International Student Assessment (PISA) and other large-scale assessment (LSA) projects and to promote the use of data to help address relevant policy issues.

Over 100 educators, researchers, policy advisors, and scholars attended the event, which included two distinguished keynote speakers, 11 presentations, a round-table discussion on uses of international LSA data by provinces and other countries, and a technical workshop on the analysis of the PISA data set. All video presentations and keynote addresses are available on the conference Web site at https://webcasts.welcome2theshow.com/prc-cpr2018.

Welcome addresses were given by Chantal C. Beaulieu, Executive Director, CMEC; Gilles Bérubé, Director, Labour Market and Skills Research Division, ESDC; and Jane Rooney, Financial Literacy Leader, Financial Consumer Agency of Canada (FCAC).

The first keynote speaker, Louis Volante from Brock University, set the stage for the conference with a speech entitled “PISA, Achievement Trends, and Educational Policy Reform.” Dr. Volante explained how LSA outcomes are being used as output measures and provided analyses of PISA data that included Canadian trends, international comparisons, and information on equity. He also summarized findings from a book he recently co-edited, *Immigrant Student Achievement and Education Policy*, and introduced an upcoming publication on socioeconomic inequality and student outcomes: *The PISA Effect on Global Educational Governance*.

The second keynote speaker was Michael Stevenson, Senior PISA Advisor at the Organisation for Economic Co-operation and Development (OECD). Dr. Stevenson presented a “PISA Long-Term Roadmap” and invited comments and suggestions on the strategic direction of the OECD in relation to PISA 2024 and beyond. He also described how the OECD could potentially enhance its support for teachers and teaching by unlocking the value of PISA and the Teaching and Learning International Survey (TALIS).

Conference presentations were organized according to four themes:

- opportunity to learn
- exploring achievement gaps
- what the data tell us about schools and classrooms
- beyond the core competencies

Under the first theme, Anna Katyn Chmielewski, from the Ontario Institute for Studies in Education (OISE) of the University of Toronto, summarized 50 years of data measuring the socioeconomic gap in Canada and across the world. Through multivariate modelling, Dr. Chmielewski demonstrated how educational achievement is an important predictor of later educational transitions and explained why Canada should be concerned about the increasing advantage of high-socioeconomic status students. Her presentation was followed by that of Ashley Pullman, from the University of Ottawa, who described her current research focusing on youths not in education, employment, or training (NEET) in Canada. Although NEET youth performed less well than non-NEET youth on measures of literacy, numeracy, and problem solving, the performance difference between these two groups could be accounted for in large part by variables related to skill-based activities at home and self-reported learning strategies. Dr. Pullman concluded that home-based activities such reading and using information and communication technology (ICT) are associated with a lower probability of becoming NEET.

Under the second theme, exploring achievement gaps, Alfred Sakyi and Janusz Zieminski presented a detailed analysis of the Trends in International Mathematics and Science Study (TIMSS) data to evaluate the extent to which student performance in Alberta is affected by differences between what TIMSS assesses and what the Alberta mathematics curriculum covers. Dr. Sakyi and Mr. Zieminski also showed that mathematics coverage in the Quebec curriculum could help explain why students in that province performed better than those in Alberta on TIMSS 2015 but noted that other factors, such as teacher characteristics and teaching practices, may have a stronger link with achievement. In the second presentation on this theme, a team from the Higher Education Quality Council of Ontario (HEQCO) presented its Essential Adult Skills Initiative, whereby a measure of literacy, numeracy, and problem-solving skills was piloted in 20 postsecondary institutions. Based on this study, Hudak and her colleagues determined that large-scale skills assessments are feasible in postsecondary education settings. Their initial results also suggest that one in four postsecondary students does not possess the skills to succeed in life and work. The third speaker was Brad Seward, who presented the Research Initiative on Education and Skills (RIES) on behalf of his team at the Mowat Centre. The objective of this new initiative is to provide access and analysis capacity for research on data related to the Programme for the International Assessment of Adult Competencies (PIAAC) and the Longitudinal and International Study of Adults (LISA) to inform skills development policy. A call for proposals was launched in October 2018 to identify research teams that would benefit from the expertise and access to data of the RIES team. Dr. Seward also provided an example of a research project looking at female involvement in the ICT sector and why increased female participation in this field might help employers fulfil labour demands in the future.

Under the third theme, what the data tell us about schools and classrooms, Alexander MacDonald explored how strong evidence derived from the Pan-Canadian Assessment Program (PCAP) and PISA, along with data from the Provincial Common Assessments, has guided the development of targeted initiatives to improve teachers' instructional skills throughout Prince Edward Island. Dr. MacDonald reported that data from the most recent rounds of PCAP and PISA confirmed the significant improving trend observed in provincial assessments. The second presentation, entitled "Exploring the Effects of Inquiry-Based Strategies and Teacher-Directed Instruction on Student Performance in Science Using PISA Data," was given by Alfred Sakyi, from Alberta Education. Using data from PISA 2015, Dr. Sakyi showed that Alberta students were frequently exposed to both inquiry-based science teaching and learning practices and teacher-directed science instruction . Analysis suggested that these factors may have contributed to a greater understanding and enjoyment of science by students, which in turn may have affected science achievement positively in the province. The third presenter under this theme was Pierre Brochu. In his paper "What's in It for Them? Generating School-Level Results Based on PISA: Possibilities and Limitations," Dr. Brochu shared new analyses of school-level results based on PISA 2015 in the province of Manitoba. This type of analysis has not been exploited to its fullest potential in the past, and it may provide useful information to guide education policy. He also presented an example of a school report that could potentially be shared with Canadian schools participating in PISA in the future.

Under the last theme, beyond the core competencies, the first team of presenters was Muriel Péguret, Dominique Scheffel-Dunand, and Liam Bekirsky, from Glendon College, York University. In their paper "A Perspective of Ontario Pre-University Teachers on Global Competencies," the authors presented results from a teacher survey they conducted through the Ontario Modern Language Teachers' Association (OMLTA) to measure the extent to which global competencies are integrated into the curriculum in language classrooms. Based on this small-scale study, they concluded that teachers were not equally prepared to integrate global competencies into their teaching; results varied significantly by geographic location, teaching experience, subjects of instruction, subjects of qualification, teaching level, language of instruction, and first language. They also noted the lack of training relating to global competencies for pre-service teachers in Ontario. The second group of presenters, David Hagen Cameron, Annie Kidder, and Christine Schandl from People for Education, presented findings from two studies related to transferable skills in a paper entitled "Tracking Resources and Infrastructure in Schools: A Proxy Measure of System Capacity to Support All Students in Creativity, Social-Emotional Development, Citizenship, and Health" The first study examined how educators are using different approaches in the development of competencies in creativity, social-emotional learning, citizenship, and health within the classroom while addressing the kind of diverse learning opportunities and resources that schools require. In the second study, results from a large-scale survey suggest that development of these competencies requires public, political, and financial investment that goes beyond simple policy rhetoric. The last paper under this theme, entitled "PISA 2015: Do Canadian Students Have the Right Skills and Attitudes toward Science?" was presented by Marie-Anne Deussing from ESDC and Tanya Scerbina and Yitian Tao from CMEC. Based on their analyses, they conclude that Canadian 15-year-olds have the science skills that will prepare them for the jobs of tomorrow. Young Canadians also enjoy learning about science, are confident in their science ability, express interest in science topics, and recognize the importance of science in their world. Finally, proportionally more students in Canada than in other PISA countries expect to pursue a career in a science-related field.

The last session of the conference was devoted to a round table on Provincial and Country Uses of International Large-Scale Assessment Data. The participants provided a broad view of the potential uses of LSA data, from teacher professional development to evidence-based decision making to inform policy issues. Dr. Elizabeth Costa and Laura Brake, from the Department of Education, Early Learning and Culture in Prince Edward Island, described their recent experience using LSA data to increase student achievement in mathematics in the province through formative assessment and a comprehensive professional development initiative. Nancy Walt from the Ministry of Education in British Columbia showed how LSAs are an integral part of the renewed provincial assessment. She explained how data from assessments at the classroom, provincial, national, and international level are used to support decision making through an iterative process based on feedback and reports from various levels in the education system. Jennifer Maw, from Manitoba Education and Training, presented research on how well PCAP results correlated with results from the Manitoba provincial middle years assessment for mathematics and reading. Maie Kitsing, from the Ministry of Education and Research in Estonia, showed examples of how quantitative data from international and national surveys, as well as national databases, can be used as quality indicators for schools in the country. Finally, Dr. Erja Vitikka and Anneli Rautiainen, from the Finnish National Agency for Education, explained how national authorities work with local administrations to implement a number of new initiatives to enhance equality in education throughout the country. I also had the opportunity to participate in the round table by providing examples of how LSAs can be used (or misused) for improvement at the student, school, and system levels. All presentations for the round table session are available on the CMEC Web site.

I hope that these *Proceedings* will provide a solid reference for those interested in using PISA and other LSAs to guide education policy for the benefit of students. I also hope that they will encourage other researchers to use the remarkable wealth of data from LSAs to advance education research in Canada. As the body of research grows, CMEC and ESDC will consider additional ways to disseminate it.

I would like to thank all keynote speakers, authors, and participants for their contributions and all my colleagues at CMEC for their outstanding support to make this conference such a success.

Dr. Kathryn O'Grady
Coordinator, Learning Assessment Programs and Conference Chair

The Widening Socioeconomic Achievement Gap in Canada and Worldwide, 1964–2015

Anna Katyn Chmielewski
Ontario Institute for Studies in Education

Abstract

The existence of a socioeconomic achievement gap—a disparity in academic achievement between students from high- and low-socioeconomic status (SES) backgrounds—is well known in Canadian education policy research. The SES achievement gap has been documented across Canada and a wide range of other countries. There has been less evidence, however, on whether SES achievement gaps might be *changing* over time. In a new study (Chmielewski, forthcoming), I show that, over the past 50 years, SES achievement gaps have increased in a large number of countries around the world. In that study, I draw on data from 30 international large-scale assessments over 50 years, representing 100 countries and about 5.8 million students. I find that achievement gaps have increased in a majority of sample countries for each of the three available measures of family SES: parental education, parental occupation, and the number of books in the home. In the current paper, I report the trend in the Canadian SES achievement gap in more detail. I find that in Canada, the SES achievement gap has increased substantially, growing at a rate that outpaces 80 per cent of other sample countries. Unlike in most other high-income countries, the Canadian achievement gap has increased primarily between the middle and the top of the SES distribution, raising concerns about an increasing concentration of advantage for high-SES students.

Introduction

The existence of a socioeconomic achievement gap—a disparity in scores on tests of academic achievement between students from high- and low-socioeconomic status (SES) backgrounds—is well known in Canadian education policy research (Brown & Sinay, 2008; Davies, Cyr, Rizk, & Janus, 2016; O'Grady & Houme, 2015; Willms, 1999). International assessments show that SES achievement gaps are present across a wide range of countries, and indeed far larger in most countries than in Canada (Mullis, Martin, Foy, & Hooper, 2016; Organisation for Economic Co-operation and Development [OECD], 2016). This suggests that, in most societies, low-SES children do not receive the same learning experiences in and/or out of school as their high-SES counterparts. There has been less evidence, however, on whether SES achievement gaps might be *changing* over time. In a new study (Chmielewski, forthcoming), I show that, over the past 50 years, SES achievement gaps have increased in a large number of countries around the world. Other recent research shows that SES achievement gaps have increased in three individual countries: the United States (Reardon, 2011), South Korea (Byun & Kim, 2010), and Malaysia (Saw, 2016). Is this trend also occurring in Canada? In the current paper, I report the trend in the Canadian SES achievement gap in more detail. I draw on evidence from the Canadian sub-samples of 15 international large-scale assessments, spanning 20 years (1995–2015) and representing about 360,000 students.

Theoretical framework

Previous research suggests several possible candidates for trends that could drive increasing SES achievement gaps in Canada and many other countries. First, the population of students enrolled in schools has become more diverse. Primary and secondary school enrolment has become increasingly universal in both developed and developing countries (Baker, Goesling, & LeTendre, 2002). Countries with rapidly expanding school access may experience growing SES achievement gaps due to the inclusion of relatively disadvantaged populations.

Second, increasing SES achievement gaps may be due to dramatic world growth in income inequality (OECD, 2015; United Nations Development Programme [UNDP], 2013). Countries with rising income inequality may experience widening SES achievement gaps due to increasing disparities in the material resources of low- and high-SES families, as well as possible corresponding increases in neighbourhood segregation by income (Hulchanski, 2010; Musterd, Marcińczak, Van Ham, & Tammaru, 2017; Reardon & Bischoff, 2011; Walks, 2013).

Third, changing education policies could cause rising SES achievement gaps. Despite growing educational access and de-streaming reforms worldwide (Ariga, Brunello, Iwahashi, & Rocco, 2005; Benavot, 1983; Manning & Pischke, 2006), market-based education policies may create new inequalities. School choice and privatization are growing in Canada and many other countries around the world (Bosetti, Van Pelt, & Allison, 2017; Davies & Aurini, 2011; Eyles & Machin, 2015; Parekh & Gaztambide-Fernández, 2017). If such policies increase SES segregation among schools, causing students of different backgrounds to experience increasingly differentiated learning environments, this may explain growth in SES achievement gaps.

Fourth, increasing SES achievement gaps could be due to increasing disparities in parental investments of time and money in children. Private household expenditures on children, such as child care, school tuition, and private tutoring, appear to be growing dramatically and becoming more unequal between SES groups in Canada and a number of other countries (Aurini, Davies, & Dierkes, 2013; Kornrich, Gauthier, & Furstenberg, 2011, April; Park, Buchmann, Choi, & Merry, 2016). Likewise, parental time-use surveys across a range of countries show both increasing time spent on child care and increasing SES disparities in child-care time (Dotti Sani & Treas, 2016; Gauthier, Smeeding, & Furstenberg, 2004).

Data sources

The data for this paper are derived from a sub-sample of 15 of the 30 international large-scale assessments analyzed in Chmielewski (forthcoming) for which Canadian data are available. All assessments test math, science, and/or reading; they span 20 test years (1995–2015) and 24 cohort birth years (1981–2005). The assessments include:

- four cycles of the Trends in International Mathematics and Science Study (TIMSS): TIMSS 1995 (Grades 4 and 8), TIMSS 1999 (Grade 8), TIMSS 2007 (Grades 4 and 8), and TIMSS 2015 (Grades 4 and 8);
- two cycles of the Progress in International Reading Literacy Study (PIRLS): 2006 and 2011; and
- six cycles of the Programme for International Student Assessment (PISA): 2000, 2003, 2006, 2009, 2012, and 2015.

Some cycles of TIMSS and PIRLS are not fully nationally representative because some provinces did not participate. I include in the analysis only cycles that represent at least 70 per cent of the Canadian student population. TIMSS 2007 (Grades 4 and 8) and PIRLS 2006 meet this standard because they include the most populous provinces, but they are not fully nationally representative. The other 13 included assessments are nationally representative.

Variables

Full descriptions of the math, science, and reading skills assessed in each study are available from the official published reports of the International Association for the Evaluation of Educational Achievement (IEA) and the OECD. Since each test is on a different scale, in the main trend analyses that combine different studies, I standardize all scores to a mean of 0 and standard deviation of 1 within each country-study-year before calculating each SES achievement gap.[1]

SES In each data set, at least one of the following three measures of family socioeconomic status is available: parental education, parental occupation, or the number of books in the household. For parental education and occupation, I use the higher level of those achieved by the two parents. All SES variables are reported in ordered categories; the number of categories varies somewhat by study. All SES variables are student reported except in three recent studies—PIRLS 2001 and 2011 and TIMSS 2015 (Grade 4)—in which they are parent reported.[2]

[1] In Chmielewski (forthcoming), I show that results are similar when trends in each test instrument (i.e., TIMSS, PISA, or PIRLS) are analyzed separately.

[2] In Chmielewski (forthcoming), I show that trends in gaps based on each SES variable are similar when analyzed separately.

Subject All assessments test math, science, and/or reading; the main trend analyses pool gap estimates of all subjects.[3]

Cohort birth year I compute the mean birth year for each country-study from student reports either of birth year and month or of age in years and months, relative to the known year and month of testing in each country. I use survey weights when calculating means.

Age at testing Students are either in Grade 4/age 10 (TIMSS and PIRLS), Grade 8/age 14 (TIMSS), or age 15 (PISA). The main trend analyses pool gap estimates of all age groups.[4] Note that cohort birth year and age at testing are not collinear because observations come from a wide range of years.

Methods

Missing data for all student-level variables except achievement are imputed using multiple imputation by iterative chained equations, creating five imputed data sets for each country-study.[5] The plausible values of achievement included in the PISA, TIMSS, and PIRLS data sets can also be understood within a multiple imputation framework, and are therefore included in this procedure.[6]

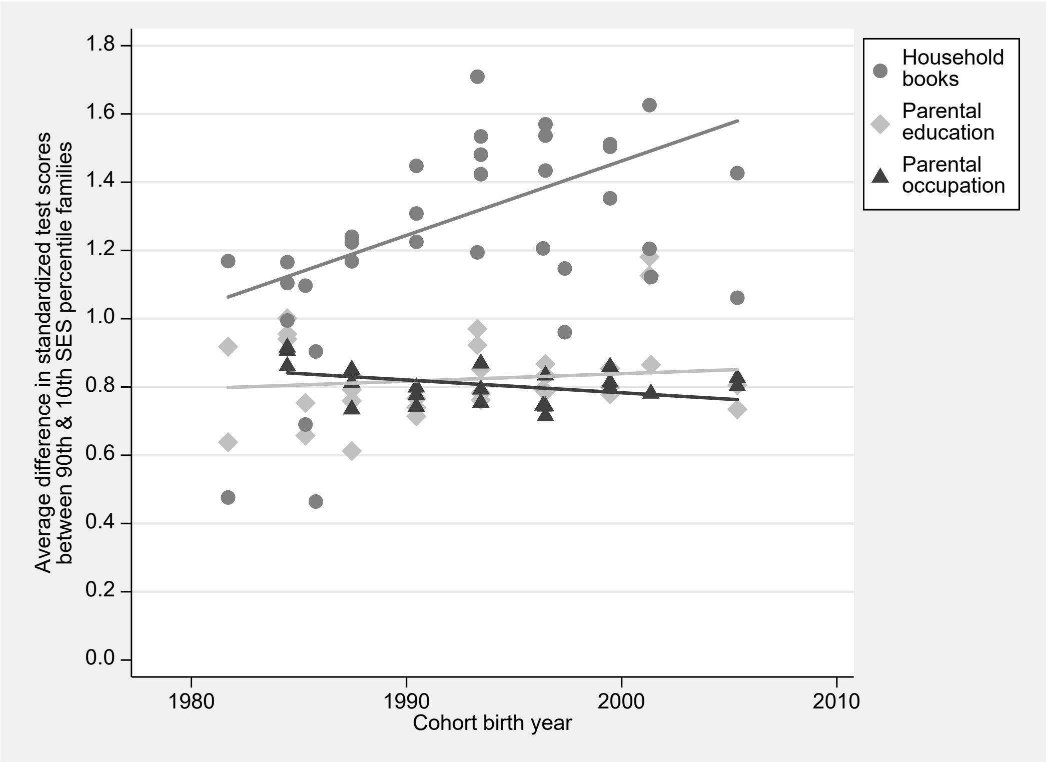
In each study, for each subject–SES variable combination, the SES achievement gap is computed as the gap in standardized achievement between the country-specific 90th and 10th percentiles of SES, following Reardon's (2011) method for income achievement gaps. I also compute gaps between the top and middle (90th and 50th percentiles) of the SES distribution and between the middle and bottom (50th and 10th percentiles) of the SES distribution. I adjust each gap for the estimated reliability of students' or parents' reports of SES, as well as for each test. I compute bootstrap standard errors for each gap. (See Chmielewski [forthcoming] for more methodological details.)

Results

Figure 1 shows estimated trends in 90/10 SES achievement gaps for Canada for each of the three SES variables: parental education, parental occupation, and household books. Each data point is the estimated achievement gap between students at the 90th and 10th percentiles of one SES variable in the Canadian sub-sample of a particular international assessment. The gaps are plotted against the birth year of sampled students, running from approximately 1981, corresponding to 14-year-old students tested in TIMSS 1995, to approximately 2005, corresponding to 10-year-old students tested in TIMSS 2015. Three fitted linear trends are estimated using weighted least squares regression analysis (weighting by the estimated inverse sampling error variance of each gap) to describe the three trends in gaps based on each SES variable.

[3] In Chmielewski (forthcoming), I show that trends in gaps for each subject are similar when analyzed separately.
[4] In Chmielewski (forthcoming), I show that trends in gaps for primary and secondary school students are similar when analyzed separately.
[5] In Chmielewski (2018), I show that trends are similar when gaps are computed after using listwise deletion rather than multiple imputation of missing data.
[6] PISA 2015 used 10 rather than 5 plausible values of achievement. Thus, 10 imputed data sets were generated and combined with the 10 plausible values of achievement.

FIGURE 1 Trends in 90/10 SES achievement gaps in Canada, by SES variable, 1981–2005 birth cohorts



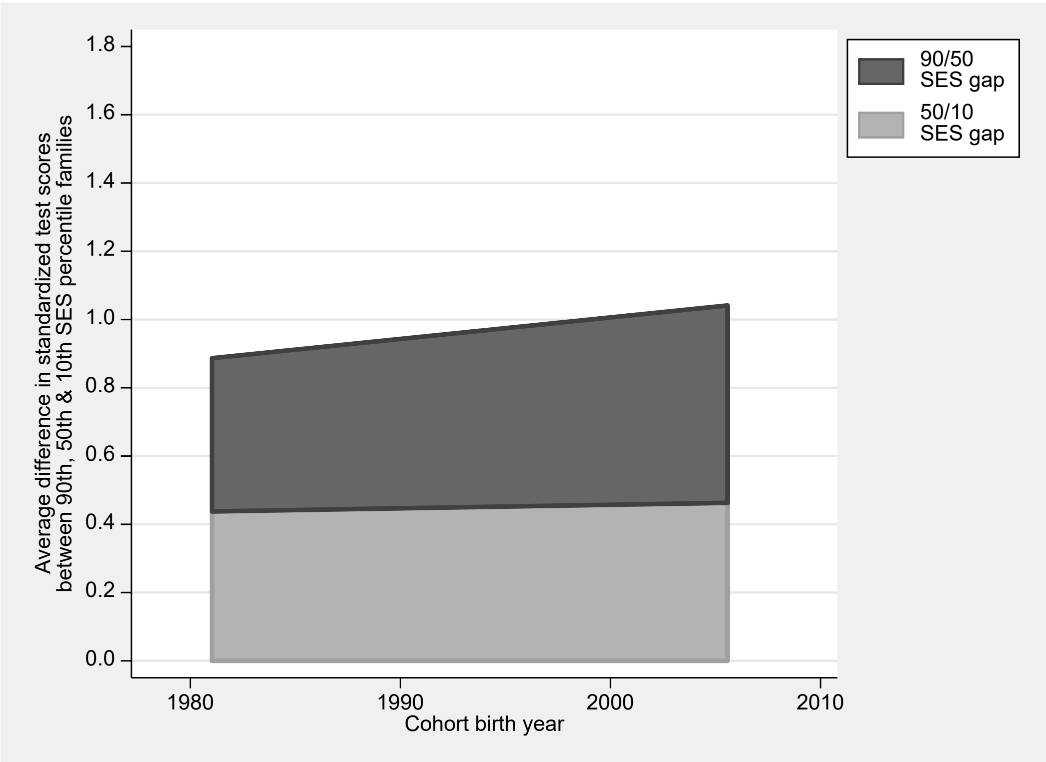
Note: Observations are raw, unadjusted 90/10 SES achievement gaps, pooling all available Canadian studies, SES variables, test subjects, and age groups. Fitted linear trends are weighted by the inverse of the sampling variances of the estimated gaps.

The achievement gap between the 90th and 10th percentiles of parental education has increased by 7 per cent, from 0.8 SD in the 1981 birth cohort to 0.85 SD in the 2005 cohort. The gap based on parental occupation has declined by about 9 per cent. By contrast, the gap based on household books has increased dramatically—by about 50 per cent. The difference in the trend for achievement gaps based on books may imply that household books are gaining salience relative to parental education and occupation in predicting children's academic achievement. However, the discrepancy is also likely to reflect differences in data quality. In later years, large proportions of Canadian students fall into the top categories of parental education and occupation, making it difficult to precisely estimate achievement at the 90th percentile of SES. This issue affects not only Canada but also several other high-income countries and appears to cause achievement gaps based on parental education and occupation but not household books to be underestimated in later years.

In the larger international study (Chmielewski, forthcoming), I run a model that pools all available gap types (based on different SES variables as well as different achievement subjects) to obtain the best estimate of the true trend in the SES achievement gap for each country. In doing so, I assume that, although achievement gaps by each SES variable do not have identical meanings (and may be differentially affected by the data quality issues mentioned above), any observed *trend* in gaps across cohorts is driven by the same underlying process. Based on such a model, the Canadian 90/10 SES achievement gap is estimated to be increasing at a rate of nearly 0.01 SD of achievement per year, or over 0.2 SD across the full span of 24 cohort birth years in the data.

Figure 2 displays trends in 90/50 and 50/10 achievement gaps. The dark grey region represents the gap between the top and middle of the SES distribution (the 90th and 50th percentile), and the light grey region represents the gap between the middle and bottom of the distribution (50th and 10th percentile). Trends are estimated via weighted least squares by pooling all available 90/50 or 50/10 gaps for all SES variables and subjects. The figure shows that, while the 50/10 SES achievement gap has increased slightly in Canada over this period, the 90/50 gap has increased far more. Approximately 85 per cent of the overall increase in the 90/10 achievement gap has occurred between the 90th and 50th percentiles of the SES distribution. Further analyses show that this result holds when gaps are analyzed separately by SES variable. For both parental education and books, 90/50 gaps have increased more than 50/10 gaps. For parental occupation (where overall 90/10 gaps are narrowing slightly), 90/50 gaps have decreased less than 50/10 gaps. Thus, it appears that in Canada, the achievement of high-SES students is rapidly pulling away from that of middle- and low-SES students.

FIGURE 2 Trends in 90/50 and 50/10 SES achievement gaps in Canada, 1981–2005 birth cohorts



Note: Trends are based on raw, unadjusted 90/50 and 50/10 SES achievement gaps, pooling all available Canadian studies, SES variables, test subjects, and age groups. Fitted linear trends are weighted by the inverse of the sampling variances of the estimated gaps.

To put the Canadian results in a global context, we can compare them to the international results from Chmielewski (forthcoming). The international data cover a longer time period than the Canadian data, spanning 51 test years (1964-2015) and 55 cohort birth years (1950-2005). In the international results, for all three SES variables, I find that SES achievement gaps have increased in the majority of sample countries. On average across all sample countries, achievement gaps based on parental education and occupation are increasing at a rate of about 0.007 SD of achievement per year and gaps based on household books by about 0.008 SD per year. For all three gap types, this is equivalent to a gap increase of about 0.4 SD of achievement between the 1950 and 2005 birth cohorts. However, there is a great deal of cross-national variation in trends in SES achievement gaps, with some countries experiencing gap increases of over 0.01 SD per year or 0.55 SD total over 55 cohort years (e.g., Belgium, Ireland, Thailand, Poland, and Hungary) and others experiencing flat or slightly declining achievement gaps (e.g., England, Finland, the United States, Japan, Mexico, and Brazil). Overall, gaps decline in about 10–20 per cent of the full sample of 100 countries. Canada's estimated annual gap increase of 0.01 SD of achievement per year is larger than about 80 per cent of the other countries in the sample.

Unlike in Canada, the SES achievement gap in most other high-income countries has increased more between the middle and the bottom of the SES distribution (50/10) than between the top and middle of the SES distribution (90/50). In middle- and low-income countries, the SES achievement gap has increased approximately equally at the top and bottom of the SES distribution.

Finally, the cross-national analyses in Chmielewski (forthcoming) examine which changing country characteristics are associated with growing SES achievement gaps. The strongest predictor of growing SES achievement gaps is expanding school enrolment. For example, among the countries with the largest increases in gaps, Ireland's secondary school enrolment has increased by over 20 percentage points, and Thailand has increased secondary enrolment by nearly 70 percentage points. Growing income inequality is associated with increasing gaps in middle- and low-income countries, a pattern that is driven by increasing income inequality in some post-Soviet countries and declining income inequality in some Latin American countries. Unexpectedly, however, growing income inequality is associated with *smaller* increases in gaps in high-income countries, driven by countries such as the United States, England, Finland, Israel, and Japan, with increasing income inequality and stable SES achievement gaps. The latter result suggests that, in wealthy post-industrial economies with high levels of educational attainment and white-collar employment, many important gradations of inequality are not captured by educational degree and occupational categories, for example status hierarchies of educational institutions or fields of study and occupational sector. It may be that household income would better capture these gradations. Indeed, Reardon (2011) finds that achievement gaps based on income have increased in the United States, while achievement gaps based on parental education have remained stable. Unfortunately, household income is not available in a large enough number of international assessments to examine trends for Canada or other countries.

Discussion

SES achievement gaps increased substantially in Canada between the 1981 and 2005 birth cohorts, growing at a rate that outpaces 80 per cent of other countries with available data. This trend is driven primarily by a dramatic increase in achievement gaps based on the number of books in the household. Achievement gaps based on parental education and occupation have been relatively stable over this period. Canada's growing SES achievement gap is occurring primarily between the top and the middle of the SES distribution, whereas in most other high-income countries in the sample, achievement gaps are growing more between the middle and bottom of the SES distribution.

Cross-national multivariate analyses in Chmielewski (forthcoming) do not definitively explain why Canada has experienced a larger-than-average increase in its SES achievement gap. Internationally, the strongest predictor of increasing SES achievement gaps is expanding school enrolment, suggesting that, in many countries, greater access to school may not cause inequality to increase but rather *reveal* inequality that was previously hidden outside the school system. Yet in the earliest available years of Canadian international assessments (the mid-1990s), the country already had almost universal school enrolment of the target populations (Grade 4, Grade 8, and 15-year-old students). Thus, expanding school enrolment cannot directly explain the increasing SES achievement gap in Canada. A moderate (by international standards) increase in enrolment in upper grades in Canada due to declining high-school dropout rates may nonetheless point to an indirect mechanism by which increasing enrolments may drive SES achievement gaps in Canada and elsewhere. A larger share of students graduating from high school and potentially entering colleges and universities may increase competition within the education system and thus widen SES achievement gaps.

Some other trends that were hypothesized to drive increasing achievement gaps are evident in Canada: increasing income inequality, private school enrolment, and investments of parental time and money in children. However, across all high-income countries in the sample, those with the largest increases in income inequality tend to experience *smaller* increases in SES achievement gaps. Increasing private-school enrolment is associated with increasing SES achievement gaps, but the association is small and not statistically significant. No good internationally comparable measures are available for increasing investments of parental time or money in children in a large enough sample of countries to enable multivariate analyses. The weak results for many predictors in multivariate models may indicate that the *causes* of increases are heterogeneous across countries even though increasing SES achievement gaps are a global trend.

The results presented in this paper nevertheless provide some clues to the causes of increasing SES achievement gaps in Canada. That achievement gaps based on household books have increased more than those based on parental education or occupation suggests that deliberate parenting behaviours may play a role. With widespread access to digital devices in recent times, the decision to own physical books may increasingly capture not only economic but also cultural capital. Moreover, the finding that 90/50 SES achievement gaps have increased more than 50/10 SES achievement gaps suggests an intensifying concentration of advantage for high-SES students. This trend may be facilitated in Canada by growing income inequality, private school enrolment, and parental expenditures on extra resources such as private tutoring, even if these factors are not associated with growing SES achievement gaps in cross-national analyses.

The results of this paper and of Chmielewski (forthcoming) suggest that cognitive skills are an increasingly important dimension of educational inequality, both in Canada and worldwide. Growing SES achievement gaps raise serious concerns about equality of opportunity. Across many countries, SES achievement gaps impede upward mobility, as they account for a large proportion of SES disparities in university attendance (Jackson, 2013; Jerrim & Vignoles, 2015), particularly attendance at selective universities (Jerrim, Chmielewski, & Parker, 2015). Via unequal access to selective higher education and lucrative jobs, SES achievement gaps in childhood may even drive future income inequality in adulthood. Indeed, a recent international study shows that a society's variability in childhood test scores predicts its level of income inequality decades later (Checchi & van de Werfhorst, 2017). The large Canadian and international data sets compiled for this paper and for Chmielewski (forthcoming) will be an important source of evidence for future educational and social policies that may help to mitigate disparities in learning opportunities for high- and low-SES children.

References

Ariga, K., Brunello, G., Iwahashi, R., & Rocco, L. (2005). *Why is the timing of school tracking so heterogeneous?* (IZA Discussion Paper No. 1854). Bonn, Germany: Institute for the Study of Labor.

Aurini, J., Davies, S., & Dierkes, J. (Eds.). (2013). *Out of the shadows: The global intensification of supplementary education (International perspectives on education and society, Vol. 22)*. Bingley, UK: Emerald Group Publishing.

Baker, D. P., Goesling, B., & LeTendre, G. K. (2002). Socioeconomic status, school quality, and national economic development: A cross-national analysis of the "Heyneman-Loxley Effect" on mathematics and science achievement. *Comparative Education Review*, *46*(3), 291–312.

Benavot, A. (1983). The rise and decline of vocational education. *Sociology of Education*, *56*(2), 63–76.

Bosetti, L., Van Pelt, D., & Allison, D. (2017). The changing landscape of school choice in Canada: From pluralism to parental preference? *Education Policy Analysis Archives*, *25*, 38.

Brown, R. S., & Sinay, E. (2008). *2006 student census: Linking demographic data with student achievement*. Toronto, ON: Toronto District School Board.

Byun, S.-y., & Kim, K.-k. (2010). Educational inequality in South Korea: The widening socioeconomic gap in student achievement. *Research in Sociology of Education, 17*, 155–182.

Checchi, D., & van de Werfhorst, H. G. (2017). Policies, skills and earnings: How educational inequality affects earnings inequality. *Socio-Economic Review*, *16*(1), 137–160.

Chmielewski, A. K. (forthcoming). The global increase in the socioeconomic achievement gap, 1964–2015. *American Sociological Review.*

Davies, S., & Aurini, J. (2011). Exploring school choice in Canada: Who chooses what and why? *Canadian Public Policy*, *37*(4), 459–477.

Davies, S., Cyr, D., Rizk, J., & Janus, M. (2016). Exploring patterns of educational achievement gaps: Where do they come from? In W. Lehmann (Ed.), *The sociology of education in Canada: Contemporary debates and perspectives*. Don Mills, ON: Oxford University Press.

Dotti Sani, G. M., & Treas, J. (2016). Educational gradients in parents' child-care time across countries, 1965–2012. *Journal of Marriage and Family*, *78*(4), 1083–1096.

Eyles, A., & Machin, S. J. (2015). *The introduction of academy schools to England's education* (IZA Discussion Paper No. 9276). Bonn, Germany: Institute for the Study of Labor.

Gauthier, A. H., Smeeding, T. M., & Furstenberg, F. F. (2004). Are parents investing less time in children? Trends in selected industrialized countries. *Population and Development Review*, *30*(4), 647–672.

Hulchanski, J. D. (2010). *The three cities within Toronto: Income polarization among Toronto's neighbourhoods, 1975–2005*. Toronto: Cities Centre, University of Toronto.

Jackson, M. (2013). *Determined to succeed? Performance versus choice in educational attainment.* Stanford, CA: Stanford University Press.

Jerrim, J., Chmielewski, A. K., & Parker, P. (2015). Socioeconomic inequality in access to high-status colleges: A cross-country comparison. *Research in Social Stratification and Mobility*, *42*, 20–32.

Jerrim, J., & Vignoles, A. (2015). University access for disadvantaged children: A comparison across countries. *Higher Education*, *70*(6), 903–921.

Kornrich, S., Gauthier, A. H., & Furstenberg, F. F. (2011, April). *Changes in private investments in children across three liberal welfare states: Australia, Canada, and the United States.* Paper presented at the annual meeting of the Population Association of America, Washington, DC.

Manning, A., & Pischke, J.-S. (2006). *Comprehensive versus selective schooling in England in Wales: What do we know?* (IZA Discussion Paper No. 2072). Bonn, Germany: Institute for the Study of Labor.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS 2015 international results in mathematics*. Boston, MA: TIMSS & PIRLS International Study Center, Boston College.

Musterd, S., Marcińczak, S., Van Ham, M., & Tammaru, T. (2017). Socioeconomic segregation in European capital cities. Increasing separation between poor and rich. *Urban Geography*, *38*(7), 1062–1083.

O'Grady, K., & Houme, K. (2015). *PCAP 2013 contextual report on student achievement in science*. Toronto, ON: Council of Ministers of Education, Canada.

Organisation for Economic Co-operation and Development. (2015). *In it together: Why less inequality benefits all*. Paris: OECD Publishing. https://doi.org/10.1787/9789264235120-en

Organisation for Economic Co-operation and Development. (2016). *PISA 2015 results: Vol. 1. Excellence and equity in education*. Paris: PISA and OECD Publishing. doi:10.1787/9789264266490-en.

Parekh, G., & Gaztambide-Fernández, R. (2017). The more things change: Durable inequalities and new forms of segregation in Canadian public schools. In W. T. Pink and G. W. Noblit (Eds.), *Second international handbook of urban education (Vol. 1)* (pp. 809–831). Cham, Switzerland: Springer International Publishing.

Park, H., Buchmann, C., Choi, J., & Merry, J. J. (2016). Learning beyond the school walls: Trends and implications. *Annual Review of Sociology*, *42*, 231–252.

Reardon, S. F. (2011). The widening academic-achievement gap between the rich and the poor: New evidence and possible explanations. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 91–115). New York: Russell Sage Foundation.

Reardon, S. F., & Bischoff, K. (2011). Income inequality and income segregation. *American Journal of Sociology*, *116*(4), 1092–1153.

Saw, G. K. (2016). Patterns and trends in achievement gaps in Malaysian secondary schools (1999–2011): Gender, ethnicity, and socioeconomic status. *Educational Research for Policy and Practice*, *15*(1), 41–54.

United Nations Development Programme. (2013). *Humanity divided: Confronting inequality in developing countries*. New York: UNDP.

Walks, A. (2013). *Income inequality and polarization in Canada's cities: An examination and new form of measurement* (Research Paper 227). Toronto: Cities Centre, University of Toronto.

Willms, J. D. (1999). *Inequalities in literacy skills among youth in Canada and the United States* (Cat. No. 89F0116X). Ottawa, ON: Statistics Canada.

Literacy, Numeracy, and Problem-Solving Skills among Youth Not in Education, Employment, or Training in Canada

Ashley Pullman
Postdoctoral fellow, University of Ottawa

Under the direction of Ross Finnie
Professor, University of Ottawa

Abstract

The purpose of this study is to explore sociodemographic background factors and skill indicators among youth not in education, employment, or training (NEET) in Canada. Through linear and logistic regression analyzing microdata from the 2012 Programme for the International Assessment of Adult Competencies (PIAAC) and the 2014 Longitudinal and International Study of Adults (LISA), this study examines how skill-based activities at home, readiness to learn, and assessed cognitive skills—namely, literacy, numeracy, and problem solving in technology-rich environments—differ between NEET and non-NEET youth. Compared to differences by age group, immigration profile, Indigenous ancestry, disability status, and parental education, the results demonstrate that in 2012 NEET youth had one of the largest cognitive skill disparities compared to non-NEET youth. Sociodemographic characteristics and lower rates of skill-based activities at home do, however, partially reduce the assessment score disparity between NEET and non-NEET youth. Although this study cannot fully disentangle the causal relationship between assessment scores and skill-based activities at home, there is strong indication that home-based activities matter, especially in relation to being NEET two years later. The longitudinal follow-up illustrates that lower rates of reading and of activities based on information and communication technology at home were associated with a lower likelihood of being in school and/or employed in 2014. Collectively, the results suggest that the promotion of home-based informal learning is important for NEET youth and may lower the probability of re-experiencing NEET status over time.

Key terms: Canada, PIAAC, youth, employment, education, cognitive skill, informal learning

Introduction

Understanding which background factors contribute to a higher likelihood of youth not being in education, employment, or training (NEET) is essential to prevention strategies. Specific individuals may be more likely than their peers to become NEET due to demographic differences such as gender, age, and parental education (Carcillo, Fernandez, Königs, & Minea, 2015; Zuccotti & O'Reilly, 2018). Certain groups may also face challenges in school and at work due to social and economic disadvantage and discrimination (Canadian Labour Congress, 2016; Holloway et al., 2018). Related but less often considered factors associated with NEET status are cognitive skill level, engagement in skill-based activities, and orientation toward learning. On one hand, an individual's skill profile connects his or her background and the opportunities for skill development it affords (Carcillo et al., 2015). On the other hand, skill level changes over time in relation to everyday activities (Reder, 1994, 2016).

NEET youth may have fewer future work and educational opportunities than youth who have never been NEET (Thompson, 2017). Prior research demonstrates that youth may experience difficulties in securing employment or re-entering education after periods of being NEET (Gregg & Tominey, 2005; Ralston, Feng, Everington, & Dibben, 2016). That is, NEET youth are more likely to remain in or re-experience unemployment and/or school non-attendance over time. If cognitive skill level and skill-based activities are associated with the likelihood of *becoming* NEET, these same indicators may also influence the recurrence of this status. Examining how variables representing skill, skill-based activities, and orientation toward learning relate to becoming and remaining NEET over time provides insights that are vital to the development of intervention strategies that aim to promote employment and school attendance among youth.

This study explores sociodemographic indicators of NEET status in Canada in 2012 and examines how assessed cognitive skill level in the domains of literacy, numeracy, and problem solving in technology-rich environments differs among NEET and non-NEET youth aged 16 to 24. To consider why cognitive skill levels may differ, we examine if rates of self-reported skill-based activities at home and orientation to learning vary between NEET youth and their non-NEET counterparts. Because the causal relationship between assessment scores and skill-based activities at home cannot be fully disentangled when measured at the same time, we go beyond the 2012 data to consider if skill-based activities at home and orientation to learning relate to the likelihood of being NEET later, in 2014, among a sub-sample of respondents. Together, three primary research questions guide this study:

- Who are NEET youth in Canada? Specifically, do NEET rates differ relative to gender, age, parental education, immigration history, Indigenous ancestry, and disability status?
- Do the literacy, numeracy, and problem-solving assessment scores of NEET and non-NEET youth differ, and if so, does accounting for skill-based activities at home and self-reported learning strategies reduce this disparity?
- How does the association of skill-based activities at home, readiness to learn in 2012, and the likelihood of being NEET in 2014 differ between youth initially identified as NEET in 2012 and non-NEET youth?

The following section provides an overview of past research. It details the factors that contribute to the likelihood of NEET status and why it is important to examine different skill and learning indicators. The next section introduces the reader to the data, variables, and analytical approach. We use both linear and logistic regression to analyze data from the 2012 Programme for the International Assessment of Adult Competencies (PIAAC) and a smaller sub-sample from the 2014 Longitudinal and International Study of Adults (LISA). The penultimate section then reports the findings, and a final section highlights the main implications of the study and avenues for future research.

Literature review

Characteristics of NEET youth

Data availability, survey differences, and research objectives often produce notable inconsistency in defining NEET status (Furlong, 2006). A conservative definition labels individuals who have not been employed or in school for at least one year as NEET (Organisation for Economic Co-operation and Development [OECD], 2013b). At the opposite end of the definition spectrum, some studies classify individuals working in informal or precarious employment as NEET (Sustainable Development Solutions Network, 2015). As operationalized in this study, the most common definition classifies NEET individuals as those neither employed nor in school within a selected reference period, typically a survey week or a period under study. NEET youth may or may not be actively seeking work and/or training (Maguire, 2015). Rather than focus on individual reasons for being NEET, this study examines overarching trends related to sociodemographic background factors and skill.

In Canada, the NEET rates differ for various sociodemographic youth groups. Young men typically have higher NEET rates, often attributed in part to greater levels of school enrolment among young women (Statistics Canada, 2018). Across Canada, 7.2 per cent of males aged 15 to 19 were NEET in 2016 (versus 5.4 per cent of females in the same age group), as were 16.4 per cent of males aged 20 to 24 (versus 13.3 per cent of females) (OECD, 2018). NEET rates increase among older youth, especially in the period following compulsory education (Statistics Canada, 2018). Age is particularly important, as older youth may be ineligible for social services aimed at children and teenagers and living away from their primary guardians for the first time.

International research demonstrates that immigrant youth are often more likely to experience periods of being NEET (Eurofound, 2016; Tamesberger & Bacher, 2014). However, the disparity between native- and foreign-born youth is smaller in Canada than in many other countries. In 2014, 11.2 per cent of Canadian-born youth were NEET compared to 13.4 per cent of foreign-born youth in the country (OECD, 2016, Figure 1.16). Youth who have recently immigrated to Canada experience higher rates of unemployment. For example, in 2011, employment rates of recent immigrant youth aged 15 to 29 were approximately 12 percentage points lower than those of youth born in

Canada (Employment and Social Development Canada [ESDC], 2017, p. 15). A known factor associated with this employment gap is proficiency in one of Canada's official languages (Grondin, 2007).

Indigenous youth in Canada have increasingly higher levels of educational attainment and employment, but education and employment inequalities persist for many Indigenous groups (Arriagada, 2016; ESDC, 2017; Moyser, 2017). For example, in 2015, Indigenous youth school enrolment was 9.1 percentage points lower and labour market participation was 14.6 percentage points lower than that of non-Indigenous youth (Moyser, 2017, Section 2.13). The 2015 Truth and Reconciliation Commission report contains several calls to action that explicitly highlight the need to support participation in schooling and employment among Indigenous youth.

Parental education—often used as a proxy indicator of family socioeconomic status—is also associated with the likelihood of becoming NEET in youth. Parental education is a known factor in educational attainment (Krahn, 2017; McMullen, 2011), which in turn influences later employment (Boudarbat & Chernoff, 2010). Greater financial resources among highly educated parents increase their children's opportunities for learning and skill development (Davies & Aurini, 2013). These and other factors culminate in lower NEET rates among youth with highly educated parents in Canada and other countries (Carcillo et al., 2015).

Education and employment outcomes are also known to differ by disability status. In 2012, youth with disabilities were more likely to be NEET compared to youth without disabilities (Till, Leonard, Yeung, & Nicholls, 2015). The likelihood of being NEET increases with age for youth with disabilities, a trend that may reflect difficulties in making the transition to employment and/or postsecondary education after leaving compulsory education (Ebersold, 2012; Statistics Canada, 2018). In 2012, for example, 24.4 per cent of 20-year-olds with disabilities were NEET compared to 10.6 per cent of 20-year-olds without disabilities (Till et al., 2015, Chart 1). Alongside job availability, 31.9 per cent of unemployed adults aged 15 to 64 with disabilities describe inadequate training and experience as a barrier to being in the labour force (Till et al., 2015, Chart 5).

Several factors that contribute to the likelihood of becoming NEET in youth are bidirectional, such as health challenges (Benjet, Hernández-Montoya, Borges, Méndez, Medina-Mora, & Aguilar-Gaxiola, 2012; Gutiérrez-García, Benjet, Borges, Méndez Ríos, & Medina-Mora, 2017; Henderson, Hawke, & Chaim, 2017, April) or motherhood (Marshall, 2012). For this reason, it is difficult to determine the causal relationship between simultaneously occurring factors and their influence on becoming NEET. There is also a risk of simultaneity when using cross-sectional data to examine the connection between one-time variables representing skill and the likelihood of being NEET. Relatedly, lower educational achievement is associated with a higher chance of being NEET (Marshall, 2012), but the causal relationship is far from straightforward. Do NEET youth have lower educational attainment because they are not in school? Or are individuals not in school more likely to be NEET? As discussed next, this is an important consideration when analyzing skill differences among NEET and non-NEET youth.

Skill-based activities among NEET youth

Across Canada, unemployed youth have lower numeracy and literacy skills, even when accounting for prior education (Lundetræ, Gabrielsen, & Mykletun, 2010). Lower skills and competencies in a range of domains contribute to education and labour-market disengagement (Calero & Choi, 2017; Coates, 2016; Falch, Nyhus, & Strom, 2014; Rivera-Batiz, 1992; Turner & Conkel, 2010). For example, English and Welsh individuals who self-report fewer soft skills (e.g., communication, decision making, and teamwork) and professional skills (e.g., in computer use, sales, and construction) are more likely to be NEET at age 18, even after accounting for childhood literacy and other background factors (Goldman-Mellor et al., 2016). Likewise, self-reported cognitive and resource-based skills predict the likelihood of NEET status among Ontario youth transitioning out of extended care programs (Flynn & Tessier, 2011).

Lower skills in key areas contribute to education and employment disengagement, a reciprocal relationship that prevents opportunities to develop skills (OECD, 2013a). Given their status, NEET youth are excluded from the skill-based activities that take place within education and employment. Nonetheless, NEET youth may still engage in skill-based activities at home, an informal avenue through which learning takes place (Livingstone, 1999). As a precursor to further skill acquisition (Wentzel & Brophy, 2014), NEET youth may also have a positive orientation toward learning, an attitude described below as "readiness to learn" (Smith, Rose, Ross-Gordon, & Smith, 2015). Thus, while we expect in our study to find that NEET youth have lower skill assessment scores, skill-based activities at home and readiness to learn may help to reduce this discrepancy.

Practice engagement theory provides insight into why home-based activities influence skill acquisition (Reder, 1994, 2016). This perspective emphasizes the contribution of everyday activities to promoting skill; in short, it stresses a "use it or lose it hypothesis" (Desjardins, 2003, p. 215). Skill-based activities do not take place only within formal education settings (OECD & Statistics Canada, 2000), although these settings are associated with such activities (Scandurra & Calero, 2017). Practice engagement theory has important implications for understanding skill level among NEET youth. Rather than describing NEET youth as "disengaged" or "excluded" from learning, we may characterize them as engaged in informal skill-based activities at home. Nevertheless, NEET youth may have lower rates of skill-based activities at home overall because formal education promotes these activities (Trautwein, 2007).

Data, variables, and analytical approach

Data sources

Canada participated in the first round of the PIAAC study, surveying more than 27,000 residents between the ages of 16 and 65, irrespective of legal status or nationality between November 2011 and June 2012. The response rate was approximately 58 per cent, a comparable level to other major surveys conducted during this time but lower than the Canadian components of the 1994 International Adult Literacy Survey (69 per cent) and the 2003 Adult Literacy and Lifeskills Survey (66 per cent) (OECD, 2013b, p. 229). In Canada, the formation of the target population was based on the 2011 short-form Census of Population and National Household Survey. Several groups were excluded from the Canadian survey, although together they fell well under the OECD's expected non-coverage rate for the target population, set at a maximum of 5 per cent.[1]

The Canadian PIAAC study also includes a longitudinal component, the LISA survey, which began concurrently in 2012 and includes a sub-sample of PIAAC participants. Like PIAAC, the target population for LISA was generated using the 2011 short-form Census of Population and National Household Survey.[2] Unlike PIAAC, the sample selection included all members of a household, and children were invited to participate as permanent sample members at age 15. The first LISA wave, in 2012, had a 72 per cent household-level and 89 per cent individual-level response rate. The total number of respondents was 23,926 and, of these, 8,598 answered the complete PIAAC background questionnaire and undertook the assessment tests. The overall sample size changed in wave two, in 2014, not only due to attrition but also to the inclusion of new respondents who had turned 15 between survey periods. In 2014, LISA had 19,178 respondents, among whom 6,162 were PIAAC respondents.

Although the majority of countries participating in PIAAC surveyed approximately 5,000 people, the Canadian 2012 PIAAC component surveyed a larger number of individuals in order to capture regional variation and also oversampled several groups, including Indigenous people and recent immigrants. The larger sample size allows for an analytical approach that focuses on specific age groups: in this case, youth aged 16 to 24. Unfortunately, the LISA sub-sample of youth participants who also participated in PIAAC is much smaller, a notable limitation of this study.

We estimate all PIAAC-based models using the final sampling and replicate weights via the jackknife repeated replication method for Canada (i.e., JK1) and the 10 plausible values necessary for correct estimates. Although the LISA-based models utilize the same 10 plausible values, the weighting strategy differs. LISA-based models use the wave two

[1] The PIAAC sampling frame excluded individuals living on First Nations reserves, in small and remote communities, in non-institutional collective dwellings, on military bases, or in institutions (OECD, 2013b). It is important to highlight these sampling restrictions, especially as this study includes measures of NEET status among Indigenous respondents, statistics that only represent those living off reserve.

[2] Like PIAAC, the LISA target population excludes individuals living on First Nations reserves, in religious and communal colonies, and in institutions (e.g., nursing homes, jails, and hospitals). Additionally, LISA excludes individuals who work for the Canadian Armed Forces and are stationed abroad, as well as foreign representatives and delegates living in Canada. With these exclusions, the LISA survey represents approximately 98 per cent of the Canadian population.

responding person weight to correct for non-response and attrition and the corresponding 1,000 bootstrap weights to estimate sampling variance. The PIAAC and LISA microdata are accessible only within securitized Canadian Research Data Centre Network (CRDCN) centres.

Variables employed in analysis

NEET status: Our study uses a standard definition of NEET status that also takes respondents' own subjective interpretations into account. We operationalize NEET status through the official OECD definition: those who are not employed (i.e., employed for less than one hour in the survey reference week), and not in education or training in the survey reference week. Given that youth may be formally enrolled in school but not attending, we also took subjective status into account when forming our 2012 indicator of NEET status. We did so through a question that asked respondents to look at a list of statements describing possible life status and choose the one that best described their current situation. We used this variable to "correct" for individuals who were marginally employed or in school but self-reported their status to be unemployed or not in the labour force.

Demographic variables: The analytical approach includes seven key demographic variables. Dummy and categorical variables measure:

- age (16–19 or 20–24)
- gender (male or female)
- immigration status (non-immigrant, second-generation immigrant, first-generation immigrant)
- Indigenous ancestry (status or non-status First Nations, Métis, or Inuit)
- disability status (chronic illness or disability lasting six months or more)
- parental education (either mother/father or female/male guardian has postsecondary degree, certificate, or diploma)
- number of books in the home at age 16 (25 or fewer, 26 to 100, or 101 or more).

The Appendix provides a complete description of these variables, the original PIAAC microdata variables used to construct them, and PIAAC and LISA sample sizes. Among sociodemographic variables, the overall level of missingness at the variable level is under 1 per cent—except for parental education, which was under 3 per cent—and we thus implement listwise deletion. Finally, in the model development stage, we also examined other variables measuring educational level, household composition, and language ability. We removed these from our analytical approach due to either endogeneity (e.g., education level or the presence of a child in a household), non-significance (e.g., number of people in a household), or multicollinearity (e.g., language ability).

Skill-based variables Given the multiple ways to consider skill, this study relies on three types of variables:

- assessment scores in literacy, numeracy, and problem solving in technology-rich environments (PS-TRE)
- a series of self-reported variables that examine reading, numeracy, and information and communication technology (ICT) skill-based activities at home
- a self-reported variable that represents "readiness to learn."

Among respondents aged 16 to 24 in Canada, missing data at the variable level is less than 1 per cent for the readiness to learn and reading at home scales and literacy and numeracy assessment tests. Missingness is 5 per cent to 7 per cent for the numeracy and ICT activities at home scales and the PS-TRE assessment module.

The PIAAC literacy, numeracy, and PS-TRE assessments scores range from 0 to 500, a standardized continuum representing ability level, and measure distinct but overlapping areas.[3] The OECD defines the PIAAC literacy assessment variable as capturing "the ability to understand, evaluate, use and engage with written texts" (OECD, 2013a, p. 59) and includes comprehending, decoding, interpreting, and evaluating digital or print-based written text. The corresponding

[3] Although there is a strong correlation among the three assessments, ranging from 0.740 to 0.868 in Canada (OECD, 2013b, Table 18.5), they measure distinct domains. In Canada, the PIAAC literacy test reliability was 0.878, the numeracy test reliability was 0.874, and the PS-TRE test reliability was 0.847 (OECD, 2013b, Table 18.4).

numeracy variable assesses "the ability to access, use, interpret and communicate mathematical information and ideas" (OECD, 2013a, p. 59). Similar to literacy, this domain includes identifying, interpreting, evaluating, and communicating numeracy-based information but differs in use, such as counting, estimating, or measuring. Finally, the PS-TRE assessment captures the skills necessary "to use digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks" (OECD, 2013a, p. 59). Like literacy and numeracy, this domain involved acquiring, evaluating, and using information. Differing from these two domains, however, this assessment area foregrounds goal setting and process monitoring, both of which take place through computers, technology, and networks.

The skill-based activities at home scales are based on PIAAC background questionnaire items that asked all respondents to self-report how often they engage in specific activities. Responses were constrained to five options: (1) never; (2) less than once a month; (3) less than once a week but at least once a month; (4) at least once a week but not every day; and (5) every day.[4]

The reading activities at home scale includes eight items, such as reading books, newspapers, manuals, directions, reference material, and professional publications. The numeracy activities at home scale contains six activities, such as calculating a budget, using a calculator, preparing a graph or table, using simple algebra or formulas, or using advanced math or statistics. Finally, the ICT activities at home scale has seven items, such as writing an email, looking up new information, conducting a transaction, using word processing or spreadsheet software, programing, or having an online discussion.

Readiness to learn refers to the "propensity to learn new things, relate these knowledge and skills to existing knowledge and life situations, and engage in problem solving and information seeking behaviors" (Smith et al., 2015, p. 3). The scale construction involved six questions from the PIAAC background questionnaire that asked respondents to self-report their learning strategies. Five response options constrained answers to each statement: (1) not at all; (2) very little; (3) to some extent; (4) to a high extent; or (5) to a very high extent. A generalized partial credit approach—an item response theory model for polytomous rating responses—formed the derived readiness to learn and skill-based activities at home scales. Median scores range from 2.3 (readiness to learn) to 2.7 (numeracy at home), and the range between the upper and lower quantile is similar across variables. For more information on the home-based activity and readiness to learn scales, see Chapter 20 of the PIAAC technical report (OECD, 2013b).

Analytical approach

Research question one considers who Canadian NEET youth are: specifically, if rates differ according to age, gender, immigration status, Indigenous ancestry, disability status, parental education, and the number of books at home at age 16. Given that NEET status is operationalized dichotomously (i.e., PIAAC participants who are or are not NEET in 2012), we use logistic regression to answer this question. Unlike descriptive statistics, logistic regression concurrently controls for multiple factors and generates the estimated probability of being NEET if all other variables were held constant (Hosmer & Lemeshow, 2013). In more technical terms, the underlying logistic regression equation for research question one is:

$$Pr(NEET\ status\ in\ 2012) = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)}{1 + (\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)}$$

In this equation, *Pr* is the estimated probability of being NEET by the distinct sociodemographic variables X_1 to X_k modelled. The intercept β_0 estimates the probability of being NEET when all explanatory variables equal 0.

To facilitate the interpretation of these results, we report the average marginal effects (AME) of the logistic regression model (Mood, 2010). This approach reports the average expected difference in the probability of being NEET in 2012 with a one-unit increase in a sociodemographic variable. Given that these variables are categorical or binary, group

[4] Self-reported skill-based activities at home have a moderate correlation with their corresponding assessment area, ranging from 0.34 to 0.27 in Canada (OECD, 2013b, Table 18.6). Thus, individuals who engage in these activities are expected to have higher skill assessment scores. Although there is notable parallel between assessment areas and skill-based activities at home, there is no direct correspondence between questions asked in the background questionnaire and assessment tests (OECD, 2013a, p. 182).

differences are intuitively represented as "the partial effect averaged across the specified population" (Wooldridge, 2016, p. 844). That is, it compares hypothetical groups in that all other variables in the model are held constant.

Research question two examines if literacy, numeracy, and PS-TRE assessment scores differ between NEET and non-NEET youth and if including variables measuring skill-based activities at home and readiness to learn reduces the overall difference in scores. Given that a continuous measurement scale ranging from 0 to 500 represents assessment scores, ordinary least squares regression offers an interpretable approach to examining skill among NEET youth.[5] The underlying equation for research question two is:

$$y_i\,(assessment\ score\ in\ 2012) = \beta_0 + \beta_1 X_1\ + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon$$

In this equation, the dependent variable y_1 represents an assessment score in either literacy, numeracy, or PS-TRE, each modelled separately. The model relates each assessment score to a set of explanatory variables—X_1 to X_k—such as NEET status and other sociodemographic variables. The intercept β_0 estimates the assessment score when all explanatory variables equal 0. Each significant demographic coefficient is interpreted as the extent to which the average score differs for that group from its baseline, holding all other variables constant.

To assess the association of skill-based activities at home and readiness to learn, we run three separate models: the first linear regression model includes only NEET status; the second model adds all sociodemographic variables; and the third model controls for readiness to learn and the skill-based activity that corresponds to the assessment area. This approach allows for an examination of how the difference between NEET and non-NEET cognitive assessment scores changes once we account for sociodemographic characteristics, learning strategies, and skill-based activities at home.

To answer research question three, we use logistic regression again to assess how skill-based activities at home and readiness to learn in 2012 are associated with being NEET in 2014 among youth who were and were not initially identified as NEET in 2012. Again, NEET status is operationalized dichotomously (i.e., LISA-PIAAC participants who are or are not NEET in 2014). Given the temporal focus of this research question, we add another background variable to the model that measures NEET status in 2012. Thus, the underlying logistic regression equation for research question three is:

$$Pr(NEET\ status\ in\ 2014) = \frac{\exp(\beta_0 + \beta_1 X_1\ + \beta_2 X_2 + \cdots + \beta_k X_k)}{1 + (\beta_0 + \beta_1 X_1\ + \beta_2 X_2 + \cdots + \beta_k X_k)}$$

Again, in this equation, *Pr* is the estimated probability of being NEET by all sociodemographic variables, an additional variable representing NEET status in 2012, skill-based activities at home, and cognitive assessment scores.

Three models separately examine each type of skill-based activity at home, controlling for its corresponding skill domain. A fourth model examines the readiness to learn index without controlling for any skill assessment area. Given that the response scale for these variables is continuous, we use line graphs with confidence intervals to report the AME results. They report the average expected difference in the probability of being NEET in 2014 with a one-unit increase in each index score. They compare hypothetical groups representing individuals who were and were not NEET in 2012 with all other variables in the model mean centred.

[5] We use the least squares method—perhaps the most widely used linear regression approach—to estimate all regression coefficients. These estimates cannot be interpreted as representing the causal effect of a given explanatory variable on the assessment score without further consideration. Notably, if any explanatory variables are correlated with the error term, the causal effects may be biased.

Findings

Research question one: Who are NEET youth in Canada?

The intent of the first stage of the analytical approach is to examine how the likelihood of NEET status in 2012 differs by sociodemographic group in Canada. We present both the rate of NEET youth in each group and a logistic regression model that includes all sociodemographic variables. Because sociodemographic groups are not mutually exclusive (e.g., respondents with more books at home at age 16 may also be those who have highly educated parents), the logistic regression model provides insight into which groups are more likely to be NEET when simultaneously controlling for all other sociodemographic variables.

Table 1 illustrates that the total rate of NEET youth in Canada was 7.5 per cent within the 2012 PIAAC survey. There was a higher percentage of older NEET youth—8.7 per cent compared to 6 per cent of those aged 16 to 19—a difference that is statistically significant in the corresponding logistic regression model. There was no statistically significant difference between overall rates of NEET status for male and female respondents or for those who self-reported a disability or illness. Approximately 10 per cent of first-generation immigrants and 8.6 per cent of second-generation immigrants were NEET in 2012, compared to 6.2 per cent of non-immigrant PIAAC respondents. However, the percentage point differences among these immigrant groups is non-significant when controlling for all other sociodemographic variables in the logistic regression model.

TABLE 1 **Descriptive statistics and logistic regression results of NEET youth in Canada (*N* = 4,561)**

	Descriptive statistics		Logistic regression (pseudo R^2 = 0.071)	
	% NEET	Proportion (%) of NEET	Average marginal effect	Standard error
Total NEET	7.5	–		
Age				
16–19	6.0	36.7	–	–
20–24	8.7	63.3	0.039**	0.013
Gender				
Male	7.6	52.1	–	–
Female	7.3	47.9	0.006	0.016
Immigration status				
Non-immigrant	6.2	50.4	–	–
Second-generation immigrant	8.6	25.3	0.046	0.022
First-generation immigrant	10.1	24.3	0.031	0.021
Indigenous background				
Non-Indigenous	7.1	91.4	–	–
Indigenous	16.1	8.6	0.088**	0.031
Disability status				
No disability/illness	7.7	85.9	–	–
Has disability/illness	6.4	14.1	–0.005	0.015
Parental education				
No postsecondary education	12.7	43.3	–	–
Has postsecondary education	5.7	56.7	–0.051**	0.018
Number of books at home at age 16				
0–25	11.8	38.0	–	-
26–100	6.7	29.0	–0.069***	0.019
101+	5.6	33.0	–0.081***	0.019
Province/territory				
Newfoundland and Labrador	11.5	2.1	–	–
Prince Edward Island	11.1	0.7	0.004	0.048
Nova Scotia	11.5	4.1	0.037	0.042
New Brunswick	11.4	3.1	0.021	0.044
Quebec	6.4	18.7	–0.069*	0.029
Ontario	8.3	44.3	–0.057	0.031
Manitoba	9.0	4.4	–0.023	0.041
Saskatchewan	7.8	3.2	–0.065	0.036
Alberta	1.8	2.8	–0.116***	0.033
British Columbia	8.9	15.8	–0.034	0.037
Yukon	10.2	0.1	–0.040	0.055
Northwest Territories	18.2	0.4	–0.017	0.038
Nunavut	28.4	0.5	–0.036	0.038

Note: Descriptive statistics and AME are calculated from Canadian PIAAC microdata for respondents aged 16 to 24, using final and replicate weights. AME is interpreted as the average expected difference in the probability of being NEET compared to the corresponding reference group, controlling for all other sociodemographic variables.

* p <.05

** p <.01

*** p <.001

There are notable differences in NEET rates by parental education, the number of books at home at age 16, and Indigenous ancestry. Youth with highly educated parents were approximately 7 percentage points less likely to be NEET, a difference that reduces slightly but remains statistically significant in the logistic regression model. The importance of social capital is also apparent when examining the estimated number of books at home at age 16. Among respondents who reported owning 25 books or fewer, 11.8 per cent were NEET, compared to 6.7 per cent among those who reported 26 to 100 books and 5.6 per cent of those who reported 101 books or more. Finally, Indigenous youth in Canada were NEET at a much greater rate (16.1 per cent) than non-Indigenous youth (7.1 per cent). Importantly, the logistic regression model illustrates that other sociodemographic characteristics account for approximately half of this percentage point difference.

Research question two: Do the literacy, numeracy, and problem-solving assessment scores of NEET and non-NEET youth differ? If so, does accounting for skill-based activities at home and self-reported learning strategies reduce this disparity?

To answer research question two, we first present average cognitive assessment scores in literacy, numeracy, and PS-TRE among NEET and non-NEET youth, and show how their average index scores in readiness to learn and skill-based activities at home differ. Next, we use linear regression to examine not only how assessment scores differ between NEET and non-NEET youth (Model 0) but also how the difference between the two groups changes when controlling for sociodemographic background factors (Model 1) and skill-based activities at home and readiness to learn (Model 2).

In preparation for the regression analysis that follows, Table 2 illustrates that average skill assessment scores in literacy, numeracy, and PS-TRE were lower among NEET youth. NEET youth had a mean literacy score that was 28 points lower than that of their non-NEET counterparts, a mean numeracy score that was 35 points lower, and a mean PS-TRE score that was 20 points lower. Index scores measuring skill-based activities at home and readiness to learn ranged in discrepancy between NEET and non-NEET youth. There was a small difference in average readiness to learn scores—NEET youth self-reported mean rates that were 9 per cent lower. Average scores in reading-based activities at home for NEET youth were 19 per cent lower than non-NEET youth, average scores in numeracy-based activities at home were 26 per cent lower, and average scores in ICT-based activities at home were 18 per cent lower.

Table 2 **Skill, readiness to learn, and skill-based activities at home assessment scores among NEET and non-NEET youth in Canada (*N* = 4,542)**

	Mean			T-test
	NEET	Non-NEET	All	
Skill assessment scores				
Literacy	250	278	276	***
Numeracy	237	272	268	***
PS-TRE	276	296	294	***
Readiness to learn	2.15	2.36	2.34	*
Home-based activities				
Reading	2.06	2.54	2.50	***
Numeracy	2.08	2.81	2.74	***
ICT	2.00	2.43	2.39	***

Note: Means and t-tests are calculated from Canadian PIAAC microdata for respondents aged 16 to 24, using final and replicate weights.
* $p < .05$
** $p < .01$
*** $p < .001$

Table 3 provides the results of three linear regression models that examine literacy assessment scores. In Model 0, which includes no other variables, NEET youth scored, on average, 26.5 points lower than non-NEET youth—a small change from the descriptive statistics above due to missing data for the covariates included in the full model. When accounting for all other sociodemographic variables in Model 1, the difference between NEET and non-NEET literacy assessment scores decreases by 22 per cent to 20.7 points. When controlling for readiness to learn and reading activities at home in Model 2, the difference in average literacy assessment scores remains statistically significant but diminishes by a further 21 per cent to 16.4 points. Together the results suggest that the difference in the literacy assessment scores of NEET and non-NEET youth was due in part to both sociodemographic factors and differences in skill-based activities at home and readiness to learn. Nevertheless, even when including all covariates, the lower average literacy assessment scores among NEET youth remain statistically significant.

TABLE 3 Regression results for literacy assessment scores in 2012 (*N* = 4,542)

	Model 0 (adjusted R^2 = 0.03)		Model 1 (adjusted R^2 = 0.14)		Model 2 (adjusted R^2 = 0.20)	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
NEET (non-NEET)	–26.5***	4.19	–20.7***	4.20	–16.4***	4.09
Age 20–24 (16–19)			13.1***	2.28	11.0***	2.33
Female (male)			1.70	2.13	1.25	2.15
Second-generation immigrant (non-immigrant)			6.27*	3.20	4.44	3.19
First-generation immigrant (non-immigrant)			–14.9***	3.45	–17.9***	3.58
Indigenous (non-Indigenous)			–12.9***	3.32	–12.5***	3.37
Has disability or illness (none)			–2.30	3.91	2.55	3.79
Mother/father has postsecondary education (no postsecondary)			15.2***	2.57	11.1***	2.54
Books at age 16: 26–100 (< 25)			12.5***	3.10	10.0**	3.27
Books at age 16: 101+ (< 25)			20.0***	2.97	14.6***	3.24
Index of reading at home					7.98***	1.72
Index of readiness to learn					6.90***	1.43
Province			Included		Included	
Intercept	278.7***	1.28	249.5***	3.96	220.7***	4.79

Note: All results are calculated from Canadian PIAAC microdata for respondents aged 16 to 24. Reference is group in parentheses.
* $p < .05$
** $p < .01$
*** $p < .001$

Similar linear regression models presented in Table 4 examine the NEET and non-NEET difference in average numeracy assessment scores. In Model 0, NEET youth scored, on average, 33.5 points lower than non-NEET youth in the numeracy assessment module. The addition of sociodemographic covariates in Model 1 reduces the difference in assessment scores by 25 per cent to 25 points. The addition of variables capturing readiness to learn and numeracy-based activities at home in Model 2 reduces the difference by a further 23 per cent to 19 points. Like literacy skills, these models demonstrate that the numeracy assessment score difference between NEET youth and their non-NEET peers was among the most pronounced of all the sociodemographic groups modelled.

TABLE 4 Regression results for numeracy assessment scores in 2012 (*N* = 4,276)

	Model 0 (adjusted R^2 = 0.04)		Model 1 (adjusted R^2 = 0.15)		Model 2 (adjusted R^2 = 0.20)	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
NEET (non-NEET)	–33.5***	4.95	–25.0***	5.22	–19.3***	5.17
Age 20–24 (16–19)			14.2***	2.80	17.4***	2.74
Female (male)			–8.53***	2.64	–7.76**	2.64
Second-generation immigrant (non-immigrant)			3.25	3.68	0.30	3.81
First-generation immigrant (non-immigrant)			–11.9***	4.01	–15.3***	4.08
Indigenous (non-Indigenous)			–18.3***	4.08	–16.3***	4.05
Has disability or illness (none)			–1.13	3.82	–0.85	3.69
Mother/father has postsecondary education (no postsecondary)			18.8***	3.11	15.5***	3.05
Books at age 16: 26–100 (< 25)			15.8***	3.93	13.8***	3.97
Books at age 16: 101+ (< 25)			23.9***	3.68	19.8***	3.80
Index of reading at home					7.35***	1.16
Index of readiness to learn					6.12***	1.61
Province			Included		Included	
Intercept	271.9***	1.55	244.5***	4.29	209.6***	5.37

Note: All results are calculated from Canadian PIAAC microdata for respondents aged 16 to 24. Reference group is in parentheses.
* $p < .05$
** $p < .01$
*** $p < .001$

The PS-TRE assessment difference between NEET and non-NEET youth was smaller than the difference in literacy and numeracy scores. Table 5 illustrates that in Model 0, NEET youth scored, on average, 19.2 points lower than non-NEET youth on the PS-TRE assessment. In Model 1, the PS-TRE skill assessment difference between NEET and non-NEET youth diminishes by 29 per cent, to 13.7 points, when accounting for sociodemographic characteristics. In Model 2, the addition of ICT activities at home and readiness to learn indices further reduces the PS-TRE assessment skill difference by 29 per cent to a small gap of less than 10 points.

TABLE 5 Regression results for PS-TRE assessment scores in 2012 (N = 3,955)

	Model 0 (adjusted R^2 = 0.02)		Model 1 (adjusted R^2 = 0.10)		Model 2 (adjusted R^2 = 0.16)	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
NEET (non-NEET)	–19.2	4.67***	–13.7**	4.34	–9.73*	4.29
Age 20–24 (16–19)			4.51	2.79	2.36	2.71
Female (male)			0.69	2.40	0.17	2.33
Second-generation immigrant (non-immigrant)			4.58	3.15	2.75	3.12
First-generation immigrant (non-immigrant)			–7.98*	3.61	–12.6***	3.69
Indigenous (non-Indigenous)			–8.82*	4.05	–7.39	4.17
Has disability or illness (none)			–1.22	3.43	–1.84	3.63
Mother/father has postsecondary education (no postsecondary)			10.9***	2.41	7.58***	2.35
Books at age 16: 26–100 (< 25)			14.2***	2.81	11.7***	2.86
Books at age 16: 101+ (< 25)			20.6***	3.11	16.3***	3.23
Index of reading at home					8.85***	1.48
Index of readiness to learn					5.03***	1.29
Province			Included		Included	
Intercept	295.7***	1.51	272.2***	4.08	246.4***	5.67

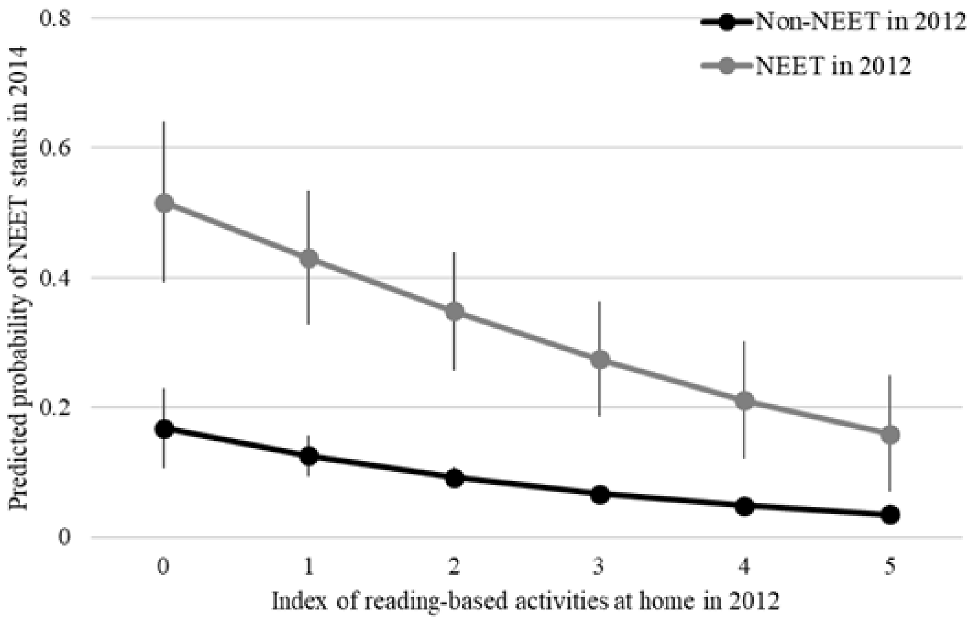
Note: All results are calculated from Canadian PIAAC microdata for respondents aged 16 to 24. Reference group is in parentheses.
* $p < .05$
** $p < .01$
*** $p < .001$

Research question three: How does the association of skill-based activities at home, readiness to learn in 2012, and the likelihood of being NEET in 2014 differ between youth initially identified as NEET in 2012 and non-NEET youth?

As discussed above, it is difficult to unravel the relationship between skill-based activities at home and NEET status with reference to a single time point, given that they occur simultaneously. Importantly, research question two is unable to determine if NEET youth engage less with skill-based activities at home because they are NEET or if having fewer skill-based activities at home is associated with an increased likelihood of being NEET. To further understand the connection between skill-based activities at home and NEET status, Figures 1 to 4 reveal how the predicted probability of NEET status in 2014 (i.e., AME generated from an underlying logistic regression model) differs by skill-based activity and readiness to learn index scores for a LISA sub-set of youth identified as NEET or non-NEET in 2012.

The results in Figure 1 reveal the importance of home-based reading activities, especially among youth who were NEET in 2012. It shows that youth who either were or were not NEET in 2012 had a lower probability of being NEET in 2014 if they had higher scores on the reading-based activities at home scale in 2012. Among youth who were NEET in 2012, those who engaged in fewer reading-based activities at home were more likely to be NEET in 2014. Notably, the probability of NEET status in 2014 begins to converge between the two groups at higher scores, suggesting that increasing reading-based activities at home may decrease the likelihood of NEET status over time.

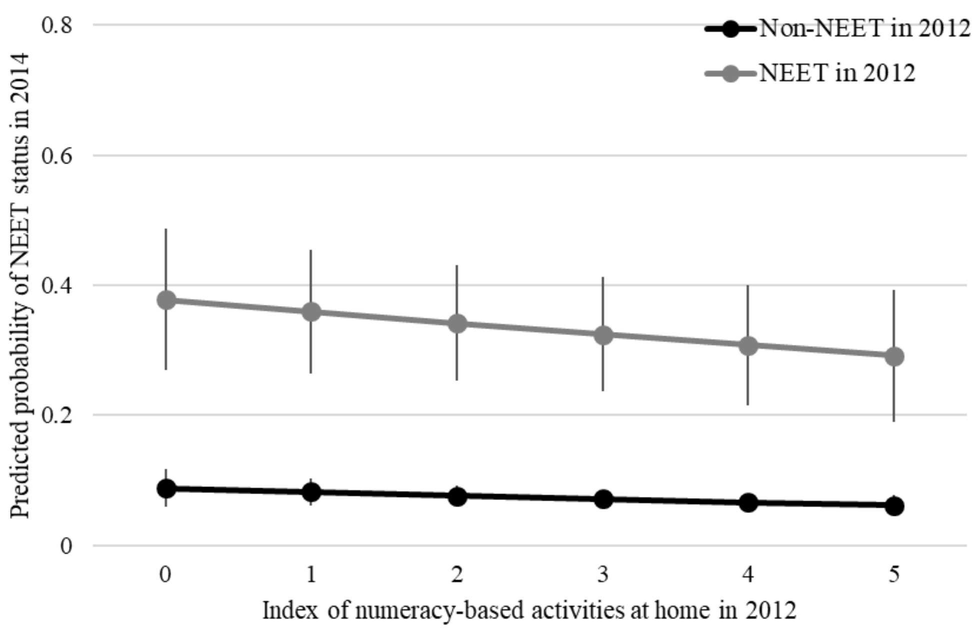
FIGURE 1 Probability of 2014 NEET status by reading-based activities at home in 2012



Note: All results are calculated from Canadian LISA microdata for respondents aged 16 to 24 in 2012. Conditional marginal effects are generated from a logistic regression that controls for age, gender, immigration status, disability status, parental education, number of books at home at age 16, and literacy assessment score, as measured in 2012. To interpret the average predicted probability of NEET status among all LISA participants, all covariates are mean centred.

The analysis of numeracy-based activities at home shows a less pronounced trend than reading activities do. As illustrated in Figure 2, the predicted probability of being NEET in 2014 decreases slightly by numeracy-based activities at home index scores among NEET and non-NEET youth in 2012. Nevertheless, overlap in the confidence intervals highlights greater uncertainty compared to reading at home index scores. That is, although the lines signal slight differences when comparing the probability of being NEET in 2014 at the low and high ends of the index, the large confidence intervals surrounding these point estimates suggest a lack of confidence in this relationship. We cannot establish if NEET youth who engaged in more numeracy-based activities at home were or were not less likely to be NEET in 2014.

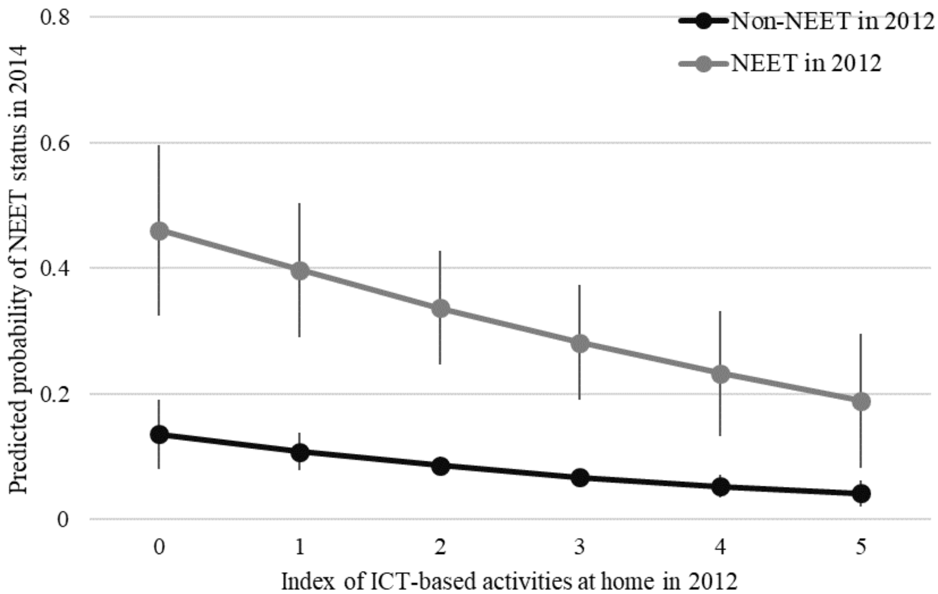
FIGURE 2 Probability of 2014 NEET status by numeracy-based activities at home in 2012



Note: All results are calculated from Canadian LISA microdata for respondents aged 16 to 24 in 2012. Conditional marginal effects are generated from a logistic regression that controls for age, gender, immigration status, disability status, parental education, number of books at home at age 16, and numeracy assessment score, as measured in 2012. To interpret the average predicted probability of NEET status among all LISA participants, all covariates are mean centred.

The analysis of ICT-based activities at home shows a similar trend to reading activities. Figure 3 demonstrates that the predicted probability of being NEET in 2014 decreases by ICT-based activities at home index scores among youth who either were or were not NEET in 2012. Although there is some overlap in the confidence intervals, differences can be seen between the low and high ends of the index, especially for youth who were NEET in 2012. This suggests that NEET youth who engaged in more ICT-based activities at home in 2012 were also less likely to be NEET in 2014.

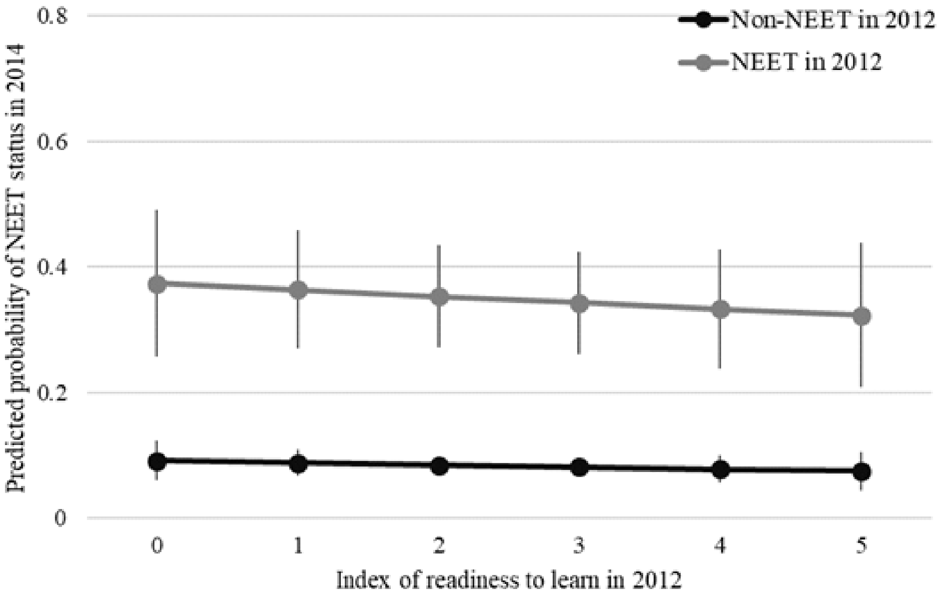
FIGURE 3 Probability of 2014 NEET status by ICT-based activities at home in 2012



Note: All results are calculated from Canadian LISA microdata for respondents aged 16 to 24 in 2012. Conditional marginal effects are generated from a logistic regression that controls for age, gender, immigration status, disability status, parental education, number of books at home at age 16, and PS-TRE assessment score, as measured in 2012. To interpret the average predicted probability of NEET status among all LISA participants, all covariates are mean centred.

Compared to skill-based activities at home, higher scores on the readiness to learn scale in 2012 are not associated with the probability of being NEET in 2014. As shown in Figure 4, both NEET and non-NEET youth in 2012 with high and low readiness to learn scores had a similar probability of being NEET in 2014. The absence of a connection between 2012 readiness to learn and 2014 NEET status may be due to the small distinction between the scores of NEET and non-NEET youth in 2012. That is, NEET youth had average readiness to learn index scores that were only marginally smaller than non-NEET youth in 2012. As the discussion highlights next, what NEET youth report doing at home is more strongly associated with the later probability of being NEET than is their self-reported readiness to learn.

FIGURE 4 Probability of 2014 NEET status by readiness to learn index score in 2012



Note: All results are calculated from Canadian LISA microdata for respondents aged 16 to 24 in 2012. Conditional marginal effects are generated from a logistic regression that controls for age, gender, immigration status, disability status, parental education, and number of books at home at age 16, as measured in 2012. To interpret the average predicted probability of NEET status among all LISA participants, all covariates are mean centred.

Summary, implications, and future research

This section revisits our main research findings and highlights their implications for policy and future research. Using PIAAC, this study examined the sociodemographic background characteristics of NEET youth in Canada and how literacy, numeracy, and problem-solving assessment scores differ between NEET and non-NEET youth. A longitudinal follow-up using LISA generated insight into whether skill-based activities at home and readiness to learn in 2012 were associated with the likelihood of being NEET in 2014 among youth who either were or were not NEET in 2012.

Supporting prior research highlighted in the literature review, the logistic regression results illustrate that older (aged 20 to 24) and Indigenous youth in Canada were more likely than other sociodemographic groups to be NEET in 2012. Highlighting the importance of cultural capital and socioeconomic status (Xu & Hampden-Thompson, 2011), youth whose parents held a high-school diploma or had less education and who had fewer than 100 books at home at age 16 were also more likely to be NEET in 2012. The descriptive statistics also show that first-generation immigrants were more likely to be NEET, although controlling for other sociodemographic variables influences the difference between recent and non-immigrants. Unlike in previous research, no difference in rates of NEET status by gender or disability status was observed among PIAAC youth participants. Nevertheless, research that uses less coarsened variables of disability status or an account of gender differences by age may reveal distinctions among these sub-groups.

Both the descriptive and the linear regression results illustrate that PIAAC literacy and numeracy assessment scores were approximately half a level lower (27 to 36 points) among NEET youth compared to their non-NEET peers. Our results suggest that there was a smaller difference in PS-TRE scores (19 points). Among all countries that participated in PIAAC, credential level and skill-related activities at home and work are the strongest predictors of assessment scores (Statistics Canada, 2013; OECD, 2013a). Indeed, there is strong indication that home-based activities matter. First, NEET youth were more likely to have fewer books at home at age 16. Second, they self-reported fewer numeracy-, literacy-, and ICT-based activities at home while being NEET. Third, greater frequency in reading- and ICT-based learning activities at home decreased the likelihood of being NEET two years later in 2014.

It is important to consider possible reasons for the smaller difference in PS-TRE assessment scores and larger difference in numeracy assessment scores between NEET and non-NEET youth. Although there is inequality in access to technology in Canada (Haight, Quan-Haase, & Corbett, 2014), PIAAC respondents aged 16 to 24—a group born between 1988 and 1996—probably interacted with technology from an early age (Gilmore, 2010). Indeed, youth had higher PS-TRE scores on average than older cohorts. In contrast, the numeracy assessment score difference between NEET and non-NEET youth was more pronounced, a finding that converges with prior research examining inequality across PIAAC cognitive assessment areas. The differences in numeracy assessment scores for NEET and non-NEET youth tend to be most marked between native- and non-native-language speaker (OECD, 2013a), even when accounting for education level (Ford, 2018).

PS-TRE skills and readiness to learn index scores showed the smallest differences between NEET and non-NEET youth. Although PIAAC assessed NEET youth as having lower average cognitive assessment scores, this group reported only slightly less interest in learning. As measured by the readiness to learn variable, NEET youth reported a willingness to learn new things, figure out how ideas fit together, and relate new information to their life at a rate similar to non-NEET youth. Although readiness to learn is important to consider, we highlight that it has no association with the likelihood of being NEET two years later in 2014. The theoretical and empirical ramifications support practice engagement theory (Reder, 1994, 2016) and indicate that orientation to learning is a distinct attribute and functions differently from skill-based activities at home.

To better understand the relationship between NEET youth and determinants of skill, we suggest several avenues for further research. First, it is important to conduct further research to generate insight into both individual and structural accounts of skill differences. As Thompson (2011) writes,

> the experiences of NEET young people cannot be accounted for purely in terms of the dispositions of the individuals and the choices they make. Subjective factors are important, but they are embedded in and arise from objective conditions, including local structures of opportunity and more general interactions between labour markets and the social distribution of educational achievement (p. 798).

Our study began this examination by demonstrating that skill level is associated with social background factors such as parental education and number of books at home at age 16. Further research must consider how experiences at school and work are associated with both the likelihood of becoming NEET and skill differences among youth.

As future waves of the LISA study are made available, longitudinal approaches may examine both individual (e.g., school grades) and structural (e.g., school type) variables and address the limitations of our study. As highlighted throughout, it is important to consider the study of education and skill among NEET youth as susceptible to simultaneous causality, especially as data collected at a single time point cannot be used to disentangle temporal ordering (Cox, 1992). It remains unknown whether lower assessment scores among NEET youth are due to not being in school or work—two possible locations for education and training—or whether the skill level of NEET youth is a barrier to education and work.

In conclusion, we would like to emphasize that all individuals and groups should have meaningful opportunities for skill acquisition. Skill level is associated with a range of social outcomes; for example, adults with lower literacy skills are more likely to have higher rates of unemployment and lower income (OECD, 2013a). These individuals are also more likely to self-report lower civic engagement, less participation in volunteer activities, and poorer overall health (OECD, 2013a). Although addressing skill differences and how they arise is important in reducing social inequality, it is necessary to provide social support to improve these and other outcomes for all individuals, regardless of their skill level.

Appendix: Sociodemographic variables

Variable	Description	Coding	Original PIAAC name		PIAAC (*N*)	LISA (*N*)
			Name	Description		
Age	Dummy variable: comparing individuals aged of 16–19 and 20–24	0 = 16–19 1 = 20–24	ageg5lfs	Derived indicator of age groups in five-year intervals	4,689	833
Gender	Dummy variable: comparing males and females	0 = male 1 = female	gender_r	Gender of respondent	4,689	833
Immigration status	Categorical variable: immigration background of respondents	1 = non-immigrant 2 = second-generation (mother/father born outside Canada) 3 = first-generation (born outside Canada)	j_q07a_t j_q06a_t j_q04a	Father or male guardian born in country Mother or female guardian born in country Born in country	4,601	828
Indigenous ancestry	Dummy variable: respondent is First Nations (status or non-status Indians), Métis or Inuit	0 = no 1 = yes	jq04fca1	Indigenous person: First Nations (including status and non-status Indians), Métis, or Inuit	4,644	833
Disability status	Dummy variable: respondent has longstanding illness or disability	0 = no 1 = yes	i_q10aca	Longstanding illness or disability lasting/lasted 6 months or more	4,640	832
Parental education	Dummy variable: either mother or father (or female or male guardian) has postsecondary degree, certificate, or diploma	0 = no 1 = yes	j_q07bca j_q06bca	Father/male guardian's highest level of education Mother/female guardian's highest level of education	4,637	809
Books at home at age 16	Categorical variable: self-reported number of books at home at age 16	1 = 25 or fewer 2 = 26 to 100 3 = 101 or more	j_q08	Background: books at home	4,617	829

References

Arriagada, P. (2016). *First Nations, Métis, and Inuit women* (Cat. No. 89-503-X). Ottawa: Statistics Canada. Retrieved from http://www.statcan.gc.ca/pub/89-503-x/2015001/article/14313-eng.htm

Benjet, C., Hernández-Montoya, D., Borges, G., Méndez, E., Medina-Mora, M. E., & Aguilar-Gaxiola, S. (2012). Youth who neither study nor work: Mental health, education and employment. *Salud Pública de México*, *54*, 410–417.

Boudarbat, B., & Chernoff, V. (2010). *The determinants of education–job match among Canadian university graduates* (IZA Discussion Paper No. 4513). Bonn, Germany: Institute for the Study of Labor. Retrieved from http://nbn-resolving.de/urn:nbn:de:101:1-20091105842

Calero, J., & Choi, Á. (2017). The distribution of skills among the European adult population and unemployment: A comparative approach. *European Journal of Education*, *52*(3), 348–364.

Canadian Labour Congress. (2016). *Diverse, engaged, and precariously employed: An in-depth look at young workers in Canada*. Ottawa: Canadian Labour Congress. Retrieved from http://documents.clc-ctc.ca/sep/YoungWorkersinCanada-ResearchPaper2016-8-04-EN.pdf

Carcillo, S., Fernandez, R., Königs, S., & Minea, A. (2015). *NEET youth in the aftermath of the crisis: Challenges and policies* (OECD Social, Employment and Migration Working Papers No. 164). Paris: OECD Publishing.

Coates, H. L. (2016). Disadvantaged youth in southern Scotland experience greater barriers to information access resulting from poor technology skills, information literacy, and social structures and norms. *Evidence Based Library and Information Practice*, *11*(4), 75–78.

Cox, D. R. (1992). Causality: Some statistical aspects. *Journal of the Royal Statistical Society: Series A*, *155*(2), 291–301.

Davies, S., & Aurini, J. (2013). Summer learning inequality in Ontario. *Canadian Public Policy*, *39*(2), 287–307.

Desjardins, R. (2003). Determinants of literacy proficiency: A lifelong-lifewide learning perspective. *International Journal of Educational Research*, *39*(3), 205–245.

Ebersold, S. (2012). T*ransitions to tertiary education and work for youth with disabilities*. Paris: Education and Training Policy, OECD Publishing.

Employment and Social Development Canada. (2017). *13 ways to modernize youth employment in Canada – Strategies for a new world of work*. Ottawa: ESDC. Retrieved from https://www.canada.ca/en/employment-social-development/corporate/youth-expert-panel/report-modern-strategies-youth-employment.html

Eurofound. (2016). *Exploring the diversity of NEETs*. Luxembourg: Publications Office of the European Union. Retrieved from https://www.eurofound.europa.eu/exploring-the-diversity-of-neets-1

Falch, T., Nyhus, O. H., & Strom, B. (2014). Performance of young adults: The importance of different skills. *CESifo Economic Studies*, *60*(2), 435–462.

Flynn, R. J., & Tessier, N. G. (2011). Promotive and risk factors as concurrent predictors of educational outcomes in supported transitional living: Extended care and maintenance in Ontario, Canada. *Children and Youth Services Review*, *33*(12), 2498–2503.

Ford, K. (2018). Persisting gaps: Labor market outcomes and numeracy skill levels of first-generation and multi-generation college graduates in the United States. *Research in Social Stratification and Mobility*, *56*, 21–27

Furlong, A. (2006). Not a very NEET solution: Representing problematic labour market transitions among early school-leavers. *Work, Employment and Society*, *20*(3), 553–569.

Gilmore, J. (2010). *Trends in dropout rates and the labour market outcomes of young dropouts* (Cat. No. 81-004-X). Ottawa: Statistics Canada. Retrieved from http://www.statcan.gc.ca/pub/81-004-x/2010004/article/11339-eng.htm

Goldman-Mellor, S., Caspi, A., Arseneault, L., Ajala, N., Ambler, A., Danese, A., et al. (2016). Committed to work but vulnerable: Self-perceptions and mental health in NEET 18-year-olds from a contemporary British cohort. *Journal of Child Psychology and Psychiatry*, *57*(2), 196–203.

Gregg, P., & Tominey, E. (2005). The wage scar from youth unemployment. *Labour Economics*, *12*(4), 487–509.

Grondin, C. (2007). *Knowledge of official languages among new immigrants: How important is it in the labour market?* (Cat. No. 89-624-XIE). Ottawa: Statistics Canada. Retrieved from http://www.statcan.gc.ca/pub/89-624-x/2007000/4123709-eng.htm

Gutiérrez-García, R. A., Benjet, C., Borges, G., Méndez Ríos, E., & Medina-Mora, M. E. (2017). NEET adolescents grown up: Eight-year longitudinal follow-up of education, employment and mental health from adolescence to early adulthood in Mexico City. *European Child & Adolescent Psychiatry*, *26*(12), 1459–1469.

Haight, M., Quan-Haase, A., & Corbett, B. A. (2014). Revisiting the digital divide in Canada: The impact of demographic factors on access to the internet, level of online activity, and social networking site usage. *Information, Communication & Society*, *17*(4), 503–519.

Henderson, J. L., Hawke, L. D., & Chaim, G. (2017, April). Not in employment, education or training: Mental health, substance use, and disengagement in a multi-sectoral sample of service-seeking Canadian youth. *Children and Youth Services Review*, *75*, 138–145.

Holloway, E. M., Rickwood, D., Rehm, I. C., Meyer, D., Griffiths, S., & Telford, N. (2018). Non-participation in education, employment, and training among young people accessing youth mental health services: Demographic and clinical correlates. *Advances in Mental Health*, *16*(1), 19–32.

Hosmer, D. W., & Lemeshow, S. (2013). *Applied logistic regression*. Hoboken, NJ: Wiley.

Krahn, H. (2017). Choose your parents carefully: Social class, post-secondary education, and occupational outcomes. In E. Grabb, J. G. Reitz, & M. Hwamg (Eds.), *Social inequality in Canada* (pp. 90–103). Don Mills, ON: Oxford University Press.

Livingstone, D. W. (1999). Exploring the icebergs of adult learning: Findings of the first Canadian survey of informal learning practices. *Canadian Journal for the Study of Adult Education*, *13*(2), 49–72.

Lundetræ, K., Gabrielsen, E., & Mykletun, R. (2010). Do basic skills predict youth unemployment (16- to 24-year-olds) also when controlled for accomplished upper-secondary school? A cross-country comparison. *Journal of Education and Work*, *23*(3), 233–254.

Maguire, S. (2015). NEET, unemployed, inactive or unknown – Why does it matter? *Educational Research*, *57*(2), 121–132.

Marshall, K. (2012). Youth neither enrolled nor employed. *Perspectives on Labour and Income*, *24*(2), 1–6.

McMullen, K. (2011). Post-secondary education participation among underrepresented and minority groups (Cat. No. 81-004-X). *Education Matters: Insights on Education, Learning and Training in Canada*, *8*(4), 1–10.

Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review*, *26*(1), 67–82.

Moyser, M. (2017). *Aboriginal people living off-reserve and the labour market: Estimates from the Labour Force Survey, 2007–2015* (Cat. No. 71-588-X). Statistics Canada: Ottawa. Organisation for Economic Co-operation and Development. (2013a). *OECD skills outlook 2013: First results from the Survey of Adult Skills*. Paris: OECD Publishing.

Organisation for Economic Co-operation and Development. (2013b). *Technical report of the Survey of Adult Skills (PIAAC)*. Paris: OECD Publishing.

Organisation for Economic Co-operation and Development. (2016). *Society at a glance 2016: OECD social indicators.* Paris: OECD Publishing.

Organisation for Economic Co-operation and Development. (2018). *Youth not in employment, education or training (NEET)* (Indicator). Paris: OECD Publishing. doi:10.1787/72d1033a-en.

Organisation for Economic Co-operation and Development & Statistics Canada. (2000). *Literacy in the information age: Final report on the International Adult Literacy Survey*. Paris and Ottawa: OECD Publishing and Statistics Canada.

Ralston, K., Feng, Z., Everington, D., & Dibben, C. (2016). Do young people not in education, employment or training experience long-term occupational scarring? A longitudinal analysis over 20 years of follow-up. *Contemporary Social Science*, *11*(2–3), 203–221.

Reder, S. (1994). Practice-engagement theory: A sociocultural approach to literacy across languages and cultures. In B. M. Ferdman, R. Weber, & A. G. Ramirez (Eds.) *Literacy across languages and cultures* (pp. 33–74). Albany: State University of New York Press.

Reder, S. (2016). Skill use: Engagement in reading, writing and numeracy practices. In A. Grotlüschen, D. Mallows, S. Reder, & J. Sabatini (Eds.), *Adults with low proficiency in literacy or numeracy* (pp. 37–59). Paris: OECD Publishing.

Rivera-Batiz, F. L. (1992). Quantitative literacy and the likelihood of employment among young adults in the United States. *Journal of Human Resources*, *27*(2), 313–328.

Scandurra, R., & Calero, J. (2017). Modelling adult skills in OECD countries. *British Educational Research Journal*, *43*(4), 781–804.

Smith, M. C., Rose A. D., Ross-Gordon J., & Smith T. J. (2015). *Adults' readiness to learn as a predictor of literacy skills*. Washington, DC: AIR-PIAAC. Retrieved from https://static1.squarespace.com/static/51bb74b8e4b0139570ddf020/t/54da7802e4b08c6b90107b4f/1423603714198/Smith_Rose_Ross-Gordon_Smith_PIAAC.pdf

Statistics Canada. (2013). *Skills in Canada: First results from the Programme for the International Assessment of Adult Competencies (PIAAC)* (Cat. No. 89-555-X). Ottawa: Statistics Canada. Retrieved from http://www.statcan.gc.ca/pub/89-555-x/89-555-x2013001-eng.pdf

Statistics Canada. (2018). *The transition from school to work – The NEET (not in employment, education or training) indicator for 15 to 19 year olds in Canada* (Cat. No. 81-599-X – No. 12). Ottawa: Statistics Canada. Retrieved from http://www.statcan.gc.ca/pub/81-599-x/81-599-x2018012-eng.pdf

Sustainable Development Solutions Network. (2015). *Indicators and a monitoring framework for Sustainable Development Goals: Launching a data evolution for the SDGs* (Revised working draft, 25 July). New York: Leadership Council of the Sustainable Development Solutions Network. Retrieved from http://unsdsn.org/wp-content/uploads/2015/05/FINAL-SDSN-Indicator-Report-WEB.pdf

Tamesberger, D., & Bacher, J. (2014). NEET youth in Austria: A typology including socio-demography, labour market behaviour and permanence. *Journal of Youth Studies*, *17*(9), 1239–1259.

Thompson, R. (2011). Individualisation and social exclusion: The case of young people not in education, employment or training. *Oxford Review of Education*, *37*(6), 785–802.

Thompson, R. (2017). Opportunity structures and educational marginality: The post-16 transitions of young people outside education and employment. *Oxford Review of Education*, *43*(6), 749–766.

Till, M., Leonard, T., Yeung, S., & Nicholls, G. (2015). *A profile of the labour market experiences of adults with disabilities among Canadians aged 15 years and older, 2012* (Cat. No. 89-654-X). Ottawa: Statistics Canada. Retrieved from http://www.statcan.gc.ca/pub/89-654-x/89-654-x2015005-eng.htm

Trautwein, U. (2007). The homework-achievement relation reconsidered: Differentiating homework time, homework frequency, and homework effort. *Learning and Instruction*, *17*(3), 372–388.

Truth and Reconciliation Commission (TRC) (2015). *Honouring the truth, reconciling for the future: Summary of the final report of the truth and reconciliation of Canada*. Winnipeg, MB: TRC.

Turner, S. L., & Conkel, J. L. (2010). Evaluation of a career development skills intervention with adolescents living in an inner city. *Journal of Counseling & Development*, *88*(4), 457–465.

Wentzel, K. R., & Brophy, J. E. (2014). *Motivating students to learn* (4th ed.). New York: Routledge.

Wooldridge, J. M. (2016). *Introductory econometrics: A modern approach* (6th ed.). Boston, ME: Cengage Learning.

Xu, J., & Hampden-Thompson, G. (2011). Cultural reproduction, cultural mobility, cultural resources, or trivial effect? A comparative approach to cultural capital and educational performance. *Comparative Education Review*, *56*(1), 98–124.Zuccotti, C. V., & O'Reilly, J. (2018). Ethnicity, gender and household effects on becoming NEET: An intersectional analysis. *Work, Employment and Society*, Advance online publication. doi:10.1177/0950017017738945.

Grade 4 Math Results: Is Alberta's Performance Affected by Curriculum Coverage?

Janusz Zieminski
Senior Manager, Research, Alberta Education

Abstract

By their very nature, international student assessment studies cannot be perfectly aligned with curricula in all participating countries. The alignment is better in some countries than in others, a difference that can be expected to affect student performance.

This research attempts to measure the impact of Alberta curriculum coverage on student performance, as measured by the Trends in International Mathematics and Science Study (TIMSS) assessment. This is done by using Quebec as a comparator province and a control factor. We set the context for this analysis by showing changes in Alberta's performance over time, and Alberta students' areas of strength and weakness in comparison to Ontario and Quebec. This is followed by a section comparing the alignment between TIMSS assessment items and the Quebec curriculum with the alignment between TIMSS items and the Alberta curriculum. We then analyze the difference between the proportion of students in Quebec and in Alberta providing correct answers. The comparison focuses on the difference in these proportions when items are in both province's curricula, and when items that are in the Quebec curriculum are not covered by the Alberta curriculum.

The paper concludes by suggesting that while curriculum coverage in some areas affects Alberta's scores—and that items not covered in its curriculum appear to be of a lower level of difficulty—other factors that are not related to curriculum coverage are also at play.

For answers to any questions, please contact the Alberta Education Research Branch, 780-415-2446 or edc.research@gov.ab.ca.

Key terms: elementary education, TIMSS, Alberta curriculum, Quebec curriculum, curriculum alignment.

Introduction

The Trends in International Mathematics and Science Study (TIMSS), conducted by the International Association for the Evaluation of Educational Achievement (IEA), monitors trends in mathematics and science achievement in participating countries. The TIMSS study is conducted every four years and assesses students in Grades 4 and 8.

A total of 57 countries took part in TIMSS 2015; in Canada, five provinces participated. Alberta, Ontario, and Quebec participated with large enough sample sizes to obtain reliable results at the provincial level. Manitoba and Newfoundland and Labrador had smaller sample sizes, and their results are therefore not reported at the provincial level. In Alberta, only Grade 4 students participated. The province's average score for Grade 4 science was similar to the Canadian average, while the average score for Grade 4 mathematics was lower than the Canadian and, for the first time, below the international average.

This report attempts to evaluate the extent to which student performance in Alberta is affected by differences between what TIMSS assesses and what the provincial mathematics curriculum covers. In addition, it provides indirect evidence of the impact of teaching and pedagogy in Alberta on student results. When student performance does not seem to be affected by curriculum coverage, it is reasonable to conclude that pedagogy, also a crucial factor in student learning, may be playing a role.

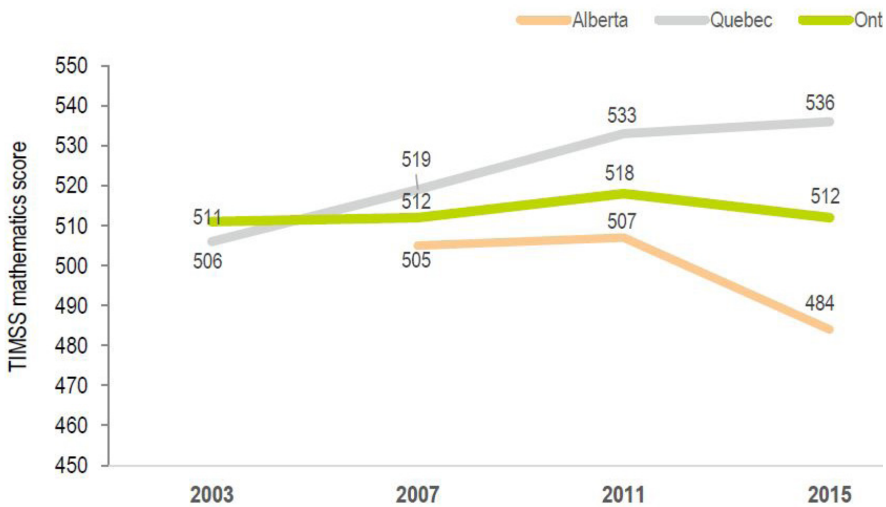
The decline in Alberta students' performance in mathematics observed between the last two cycles of TIMSS has prompted Alberta Education to investigate the factors that could have caused it. Curriculum determines what students are expected to learn and is one of the key tools at the government's disposal to influence student learning outcomes. Consequently, exploring the impact of curriculum coverage on student performance is of keen interest to Alberta Education.

TIMSS mathematics score trends

Trends in total scores

Alberta has participated in Grade 4 mathematics assessment in every TIMSS cycle since 2007, and Quebec and Ontario have participated since 2003. Alberta's score was consistent between 2007 and 2011 but declined in 2015 to below the TIMSS scale centre point/international average of 500. Comparatively, Quebec's mathematics score has improved every year since 2003, while Ontario's scores have remained stable over the same period. Figure 1 compares the Grade 4 TIMSS mathematics scores of Alberta, Quebec, and Ontario.

FIGURE 1 Provincial trends in Grade 4 TIMSS mathematics scores



Trends by content domains

TIMSS 2015 mathematics assessment covered three content domains: *number*, *geometric shapes and measures*, and *data display*. The content domains, the topic areas under each domain, and the proportion of these domains in the total assessment are presented in Table 1.

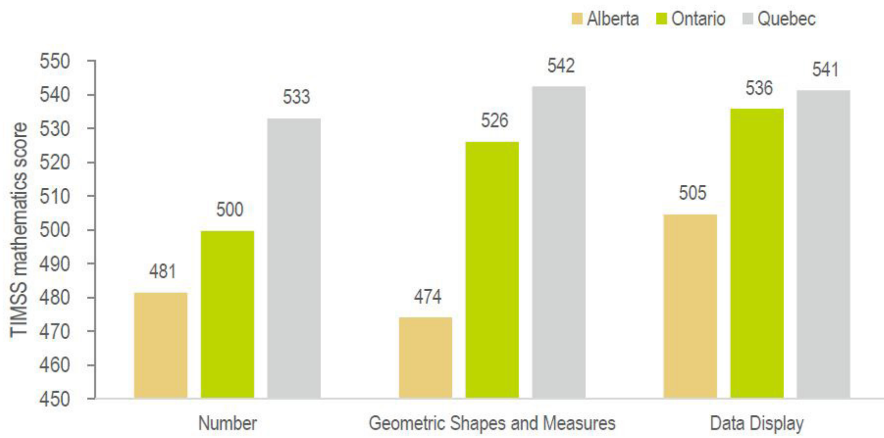
TABLE 1 TIMSS Grade 4 content domains

Content domain	Topic area	Proportion of assessment (%)
Number	Whole numbers Fractions and decimals Expressions, simple equations, and relationships	50
Geometric shapes and measures	Points, lines, and angles Two- and three-dimensional shapes	35
Data display	Reading, interpreting, and representing	15

As Figure 2 shows, *data display* was a domain of relatively strong performance in Alberta (above the international average), but it accounts only for 15 per cent of the score, while scores in *geometric shapes and measures* and *number*, which together account for 85 per cent of the assessment, were similar to each other and well below the international average.

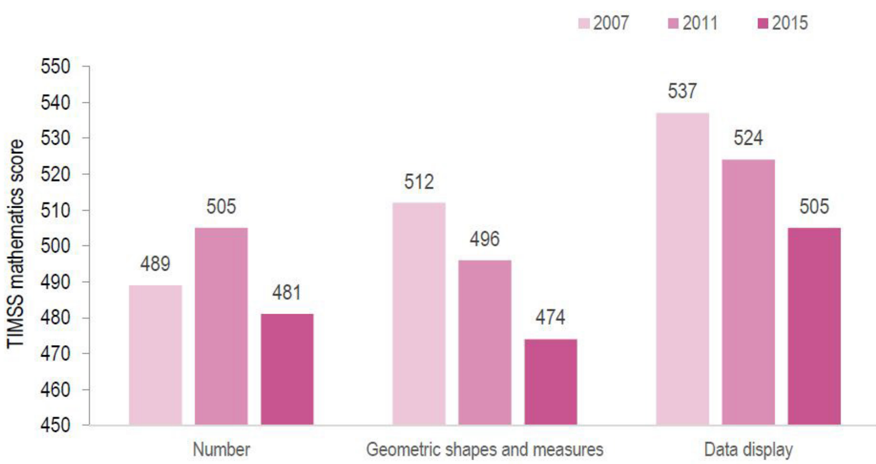
Quebec's score was consistently high in all three content domains. Ontario's scores in *geometric shapes and measures* and *data display* were relatively high and comparable to one another, while the score in the *number* domain was significantly lower. Quebec and Ontario had higher average scores than Alberta in all content domains.

FIGURE 2 Provincial Grade 4 TIMSS 2015 mathematics scores by content domains



Examining Alberta's content scores over time reveals declines in all three content domains between 2007 and 2015 (Figure 3). Moreover, results in *data display* and in *geometric shapes and measures* have been declining steadily since 2007. There was a slight increase in the *number* score between 2007 and 2011, but a decline in 2015. Scores in *geometric shapes and measures* show the steepest decline between 2007 and 2015.

FIGURE 3 Alberta trends in Grade 4 TIMSS mathematics scores by content areas



Analysis of results by content domains

Methodology

All international student assessment studies face the difficult challenge of establishing international consensus on what should be assessed, while having to accept that in each participating country there will be a certain (and different) degree of misalignment between curriculum coverage and assessment coverage. To get a better sense of the scope of this misalignment, TIMSS asked each country to indicate which test items are, and which are not, covered in its curriculum at the grade being tested.

This paper seeks to determine whether curriculum coverage is affecting Alberta's score. This is done by comparing Alberta students' performance (percentage of students answering correctly) on test items that are included in the provincial curriculum with their performance on items that are not covered by the curriculum.

A key issue that needed to be addressed in the analysis is the varying difficulty of test items. Consequently, it is not appropriate to compare results for test items included and not included in curriculum directly, without taking into account the possibility that the difficulty of the items may differ. In order to isolate the effect of test item difficulty, Quebec is used as a control factor. Alberta's results are first compared to Quebec's results for the same test items. The difference between Quebec and Alberta results for items within Alberta's curriculum is then compared to the difference for items that are not included in the Alberta curriculum. A growing gap between Alberta and Quebec that can be explained by the increase in the performance of students in Quebec, rather than by the decline in performance of students in Alberta, is an indicator that the items not included in the Alberta curriculum are of lower difficulty.

The analysis is completed at the overall level, by each content domain, and by each topic area within each content domain. A comparison of the performance gap between the two provinces is made using the difference in the proportion of students who answered each test item correctly (per cent correct in Alberta minus per cent correct in Quebec).

Only items included in Quebec's curriculum are included in this analysis. The analysis relies on mapping test items onto those covered under Alberta curriculum and onto those not covered. The mapping was completed by mathematics consultants in Alberta Education's Curriculum Branch.

Curriculum coverage

Overall, Quebec's curriculum appears to have a higher degree of alignment with content domains and topics presented in the TIMSS assessment than Alberta's curriculum does. Of the 169 items in the TIMSS test, 83 per cent were covered by the Quebec curriculum, compared to 69 per cent in the Alberta curriculum (see Table 2). The most significant gap between Alberta's curriculum and the TIMSS assessment was in the content domain of *geometric shapes and measures.* In particular, in the topic area of *points, lines, and angles*, Quebec's curriculum covered more than twice as many items included in TIMSS as Alberta's curriculum.

Quebec's curriculum was more closely aligned with TIMSS assessment than Alberta's in five of the six topic areas. *Fractions and decimals*, in the content domain of *numbers*, was the only topic area in which Alberta's curriculum was more closely aligned with TIMSS.

Table 2 Comparison of Alberta and Quebec curriculum coverage based TIMSS 2015 test items

Content domain and related topic areas	Number of test items*	Proportion of test items covered in Alberta curriculum (%)	Proportion of test items covered in Quebec curriculum (%)
Number	**89**	**76**	**84**
Expressions, simple equations, and relationships	20	75	100
Fractions and decimals	25	88	64
Whole numbers	44	70	89
Geometric shapes and measures	**56**	**55**	**82**
Points, lines, and angles	24	38	83
Two- and three-dimensional shapes	32	69	81
Data display	**24**	**75**	**79**
Reading, interpreting, and representing	24	75	79
Total	**169**	**69**	**83**

* Only items included in Quebec's curriculum are included in this analysis.
Note: Of 169 total test items, the Alberta curriculum covered 117 and the Quebec curriculum covered 140.

Item analysis

As mentioned above, only test items that were covered by Quebec's curriculum were included in the analysis, in order to control for possible differences in item difficulty. Of these 140 test items, 106 were covered by Alberta's curriculum and 34 were not.

Reflecting Alberta's lower scores compared to Quebec in all three content domains, the proportion of students who correctly answered questions was lower in Alberta than in Quebec in all three content domains—*number*, *geometric shapes and measures*, and *data display*—as shown in Table 3.

TABLE 3 Comparison of Alberta and Quebec scores on TIMSS 2015 test items

Content domain	Number of test items*	Difference in % correct (Alberta minus Quebec)
Number	**75**	**–14**
In Alberta curriculum	61	–14
Not in Alberta curriculum	14	–12
Geometric shapes and measures	**46**	**–16**
In Alberta curriculum	27	–10
Not in Alberta curriculum	19	–24
Data display	**19**	**–9**
In Alberta curriculum	18	–10
Not in Alberta curriculum	1	–8
Total	**140**	**–14**

* Only items included in Quebec's curriculum are included in this analysis.

The following general observations can be made from the results of the analysis:

I. Alberta's performance in the content domain of *number* does not appear to be affected adversely by curriculum coverage. In fact, Alberta students did *relatively* better than their Quebec peers when dealing with items not covered in Alberta's curriculum than they did on items that are covered.

II. Alberta's performance in the content domain of *data display* is also not affected adversely by curriculum coverage. Only one item in *data display* was not covered in the Alberta curriculum, and there was no significant difference between Alberta's performance on test items that were covered and the one that was not covered, when compared to Quebec's performance.

III. Alberta students' performance in *geometric shapes and measures* appears to be affected by the fact that a large proportion of TIMSS items are not in the province's curriculum. In this content domain, the difference between the proportion of students with correct answers in Alberta and the proportion with correct answers in Quebec is more than twice as large for test items not in Alberta's curriculum (–24 per cent) as for test items that are in Alberta's curriculum (–10 per cent).

The following are the results for the various topic areas under the number and geometric shapes and measures content domains.

Number The content domain of *number* assessed three topic areas: *fractions and decimals*; *whole numbers;* and *expressions, simple equations, and relationships* . Although the performance of Alberta students on the overall content domain does not seem to have been affected by curriculum coverage, an analysis of their performance in the three topic areas that together make up the *number* domain uncovered a more nuanced picture (see Table 4).

First, curriculum coverage of *fractions and decimals* was broader in Alberta than in Quebec. All items that were covered in Quebec were also covered by Alberta's curriculum. Consequently, the performance gap between students' performance in the two provinces with respect to this topic area has nothing to do with differences in curriculum coverage. Interestingly, this topic area in the *number* content domain had the largest difference in performance between Alberta and Quebec in any of the content domains.

Second, in the *whole numbers* topic area, Alberta students' performance was relatively weaker than that of their Quebec peers on test items covered in Alberta's curriculum than on items not covered. Data in the Appendix table clearly show that this is fully attributable to relatively weaker performance of students in Quebec on these items rather to an improvement of performance of students in Alberta. While we do not have an explanation for this finding, it is clear that Alberta's performance in this topic area is not adversely affected by curriculum coverage.

Finally, *expressions, simple equations, and relationships* is the only topic area within the *numbers* domain in which curriculum coverage appears to play a role in the weaker performance of students in Alberta . In this topic area, the difference between Alberta and Quebec in the proportion of students correctly responding to an item was almost twice as large for test items not included in Alberta curriculum (–17 per cent) as for the items that were included in Alberta' curriculum (–9 per cent). As data in the Appendix table show, students in both provinces performed better on items not in Alberta curriculum, suggesting that these items are of a lower level of difficulty. Improvement is much more pronounced in Quebec than in Alberta, which explains the increase in the performance gap.

Table 4 Comparison of scores in the number content domain

Topic areas within number domain	Number of test items*	Difference in % correct (Alberta minus Quebec)
Fractions and decimals	**16**	**–19**
In Alberta curriculum	16	–19
Whole numbers	**39**	**–13**
In Alberta curriculum	30	–15
Not in Alberta curriculum	9	–9
Expressions, simple equations, and relationships	**20**	**–11**
In Alberta curriculum	15	–9
Not in Alberta curriculum	5	–17
Total	**75**	**–14**

*Only items included in Quebec's curriculum are included in this analysis.

Geometric shapes and measures The content domain of *geometric shapes and measures* included two topic areas: *points, lines, and angles*; and *two- and three-dimensional shapes*. On both of these topics, the performance of students in Alberta was relatively weaker than students in Quebec on items that were in the Alberta curriculum. This indicates that differences between Alberta and Quebec in terms of curriculum coverage are not the only reason for poorer performance of students in Alberta. However, in *points, lines, and angles*, the difference between Alberta and Quebec was larger for test items that were not included in Alberta's curriculum (–29 per cent) than for items that were included (–7 per cent) (see Table 5). This indicates that Alberta's performance in this topic area is associated with lower curriculum coverage. Similar to the findings for *expressions, simple equations, and relationships,* the larger gap is almost fully explained by the significant increase in the proportion of students in Quebec with correct answers. (There was only a small decline in the proportion of students in Alberta with correct answers.) This indicates that items not covered in Alberta curriculum are of a lower level of difficulty.

The topic area of *two- and three-dimensional shapes* also revealed a gap in relative performance, but it was not as pronounced as that for *points, lines, and angles.* The difference between Alberta and Quebec was larger for test items not included in Alberta's curriculum (–18 per cent), compared to the items that were included in Alberta's curriculum (–11 per cent). This indicates that Alberta's performance in *two- and three-dimensional shapes* is also associated with lower curriculum coverage.

TABLE 5 Comparison of scores in geometric shapes and measures

Topic areas within geometric shapes and measures	Number of test items*	Difference in % correct (Alberta minus Quebec)
Points, lines, and angles	**20**	**–19**
In Alberta curriculum	9	–7
Not in Alberta curriculum	11	–29
Two- and three-dimensional shapes	**26**	**–13**
In Alberta curriculum	18	–11
Not in Alberta curriculum	8	–18
Total	**46**	**–16**

*Only items included in Quebec's curriculum are included in this analysis.

Conclusion

Alberta's performance is lower than Quebec's in all content domains, and Quebec's curriculum covers more TIMSS items at Grade 4 than Alberta's in most topic areas. However, differences in curriculum coverage seem to be linked to differences in student performance only in the *geometric shapes and measures* domain and in the topic area of *expressions, simple equations, and relationships.* There is also a topic area (*whole numbers)* in which Alberta students' performance is closer to the performance of their Quebec peers on items that are not in Alberta's curriculum than on those that are.

All of this suggests that factors other than curriculum play a more significant role in explaining why students in Quebec perform better then students in Alberta. In particular, looking at teaching and learning processes in Quebec (as opposed to examining what teachers are supposed to cover: the curriculum) seems a promising direction for further analysis of TIMSS data.

Topic areas in which the curriculum coverage is linked to student performance appear to contain items that are at a relatively lower level of difficulty. Changes to Alberta curriculum in these topic areas could offer opportunities for quick improvement in results. Specifically, Alberta's performance in TIMSS Grade 4 math should improve if curriculum coverage were broadened in the following areas:

- points, lines, and angles
- two- and three-dimensional shapes
- expressions, simple equations, and relationships

Further analysis may shed light on whether the change in Alberta's results between 2007 and 2015 can be linked to any changes in curriculum coverage in both Alberta and Quebec. This analysis could enhance our understanding of the reasons for the significant decline in Alberta's scores in the *geometry* content domain.

Appendix: Correct responses by content domain/topic area for TIMSS 2015 items

Content domain and topic area	Number of test items*	Average % correct Alberta	Average % correct Quebec	Difference in % correct (Alberta minus Quebec)	Proportion of test items on which % correct is …		
					Higher in Alberta	Similar	Higher in Quebec
Number	**75**	**43**	**57**	**–14**	**0**	**33**	**67**
Fractions and decimals	16	41	60	–19	0	13	88
In Alberta curriculum	16	41	60	–19	0	13	88
Whole numbers	39	44	58	–13	0	36	64
In Alberta curriculum	30	44	59	–15	0	33	67
Not in Alberta curriculum	9	44	53	–9	0	44	56
Expressions, simple equations, and relationships	20	41	52	–11	0	45	55
In Alberta curriculum	15	40	49	–9	0	53	47
Not in Alberta curriculum	5	45	62	–17	0	20	80
Geometric shapes and measures	**46**	**44**	**60**	**–16**	**0**	**33**	**67**
Points, lines, and angles	20	38	57	–19	0	25	75
In Alberta curriculum	9	39	46	–7	0	56	44
Not in Alberta curriculum	11	36	66	–29	0	0	100
Two- and three-dimensional shapes	26	49	62	–13	0	38	62
In Alberta curriculum	18	50	61	–11	0	44	56
Not in Alberta curriculum	8	47	64	–18	0	25	75
Data display	**19**	**59**	**68**	**–9**	**0**	**37**	**63**
Reading, interpreting, and representing	19	59	68	–9	0	37	63
In Alberta curriculum	18	59	69	–10	0	39	61
Not in Alberta curriculum	1	50	57	–8	0	0	100
All items	**140**	**45**	**59**	**–14**	**0**	**34**	**66**

* Test items covered by curricular outcomes in Quebec

Sample questions

Mary left Apton and rode at the same speed for 2 hours.

She reached this sign.

Mary continues to ride at the same speed to Brandon.

How many hours will it take her to ride from the sign to Brandon?

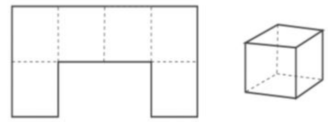
Ⓐ $1\frac{1}{2}$ hours

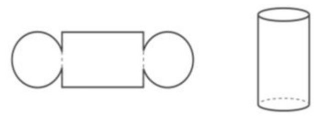
Ⓑ 2 hours

● 3 hours

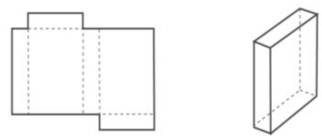
Ⓓ $3\frac{1}{2}$ hours

Ina found the following patterns to make containers. Which pattern actually makes the container shown beside it?

Ⓐ 

Ⓑ 

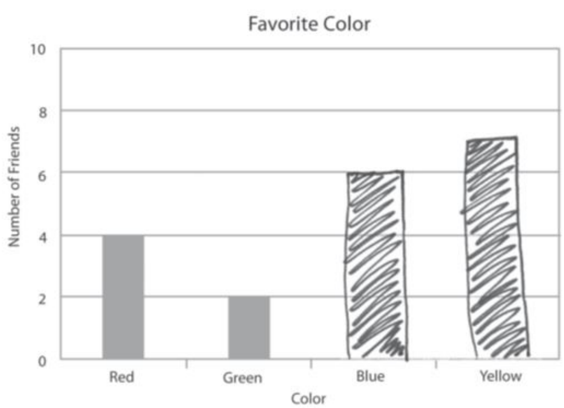
Ⓒ 

● 

Darin asked his friends to name their favorite color. He collected the information in the table shown below.

Favorite Color	Number of Friends
Red	4
Green	2
Blue	6
Yellow	7

Then Darin started to draw a graph to show the information. Complete Darin's graph.



Joan had 12 apples. She ate some apples, and there were 9 left. Which number sentence describes what happened?

Ⓐ $12 + 9 = \square$

Ⓑ $9 = 12 + \square$

● $12 - \square = 9$

Ⓓ $9 - \square = 12$

Tom ate $\frac{1}{2}$ of a cake, and Jane ate $\frac{1}{4}$ of the cake. How much of the cake did they eat altogether?

Answer: $\frac{3}{4}$

$\frac{1}{2} + \frac{1}{4}$

$\frac{2}{4} + \frac{1}{4}$

Measuring Learning Gain in Postsecondary Education: Lessons from HEQCO's Large-Scale Assessment Project

Harvey P. Weingarten
President and CEO, HEQCO

Lauren Hudak
Senior Manager, Research, HEQCO

Sarah Brumwell
Senior Researcher, HEQCO

Ken Chatoor
Researcher, HEQCO

Abstract

We know little about the skill and competency levels of postsecondary graduates for the simple reason that we do not measure them. Without these measurements, we cannot assess the quality of postsecondary education or engage in informed discussion about a purported skills gap. In one large trial, the Higher Education Quality Council of Ontario (HEQCO) partnered with 19 Ontario postsecondary institutions to: (1) assess the feasibility of large-scale skills assessment in a postsecondary setting; (2) explore the value of Education and Skills Online (ESO), the online version of the Programme for the International Assessment of Adult Competencies (PIAAC), to measure employment-related skills directly; and (3) explore changes in skill levels from the start of programs to the end.

We found that large-scale assessment in postsecondary settings is feasible; all logistical, methodological, and technical challenges were met. We also found that the majority of students demonstrate adequate skills. Notably, however, a significant number of students were graduating with below-average skill levels and a minority of students were graduating with superior skills. Finally, we saw skills improve in some situations but not others. This trial shows that skills testing is easily scalable to provincial and national levels and provides important information about how to conduct these trials for maximum benefit and utility.

Key terms: postsecondary education, PIAAC, skills measurement, learning gain, assessment, ESO

Introduction and theoretical framework

It will come as no surprise that postsecondary students expect their education to hone their abilities and enhance their job prospects. Similarly, employers expect postsecondary graduates to possess the skills needed to make positive, productive contributions to their organizations. In the coming decades, Canadian workers will need a transferable set of skills if they are to succeed in their jobs, and if the country is to be economically competitive and attract the industries and jobs that are the hallmark of a modern, knowledge-based, innovative economy.

The curricula of the K–12 system in every province have long emphasized basic skills like literacy, numeracy, and problem solving. These skills are assessed at regular intervals through Canada's participation in international tests, such as the Programme for International Student Assessment (PISA). Canada has also participated in international assessments that test similar skills in adults, such as the Programme for the International Assessment of Adult Competencies (PIAAC).

Innumerable employer, student, and industry surveys have demonstrated clearly that the greatest concern of employers and students is not the content Canadians acquire in their postsecondary programs but rather an apparent shortcoming in the cognitive and behavioural skills needed for success in volatile, changing, and unpredictable job markets (Business

Council of Canada, 2018, Spring). The areas most frequently cited as necessary for job success are: in terms of cognitive skills, an adequate level of literacy and numeracy, problem solving, and critical thinking; and, at a behavioural level, effective communication skills, resourcefulness, and adaptability. From an accountability perspective, a public-policy perspective—and, most important, a learning-gain perspective—skills are now synonymous with quality in postsecondary education.

However, we do not know whether these concerns are justified: whether postsecondary programs are doing an adequate job of fostering these cognitive and behavioural skills. Why? Because we do not measure them in a consistent way. Postsecondary education to a large extent still teaches, evaluates, and credentials information and content. Statements about the employment-related skills of graduates from academic institutions are based largely on inference, opinion, gut feelings, or aspirations.

There is no substitute for the direct measurement of job-related skills to answer important questions about the skills gap, to determine the effectiveness of our investments in programs designed to reduce this gap, and most significantly to determine the most effective ways of teaching these desired skills and competencies to a variety of students.

While measurement can take many forms, large-scale skills assessments can be particularly effective in helping us understand students' learning gain at the systemic and institutional levels. Testing students' skill levels when they enter postsecondary and again when they leave should help us to determine whether their education contributed to an increase in their skill level.

The Essential Adult Skills Initiative (EASI) is an ambitious, large-scale research project undertaken by the Higher Education Quality Council of Ontario (HEQCO) with 19 Ontario postsecondary partners, with funding from both the federal and provincial governments. EASI was designed to measure the literacy, numeracy, and problem-solving skills of incoming and graduating college and university students, and to evaluate the feasibility of administering assessments on a large scale in Ontario's postsecondary sector. In this respect, EASI represents an important first step toward the measurement of learning gain—the degree to which students' skills change over the course of their program of study—across multiple postsecondary institutions.

Research questions, methodology, and data

EASI's central research questions are as follows:

- What are the practical implications of implementing a project like EASI in a postsecondary institution? How feasible is it to scale up this project to a full provincial or national level?
- Is the Education and Skills Online (ESO) assessment a suitable measure of postsecondary students' literacy, numeracy, and problem-solving skills?
- Are there observable differences between incoming and graduating students' literacy, numeracy, and problem-solving skills?

Research design

The EASI pilots were cross-sectional and voluntary, testing first- and final-year students from the same programs in a single academic year. This design allowed for aggregate-level comparisons of student skill levels at the start and conclusion of their postsecondary careers. During the data analysis phase, assessment results were linked to de-identified administrative variables provided by the institution to contextualize the EASI data set.

The EASI college pilot took place in the 2016–17 academic year. First-year college students were tested in the fall 2016 semester and final-year college students in the winter 2017 semester.[1] The EASI university pilot took place in the fall 2017 semester, with first- and final-year university students tested simultaneously. Each testing window was scheduled

[1] One college wished to test both first- and final-year cohorts in fall 2016, so only 10 Ontario colleges opened winter 2017 testing windows.

for early in the semester to avoid conflict with end-of-term assignments and examinations. Institutions selected start dates that suited their academic calendars. Testing windows were generally closed after four to six weeks.

Table 1 provides a high-level view of the design of EASI's college and university pilots. For more detailed information about the items in this table, such as sampling, please refer to HEQCO's EASI report (Weingarten, H. P., Brumwell, S., Chatoor, K. & Hudak, L., 2018). The table also reflects slight variations in study design between institutions, as well as adjustments to the study design between the first- and final-year college testing windows and between the college and university pilots. In each instance, these adjustments were made with the goal of securing student and institutional participation, and were approved by the relevant institutional research ethics boards before being put into practice.

TABLE 1 Research design summary: EASI college and university pilots

	College pilot, fall 2016–winter 2017		University pilot, fall 2017
	First-year students	**Final-year students**	**First- and final-year students**
Number of institutions	11	11	8
Design	Cross-sectional, testing first- and final-year students in selected programs in the same academic year		
Assessment	ESO		
Participation	Voluntary		
Setting	Participants take the test online, on their own time		
Testing launch (rolling)	Late September– early October 2016	Late January–early February 2017*	Late September– early October 2017
Testing window	4–6 weeks		
Recruitment	Email invitation and weekly email reminders to all eligible participants, with additional recruitment strategies that varied by institution		
Individual participant incentives	• Personalized ESO score report • C$10 Amazon.ca gift card†	• Personalized ESO score report • C$20 Amazon.ca gift card • Access to Paddle‡	• Personalized ESO score report • C$20 gift card to Amazon.ca • Access to Paddle‡
Lottery participant incentives	Entry into draw for one C$500 and five C$100 Amazon.ca gift cards per institution per cohort		
Data linkage	Individual-level linkage of ESO results, EASI registration data, and institutional administrative data, e.g., entry category, domestic/international student status, program of study		

Notes:
* One institution tested its first- and final-year students simultaneously in fall 2016.
† One college paid out of pocket to increase the value of the individual incentive for first-year students from C$10 to C$20. Following the success of this strategy, HEQCO raised the individual incentive to C$20 for all final-year college students tested in winter 2017. HEQCO amended EASI's initial research ethics application to reflect both this change and the optional addition of Paddle to the incentive package. The change was approved by each college's research ethics board prior to the start of the winter 2017 testing window.
‡ Paddle is a Canadian online career-exploration platform.

Implementation process

The implementation process was shaped by three core principles. First, the process should be flexible enough to adapt to each institution's unique culture and resources. Second, it should minimize demands on the time of the institutional EASI teams. The process, which was tested and tweaked during the fall 2016 testing windows, allowed us to identify and better handle issues that might arise during subsequent testing. To this end, a registration website was created along with a suite of materials, including email templates, data-management files, and how-to guides, which each institution used to run EASI. Members of HEQCO's EASI team also conducted on-site visits at each institution to ensure that the institutional teams were comfortable with the protocol and materials before testing got under way, and maintained regular phone and email communication with the teams throughout the testing period.

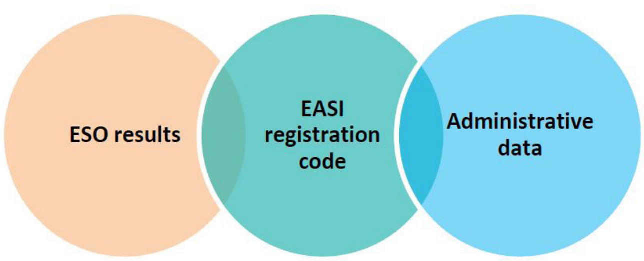
The third principle shaping the implementation process was student privacy. HEQCO at no time had access to student identities, so the institutional teams played an active part in the day-to-day administration of the testing windows. HEQCO handled data tracking and reporting, while the institutional EASI teams communicated with participants, delivered incentives, and sent reminders. The process hinged on the unique codes assigned to each invited student, which were known as EASI registration codes. The registration codes served as proxies for student identity, facilitating data linkage between HEQCO and the institutional teams without requiring personal information to be shared outside of the institution (Figure 1).

Students used the EASI registration codes to complete the participant consent form, claim an ESO authorization code, take the ESO test, and receive rewards for participating—all without HEQCO knowing their identity, and without their home institution knowing their test results until after the testing window had closed. This process was facilitated by the EASI registration website, which hosted participant consent forms for each institution and assigned an ESO authorization code to each consenting student. EASI testing windows remained opened for four to six weeks. Students could join EASI at any point during their institution's testing window and could log on and off the ESO test portal at will.

Data linkages

Data from the EASI registration site and the ESO test portal were available to HEQCO in real time. HEQCO linked these two data sets using the EASI registration codes to determine which students needed to receive their incentives, and to send targeted reminders. Once the testing windows closed, institutions provided limited administrative variables for each consenting participant, such as entry category (i.e., direct-entry from high school versus indirect entry) and program of study. HEQCO used the EASI registration codes to link this data back to the EASI registration site data and ESO test results to assemble the broader EASI data set. Figure 1 displays the central role of the EASI registration code in these linkages.

FIGURE 1 EASI data linkages



Choice of assessment tool

EASI's primary data collection tool was the ESO assessment.[2] ESO is the commercial version of the Programme for the International Assessment of Adult Competencies (PIAAC) of the Organisation for Economic Co-operation and Development (OECD). Like PIAAC, this instrument is delivered in over 40 countries. Both ESO and PIAAC were validated for populations between the ages of 16 and 65. ESO is administered by the Educational Testing Service (ETS) on behalf of OECD and its partners, including the Council of Ministers of Education, Canada. The test has been adapted for Canadian audiences in both official languages.

Like PIAAC, ESO consists of three major components: literacy and numeracy (the core assessment); problem solving in technology-rich environments (PS-TRE); and a background questionnaire. It also includes a brief remedial section (the reading components subtest) for test takers who score low on the core assessment. ESO is an adaptive assessment

[2] Development of ESO has been supported by the European Commission and by the Council of Ministers of Education, Canada.

tool, so questions become progressively easier or more difficult depending on the test taker's performance. The number of questions therefore differs for each test taker.

HEQCO selected ESO for EASI for several reasons, as outlined below.

Quality and comparability ESO underwent a rigorous development and validation process, which is described fully in the *Education and skills online technical documentation* (OECD, 2015). This process included the following steps:

- an update of existing PIAAC items and development of new items based on the PIAAC framework
- review of test items by PIAAC's Literacy and Numeracy Expert Groups
- adaptation and translation of ESO for each participating country
- extensive field testing that included a large Canadian sample
- further revisions to the assessment based on statistical analysis of the field test results

Roughly half of ESO's respective literacy and numeracy items were newly developed, and the remainder were drawn from PIAAC. All of ESO's PS-TRE and reading components items were drawn directly from PIAAC's item pool (OECD, 2015, pp. 10–11). This means that ESO data collected by EASI can be compared to the provincial, national, and international PIAAC results.

User friendliness ESO is simple for both test takers and investigators to navigate. Its intuitive user interface is complemented by the high-quality technical infrastructure provided by ETS. Based on ETS's record of managing high-volume, online, and computer-based assessments such as the Graduate Record Examinations, HEQCO was confident that ESO would be a consistent, secure, and well-functioning tool. Contributing to its user friendliness, test takers received their personal score reports as soon as they completed each component, and the ESO administrator portal allowed HEQCO to access ESO data in real time.

ESO's user friendliness is also enhanced by the control that test takers are able to exercise over their test-taking experience. ESO is delivered online and administered without proctors, permitting some flexibility in terms of where and when students take the test. Although ESO typically requires 90–120 minutes to complete, students may log on and off as needed. These features in particular contributed to the feasibility of the EASI testing windows.

Subject matter and design ESO's adaptive nature sets it apart from many other assessments of adult skills currently available, with test items increasing or decreasing in difficulty according to the test taker's performance. Although this makes it unsuitable for screening test takers who need remedial support or for other developmental education purposes, the test does include a remedial section (the reading components subtest) for test takers who score very low on the first part of the core assessment.

Rather than focusing simply on mastery of the mechanics of vocabulary or arithmetic operations, ESO assesses the real-world application of literacy, numeracy, and problem solving in technology-rich environments. That is, it provides a snapshot of how effectively test takers use essential skills to navigate and engage with the world around them (OECD, 2016).

Understanding ESO scoring and proficiency levels

When test takers complete each major component of ESO, they receive a raw numerical score rounded to the nearest 10 points. The raw scores correspond to a series of proficiency levels that describe task difficulty and contextualize a test taker's range of skills.

The literacy and numeracy components are scored on the same scale. Because the extreme ends of the scale are less precise, no test taker will receive a score below 150 or above 400. The proficiency levels for literacy and numeracy range from below Level 1 to Level 4/5. The PS-TRE component is scored on its own scale. Again, the extreme ends are not sensitive enough to provide a score below 150 or above 400. Note that the range of proficiency for PS-TRE—from below Level 1 to Level 3—is smaller than for the literacy and numeracy components.

It is important to understand that the descriptions of proficiency level do not capture the totality of a test taker's skill set. For example, a test taker who receives an ESO literacy score at Level 2 may very well possess some of the skills needed to complete tasks at Level 3 or higher, both on the assessment and in everyday life. In other words, the ESO and PIAAC proficiency levels indicate the complexity of the tasks that the test taker can reliably and successfully solve (OECD, 2016).

Literacy proficiency levels The ESO scoring mechanism captures literacy's multidimensional nature by including a wide array of literacy difficulty factors. HEQCO has found that literacy tasks at Level 3 and above tend to require increasingly complex, abstract thinking, while tasks at Level 2 and below involve content and activities that are more literal.

Table 2 presents several examples of the types of literacy tasks that test takers scoring at Levels 2, 3, and 4/5 can consistently complete successfully.

TABLE 2 Examples of literacy proficiency at ESO Levels 2, 3, and 4/5

Test takers scoring at this level are likely to be able to…	
Level 2	Submit a vote for or against a new workplace dress code on an employer's web page.
	Identify information on a single web page of a camera store, explaining how this year's photo contest rules differ from those in previous years.
	Name two reasons stated in an employee newsletter for an increase in company sales.
Level 3	Find out whether a utility company accepts the same type of payment by mail and online, using information from a monthly billing statement.
	Use a music store's web page to compare and contrast several reviews to determine which song to download, based on the price and the type of music they like.
	Search several web pages of a national health organization for evidence supporting the claim that exercise can lead to greater work productivity.
Level 4/5	Evaluate posts in a discussion forum on health remedies by comparing the information they provide with that on the Web site of a well-known medical centre.
	Use several links in a city's transportation web page to locate information about special fares or services on holidays.
	Determine which claims in a newspaper article about the benefits of sleep are supported by information and graphs in two long research articles.

Source: Adapted from OECD (2015, pp. 65–68).

Numeracy proficiency levels For ESO, proficiency in numeracy depends both on mathematical skills and on the extent to which an individual can integrate those skills with their "broader reasoning, problem-solving skills and literacy skills" to successfully respond to numeracy-related problems in real-life situations (OECD, 2012, p. 38). HEQCO has also found that numeracy tasks at Level 3 and above tend to require test takers to determine, apply, and evaluate the appropriate method of solving a problem, especially when the relevant mathematical information is not immediately apparent because of either the complexity of the accompanying text, the ways in which numerical information is presented, or the presence of competing information. For this reason, test takers generally need good literacy and problem-solving skills in order to solve numeracy tasks at Levels 3 and 4/5 (OECD, 2012, 2015, 2016).

Table 3 presents several examples of the types of numeracy tasks that test takers scoring at Levels 2, 3, and 4/5 can consistently solve.

TABLE 3 Examples of numeracy proficiency at ESO Levels 2, 3, and 4/5

Test takers scoring at this level are likely to be able to…	
Level 2	Determine the price of a shirt that will be discounted by 25%.
	Determine the price of a single bottle of water when given the cost of an entire case of bottles.
	Determine how many months in a year had sales above the mean sales for the year from a table of monthly sales.
Level 3	Identify which predicted monthly gasoline price is most accurate, based on line graphs of predicted and actual gasoline prices for a year.
	Determine the amount of concentrated lemonade flavouring and water needed to make a large container of lemonade that is in the same ratio of flavouring to water as a smaller amount of lemonade.
	Read a complex graph, comparing the amount of salt, sugar, and fat in a typical diet for men versus a typical diet for women, to determine the amount of sugar consumed by men.
Level 4/5	Convert the number of students enrolled in a university each year into percentages, and then compute the change in the percentage of students enrolled each year.
	Determine how much medicine to give to a child when the dosage is based on the child's body weight.
	Calculate profit from a table containing lists of income and expense sources.

Source: Adapted from OECD (2015, pp. 69–73).

PS-TRE proficiency levels ESO's problem solving in technology-rich environments (PS-TRE) component measures how well test takers use different types of technology and how well they understand and use information in various environments in order to solve problems.

The PS-TRE component uses a different scale than the literacy and numeracy components. Table 4 presents several examples of the types of PS-TRE tasks that test takers scoring at Levels 1, 2, and 3 can consistently solve.

TABLE 4 Examples of PS-TRE proficiency at ESO Levels 1, 2, and 3

Test takers scoring at this level are likely to be able to…	
Level 1	Open, read, and respond to email using an unfamiliar email program.
	Locate specific information on the homepage of a Web site that a friend has recommended.
	Set up a system of computer folders that allows files or emails to be organized and easily retrieved.
Level 2	Find an email or file that has been "lost" somewhere on a computer hard drive.
	Use a sorting tool to make it easier to locate sales numbers for a specific product in a company spreadsheet.
	Conduct a Web search to find out how to solve a software problem, such as how to view a column that won't display properly in a spreadsheet.
Level 3	Evaluate a number of Web search results to determine which has the most relevant and reliable information. Part of this process includes evaluating and refining a search to determine if additional or different types of Web sites should be considered.
	Use a software program that they have never seen before, after limited or unclear direction. Success may be based on a user's general experience with technology or ability to gather information by consulting other online resources including Web sites or user blogs.
	Select from among a number of choices the best software to use for a particular task.

Source: Adapted from OECD (2015, pp. 77–79).

Limitations of the study

Before we turn our attention to the pilot results, we must first review some parameters of this study. These limitations relate primarily to sample selection issues and affect not only the types of analysis we can conduct on the EASI data set but also the types of conclusions we can draw.

First, it is important to remember that *by design*, the EASI sample is neither random nor representative. EASI participants were not randomly selected, and as EASI was a voluntary study, we cannot rule out the possibility of self-selection and non-response bias in the sample. HEQCO worked with its partner institutions to recruit as many students as possible from a wide array of disciplines, but the institutions were able to exercise their discretion in choosing which programs to include in the EASI sample pool.

Second, HEQCO allowed colleges and universities to tweak the design and delivery of EASI to suit their campus culture. These tweaks, which took the form of different recruitment strategies, reminder email schedules, the start and end dates of the testing windows, and in one case the delivery of testing in a tutorial setting, were made with a view to attracting as many participants as possible while creating a smooth experience for both students and staff. That being said, the differences in how students were recruited and tested at different institutions—and in some cases, the differences in how students in different cohorts were recruited and tested at the same institution—present further possibilities for bias in the sample.

HEQCO's decision to prioritize sampling strategies that optimize participation was necessarily accompanied by the possibility of sample bias. Both the EASI college and university pilots obtained response rates that are good for voluntary studies, especially considering the length of time required to complete the ESO assessment. Though the sample was small, a sufficient number of students provided usable test results to permit HEQCO to respond to the research question pertaining to feasibility. Because of the sampling methodology, we did not conduct any detailed analysis of test results—for instance, any direct comparisons of college and university students' results—or any analysis of the relationships between results and sample characteristics such as gender, program, or high school GPA.

Results and analysis

This section provides a general overview of the results of the EASI college and university pilots. Specifically, we present data related to the participation, ESO assessment results, and feedback survey responses for first- and final-year students from EASI's 19 Ontario postsecondary partners.

HEQCO classifies test takers who have completed at least the ESO literacy and numeracy components as having provided usable data. These students have completed enough of the ESO to have received raw numeric scores for both components. Only students who provided usable data were included in the analyses and figures in this section.

All data analyses were conducted using Stata 15 software. Descriptive statistics were used to analyze overall trends in data and assess for distribution of scores and participant characteristics. Given the cross-sectional design of the EASI pilots and the differences in the ways in which each institution administered testing, we have refrained from exploring the statistical significance of the results in any great detail.

Participation

Figure 2 reviews student participation in the EASI college pilot. Nearly 90 per cent of students who provided usable data completed the entire ESO. These students (N = 2,483) were evenly divided between the first- and final-year cohorts. Lastly, 63 per cent of college students who completed the ESO required at least one reminder email to prompt them.

FIGURE 2 **First- and final-year college student participation highlights**

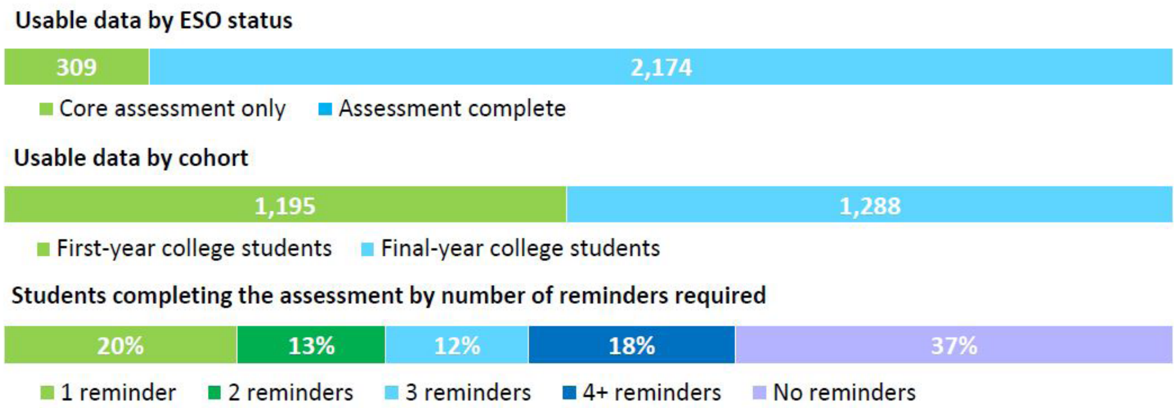
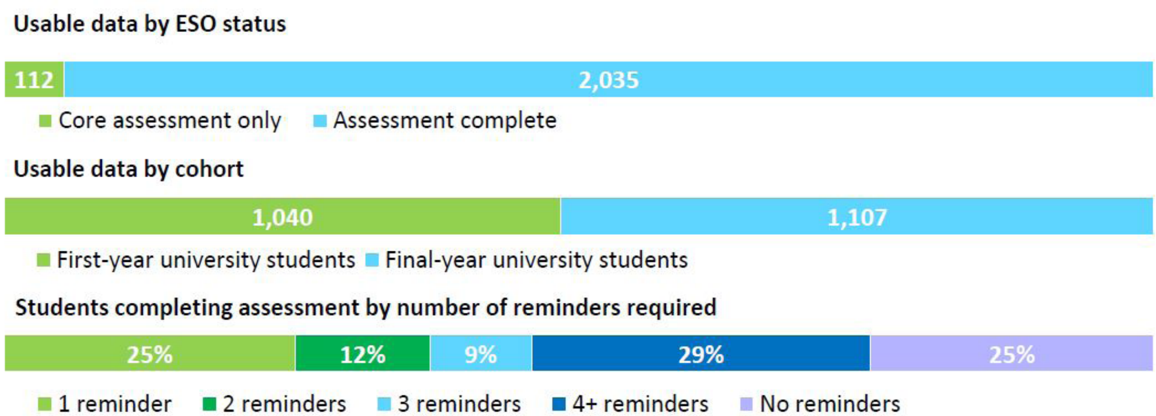


Figure 3 displays the corresponding data from the EASI university pilot, which closely mirrored the participation trends of the college pilot. Nearly 95 per cent of students who provided usable data completed the entire ESO. These students (*N* = 2,147) were evenly divided between first- and final-year cohorts. Finally, 75 per cent of university students who completed the ESO required at least one reminder email.

FIGURE 3 **First- and final-year university student participation highlights**



Sample characteristics

EASI college sample In all, 1,195 first-year and 1,288 final-year students from 11 Ontario colleges provided usable assessment results. Table 5 draws on a combination of institutional administrative data and responses to the background questionnaire portion of the ESO assessment. The median age of first-year college students was 21 years old, and the median age of final-year students was 22; there was little difference between cohorts for gender, having English as a first language, being born in Canada, or holding international status; 41 per cent of first-year students reported having some form of employment, in comparison with 54 per cent of final-year students. Sixty-five per cent of first-year students were identified as indirect entry students, in comparison with 48 per cent of final-year students.

Table 5 **Overview of EASI college sample characteristics**

Variable	First-year students	Final-year students
Background questionnaire		
Number of students with usable data	1,195	1,288
Median age	21	22
% Female	48	50
% Born in Canada	68	66
% English as first language	69	70
% Employed	41	54
Administrative data		
% Indirect entry*	65	48
% International	14	15
% Enrolled in diploma (2 year)	63	65
% Enrolled in advanced diploma (3 year)	33	30
% Enrolled in applied degree (4 year)	4	5

* Indirect entry students were not enrolled in an Ontario high school at the time of application to postsecondary education.

EASI university sample In all, 1,040 first-year and 1,107 final-year students from eight Ontario universities provided usable assessment results. Table 6 draws on a combination of institutional administrative data and responses to the background questionnaire portion of the ESO assessment. The median age of first-year university students was 18 years old, and the median age of final-year students was 21; there was little difference between cohorts for gender, having English as a first language, being born in Canada, holding international status, or entering university indirectly; 26 per cent of first-year students reported having some form of employment, in comparison with 47 per cent of final-year students.

Table 6 **Overview of EASI university sample characteristics**

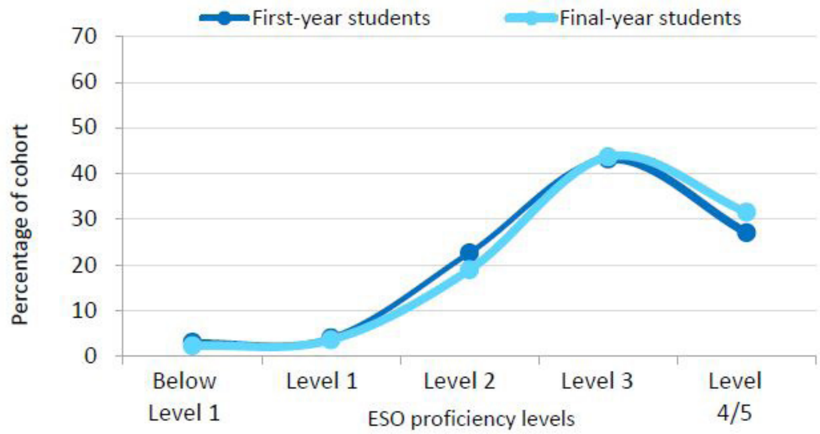
Variable	First-year students	Final-year students
Background questionnaire		
Number of students with usable data	1,040	1,107
Median age	18	21
% Female	66	66
% Born in Canada	73	76
% English as first language	72	73
% Employed	26	47
Administrative data		
% Indirect entry*	10	13
% International	5	4

* Indirect entry students were not enrolled in an Ontario high school at the time of application to postsecondary education.

Aggregate performance

The ESO provides proficiency levels that correspond to the numerical scales for each test component. When the literacy scores of college and university students were distributed by proficiency level, 70 per cent of first-years and 75 per cent of final-years were observed to have scored at or above Level 3 (Figure 4).

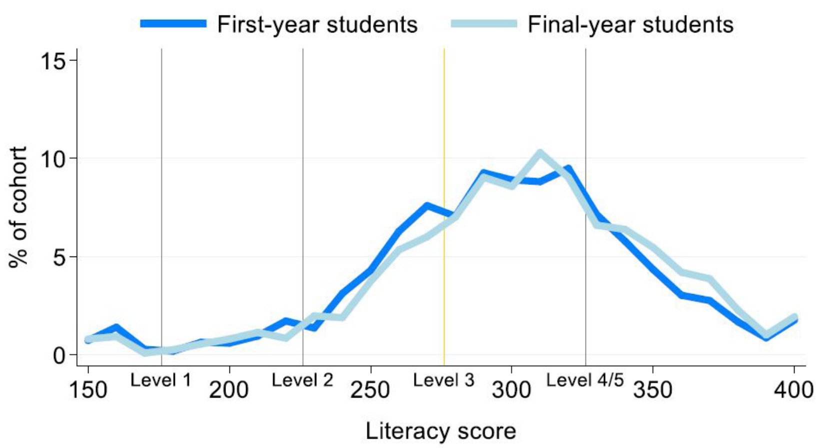
FIGURE 4 **Literacy performance by cohort and proficiency level: College and university students**



Note: First-year students, *N* = 2,235; final-year students, *N* = 2,395.

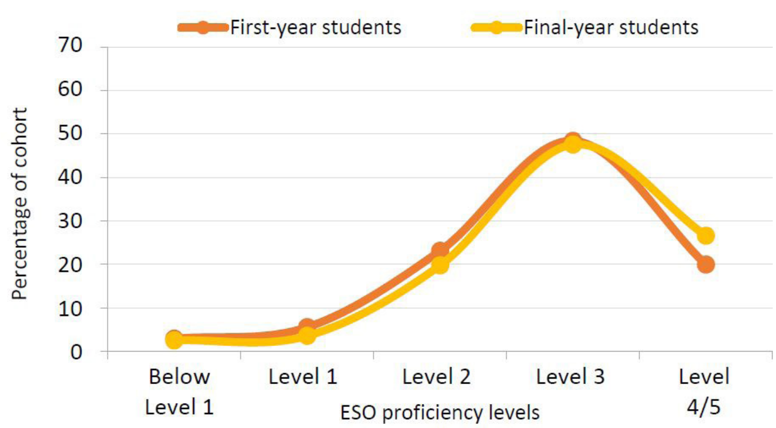
Figure 5 distributes the literacy performance of college and university students by cohort and numerical score. The vertical lines indicate the proficiency level cut-off points, with the gold-coloured vertical line marking the cut-off between Level 2 and Level 3.

FIGURE 5 **Literacy performance of college and university students by cohort and score**



When the numeracy scores of college and university students were distributed by proficiency level, 69 per cent of first-years and 74 per cent of final-years were observed to have scored at or above Level 3 (Figure 6).

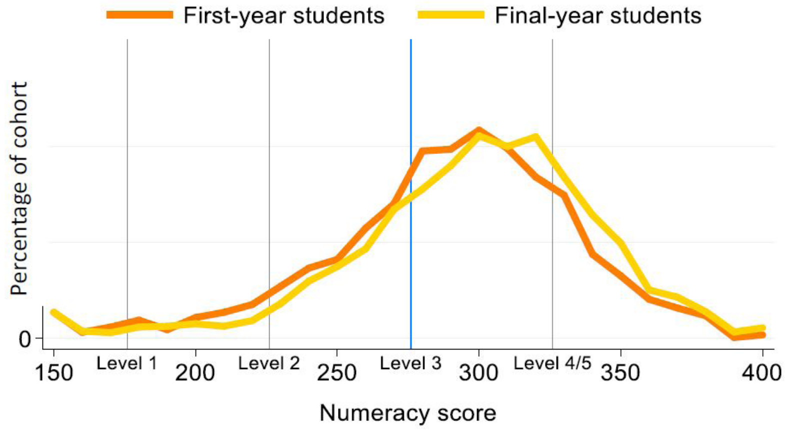
FIGURE 6 **Numeracy performance of college and university students by cohort and proficiency level**



Note: First-year students, *N* = 2,235; final-year students, *N* = 2,395.

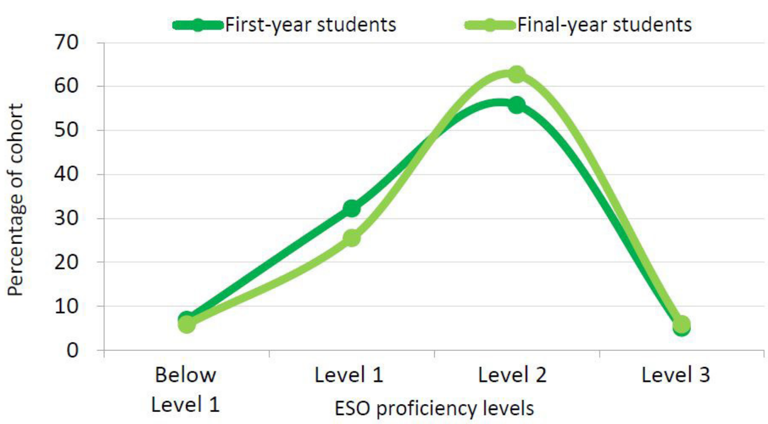
Figure 7 distributes the numeracy performance of college and university students by cohort and numerical score. The vertical lines indicate the proficiency level cut-off points, with the blue-coloured vertical line marking the cut-off between Level 2 and Level 3.

FIGURE 7 Numeracy performance of college and university students by cohort and score



When the PS-TRE scores of college and university students were distributed by proficiency level, 61 per cent of first-years and 69 per cent of final-years were observed to have scored at or above Level 2 (Figure 8).

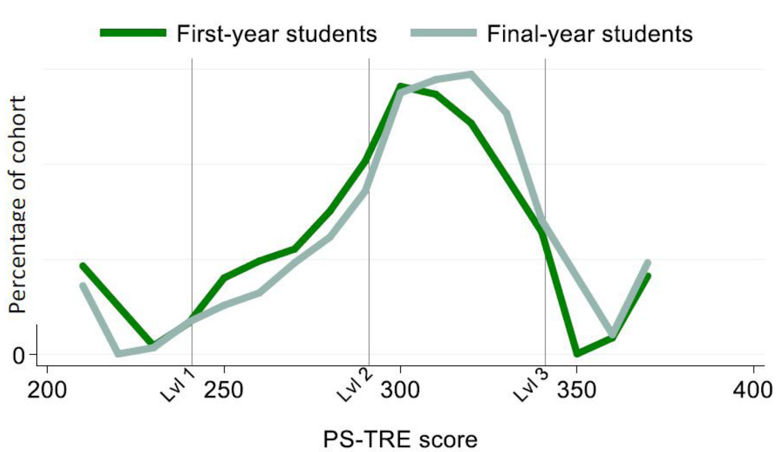
FIGURE 8 PS-TRE performance of college and university students by cohort and proficiency level



Note: First-year students, *N* = 1,912; final-year students, *N* = 2,151.

Figure 9 distributes the PS-TRE performance of college and university students by cohort and numerical score. The vertical lines indicate the proficiency level cut-off points.

FIGURE 9 PS-TRE performance of college and university students by cohort and score



Other performance comparisons

While not discussed in this report, performance based on 2012 PIAAC comparator groups, college cohort, university cohort, program length, and gender are covered in HEQCO's final report on EASI (Weingarten et al., 2018). It is important to note that the trend is consistent for all comparisons.

Discussion and conclusions

This section reviews the key findings of the EASI college and university pilots and answers the original research questions.

Large-scale skills assessment is feasible in postsecondary contexts.

The EASI implementation process has been refined and streamlined over the course of the college and university pilots. In effect, the process managed testing windows at 19 separate institutions, often simultaneously. HEQCO attributes the success of the process to its flexibility, which allows testing windows to be tailored to suit institutions of all sizes and capacities. Additionally, because HEQCO coordinates the administration of the testing windows, the demands placed on participating institutions are reduced. Throughout both trials, HEQCO resolved important logistical and methodological issues regarding student recruitment, test administration, and participant privacy, among others. These will be crucial issues as we consider scaling up the project.

While the EASI model succeeded in simplifying the logistics of administering large-scale assessments, it must be noted that institutions contributed a considerable amount of resources, primarily in the form of staff time on the project. Local staff members served as the bridge between HEQCO and students, spearheading recruitment activities on campus and fielding student inquiries about the project, in addition to administering the scheduled distribution of invitation emails, reminders, and incentives. They were also responsible for gathering institutional data and contact information for the sample, which, depending on an institution's information technology infrastructure, can be a complex task. HEQCO acknowledges that further rounds of testing will require greater support for institutions in these areas, whether in the form of providing funds to alleviate the personnel costs associated with EASI or in further streamlining the logistics of test delivery on campus.

Institutions, to a far greater extent than anticipated, were interested in participating in the college and university pilots. The importance of institutional partnership to EASI's success cannot be overstated. The administrative data provided by institutions added invaluable context to the data set, and the institutional investigators were essential to the smooth administration of the testing windows. There is no better substitute for the direct measurement of student skills in demonstrating institutional quality, and the level of interest and support EASI has received from Ontario's colleges and universities confirms that postsecondary institutions share our goal of developing methodologies that maximize the benefit and utility of large-scale assessment.

In light of these findings, we are confident that the EASI process can easily be scaled up to a provincial or national level.

ESO assessment is a suitable measure of postsecondary students' literacy, numeracy, and problem-solving skills.

The results indicate that ESO is an efficient measure of postsecondary student skills. From a performance standpoint, the fact that the average scores of EASI participants were similar to those of their OECD, Canadian, and Ontarian comparators from PIAAC 2012 reaffirmed the instrument's suitability for this audience. For more information on the PIAAC 2012 comparisons, please refer to HEQCO's EASI final report (Weingarten et al., 2018). We recognize that this assumption needs to be analyzed further, with better sampling controls in place, to help us understand the influence on the performance data of possibly confounding factors such as student motivation and institutional differences in test delivery. The assessment results were normally distributed. This indicates that when steps are taken to prevent sample bias, the data collected by ESO can be used for advanced statistical analyses.

The distribution of skill levels and relatively minor skill gain demand further research.

The distribution of skill levels among graduating students raises concerns. While the majority of graduating students demonstrate average skills, too many present below-average skill levels and too few present superior skill levels.

These findings demand further investigation through larger trials, and consideration has been given to the optimal design of these projects. Skill development is best measured when large-scale assessments are integrated with routine evaluations of student performance. We are confident that future large-scale assessment projects could explore skill gain in more robust ways by instituting better sampling controls, employing a longitudinal design, and including large-scale assessment among the standard assessment activities students encounter in postsecondary education.

EASI exceeded HEQCO's expectations. We developed a flexible, efficient methodology for measuring student skills at multiple institutions, established the suitability of the ESO assessment for postsecondary audiences, and, with the support of our partner colleges and universities, demonstrated that large-scale assessment is eminently feasible. The data collected through the EASI pilots provided invaluable insight not only into student skill development but also into the design of future studies of learning gain. The results are consistent with other skills measurement research that HEQCO and others have conducted. Nevertheless, questions of interpretation remain. We are struck by the number of important questions that knowledgeable people ask about the interpretation of our findings because they see them as relevant to their policy development, financial investment, or program design. These questions can be answered, but only with increased funds and supports for the administration of testing on campuses, larger trials that provide better control over sampling, and a longitudinal experimental design.

References

Business Council of Canada. (2018, Spring). *Navigating change: 2018 Business Council Skills Survey*. Retrieved from http://thebusinesscouncil.ca/wp-content/uploads/2018/04/Navigating-Change-2018-Skills-Survey-1.pdf

Weingarten, H. P., Brumwell, S., Chatoor, K. & Hudak, L. (2018). *Measuring Essential Skills of Postsecondary Students: Final Report of the Essential Adult Skills Initiative.* Toronto: Higher Education Quality Council of Ontario.

Organisation for Economic Co-operation and Development (OECD). (2012). *Literacy, numeracy and problem solving in technology-rich environments: Framework for the OECD survey of adult skills*. Paris: OECD Publishing. doi:10.1787/9789264128859-en.

Organisation for Economic Co-operation and Development. (2015). *Education and skills online technical documentation*. Updated June 2018. Paris: OECD Publishing. Retrieved from http://www.oecd.org/skills/ESonline-assessment/assessmentdesign/technicaldocumentation

Organisation for Economic Co-operation and Development. (2016). *The survey of adult skills: Reader's companion* (2nd d.). Paris: OECD Publishing. doi:10.1787/9789264258075-en.

The Research Initiative on Education and Skills (RIES): Using PIAAC/LISA to Inform Skills Development Policy

Brad Seward
Practice Lead, Education and Skills, Mowat Centre

Khuong Truong
Policy Associate, Mowat Centre

Deepti Kapadia
Policy Associate, Mowat Centre

Andrew Parkin
Director, Mowat Centre

Initiative summary

The Mowat Centre and the Higher Education Quality Council of Ontario (HEQCO) have entered into a partnership to implement an innovative collaborative policy research initiative, the Research Initiative on Education and Skills (RIES). The objective of the RIES is to access, analyze, and mobilize data relating to the education, skills, and labour market outcomes of Canadians and to disseminate the findings in order to inform policy development. The initiative makes use of the following Statistics Canada data:

- the Programme for the International Assessment of Adult Competencies (PIAAC), a direct assessment of the competencies of adult Canadians between the ages of 16 and 65
- the Longitudinal and International Survey of Adults (LISA), a longitudinal survey that includes a subsample of PIAAC respondents and that collects information on employment, education, training, health, and family
- linkages of LISA data to historical tax data from the T1 Family File (T1FF) and employers' files (T4), as well as to other administrative data relating to employment, income, and immigration status

The RIES is designed to achieve three goals. First, it will develop an infrastructure that makes the data accessible to researchers both inside and outside of the Mowat Centre and HEQCO. Second, a directed program of research will use the data to produce a series of studies on key policy issues relating to the education, skills, and labour market outcomes of Canadians. Third, a dissemination strategy, which is already in place, ensures that the findings of the research papers are made accessible to the public policy community in Canada, increasing the long-term impact of this research initiative.

Accessing Statistics Canada data

Normally, access to data through research data centres (RDCs) needs to be arranged for each individual researcher on a project-by-project basis. This requirement can make the commitment to survey data unwieldy for researchers. To correct this limitation, a special arrangement with Statistics Canada has been negotiated for the RIES. Under the arrangement, RIES data technicians will be located at the University of Toronto RDC and will access the data on behalf of successful applicants. This ensures that access to and use of the data conform to Statistics Canada's requirements. The RIES technicians will have access to the data sets through the RDC during the term of the project and will produce the data products required to support successful proposals.

As noted, the RIES draws on two surveys: PIAAC and LISA. PIAAC is an international survey initiated by the Organisation for Economic Co-operation and Development (OECD); it was administered between 2011 and 2016 in 34 countries. PIAAC offers an objective assessment of the skills and competencies of adults between the ages of 16 and 65 across three domains: literacy, numeracy, and problem solving in technology-rich environments. It is the first large-scale international assessment of adults to be conducted by computer and to measure individuals' proficiency in using information and computer technology. In addition to the direct assessment of skills, the PIAAC survey used an extensive questionnaire that gathered information on respondents' backgrounds, as well as on their use of skills at work and in everyday life. Canada's sample of PIAAC respondents (25,267) was the largest in the OECD. Within the sample are 5,165 Ontario respondents as well as oversamples of immigrants, youth, off-reserve Indigenous persons, and official-language minorities. PIAAC also incorporates some linkages to the previous version of the survey, the Adult Literacy and Life Skills (ALL) survey, allowing for some trend analyses using a number of the skills measures contained in PIAAC.

LISA has been conducted every two years since 2012. It contains information on roughly 24,000 Canadians about their jobs, education, health, and family, and how changes in these areas affect different aspects of their lives. LISA also contains a sub-set of 9,000 PIAAC respondents. This link permits the skills assessments available in PIAAC to be applied directly to respondents who are also in the LISA survey. The biennial LISA survey waves produce content relating to the following areas: education history; employment and income; skills training; personality traits; health, mental health, and disability; retirement; assets and debt; and family composition. The latest wave of LISA data, released in December 2018, has a close focus on elements of postsecondary education.

Perhaps most important, Statistics Canada has enhanced the LISA data by linking the waves of the survey to historical tax data from the T1 Family File (T1FF) and from employers' files (T4), as well as to administrative data relating to self-employment, pension plans, and immigration status. This allows the information collected through the surveys and skills assessments to be connected to longitudinal information regarding, among other things, employment status and earnings. The linked tax data covers the period from 1982 to the present. This linkage offers an opportunity to study demographic groups in a way that has been relatively unexplored in the wider research community.

Why PIAAC/LISA data?

Effectively producing skills research depends on the availability of data containing measures of the labour market experiences, educational history, and demographic background of Canadians. Traditionally, however, available data have had three significant limitations. First, the data typically do not contain direct information about the skills that respondents possess; skill levels are generally inferred from either educational attainment or occupation. Second, data are rarely longitudinal, meaning that observations are made at a single point of time, leaving few opportunities to explore how different factors and experiences affect outcomes over time. Third, information about respondents is limited to what can be asked and answered in the context of each separate survey, since there are no linkages between the survey and other data sources.

The availability of PIAAC and LISA overcomes these limitations. We have an unprecedented opportunity to address crucial policy research questions and deepen policy-makers' understanding of education, skills, and the labour market in Canada. At the same time, the complexity of working with these data means that relatively little research utilizing the surveys has been published. So, while the PIAAC–LISA data represent an exceptional opportunity, they also risk becoming a missed opportunity. The objective of the RIES is to ensure that the investments made to date in the creation of the data are not lost.

Existing work with PIAAC/LISA

Prior Canadian research using the PIAAC and LISA surveys has typically focused on a small number of key thematic areas. Some examples are the labour market outcomes of various demographic groups such as Indigenous populations; immigrants and second- and third-generational groups; the effect of age on economic outcomes using cohort effects within the population; and outcome differences within the population based on both educational attainment and objective measures of literacy, numeracy, and technology-based problem solving.

Most of these efforts have sought to assess not only proficiencies in literacy, numeracy, and problem solving but also the return on educational investment. Wider scopes of analyses have also sought to observe the labour market value of cognitive and non-cognitive skill development. Findings generated using PIAAC and LISA typically present narratives that underline disadvantages in labour market experiences (i.e., wage penalties and barriers to occupational mobility) for traditionally marginalized groups. The surveys also contain variables that allow researchers to study mismatches between skills and employment within the Canadian labour market. Such studies have been produced primarily to contribute to policy discussions pertaining to immigration, formal and non-formal education return, retraining, and the labour market integration of lagging demographic groups.

PIAAC has been used largely to assess the economic and social integration of immigrants. While international research has assessed skills, particularly in relation to educational mismatches (overeducation), the Canadian literature has for the most part not followed suit. Additionally, researchers have been more concerned with the cognitive skills of aging demographic groups than with those of younger segments of the population in Canada. The literature on younger demographic groups that does exist has been produced by international researchers and compares the labour market outcomes of millennials abroad to those in Canada. This research has focused primarily on access to social capital and differences in opportunities to develop human capital in order to foster understanding of demographic disparities in labour market and educational attainment.

LISA, on the other hand, has been used predominantly to assess the distribution of skills across demographic groups. These avenues include family status, income level, and measures of employment status. It is worth noting that a considerable number of studies have used Quebec rather than Ontario as the reference point in models. Still, the longitudinal advantage offered by LISA has been underutilized in the research community and is a valuable angle that could be explored by the RIES (particularly while also utilizing the administrative linkage files).

Scope of the RIES

While both the Mowat Centre and HEQCO are based in Ontario, the RIES is a national initiative that engages researchers across the country and presents data at both national and provincial/territorial levels. Given the extent of provincial responsibilities in education and skills development, as well as the regional nature of labour market dynamics in Canada, it is important to present and analyze data at the subnational level whenever possible. At the same time, understanding of the situation in any one region or province, including Ontario, is strengthened by presenting results in a comparative and national context.

Relatively few studies have utilized PIAAC/LISA data to study the experiences of Canadians, leaving a wealth of opportunities to make significant contributions to policy using these surveys. The following paper therefore offers a preview of an upcoming report being produced by the Mowat Centre to demonstrate the types of research opportunities that the data make possible. This paper offers a descriptive overview of information and communication technology occupations in Canada and highlights issues of representation that may lead to future employment problems in the field.

Untapped Pool or Leaky Pipeline? Female Involvement in the ICT Sector

Abstract

There are fewer women than men in Information and Communication Technology (ICT) professions in Canada. What drives this trend? Why is this a policy problem, and what can be done to address it? This report, the first from the Research Initiative on Education and Skills (RIES), tackles these topics.

Introduction

Skills related to information and communication technology (ICT) occupations (OECD, 2016):

- ICT specialist skills (e.g., Java programming, phone application creator, etc.)
- ICT generic skills (general-purpose technology skills in daily life and at the workplace)
- ICT complementary skills (necessary skills to solve given problems in technology-rich environments)

Innovations to digital technologies have altered the Canadian economy, leading to new industries and occupations. As a result, workers must constantly update their technical skills to keep up with these advances (Organisation for Economic Co-operation and Development [OECD], 2016). Among emerging fields, information and communication technology (ICT) offers some of the fastest-growing occupations available to Canadians.

To excel in ICT careers, workers must develop skill sets that are crucial to the digital economy. Moreover, while some ICT skills are job specific, most are required in many occupations. The Information and Communications Technology Council has projected that Canada will need approximately 182,000 highly skilled ICT workers in 2019 (ICTC, 2016). Yet while demand for ICT professionals continues to grow, the supply of these workers seems to lag.

One issue is that ICT occupations typically fail to attract and retain women. Even though more women are employed in the field than ever before, with an annual employment growth rate of approximately 13 per cent, the ICT workforce remains predominantly male. Some estimates put the proportion of men in these occupations at three times the number of women (OECD, 2016)). In an era when the economy requires highly skilled workers, women are an untapped pool of talent.

Some authors suggest rectifying labour shortages by recruiting more skilled immigrants (Tech Toronto, 2016, p. 26). However, it is more efficient in both the short and the long term to ensure that skills already available but underutilized in the workforce—such as those of recent immigrants who are not working in the occupations for which they are trained or of populations currently under-represented in the ICT field—are adequately tapped before we look outward for new labour. Policy-makers are therefore tasked with ensuring that Canada's current workforce is properly mobilized. To do so, policy-makers and non-governmental organizations must have all the necessary information at their disposal to develop effective policies and programs that will encourage female participation in ICT careers.

Problematically, there is no consensus on ICT occupational classifications. We therefore offer a novel approach to studying women's representation in ICT, by examining three existing classifications. This report analyzes female involvement across the following: a 15-category list of core ICT occupations defined by ICTC (2016), a 25-category measure of ICT occupations updated by the ICTC in 2017, and a broader group of 67 ICT occupations outlined in Mueller, Truong, and Smoke (2018). Our purpose is to determine if differences in participation arise from how researchers classify information and communication technologies. We also assess whether any divisions of work tasks between men and women in the ICT industry are contingent on their sex. Lastly, we consider differences in men's and women's ICT employment by full-time and part-time status and by the public and private sector.

This report is driven by the following research questions:

1. What is the representation of women working in the ICT occupations, and does this representation vary according to how the occupations are classified? Does female representation differ between the private and public sectors?
2. Are differences in the ICT participation of men and women driven by disadvantages in general technology skills?
3. Are there differences in the division of labour for men and women in ICT occupations?

Data sources and method of analysis

Data sources

The data are drawn from the 2012 wave of the Programme for the International Assessment of Adult Competencies (PIAAC). PIAAC assesses individuals aged 16 to 65 in three main skills domains: literacy, numeracy, and problem solving in technology-rich environments (PS-TRE). These are assessments of cognitive skills used in everyday life and workplace activities. These are foundational, basic skills that are necessary to build other complex skill sets.

In this study, PS-TRE is used as a proxy for general-purpose technology and foundational ICT skills in order to explore differences in ICT usage and the subsequent effect on labour market performance. PS-TRE scores range from 0 to 500 and are used to measure the spectrum of problem-solving skills in technology-rich environments. Thus, the test scores are not indicative of any single individual's ability or skills but are representative of a population of interest: working-age individuals aged 25 to 60, with positive self-reported hourly wages at the time of survey. Excluded from the sample are self-reported self-employed individuals, those who are under 25, and those whose reasons for not completing either the ICT core test or the computer-based assessments are uncategorized. Respondents who did not provide information on key demographic variables, such as education, place of birth, or did not include test scores for problem solving in technology-rich environments, have also been removed from the analyses.

Method of analysis

The analyses presented in this report use various multivariate methodologies (e.g., ordinary least squares and linear probability models), in addition to descriptive statistics. The purpose of these statistical techniques is to observe any differences in the general-purpose technology and problem-solving skills of women and men, the proportion of women and men represented in ICT occupations, and the labour market experiences of women and men in these occupations. PS-TRE test scores have been adjusted to have a mean of 0 and a standard deviation of 1.[1] The unit of measurement is in the standard deviation of test scores and has a scale from 0 to 1. Analyses are presented in the order described below.

First, differences in the prospect of working in the ICT occupations are estimated using linear probability modelling. In all analyses, several basic demographic factors are controlled for:

- age, province of residence, and the highest level of education attained by respondents: below high school, high school, postsecondary education below a bachelor's degree, a bachelor's degree (the reference group), and postsecondary education above a bachelor's degree
- whether respondents completed a STEM degree
- whether respondents completed their highest education level abroad

General-purpose technology skills scores are accounted for to help us understand whether the lack of female representation is driven by deficiencies in these skills. We are also interested in understanding the impact of working full or part time and of working in the private sector on the likelihood of employment in ICT occupations.

[1] Each individual's test scores were deducted from the overall weighted sample mean and divided by the standard deviation of the sample's test score.

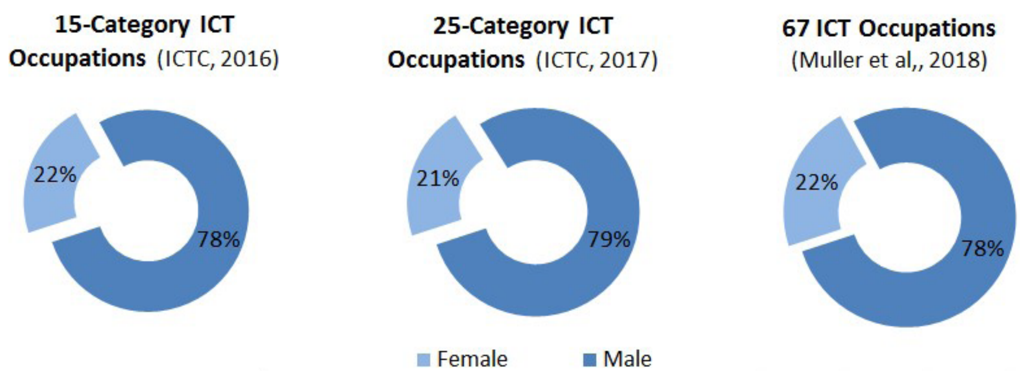
Second, we discuss the PS-TRE scores of men and women. These analyses help shed light on whether certain mechanisms may contribute to differences in pursuing a technology-related career. The analyses control for the same sociodemographic factors as in the previous set of analyses.

Third, we examine differences in ICT-related tasks at work. Using the same sociodemographic controls, these analyses examine whether the under-representation of women is due to differences in career responsibilities that may lead some women to leave or avoid the field.

Discussion of results

We begin by comparing male and female employment across three classifications of information and communication technology occupations (Figure 1): a 15-category list of the core ICT occupations defined by Information and Communications Technology Council (ICTC) in 2016, a 25-category measure of ICT occupations updated by the ICTC in 2017, and a broader group of 67 ICT occupations outlined in Mueller, Truong, and Smoke (2018)[2]. The results of our analyses indicate that even when women have similar ICT skills, they are under-represented in ICT occupations across all three types of classification. As shown in Figure 1, no matter what definition of ICT is used, women are less likely to hold employment in these occupations.[3]

FIGURE 1 Representation of women across different definition of ICT occupations



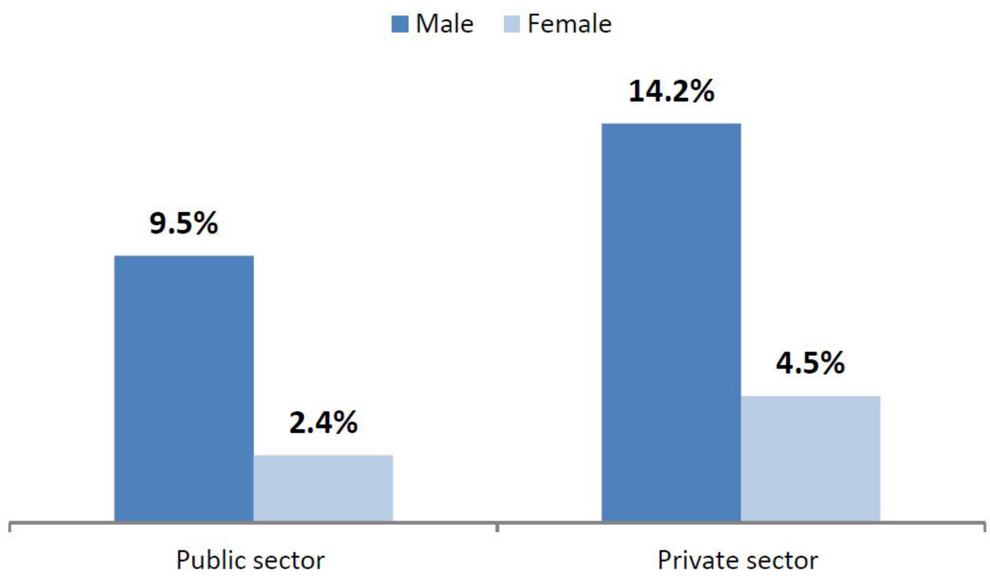
How much of the occupational effect is due to gender differences in full- and part-time employment? We explore the impact of these differences by running models that restrict the sample to individuals who are working full time. We also include models that consider differences in full- and part-time employment in the full sample regression analysis. The results are very similar: women remain significantly less likely to hold occupations in the ICT sector. We repeat this exercise to look at the chances of being employed across the occupational definitions of ICT. Again, women are less likely to be in the ICT occupations.

How much of an effect does the employment sector have on gender representation? We examine the proportion of men and women in ICT occupations within the wider labour force, controlling for employment in the public or private sector. Figure 2 shows us that in the workforce overall, a greater proportion of women in ICT jobs work in the private sector than in the public sector, though this is also true of men and is likely to be driven by the increased availability of ICT-related jobs in this sector of the labour market.

[2] This list is created by combining data from the Occupational Information Network (O*NET), the Standard Occupational Classification (SOC), and the 2011 National Occupational Classification. The O*NET-SOC database is available at www.onetonline.org.

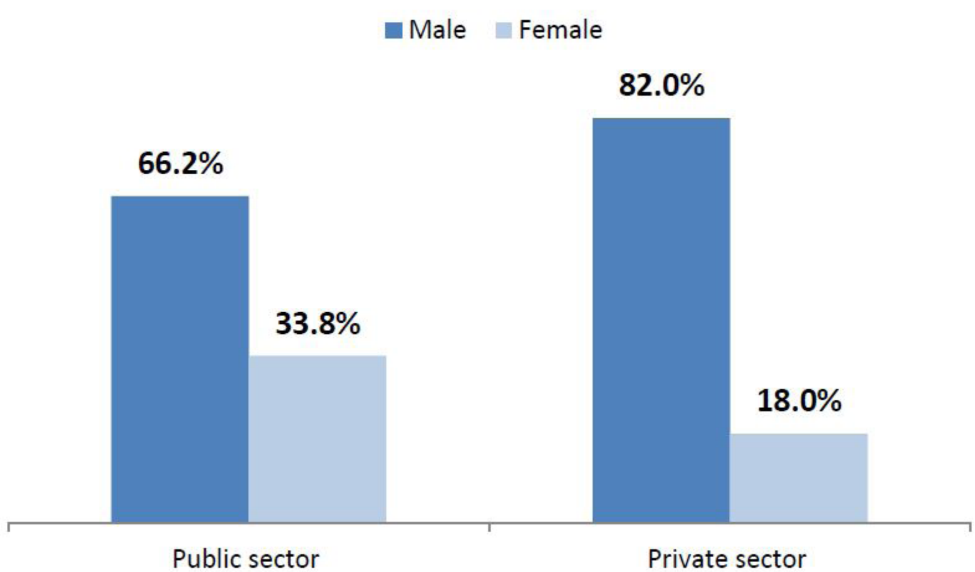
[3] In subsequent analyses, we will discuss the gender inequality in ICT occupations using the ICTC's 25 core digital economy occupations.

FIGURE 2 **Public and private sector men and women working in ICT occupations as a proportion of the labour market as a whole**



We therefore need to investigate this trend further by observing the proportions of men and women employed in the public and private sector while limiting our focus to those in ICT occupations. In Figure 3, we see that the proportion of women to men with ICT careers is higher in the public sector than in the private sector. It is likely that a greater emphasis on diversity hiring within the provincial and federal governments is encouraging female involvement.

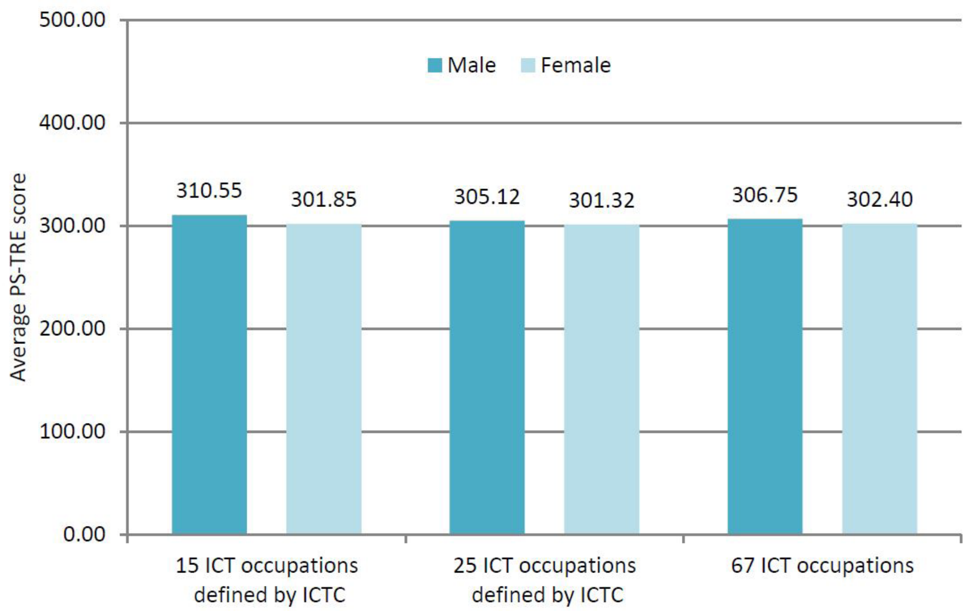
FIGURE 3 **Female and male representation in ICT occupations by public and private sector**



Does the quality of workers' basic ICT skills influence their likelihood of holding employment in ICT occupations? As one would expect, in the general population, the chances of being employed in ICT occupations increases with respondents' skills in ICT. On the other hand, we do not find support for the myth that women's lack of skills renders them unfit for occupations in the ICT sector. see Kindsiko and Türk, 2016 Summary tables of these regressions are available in Appendix C.

Figure 4 shows differences between men and women in general-purpose technology and problem-solving skills, as measured by problem solving in technology-rich environments (PS-TRE) scores. Our findings do not indicate that women and men do not score significantly differently in basic ICT skills.[4] It is therefore surprising that women are less likely to be represented in ICT occupations even when their skills are roughly equal to their male counterparts. These results suggest the existence of a pipeline issue; that is, fewer women may be enrolling in the fields of study that are necessary for ICT careers.

FIGURE 4 Comparison of PS-TRE scores by gender



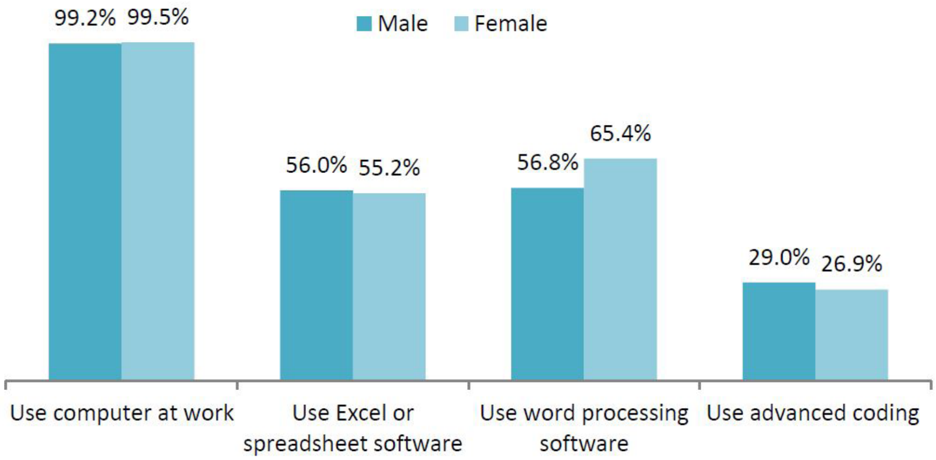
Indeed, educational pathways do seem to have some effect on occupational decisions. For example, in the general population, receiving training in the science, technology, engineering, or mathematics (STEM) disciplines increases the propensity to acquire employment in an ICT occupation by as much as 15 percent (see Appendix C). At the same time, women and men who hold degrees in STEM disciplines have similar probabilities of being employed in ICT occupations . Therefore, part of the issue appears to be that women are less likely to pursue these more technical fields of study.

Perhaps systemic factors affect women's choice to participate in ICT-related fields of study and subsequently to work in ICT occupations. The wider discussion often describes a labour market in which women's skills are "gendered," referring to how tasks are distributed between women and men. This separation of tasks can lead women to receive less specialized or less meaningful work assignments than men.

We explore several different measures of self-reported ICT applications used daily at work. Figure 5 shows the use of these applications in ICT occupations. Our findings indicate that the daily tasks of women in ICT do not differ significantly from those of their male counterparts with respect to either their use of processing software such as Word and Excel or their advanced programming and other computer-related responsibilities. This is surprising, given that we find evidence to suggest divisions of tasks based on sex within the wider population, where women are more likely than men to take on less rigorous assignments. It is therefore even more remarkable that women have lower levels of participation in ICT careers, considering that this sector practises more equitable divisions of labour.

[4] Even though the test scores used in the analysis capture only ICT generic and compleimentary skills, these skill sets are foundational skills to for building advanced ICT skills. Thus, it is hard to imagine that persons people with specialist ICT skills would not perform equally well on these "general-purpose" technology and problem-solving skills tests.

FIGURE 5 Applications used daily at work by women and men working in 25 ICT occupations



The results presented above suggest that as the ICT sector continues to grow, employers may struggle to fulfil labour demands unless a greater proportion of women pursue careers in the field. Female involvement may be affected by two related issues. First, there appears to be a pipeline problem, whereby female participation in the field seems stunted. Second, continued employment-side barriers may hinder women's access to and navigation through the field.

Appendix A: Information and communication technology occupations

2011 NOC code	Description
0131	***Telecommunication carriers' managers***
0211	**Engineering managers**
0212	Architecture and science managers
0213	***Computer and information systems managers***
0911	**Manufacturing managers**
1252	**Health information management occupations**
1254	Statistical officers and related research support occupations
2111	Physicists and astronomers
2114	Meteorologists and climatologists
2115	Other professional occupations in physical sciences
2131	Civil engineers
2132	Mechanical engineers
2133	***Electrical and electronics engineers***
2134	Chemical engineers
2141	Industrial and manufacturing engineers
2142	Metallurgical and material engineers
2143	Mining engineers
2144	Geological engineers
2145	Petroleum engineers
2146	Aerospace engineers
2147	***Computer engineers (except software engineers and designers)***
2148	Other professional engineers, n.e.c.
2151	Architects
2152	Landscape architects
2154	Land surveyors
2161	Mathematicians, statisticians, and actuaries
2171	***Information systems analysts and consultants***
2172	***Database analysts and data administrators***
2173	***Software engineers and designers***
2174	***Computer programmers and interactive media developers***
2175	***Web designers and developers***
2211	Chemical technologists and technicians
2212	Geological and mineral technologists and technicians
2221	Biological technologists and technicians
2223	Forestry technologists and technicians
2231	Civil engineering technologists and technicians
2232	Mechanical engineering technologists and technicians
2233	Industrial engineering and manufacturing technologists and technicians
2241	***Electrical and electronics engineering technologists and technicians***
2242	**Electronic services technicians (household and business equipment)**
2243	**Industrial instruments technicians and mechanics**
2244	Aircraft instrument, electrical, and avionics mechanics, technicians and inspectors
2251	Architectural technologists and technicians
2253	Drafting technologists and technicians
2255	Technical occupations in geomatics and meteorology
2261	Non-destructive testers and inspection technicians
2262	Engineering inspectors and regulatory officers

2281	***Computer network technicians***
2282	***User support technicians***
2283	***Information systems testing technicians***
5222	**Film and video camera operators**
5223	**Graphic arts technicians**
5224	*Broadcast technicians*
5225	**Audio and video recording technicians**
5226	Other technical and co-ordinating occupations in motion pictures, broadcasting and the performing arts
5241	***Graphic designers and illustrators***
6221	Technical sales specialists— wholesale trade
7315	Aircraft mechanics and aircraft inspectors
9222	**Supervisors, electronic manufacturing**
9523	**Electronic assemblers, fabricators, inspectors, and testers**

Sources: This list is amended from Mueller, Truong, and Smoke (2018) to include missing occupations that have been added by the Information and Communications Technology Council (ICTC, 2017b). **Bolded** items are 25 core digital economy occupations in ICTC (2017b). ***Italicized and bolded*** items are 15 core digital economy occupations included in ICTC (2016). These two lists of occupations are sub-sets of the broader group of 67 ICT occupations defined by Mueller, Truong, and Smoke (2018).

Appendix B: Multivariate analyses

TABLE B1 Probability of being employed in ICTC's 25 ICT occupations

	(1)	(2)	(3)	(4)	(5)
Female indicator	–0.091***	–0.047***	–0.042***	–0.041***	–0.038***
	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)
Basic ICT scores	0.033***	0.035***	0.035***	0.034***	0.034****
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Indicator for individuals with a STEM degree		0.146***	0.142***	0.145***	0.141***
		(0.014)	(0.014)	(0.014)	(0.014)
Working in the private sector			0.028**		0.027**
			(0.009)		(0.009)
Working full time				0.032***	0.030***
				(0.008)	(0.008)
Constant	–0.253***	–0.278***	–0.308***	–0.276***	–0.305***
	(0.070)	(0.072)	(0.075)	(0.072)	(0.075)
Controlling for other covariates					
Age and its square, educational variables	Y	Y	Y	Y	Y
Residential location	N	Y	Y	Y	Y

Note: The linear probability model is used to obtain these estimates. Estimated coefficients are an equally weighted average of coefficients derived from each plausible value. All 80 replicate weights are used to account for the sample stratification. Estimated coefficients are interpreted as the percentage of being employed in the ICTC's classification of ICT occupations. Standard errors are given in parentheses.

*$p < 0.5$

**$p < .01$

***$p < .001$.

TABLE B2 Women's probability of working in ICTC's 25 ICT occupations

	(1)	(2)	(3)	(4)
Female indicator	–0.090***	–0.091***	–0.047***	–0.027**
	(0.010)	(0.010)	(0.009)	(0.008)
Basic ICT scores	0.057***	0.052***	0.051***	0.050***
	(0.008)	(0.008)	(0.008)	(0.008)
Female x Basic ICT scores	–0.040***	–0.039***	–0.035***	–0.033***
	(0.008)	(0.008)	(0.008)	(0.008)
Indicator for individuals with a STEM degree			0.146***	0.156***
			(0.014)	(0.018)
Female x STEM degree				–0.053
				(0.031)
Working in private sector				0.025**
				(0.009)
Female x Working in private sector				
Working full time				0.033***
				(0.008)
Female x Working full time				
Constant	0.126***	–0.255***	–0.281***	–0.315***
	(0.008)	(0.071)	(0.073)	(0.075)
Other covariates				
Age and educational factors	N	Y	Y	Y
Immigration status	N	N	Y	Y
Residential location	N	N	Y	Y

Note: The linear probability model is used to obtain these estimates. Estimated coefficients are an equally weighted average of coefficients derived from each plausible value. All 80 replicate weights are used to account for the sample stratification. Estimated coefficients are interpreted as the percentage of being employed in the ICTC's classification of ICT occupations. Standard errors are given in parentheses.

*p < 0.5
**p < .01
***p < .001

TABLE B3 Probability of being employed in ICTC's 25 ICT occupations for the sub-sample of full-time employed individuals

	(1)	(2)	(3)
Female indicator	-0.093^{***}	-0.047^{***}	-0.044^{***}
	(0.011)	(0.010)	(0.011)
Basic ICT scores	0.038^{***}	0.040^{***}	0.040^{***}
	(0.007)	(0.007)	(0.007)
Indicator for individuals with a STEM degree		0.156^{***}	0.153^{***}
		(0.015)	(0.015)
Working in the private sector			0.023^{*}
			(0.011)
Constant	-0.241^{**}	-0.262^{**}	-0.291^{**}
	(0.087)	(0.089)	(0.091)
Controlling for other covariates			
Age and its square, educational variables	Y	Y	Y
Residential location	N	Y	Y

Note: The linear probability model is used to obtain these estimates. Estimated coefficients are an equally weighted average of coefficients derived from each plausible value. All 80 replicate weights are used to account for the sample stratification. Estimated coefficients are interpreted as the percentage of being employed in the ICTC's classification of ICT occupations. Standard errors are given in parentheses.

*p < 0.5

**p < .01

***p < .001

TABLE B4 Differences in basic ICT scores of women and men

	(1)	(2)	(3)	(4)
Female indicator	–0.053	–0.055	–0.036	–0.030
	(0.037)	(0.040)	(0.041)	(0.041)
Age	0.009	0.017	0.009	0.010
	(0.016)	(0.014)	(0.014)	(0.014)
Age	–0.000*	–0.000*	–0.000*	–0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Highest educational attainment (reference group = graduates with high-school diploma)				
Below high-school diploma		–0.723***	–0.717***	–0.720***
		(0.095)	(0.095)	(0.094)
Postsecondary below a bachelor’s degree		0.137**	0.148**	0.149**
		(0.050)	(0.050)	(0.050)
Bachelor’s degree		0.583***	0.572***	0.565***
		(0.061)	(0.061)	(0.061)
Postsecondary above a bachelor’s degree		0.700***	0.710***	0.705***
		(0.058)	(0.058)	(0.058)
Indicator for individuals who obtained education abroad		–0.289**	–0.262**	–0.262**
		(0.088)	(0.088)	(0.088)
Indicator for individuals whose location of education is unknown		0.057	0.107	0.096
		(0.268)	(0.280)	(0.296)
Indicator for individuals with a STEM degree		0.158**	0.086	0.050
		(0.054)	(0.052)	(0.051)
Female x STEM degree		0.049	0.091	0.111
		(0.085)	(0.085)	(0.087)
Working in ICTC’s ICT occupations			0.429***	
			(0.068)	
Female x ICTC’s ICT occupations			–0.156	
			(0.113)	
Working in ICT occupation (broad)				0.416***
				(0.060)
Female x ICT occupations				–0.181
				(0.125)
Constant	0.346	0.064	0.166	0.149
	(0.327)	(0.303)	(0.306)	(0.304)
Other control variables				
Age and its square	Y	Y	Y	Y
Education variables	N	Y	Y	Y
Residential locations	N	Y	Y	Y

Note: Ordinary linear regression is used to obtain these estimates. Estimated coefficients are an equally weighted average of coefficients derived from each plausible value. All 80 replicate weights are used to account for the sample stratification. Estimated coefficients are interpreted in terms of a standard deviation. Standard errors are given in parentheses.

*p < 0.5

**p < .01

***p < .001

Table B5 Differences between women and men in daily ICT usage at work

	The probability of self-reported daily usage at work		
	Excel	Word	Coding/ programming
All individuals in the sample			
Female	–0.0595**	0.0712***	–0.0450***
	(0.0190)	(0.0189)	(0.0092)
Individuals in ICTC's ICT occupations			
Female	0.0010	0.0859	–0.0313
	(0.0689)	(0.0564)	(0.0683)
Individuals in ICT occupations			
Female	0.0026	0.0637	–0.0318
	(0.0505)	(0.0557)	(0.0560)

Note: The linear probability model is used to obtain these estimated coefficients. These are weighted differences between women and men in the probability of self-reported daily usage of certain ICT applications at work. *Excel* indicates use of applications that are similar to Microsoft Excel. *Word* means use of word-processing applications similar to Microsoft Word. *Coding/programming* indicates use of applications that allow respondents to perform coding at work. In all regressions, standardized PS-TRE scores and levels of education are accounted for. The reference group is men with a bachelor's degree as the highest educational attainment. Standard errors are given in parentheses.

*p < 0.5
**p < .01
***p < .001

Table B6 Differences between women and men in self-reported level of ICT skills used at work

	Probability of self-reported level of ICT skills used at work		
	Complex	Moderate	Straightforward
All individuals in the sample			
Female	–0.0806***	0.0778***	0.0028
	(0.0117)	(0.0175)	(0.0157)
Individuals in ICTC's ICT occupations			
Female	–0.0811	0.0909	–0.0097
	(0.0699)	(0.0644)	(0.0272)
Individuals in ICT occupations			
Female	–0.0751	0.0928	–0.0177
	(0.0639)	(0.0624)	(0.0252)

Note: The linear probability model is used to obtain these estimated coefficients. These are weighted differences between women and men in the probability of self-reported level of ICT skills used at work. *Complex* indicates individuals who use complex computer skills at work. *Moderate* indicates individuals who use moderate-level skills at work. *Straightforward* requires individuals to have basic computer skills to perform their work. In all regressions, standardized PS-TRE scores and levels of education are accounted for. The reference group is men with a bachelor's degree as the highest educational attainment. Standard errors are given in parentheses.

*p < 0.5
**p < .01
***p < .001

TABLE B7 Differences between men and women in self-perceived computer skills and job performance

	Probability of individuals self-reporting	
	Enough skills to perform current job	Lack of computer skills hurting chance of promotion, a pay raise, or being hired
In all occupations		
Female	0.0068	0.0106
	(0.0066)	(0.0089)
In ICTC's ICT occupations		
Female	0.0161	0.0311
	(0.0119)	(0.0473)
In ICT occupations		
Female	0.0195*	0.0344
	(0.0090)	(0.0384)

Note: The linear probability model is used to obtain these estimated coefficients. These are weighted differences in how women perceive their computer skills, and how these skills influence their job performance at their current employment. The first column indicates individuals who self-reported having enough skills to perform well at work. The second column indicates individuals who think that lack of computer skills hurts their chance of promotion, a pay raise, and being hired at the current job. In all regressions, standardized PS-TRE scores and levels of education are accounted for. The reference group is men with a bachelor's degree as the highest educational attainment. Standard errors given are in parentheses.
*p < 0.5
**p < .01
***p < .001.

References

Information and Communications Technology Council. (2016). *A National Strategy To Develop Canada's Talent In a Global Digital Economy*. https://www.ictc-ctic.ca/

Information and Communications Technology Council. (2017a). *Digital economy annual review 2017*. Retrieved from https://www.ictc-ctic.ca/wp-content/uploads/2018/02/ICTC-Annual-Review-2017-EN.pdf

Information and Communications Technology Council. (2017b). *The next talent wave: Navigating the digital shift – outlook 2021*. Retrieved from https://www.ictc-ctic.ca/wp-content/uploads/2017/04/ICTC_Outlook-2021.pdf

Kindsiko, E., & Türk, K. (2016). Detecting Major Misconceptions about Employment in ICT: A Study of the Myths about ICT Work among Females. World Academy of Science, Engineering and Technology, *International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering, 11*(1), 107-114.

Mueller, R. E., Truong, K., & Smoke, W. (2018). The underrepresentation of women in Canada's ICT sector: What can we learn from a Canadian survey of adult skills? *Canadian Public Policy*, *44*(S1), S73–S90. doi:10.3138/cpp.2017-073.

Organisation for Economic Co-operation and Development. (2016). *Skills for a digital world* (OECD Digital Economy Papers No. 250). Paris: OECD Publishing.

Tech Toronto. (2016). *How technology is changing Toronto employment: 400,000 jobs and growing*. Retrieved from https://www.techtoronto.org/Report2016.

PISA, PCAP, and PEI: How Canada's Smallest Province Employed International and National Data to Transform Public Education

Alexander MacDonald
Vice President and Chief Learning Officer, Holland College
(former Deputy Minister, Prince Edward Island Department of Education and Early Childhood Development)

Abstract

In 2003, Program for International Student Assessment (PISA) results for Prince Edward Island (PEI) indicated that students in the province performed well below the Canadian average in reading, mathematics, and science. Concerned about the implications of these results and lacking any provincial data, the PEI government commissioned a review of the Island's public education system. Among the most controversial recommendations of the Kurial report (2005) were establishing common assessments at Grades 3 and 9, lifting a ban on closing small schools, and putting an emphasis on students' learning to read by the end of Grade 3.

The 2007 Pan-Canadian Assessment Program (PCAP) results indicated that PEI students continued to struggle. Alarmingly, PEI students fell well below all other provinces in reading. The initial results of the Provincial Common Assessments were unfortunately similar; a significant number of Island students were reading well below grade level.

By 2015–16, PEI's results on PISA, PCAP, and Provincial Common Assessments had improved significantly. Island students were among the top performing in the country. These developments were the result of a process whereby data from national and international tests were analyzed together with data from locally administered common assessments, and the findings used to drive operational and structural initiatives funded by additional educational resources. This resulted in systemic improvement in PEI's public education system.

Key to this change was the unique manner in which Provincial Common Assessments were introduced. Classroom teachers were directly involved in the development and implementation of the assessment tools and in the analysis of assessment results. This addressed much of the resistance individual teachers and teacher organizations had to standardized measurement. Teacher involvement in the development and implementation of the professional development initiatives emerging from the data was also important. These initiatives were focused on improving teachers' skills in curriculum organization, formative and summative assessment, and high-yield teaching strategies. They were integrated into the instructional calendar and teachers' communities of practice and carried over year to year.

Key terms: professional development, provincial assessment, standardized testing, system alignment, teacher skills, system reform

Introduction

Prince Edward Island (PEI) is Canada's smallest province, with a population of about 140,000. Its public school system serves approximately 19,000 students, 95 per cent of whom are enrolled in English schools and the remaining 5 per cent in the province's small French-language school board. There are 56 English and 6 French schools of varying sizes and grade configurations. Student numbers were bolstered in 2010 by the introduction of full-day kindergarten.

From 2004 to 2015, PEI's public school system underwent significant structural and operational change. Some small schools were closed, English school boards were eliminated, and the Department of Education assumed all responsibility for student assessments and teachers' professional development. Important among these changes was the introduction of Provincial Common Assessments in 2007. These changes were, in part, a result of public pressure driven by Program for International Student Assessment (PISA) and Pan-Canadian Assessment Program (PCAP) results and a lack of reliable information on Island student achievement.

The purpose of this paper is to outline how PEI leveraged PISA and PCAP data to drive educational change. More specifically, the paper explores the importance of teachers' engagement and collaboration in all areas of education reform and the central role of their ongoing skill development with regard to formative and summative assessment, the organization (scope and sequence) of curricular outcomes, the role of foundational outcomes, and the use of high-yield teaching strategies. Further, the importance of aligning educational resources and focusing on the continuous improvement of student achievement are discussed.

Theoretical framework

The Kurial report, *Excellence in education: A challenge for Prince Edward Island*, provided clear direction on where PEI's education reform initiative should focus its efforts (Kurial, 2005). Like many reform initiatives, this one began with a set of assumptions drawn from the research literature (Earl, Ali, & Lee, 2005). Chief among them was the central role of the classroom teacher in improving student achievement. Moreover, the role of classroom instruction as the most important school-based factor in improving student achievement helped guide the reform process (Fullan, Hill, & Crévola, 2006).

But what aspects of classroom instruction were most promising in terms of system improvement? Much of the literature indicated that focused instruction, consistent feedback to students, and the continuous refinement of instruction were all crucial to successful teaching (Fullan, Hill, & Crévola, 2006). When combined with instructional leadership by the school administration that focused on improving student outcomes, results were maximized (Bryk, Bender-Sebring, Allensworth, Luppescu, & Easton, 2010).

Following the examples of other provinces, notably Alberta and Ontario, PEI also focused on two principles: that all children can learn; and that emphasizing a few core priorities is an effective way to proceed with reform (Fullan, 2011). Consequently, it was decided that improving the reading, writing, and numeracy skills of PEI students–and ensuring that as many students as possible were reading at grade level by the end of Grade 3—was a good place to start (Kurial, 2005).

In focusing on the skill development of teachers and the instructional leadership abilities of principals, PEI was forced to confront the issue of accountability, always a sensitive topic in public education. Ideally, remedial interventions would be targeted and non-punitive (Hargreaves & Shirley, 2009). Common sense also dictated that if teachers and administrators were to work collaboratively with school board and Department staff, a culture of collective responsibility needed to be created (Organisation for Economic Co-operation and Development [OECD], 2010).

Focusing on classroom instruction was key to this reform initiative, but teacher leadership has also proven to be a powerful force for educational change (Lieberman & Miller, 2004). Consequently, the Eastern School District surveyed Grade 1–9 teachers in 2005 to ask how it could best support teachers in their daily classroom efforts. Two themes emerged. Many teachers indicated that they wanted standards and benchmarks, along with some consistent way to measure student achievement. They also indicated that they did not want standardized tests.

Concern about educators' sometimes visceral reaction to standardized tests, and recognition not only of the importance of building PEI teachers' capacity but also of the value and power of deep collaboration among all education stakeholders (Fullan, 2008, 2011), drove a decision to invite teachers to help lead the reform process. A major Curriculum Analysis project was implemented and tasked with examining curricular outcomes and creating common assessments in Grade 3 reading and writing and Grade 9 mathematics. Dozens of teachers volunteered to join the grade-level working groups. Marking boards were also established to help score the tests, analyze the data, and develop the classroom interventions that were at the centre of the reform process.

The remaining issue, which helped tie together the focus on instruction and the emphasis on teacher leadership and collaboration, was alignment. Certainly, effective resource alignment is directly linked to successful leadership and effective management of teaching and learning (Miles & Frank, 2008). A key aspect of the PEI initiative was the alignment of both structural and instructional resources aimed solely at improving student achievement.

Analysis

Four central themes emerged from a close analysis of this reform initiative.

Collaboration

Teacher engagement in the reform began with their involvement in the development, implementation, and analysis of the Island's common assessment measures, initially with Grade 3 reading and writing and Grade 9 math. Teachers first volunteered for a major curricular initiative involving Grade 3 literacy and Grade 9 math. They worked at standardizing the scope and sequence of curricular outcomes and the development of formative and summative assessment tools. They also participated in marking boards and helped to analyze the results. This resulted in various professional development initiatives, again created by and implemented through working groups of classroom teachers and board and Department staff.

This work was essential in that the collaboration helped mute the typical resistance of teacher organizations to the introduction of standardized tests. Given that the introduction of common assessments prompted a renewed emphasis on covering curricula, its non-punitive nature encouraged teachers and principals to open up about the challenges they faced. More specifically, many teachers indicated that they struggled with organizing the curricula, tracking students' performance, and adjusting their teaching methods. It was also not uncommon to hear teachers complain they could not cover the grade-level content over the academic year. This collaboration and resulting openness was vital to the analysis of the common test results and ensuing development of remedial initiatives.

Focus

Starting with the assumption that the classroom teacher is key to student achievement, and that it is effective to concentrate on a small number of objectives, PEI was able to focus its reform efforts on improving the instructional skills of teachers. The province initially emphasized Grade 3 literacy, quickly seeing improvement in reading comprehension but less in writing. Later, the emphasis shifted to intermediate math, and a province-wide initiative was implemented to focus on curricular outcomes, assessment, and high-yield teaching strategies.

There was significant variation across the province in terms of the scope and sequence of curricular outcomes, regardless of the subject area. Much of the initial professional development explored effective ways to organize curricular outcomes and to support the central role of foundational outcomes in teaching and assessment processes. At the same time, much work was done in the areas of formative and summative assessment. Some teachers expressed frustration about their lack of training and ongoing professional support in understanding assessment, and in developing and implementing assessment tools to help drive instruction. Many teachers had gradually developed a personal approach to or style of teaching that, in the absence of formative and summative data, was less than ideal. Much effort was expended on the study of high-yield teaching strategies and their integration with formative and summative data. This single-minded focus on the interaction of outcomes, teaching strategies, and assessment proved very helpful.

Alignment

In the early stages of the initiative, a lack of alignment across the education system quickly became apparent. Responsibility for teachers' professional development was shared by the teacher association, the various boards, and the Department. Both the Department and the boards had strategic plans but no mechanism to integrate them. Each school had strategic goals that were not directly tied to board or Department plans. Simply put, nothing was in place to connect work in a classroom to the Department's objectives.

With the introduction of common assessment data, students' academic performance could be analyzed province wide. Professional development was centralized in the Department. Five days per year were set aside for this work, which was integrated with the work of the curriculum committees during the assessment process. School literacy goals were established, and 90-minute literacy blocks were initiated in each school. The two English school boards were amalgamated to form a single English language school board which was later replaced by a three-person committee.

Other efforts to ensure alignment included the introduction of new standards for principal and teacher certification and a new principal training program. Common assessments were also expanded to include math at Grades 3, 6, and 11, and literacy at Grade 6, 9, and high school. An innovative graduation planner project aimed at Grade 9–12 students was also introduced.

As the goal of the initiative was to create a culture of continuous improvement in student achievement, it was decided that this could be accomplished only through both structural and instructional alignment of PEI's educational resources.

Resources

New and redirected resources also played a role in the initiative. Instead of taking the traditional remedial approach of simply adding classroom teachers to reduce class sizes, the province used test data to determine how to provide new human resources in a more targeted and efficient manner. Initially, literacy coaches, new K–9 literacy resources in English, and updated curriculum guides in French were part of the new spending. Math mentors and an instructional development team focused on math achievement were added later. In total, about 25 new literacy and numeracy coaches were employed. New money was also set aside to expand the Provincial Common Assessment, and 10 high-school department head positions were reconfigured to focus on literacy and numeracy instruction. This dynamic proved very effective, as it was created through a deep collaboration among stakeholders who were focused on improving classroom instruction, and supported by new targeted resources in a system aligned to achieve maximum efficiency.

Results

The past 15 years have seen improvements in the Provincial Common Assessment, PCAP, and PISA results. In the 2007 Provincial Common Assessment in reading comprehension, approximately 62 per cent of Grade 3 students in PEI were meeting expectations. By 2013, the proportion had grown to 87 per cent. Over the same period, the percentage of students experiencing difficulty dropped from 27 per cent to 9 per cent (Table 1).

Table 1 Provincial Common Assessment results: Grade 3 English reading comprehension (%)

	2007	2008	2009	2010	2011	2012	2013
Meets expectations	62	72	85	82	85	87	87
Approaches expectations	11	8	5	6	5	5	4
Experiences difficulty	27	20	10	12	10	8	9

Source: Government of PEI (2014). Student Achievement: Presentation to Greater Charlottetown Chamber of Commerce. Department of Education and Early Childhood Development. Charlottetown, PEI.

The results of the common math assessments also improved. In 2011, a revamped Grade 9 math assessment was implemented. Results indicated that approximately 31 per cent of students met expectations, while about 64 per cent were experiencing difficulty. After two years of an intensive and focused professional development initiative that supported teachers with 10 days of training, results improved markedly. By 2013, 64 per cent of students were meeting expectations, while 32 per cent experienced difficulty. By 2014–15, 71 per cent of Grade 9 math students were at or approaching expectations (Table 2).

Table 2 **Provincial Common Assessment results: Grade 9 math (%)**

	2011	2012	2013	2014	2015
Meets expectations	31	52	64	65	67
Approaches expectations	5	8	4	6	7
Experiences difficulty	64	40	32	29	6

Source: Government of PEI (2016) p. 17. *Department of Education and Early Childhood Development, Annual Report*. Charlottetown, PEI

PCAP results at Grade 8 were similar (Table 3). In 2007, PEI's performance in reading was lowest in Canada. Similar results were realized on the 2010 assessment, but by 2013 and 2016, reading, math, and science scores had improved. In 2016, Island students received the highest average score in reading in Canada. A greater proportion of students in PEI reached the expected level of performance in reading than in any other province, and they showed the greatest provincial improvement in reading from 2010 to 2016. In math, PEI students achieved at the Canadian average by 2016 and close to double the improvement of any other province from 2010 to 2016. In science, PEI students were also at the Canadian average in 2016 and achieved a significantly higher mean score than in 2013.

Table 3 **PCAP: PEI results in reading, math, science by mean score**

	2007	2010	2013	2016
Reading	471	481	494	513
Math		460	492	503
Science			491	516

Source: O'Grady, Fung, Servage, & Khan (2018) pp. 170,174,178.
Note: In PCAP, the major domains were reading in 2007 and 2016, mathematics in 2010, and science in 2013.

Similar to PCAP, PEI's early results on the PISA assessments were poor, often well below the Canadian average (Table 4). On the PISA 2015 assessment, however, Island 15-year-olds performed at the Canadian average in reading for the first time, a statistically significant improvement. Island students also demonstrated a significantly improved performance in mathematics, from below the OECD average in 2012 to reaching the OECD average in 2015.

Table 4 **PISA: PEI results in reading, math, science by mean score**

	2003	2006	2009	2012	2015
Reading	495*	497*	486*	490*	515
Science	500*	509*	495*	490*	515
Math	500*	501*	487*	479*	499

Source: O'Grady, Deussing, Scerbina, Fung, & Muhe (2016) pp. 79–81.
*denotes the lowest mean score among participating provinces.

Discussion and conclusions

What can be concluded from the PEI initiative? Clearly the introduction of Provincial Common Assessments proved a turning point in the education reform process. Previously, PISA and PCAP results had been dismissed by many stakeholders as merely "snapshots" rather than true reflections of PEI's public school system. That the initial results of the common assessments were consistent with both international (PISA) and national (PCAP) testing brought home to many Island educators that the system needed to improve.

These data also gave stakeholders a focus for their common efforts. The involvement of classroom teachers in the development and implementation of the Provincial Common Assessments helped deepen teachers' ownership of the reform process, as did voluntary participation on marking boards and their work in analyzing test results and developing remedial interventions.

Certainly the initiative changed the discourse about public education. Discussion of topics such as foundational outcomes, formative and summative assessment, high-yield teaching strategies, and the scope and sequence of curricular outcomes was uncommon in 2004. That is a sweeping statement to be sure, but PEI's professional development initiatives—which were informed by the initial results of the common assessments—clearly indicated that a focus on teaching strategies, curricular outcomes, and assessment tools and techniques was warranted. In turn, this resulted in a broadening and deepening of the public discourse about the most effective ways to improve student achievement.

Not only the content but also the process of professional development initiatives proved important. These activities had a single focus: they were designed to enhance teachers' instructional skills. They were collaborative, in that classroom teachers led remedial efforts in conjunction with staff from the boards and the Department of Education, and the interventions were facilitated by communities of practice spread across the system in the various grades and subject areas. They were also incorporated into the system; three new professional development days were added to the existing two and embedded in the teaching calendar. The ensuing continuity and predictability supported the collaborative nature of the work, as groups of educators could plan initiatives to run throughout the year and inform ongoing instruction. Lastly, the professional development initiatives were data driven. PISA, PCAP, and common assessment data were gleaned to determine where students were struggling, and effective remedial strategies were then integrated into teachers' professional development.

The common assessment data also highlighted the importance of structural change. The two English-language school boards were amalgamated to improve system efficiency and effectiveness and ultimately replaced by a three-person committee. Teacher and principal standards were improved. A 90-minute literacy period was introduced across the province, and kindergarten was introduced into the public school system. New resources were introduced. Curricular guides were purchased, along with reading, writing, science, and math support materials. The introduction of national and Provincial Common Assessments helped bring about a key change in spending on human resources. As Table 5 indicates, as enrolment dropped and teaching positions were removed through attrition average class size remained stable. At the same time, other resources were added in a more targeted and efficient manner.

Table 5 Student–educator ratio, 2010–15

Year	Enrolment	Full-time educators	Student–educator ratio
2010–11	21,169	1,670	12.7
2011–12	20,831	1,670	12.5
2012–13	20,406	1,634	12.5
2013–14	20,131	1,596	12.6
2014–15	19,938	1,565	12.7

Source: Government of PEI: Department of Education and Early Childhood Development. Annual Report 2014–2015. Charlottetown, PEI, p. 23.

Taken together, the new resources and structural and operational changes described above created a dynamic that resulted in the improvement of student achievement across various levels of PEI's public education system. Although much work remains, this initiative provides one example of the initial stages of a successful school reform process.

References

Bryk, A., Bender-Sebring, P., Allensworth, E., Luppescu, S., & Easton, J. (2010). *Organizing schools for improvement.* Chicago, IL: University of Chicago Press.

Earl, L., Ali, A.S., & Lee, L. (2005). Evaluating school improvement in Canada: A case example. *Canadian Journal of Program Evaluation*, *20*(3), 151–174.

Fullan, M. (2008, January). *From school effectiveness to system improvement: An inevitable conceptual evolution.* Paper prepared for *Journal für Schulentwicklung*. Retrieved from https://michaelfullan.ca/wp-content/uploads/2016/06/Untitled_Document_25.pdf

Fullan, M. (2011). *Choosing the wrong drivers for whole system reform* (Seminar Series No. 204). Centre for Strategic Education, Melbourne, Australia.

Fullan, M., Hill, P., & Crévola, C. (2006). *Breakthrough*. Thousand Oaks, CA: Corwin.

Government of PEI. (2014). *Student achievement: Presentation to the Greater Charlottetown Chamber of Commerce*. Summerside, PE: Department of Education and Early Childhood Development.

Government of PEI. (2016). *Annual report: Education and Early Childhood Development, 2014–2015*. Summerside, PE: Department of Education and Early Childhood Development. Retrieved from https://www.princeedwardisland.ca/sites/default/files/publications/eelc_annual_report_2014-15_2.pdf

Hargreaves, A., & Shirley, D. (2009). *The fourth way: The inspiring future for educational change*. Thousand Oaks, CA: Corwin.

Kurial, R. (2005). *Excellence in education: A challenge for PEI*. Charlottetown: Government of Prince Edward Island.

Lieberman, A., & Miller, L. (2004). *Teacher leadership*. San Francisco, CA: Jossey-Bass.

Miles, K.H., & Frank, S. (2008). *The strategic school: Making the most of people, time, and money*. Thousand Oaks, CA: Corwin.

O'Grady, K., Deussing, M.A., Scerbina, T., Fung, K., & Muhe, N. (2016). *Measuring up: Canadian results of the OECD PISA study*. Toronto: Council of Ministers of Education, Canada. Retrieved from https://www.cmec.ca/publications/lists/publications/attachments/365/pisa2015-cdnreport-en.pdf

O'Grady, K., Fung, K., Servage, L., & Khan, G. (2018). *PCAP 2016: Report on the pan-Canadian assessment of reading, mathematics, and science*. Toronto: Council of Ministers of Education, Canada. Retrieved from https://www.cmec.ca/Publications/Lists/Publications/Attachments/381/PCAP-2016-Public-Report-EN.pdf

Organisation for Economic Co-operation and Development. (2010). *Strong performers and successful reformers in education*. Paris: OECD Publishing. doi:10.1787/9789264096660-en.

Exploring the Effects of Inquiry-Based Strategies and Teacher-Directed Instruction on Student Performance in Science, Using PISA Data

Alfred Sakyi
Senior Manager Corporate Research, Alberta Education

Abstract

The study was conducted to determine how inquiry-based strategies and teacher-directed instruction contribute to Alberta students' high performance in science, using the Programme for International Student Assessment (PISA) 2015 Canadian data set. The analyses include descriptive statistics and regression analysis. The results show that Alberta teachers use both inquiry-based teaching practices and teacher-directed instruction in their science classes. Both "inquiry" and "teacher-directed" constructs, as defined by PISA, had a positive effect on students' self-efficacy and enjoyment of science, as did "teacher support" and "adaptive instruction." In addition, teacher-directed instruction had a direct positive effect on classroom discipline and science scores. The results show that despite the reported importance of inquiry-based strategies, teacher direction is very important for the effectiveness of any inquiry strategies. If teachers effectively incorporate teacher-directed strategies into science lessons, as well as strategies that improve students' sense of self-efficacy and enjoyment of science, students' science results can improve.

Introduction and theoretical framework

The purpose of this study was to determine whether and to what extent the use of inquiry-based strategies and teacher-directed instruction contributes to Alberta students' high performance in science. The study was guided by the following research questions:

- To what extent do Alberta teachers use inquiry-based teaching and learning practices and teacher-directed instruction in science lessons?
- To what degree does emphasis on inquiry learning and/or directed instruction affect Alberta students' performance in science?

The study is based on the following assumptions, as presented in the *PISA 2015 assessment and analytical framework* (Organisation for Economic Co-operation and Development [OECD], 2017):

- A student's learning outcome can be predicted by the pattern of teaching practices to which he or she is exposed. Some patterns are related to high student performance, while others are related to high student interest and motivation.
- A comparison of teaching patterns allows for detailed analysis of students' science performance across countries and of their interest in science topics.
- The teaching-practices items used in the Programme for International Student Assessment (PISA) 2015 were developed and chosen to discriminate between different patterns of teaching.

This study focuses on two teaching practices: inquiry-based and teacher-directed science teaching. A body of research suggests that inquiry-based teaching practices can motivate students to learn and to advance their problem-solving and critical-thinking skills, provide them with opportunities to gain a deeper understanding of the concepts, and improve their engagement with the cognitive dimensions of the subject and their attitudes toward it (Blanchard, Southerland, Osborne, Sampson, Annetta, & Granger, 2010; Furtak, Seidel, Iverson, & Briggs, 2012; Hattie, 2009; Minner, Levy, & Century, 2010; Oliver, 2007; Prince & Felder, 2007). Teacher-directed instruction provides a well-structured, clear, and informative lesson that usually includes teacher explanations, classroom debates, and student questions. This teaching approach often includes modelling a variety of examples and guiding students during their review and practice (OECD, 2016). Driver (1995) asserts that even if these strategies render students passive during class, some teacher direction is essential if students are expected to acquire generally accepted science knowledge.

Methodology

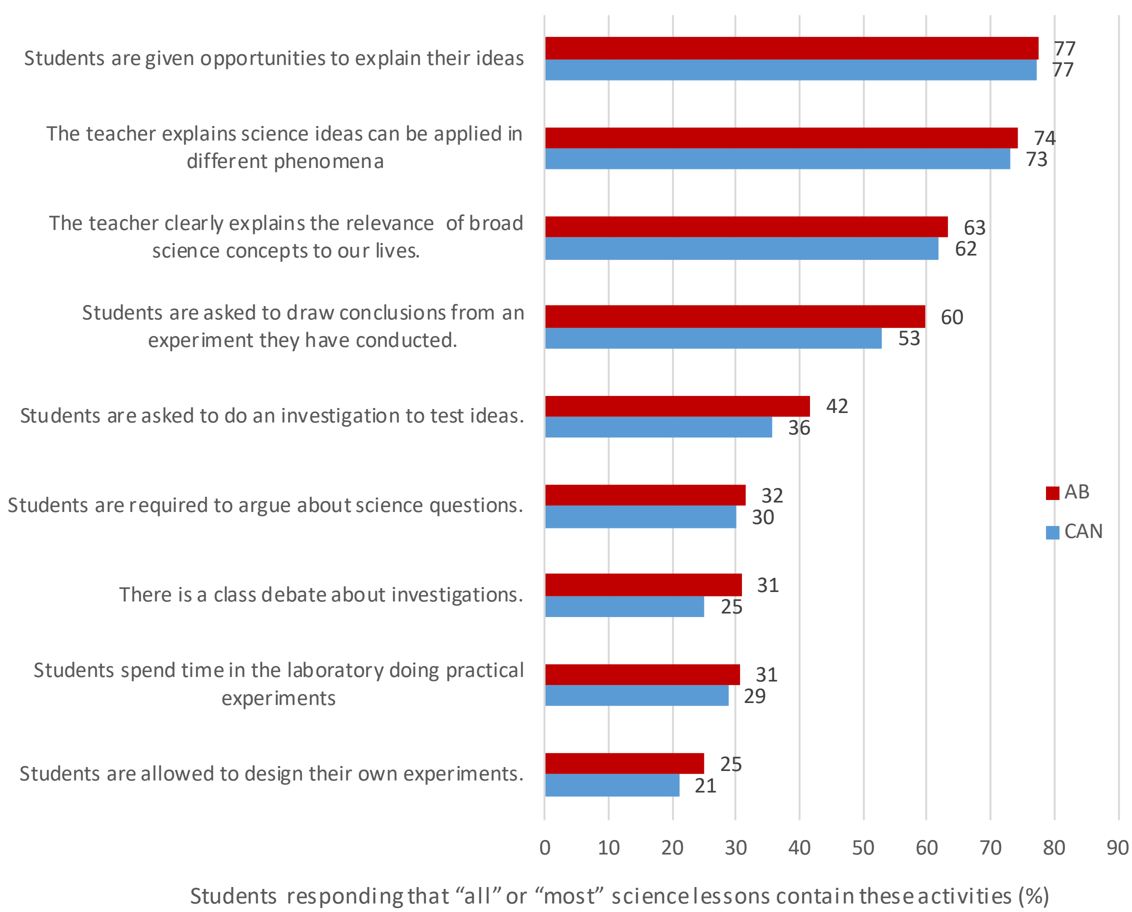
The Canadian data set from PISA 2015 was used for the analysis. Approximately 18,000 students across Canada, including 1,900 from Alberta, participated. Both teaching and learning variables were included: inquiry-based science teaching and learning practices, teacher-directed science instruction, teacher support in science classes, adaptive instruction, and disciplinary climate in science classes. Student characteristics were also considered: gender, socioeconomic status, science self-efficacy, and enjoyment of science. See Appendix A for details on the variables. Descriptive statistics (distribution of student responses and average scale scores) were used to determine the degree of emphasis on inquiry-based learning and teacher-directed instruction. Path analysis was used to determine the relationship between student characteristics and other derived teaching and learning variables and student performance in science.

Results

Inquiry-based science teaching and learning practices

To measure inquiry-based science teaching and learning practices, PISA used students' responses to a set of nine statements. For each statement, students indicated whether the activity described occurred in all lessons, in most lessons, in some lessons, or never/hardly ever. Figure 1 shows the list of statements and the proportion of Alberta students who indicated that the activities occurred in all or most of their science lessons. The results show that a slightly higher proportion of Alberta students than the national average participated frequently in inquiry-based science teaching and learning activities.

FIGURE 1 Inquiry-based science teaching and learning practices

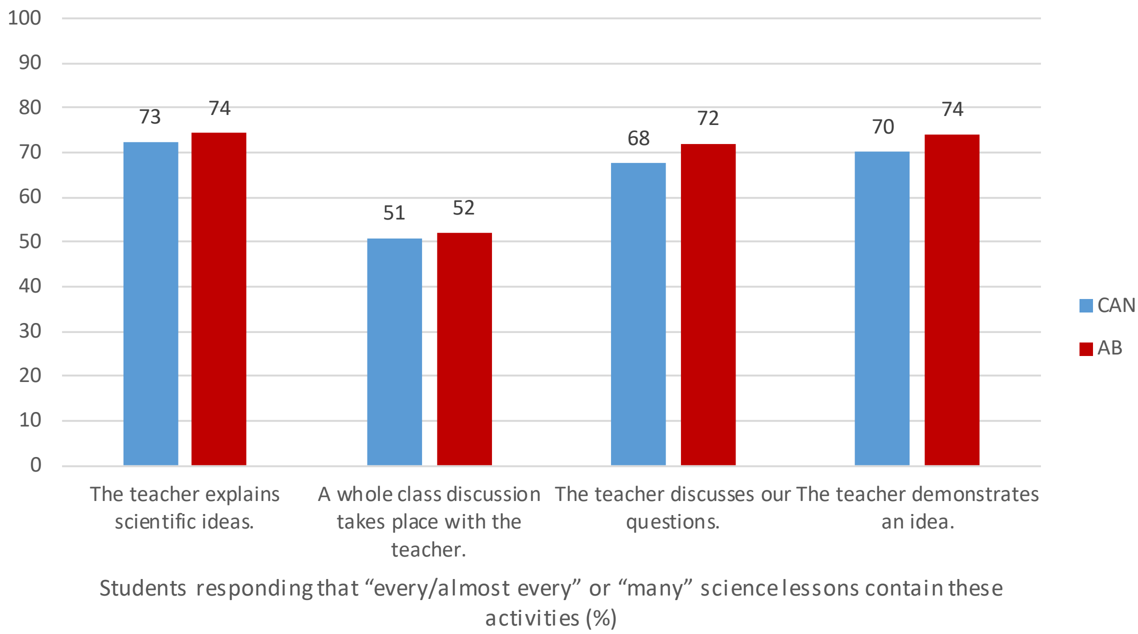


Source: OECD PISA 2015 Database (http://www.oecd.org/pisa/data/2015database)

Teacher-directed science instruction

PISA used a set of four items to determine teacher-directed science instruction. The statements focus on what the teacher does in the science class. For each action, students were asked to indicate whether it occurred in many lessons, every/almost every lesson, some lessons, or never/almost never. Figure 2 shows the list of statements and the proportion of students who indicated that those instructional practices happened in many lessons or every lesson. Similar to inquiry-based strategies, the proportion of Alberta students who indicated frequent exposure to teacher-directed science instruction practices was generally close to or slightly higher than the proportion of Canadian students overall.

FIGURE 2 Teacher-directed science instruction



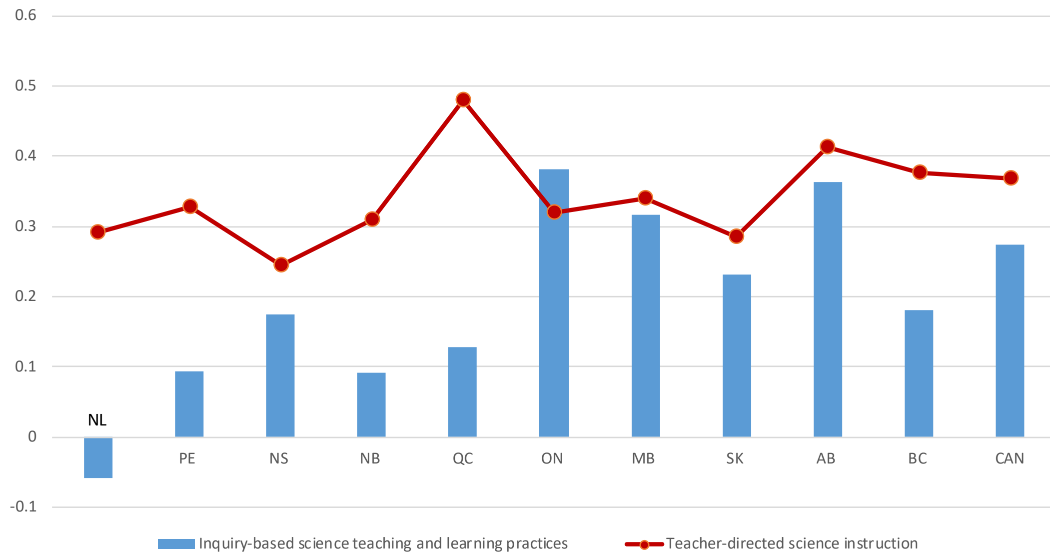
Source: OECD PISA 2015 Database (http://www.oecd.org/pisa/data/2015database)

Composite scale scores

Students' responses to the various questions were used to derive composite scale scores and then create indices for "inquiry-based science teaching and learning practices" and "teacher-directed science instruction." The computed scale scores for Canada achieved Cronbach's alpha reliability coefficients (Cronbach, 1951) of 0.87 for inquiry-based science teaching and learning practices and for teacher-directed science instruction, suggesting that the items have a relatively high internal consistency. (A reliability coefficient of 0.70 or higher is considered "acceptable" in most social science research contexts.) The scale scores were used in the regression analysis to determine their relationship to PISA science scores.

Figure 3 shows the composite scores and the gap between emphasis on science inquiry strategies and on teacher-directed science instruction for each of the 10 Canadian provinces. There appears to be a greater emphasis on teacher-directed instruction across the Canadian provinces except for Ontario. Alberta results showed that both strategies were widely used and there is a smaller difference in emphasis between them.

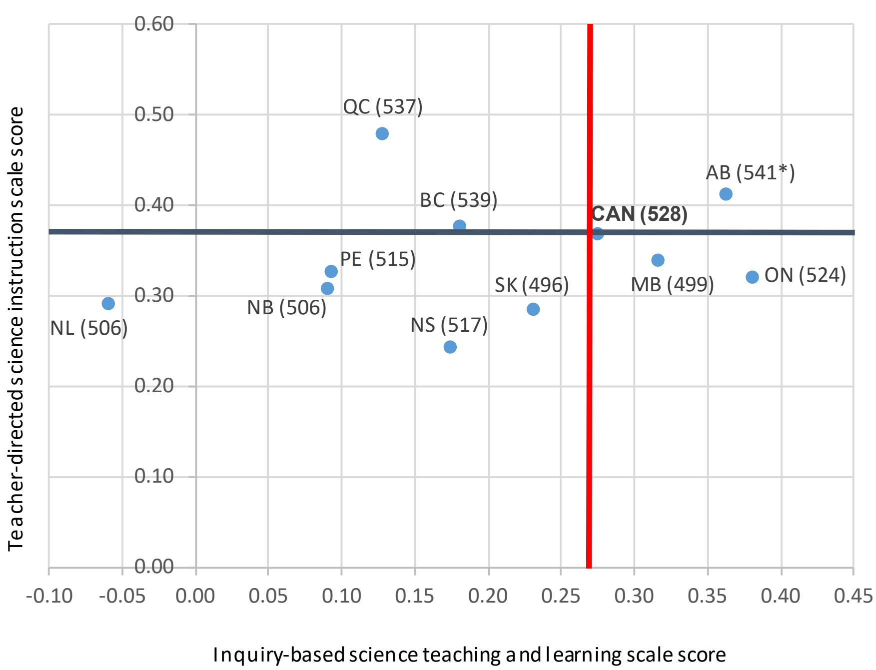
FIGURE 3 PISA 2015 average composite scores for science inquiry strategies and teacher-directed science instruction for Canadian provinces



Source: OECD PISA 2015 Database (http://www.oecd.org/pisa/data/2015database)

Figure 4 elaborates on the relative emphasis of inquiry-based and teacher-directed science instruction for Canadian provinces and their corresponding PISA science scores. It should be noted that Alberta is the only province that appears in the top right quadrant of higher emphasis than the Canadian average with respect to both instruction strategies. Another interesting observation is that provinces above the Canadian average on the PISA-defined "teacher-directed science instruction" scale have higher science scores than those below the average. However, some provinces below the Canadian average on the PISA-defined "inquiry" scale had higher science scores than those above the average.

FIGURE 4 Relative emphasis on PISA-defined teacher-directed and inquiry-based science teaching and learning



Source: OECD PISA 2015 Database (http://www.oecd.org/pisa/data/2015database)

Results on individual items

Individual items included in the constructs for inquiry and teacher-directed instruction were each tested for their relationship to science performance by comparing the average scores between two groups of students: those who observed or experienced each strategy frequently, and those who were less frequently exposed. Results indicated that the three specific inquiry-based science teaching and learning practices and three specific teacher-directed science instruction practices were associated with higher science scores. Five other inquiry-based strategies were associated with lower science scores (see Table 1).

Table 1 Inquiry-based science teaching and learning practices and teacher-directed science instruction associated with higher and lower PISA 2015 science scores, Alberta and Canada

Higher science scores*		Lower science scores**	
Activity	**INQ / TD**	**Activity**	**INQ / TD**
The teacher explains scientific ideas.	TD	Students spend time in the laboratory doing practical experiments.	INQ
The teacher discusses our questions.	TD	Students are allowed to design their own experiments.	INQ
The teacher demonstrates an idea.	TD	Students are required to argue about science questions.	INQ
The teacher explains how a science idea can be applied to different phenomena.	INQ	Students are asked to do an investigation to test ideas.	INQ
The teacher clearly explains the relevance of broad science concepts to our lives.	INQ	There is a class debate about investigations.	INQ
Students are given opportunities to explain their ideas.	INQ		

Source: OECD PISA 2015 Database (http://www.oecd.org/pisa/data/2015database)
Note: INQ = Inquiry-based teaching and learning practices; TD = Teacher-directed instruction.
*Average science scores for classes in which these practices occurred more often were higher than for classes in which these practices occurred less often.
**Average science scores for classes in which these practices occurred more often were lower than for classes in which these practices occurred less often.

Regression analysis

Path analysis was used to allow the simultaneous modelling of several related regression relationships in which a variable can be dependent in one relationship and independent in another. Two path analysis models were tested using Alberta data to determine the direct and indirect effects of inquiry-based science strategies and teacher-directed science instruction on science performance (Table 2). The first model has other teaching variables (classroom discipline, adaptive instruction, and teacher support) as mediating variables, while the second model has student-related mediating variables (enjoyment of science and science self-efficacy).

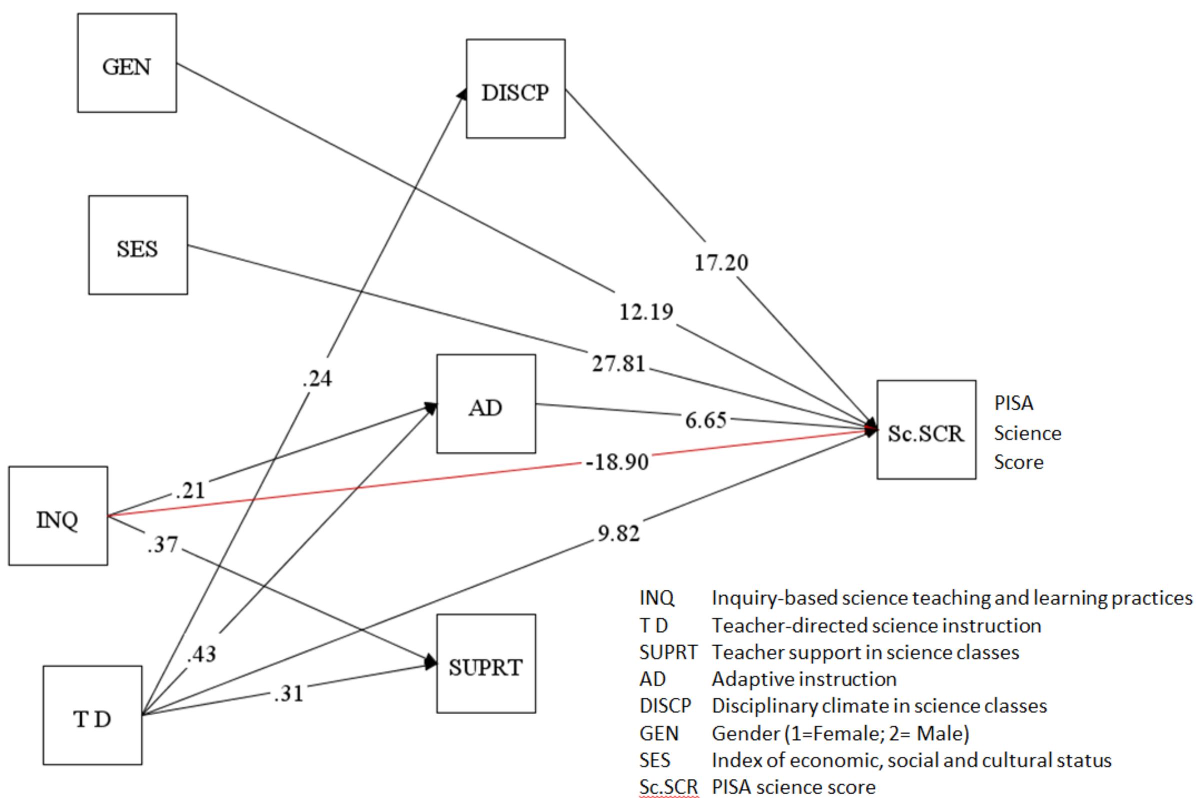
TABLE 2 Path analysis models and variables

Dependent variable	PISA 2015 science score	
	Model 1	Model 2
Predictors	• Inquiry-based* • Teacher-directed** • Socioeconomic status • Gender	• Inquiry-based* • Teacher-directed** • Socioeconomic status
Mediating variables	• Disciplinary climate in science classes • Adaptive instruction • Teacher support in science classes	• Enjoyment of science • Science self-efficacy

*Inquiry-based science teaching and learning practices
**Teacher-directed science instruction

Figures 5 and 6 are path diagrams showing the regression results of the two models that were tested.

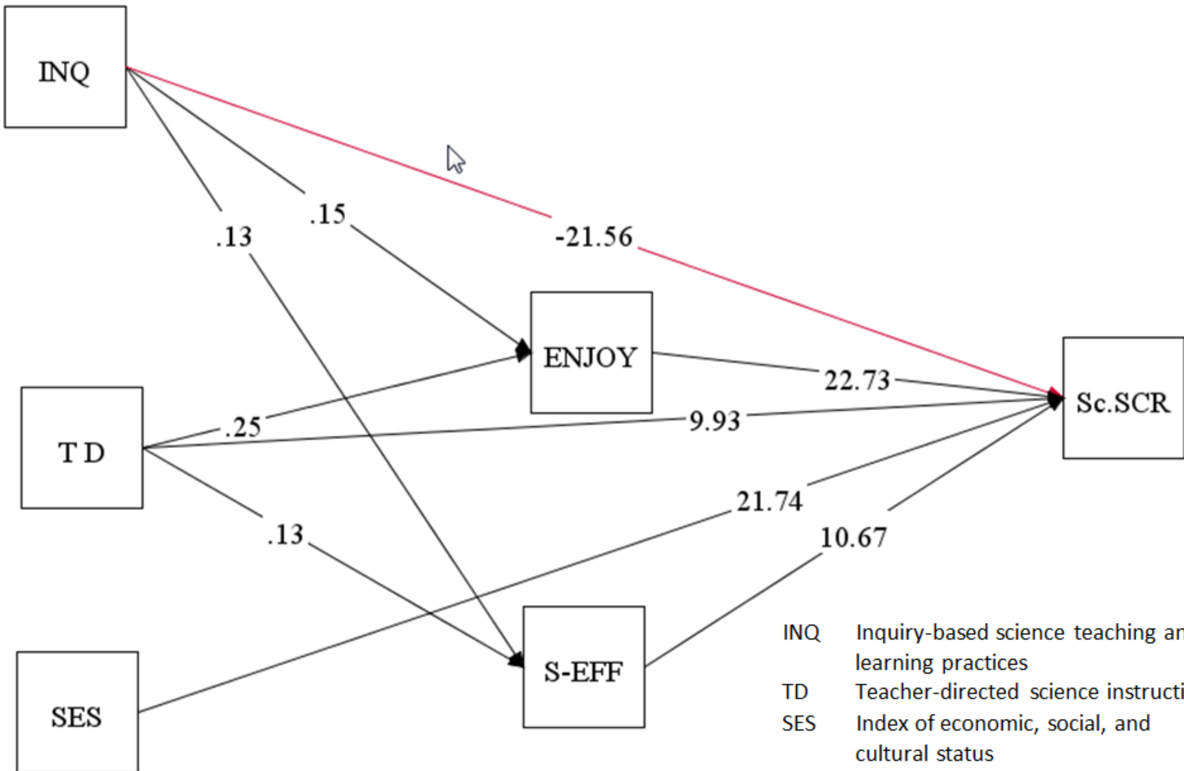
FIGURE 5 Path analysis results for model 1: Effects of inquiry-based and teacher-directed instruction on science scores with other teacher-related mediating variables



Source: OECD PISA 2015 Database (http://www.oecd.org/pisa/data/2015database)
Note: Path analysis results show only significant parameter estimates; standardized root mean square residual = 0.066.

The results for model 1 indicate that both inquiry and teacher-directed instruction are positively related to the PISA-defined variables of teacher support and adaptive instruction. In addition, teacher-directed instruction has a positive relationship with classroom discipline and science scores. Both classroom discipline and adaptive instruction are positively related to science scores, but teacher support does not have any direct significant relationship with the scores.

FIGURE 6 **Path analysis results for model 1: Effects of inquiry-based and teacher-directed instruction on science scores with other student-related mediating variables**



INQ	Inquiry-based science teaching and learning practices
TD	Teacher-directed science instruction
SES	Index of economic, social, and cultural status
S-EFF	Science self-efficacy
ENJOY	Enjoyment of science
Sc.SCR	PISA science score

Source: OECD PISA 2015 Database (http://www.oecd.org/pisa/data/2015database)

Model 2 results show both inquiry-based and teacher-directed constructs, as defined by PISA, have positive effects on students' self-efficacy and enjoyment of science, which in turn are positively related to PISA science scores. The effects of inquiry and teacher-directed instruction and the variables in the two regression models are summarized in Table 3.

TABLE 3 **Summary of regression showing the relationship between PISA-defined inquiry-based practices and teacher-directed science instruction and other teacher and student variables**

	Inquiry-based	Teacher-directed
Teacher support	+	+
Adaptive instruction	+	+
Classroom discipline	Not significant	+
Science self-efficacy	+	+
Enjoyment of science	+	+
Science score	–	+

Source: OECD PISA 2015 Database (http://www.oecd.org/pisa/data/2015database)

Summary and conclusions

The results show that Alberta teachers frequently use both inquiry-based teaching practices and teacher-directed instruction in their science classes. The higher performance of Alberta students may indicate a greater understanding and enjoyment of science, possibly due to their teachers' ability to effectively combine the two types of strategies. The results also show that despite the reported importance of inquiry-based strategies, teacher direction is very important for the effectiveness of any inquiry strategies. Science results can improve if teachers effectively incorporate teacher-directed strategies into science lessons along with strategies that improve students' sense of self-efficacy and enjoyment of science. This finding is consistent with OECD results showing that in all but three education systems, more frequent use of teacher-directed instruction is associated with higher science achievement, after accounting for the socioeconomic status of students and schools; conversely, in 56 countries and economies, greater exposure to inquiry-based instruction is negatively associated with science performance (OECD 2016).

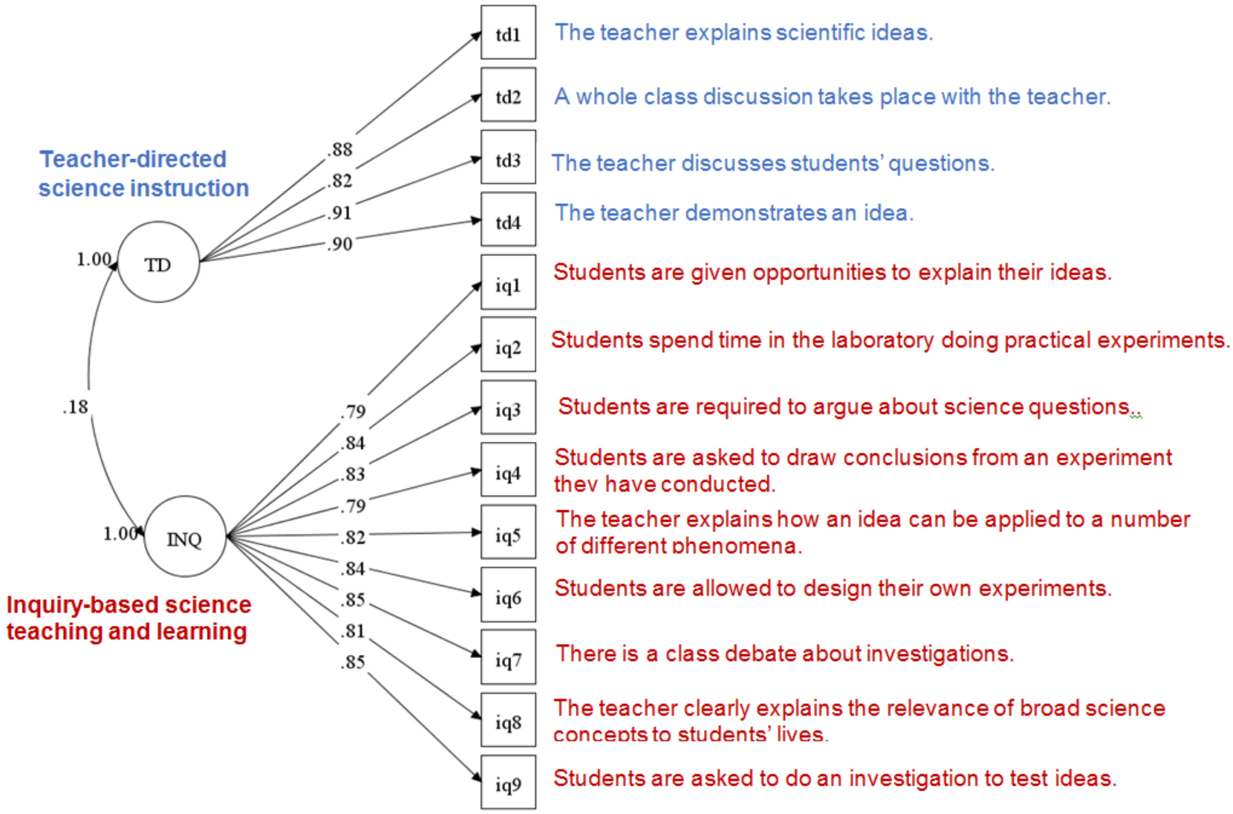
Appendix A: Teaching variables and student characteristics

	Description / Question with associated items	Response categories
Teaching variables		
Inquiry-based science teaching and learning practices	When learning science topics at school, how often do the following activities occur? • Students are given opportunities to explain their ideas. • Students spend time in the laboratory doing practical experiments. • Students are required to argue about science questions. • Students are asked to draw conclusions from an experiment they have conducted. • The teacher explains how a science idea can be applied to different phenomena. • Students are allowed to design their own experiments. • There is a class debate about investigations. • The teacher clearly explains the relevance of broad science concepts to our lives. • Students are asked to do an investigation to test ideas.	(1) in all lessons (2) in most lessons (3) in some lessons (4) Never or hardly ever
Teacher-directed science instruction	How often do these things happen in your lessons for this science course? • The teacher explains scientific ideas. • A whole class discussion takes place with the teacher. • The teacher discusses our questions. • The teacher demonstrates an idea.	(1) Never or almost never (2) Some lessons (3) Many lessons (4) Every lesson or almost every lesson
Teacher support in science classes	How often do these things happen in your school science lessons? • The teacher shows an interest in every student's learning. • The teacher gives extra help when students need it. • The teacher helps students with their learning. • The teacher continues teaching until the students understand. • The teacher gives students an opportunity to express opinions.	(1) Every lesson (2) Most lessons (3) Some lessons (4) Never or hardly ever
Adaptive instruction	How often do these things happen in your lessons for this science course? • The teacher adapts the lesson to my class's needs and knowledge. • The teacher provides individual help when a student has difficulty understanding. • The teacher changes the structure of the lesson on a topic that most students find difficult.	(1) Never or almost never (2) Some lessons (3) Many lessons (4) Every lesson or almost every lesson
Disciplinary climate in science classes	How often do these things happen in your science lessons? • Students don't listen to what the teacher says. • There is noise and disorder. • The teacher has to wait a long time for students to quiet down. • Students cannot work well. • Students don't start working for a long time after the lesson begins.	(1) Every lesson (2) Most lessons (3) Some lessons (4) Never or hardly ever

Student characteristics		
Gender	Gender was determined by students' response to the question "Are you male or female?" The study population included equal proportions of males and females.	
Socioeconomic status	PISA uses an index of economic, social, and cultural status (ESCS) to determine students' socioeconomic status. The ESCS is a composite score built via principal component analysis using the following indicators: parental education; highest parental occupation; and home possessions, including books in the home. The rationale for using these three components is that socioeconomic status has usually been seen as based on education, occupational status, and income. As no direct income measure is available from the PISA data, the existence of household items has been used as a proxy for family wealth.	
Science self-efficacy	How easy do you think it would be for you to perform the following tasks on your own? • Predict how changes to an environment will affect the survival of certain species. • Recognize the science question that underlies a newspaper report on a health issue. • Explain why earthquakes occur more frequently in some areas than in others. • Identify the science question associated with the disposal of garbage. • Interpret the scientific information provided on food labels. • Describe the role of antibiotics in the treatment of disease. • Identify the better of two explanations for the formation of acid rain. • Discuss how new evidence can lead you to change your understanding about the possibility of life on Mars.	(1) I could do this easily. (2) I could do this with a bit of effort. (3) I would struggle to do this on my own. (4) I couldn't do this.
Enjoyment of science	How much do you disagree or agree with the statements about yourself below? • I enjoy acquiring new knowledge in science. • I am interested in learning about science. • I generally have fun when I am learning science topics • I am happy working on science topics. • I like reading about science.	(1) Strongly disagree (2) Disagree (3) Agree (4) Strongly agree

Note: Students responded to a series of four-point Likert scale questions with set items and response categories.

Appendix B: A confirmatory factor analysis of results using Alberta data for PISA-defined teacher-directed science instruction and inquiry-based science teaching and learning.



Source: Based on Alberta data from OECD PISA 2015 Database.
Note: Confirmatory factor analysis results were derived from MPLUS statistical modelling software. All factor loadings are significant; standardized root mean square residual = 0.035

References

Blanchard, M.R., Southerland, S.A., Osborne, J.W., Sampson, V.D., Annetta, L.A., & Granger, E.M. (2010). Is inquiry possible in light of accountability? A quantitative comparison of the relative effectiveness of guided inquiry and verification laboratory instruction. *Science Education*, *94*(4), 577–616.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334.

Driver, R. (1995). Constructivist approaches to science teaching. In L.P. Steffe & J. Gale (Eds.), *Constructivism in Education* (pp. 385–400). Hillsdale, NJ: Lawrence Erlbaum.

Furtak, E.M., Seidel, T., Iverson, H., & Briggs, D.C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research*, *82*(3), 300–329.

Fullan, M., Hill, P., & Crévola, C. (2006). Breakthrough. Thousand Oaks, CA: Corwin Press.

Hattie, J.A.C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.

Minner, D.D., Levy, A.J., & Century, J. (2010). Inquiry-based science instruction—What is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching*, *47*, 474–496.

Oliver, R. (2007). Exploring an inquiry-based learning approach with first-year students in a large undergraduate class. *Innovations in Education and Teaching International*, *44*, 3–15.

Organisation for Economic Co-operation and Development (2016). *PISA 2015 results: Vol. 2. Policies and practices for successful schools*. Paris: OECD Publishing. doi:10.1787/9789264267510-en.

Organisation for Economic Co-operation and Development (2017). *PISA 2015 assessment and analytical framework: Science, reading, mathematics, financial literacy and collaborative problem solving* (Rev. ed.). Paris: OECD Publishing. doi:10.1787/9789264281820-en.

Prince, M., & Felder, R. M. (2007). The many faces of inductive teaching and learning. *Journal of College Science Teaching*, *36*, 14–20.

What's in It for Them? Possibilities and Limitations of Generating School-Level Results Based on PISA

Pierre Brochu
Psychometrician and assessment consultant, Consultation MEP Consulting Inc.

Abstract

In Canada, schools participating in national and international large-scale assessments do not receive any feedback about their participation, based either on questionnaire data or on achievement results. This paper looks at the possibility of generating some school-level summary information from the Programme for International Student Assessment (PISA) 2015 data set. Data from the student and school questionnaires and achievement results are summarized at the school level to demonstrate some of the possibilities and limitations in releasing such data. In addition, a proposed format for a school-level report is presented for illustrative purposes. Issues related to validity, reliability, confidentiality, and usefulness are discussed throughout the paper.

Key terms: PISA 2015, Canada, school reports, questionnaire analysis, contextual variables

Introduction and theoretical framework

The objective of this study is to see how school-level results from large-scale assessments can be used to inform education policy and practice. It also discusses some options for reporting school-level results, considering possibilities and limitations of such an approach.

Two research questions guide this study:

- In what new ways can school-level results from large-scale national and international studies be used to inform education policy and educational practice at the provincial level?
- What data can be aggregated and analyzed at the school level for the benefit of schools participating in large-scale assessments? What are some possible issues associated with these analyses?

International assessments such as the Programme for International Student Assessment (PISA), the Progress in International Reading Literacy Study (PIRLS), or the Trends in International Mathematics and Science Study (TIMSS) are designed to provide internationally comparable system-level indicators. In all cases, these assessments measure student achievement in one or more subject areas and collect contextual information from questionnaires that can be related to achievement (Campbell, Kelly, Mullis, Martin, & Sainsbury, 2001; Mullis, Martin, & Loveless, 2016; Organisation for Economic Co-operation and Development [OECD], 1999). Results from these assessments are reported at the education-system level. Most of these education systems exist at the country level but in some cases, such as Canada, they operate at the sub-national (i.e., provincial) level.

In the case of PISA, both the OECD and the Council of Ministers of Education, Canada (CMEC) provide provincial-level results (OECD, 2016; CMEC, 2016). Because of the sampling design (a sample of students is selected from a sample of schools) and the test design (matrix sampling whereby not all students respond to the same set of test questions), results are not disaggregated at a level lower than the province (e.g., by region, school program, or school). However, the PISA data set provides plausible values (multiple imputations estimating individual student achievement based on item response theory) for each participating student, along with responses to the student questionnaire. Although the public data set anonymizes participating students and schools, the PISA data set makes it possible to link students to schools, in order to aggregate results at the school level.

Although the school-level aggregated results are generally not published because of the technical limitations just described, a few countries are providing school-level indicators to participating schools (e.g., United States, United

Kingdom). As long as some statistical quality criteria are met, participating schools receive a confidential report about key school characteristics (e.g., school size, location, socioeconomic status indicators), average achievement by subject area, and other selected indicators of interest (e.g., time spent learning, student absences, student motivation and aspirations, and assessment practices in the school). It is important to note that because of the technical limitations stated above, these reports do not include actual scores or percentages. Instead, they compare the school to other similar schools, the national average, or other countries, based on the distribution of students (Westat, 2016a).

In addition, the OECD has launched the PISA for Schools project (http://www.oecd.org/PISA/pisa-for-schools), in which individual schools may wish to participate by assessing students using a PISA-comparable test that will yield reliable estimates at the school level. About 10 countries have participated in the project since its inception in 1997. Participating schools pay a fee to the OECD which coordinates testing. The OECD also prepares a comprehensive report comparing the school performance to that of PISA-participating schools. The report also includes comparative contextual data on socioeconomic context, learning environment, student engagement, and equity indicators (OECD, 2017a).

Data source

The following study uses the PISA 2015 Canadian data set obtained from CMEC. This data file is similar to the public use file (PUF) available from the OECD's PISA 2015 Database (http://www.oecd.org/pisa/data/2015database) but includes a provincial identifier to facilitate provincial analysis. In this data set, participating students and schools are assigned random identification numbers to ensure confidentiality while allowing the linkage of students with schools (and province).

For the purpose of this study, the data set for the province of Manitoba will be used. The Canadian data set includes 30,914 students; for Manitoba, 3,134 students, representing 13,554 15-year-olds, are included. As the PISA 2015 Canadian report (O'Grady, Deussing, Scerbina, Fung, & Muhe, 2016) indicates, 112 schools were selected to participate in Manitoba, of which 85 participated. This resulted in a weighted school response rate of 92.4 per cent and a weighted student participation rate of 83.4 per cent. In Manitoba, the weighted proportion of students excluded from participating in PISA by the school was 8.1 per cent.[1]

The student file and the school file were merged using IDB Analyzer software version 4.0 (International Association for the Evaluation of Educational Achievement [IEA], 2016). In addition to a number of system variables, the data set includes weights, plausible values for each subject area tested, responses to the student and school questionnaires, and various PISA-derived indices. The international data sets, code books, questionnaires, and compendia are available from the PISA 2015 Database.

Analysis

For illustrative purposes, only descriptive analyses are performed for this study. Although correlational and regression analyses could be performed on the PISA data set, these go beyond the scope of the study. The analyses focus on the following aspects of the assessment:

- average school achievement by subject area (both mean scores and levels of performance)
- demographic variables
- socioeconomic and equity indicators
- attitudinal variables (motivation and aspirations)

It is important to note that all data from the student and school questionnaires are self-reported by the student and the school principal, respectively. As such, the school average for students is calculated based on the average of responses from all students who participated in PISA, not all students in the school. Similarly, responses from the school principal provide the school average for the school questionnaire.

[1] In Canada overall, the weighted exclusion rate was 6.9 per cent, which was higher than the international standard of 5 per cent.

Results

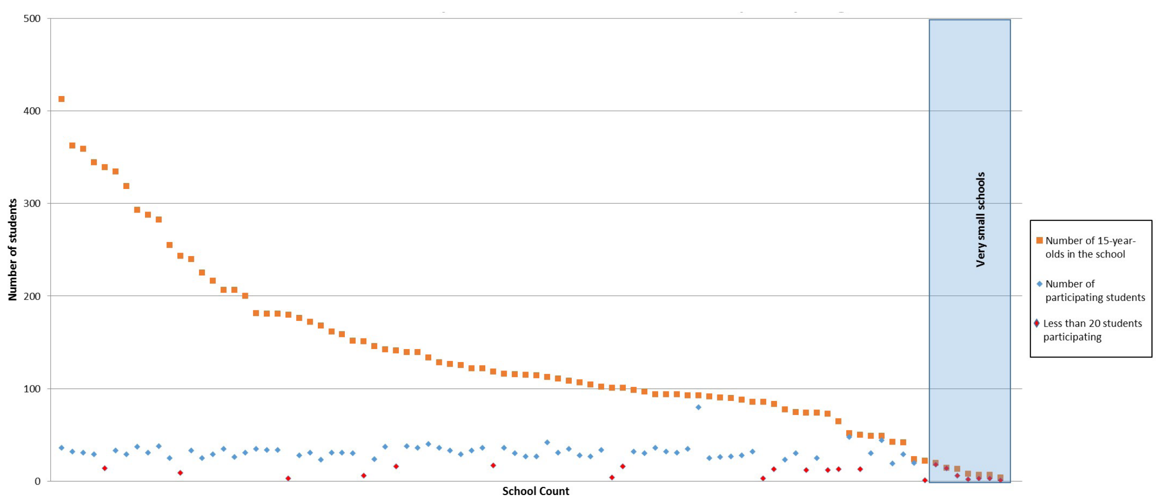
The PISA 2015 Canadian report (O'Grady et al., 2016) provided achievement results by province. These results are integrated into our analyses wherever pertinent.

A portrait of school achievement in Manitoba

Due to the way PISA calculates response rates, more schools are included in the data set than were considered in the participation rate (Westat, 2016b).[2] Thus, although 85 Manitoba schools participated, 88 schools are included in the analysis for the province. PISA normally selects a target cluster size (TCS) of 42 students per school to participate, chosen randomly. However, given that schools are sampled proportionally to size and that several stratification variables are included in the sampling design, many small schools in Manitoba participated. Figure 1 presents the number of 15-year-old students in each selected school, as well as the number of students participating.

Based on PISA criteria, schools are considered very small when their TCS is 20 students or fewer (OECD, 2017b). Although only seven participating Manitoba schools could be considered very small, with 20 or fewer 15-year-olds in the student body, a total 22 schools had fewer than 20 students participating.

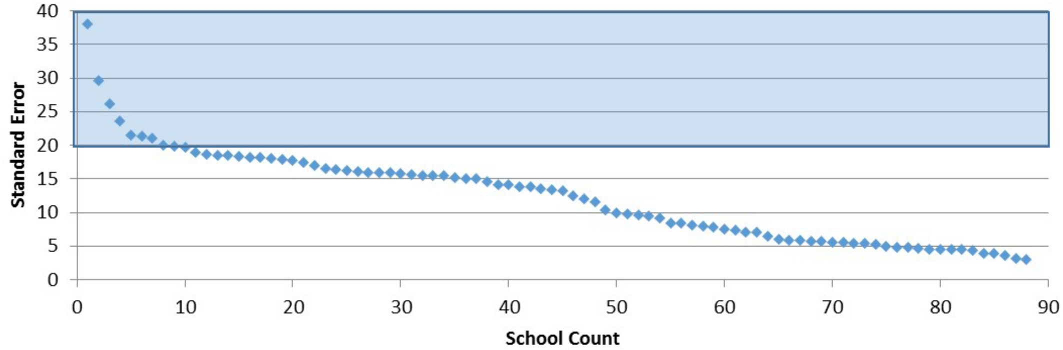
FIGURE 1 Number of 15-year-old students in the school and participating in PISA



When computing average school performance, standard error is affected by several factors, including within-school variance and school size. In PISA 2015, eight schools had standard errors that were more than one-fifth of the PISA standard deviation of 100, or more than 20 score points (Cohen, 1992). For these schools, any interpretation of school performance should be treated with extreme caution. As such, both sample size and standard error should be considered when determining the validity and reliability of individual school results. Figure 2 presents the standard errors for Manitoba schools in the major domain of science for PISA 2015.

[2] This is because some schools did not meet the minimal participation rate and were counted as non-participating. Data from students in these schools were kept in the data set and used for analyses.

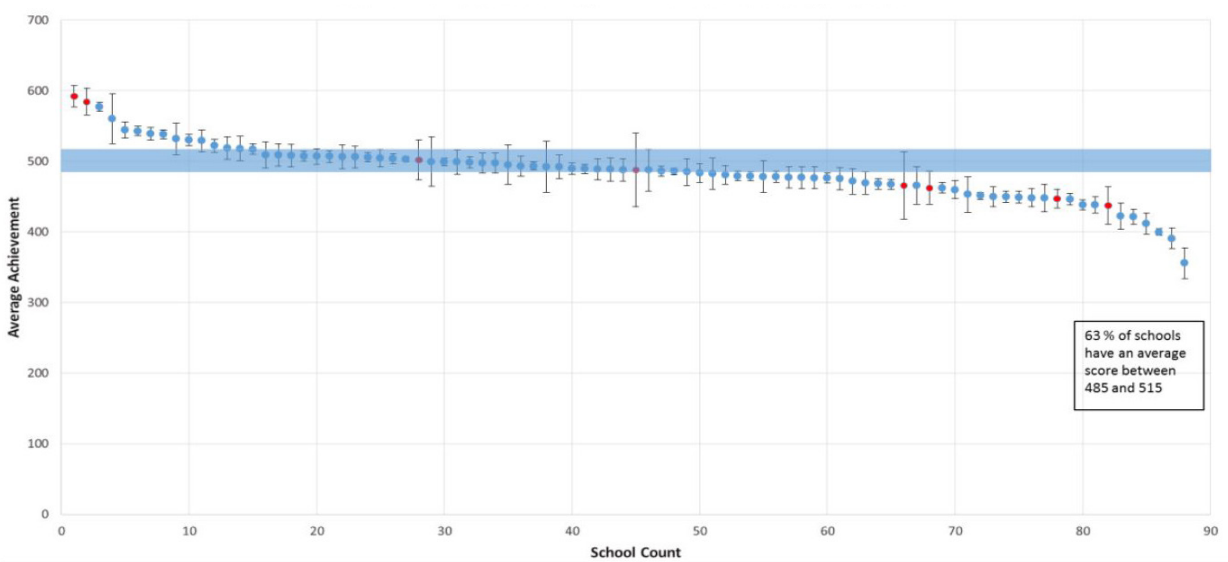
Figure 2 **School standard error in science scores**



School results by subject area provide interesting insights. In these figures, schools are shown in descending order based on school average, and schools with fewer than six students participating in the assessment are shown in red. With provincial averages of 499, 498, and 489 in science, reading, and mathematics, respectively (O'Grady et al., 2016), Figures 3 to 5 show large differences between schools.

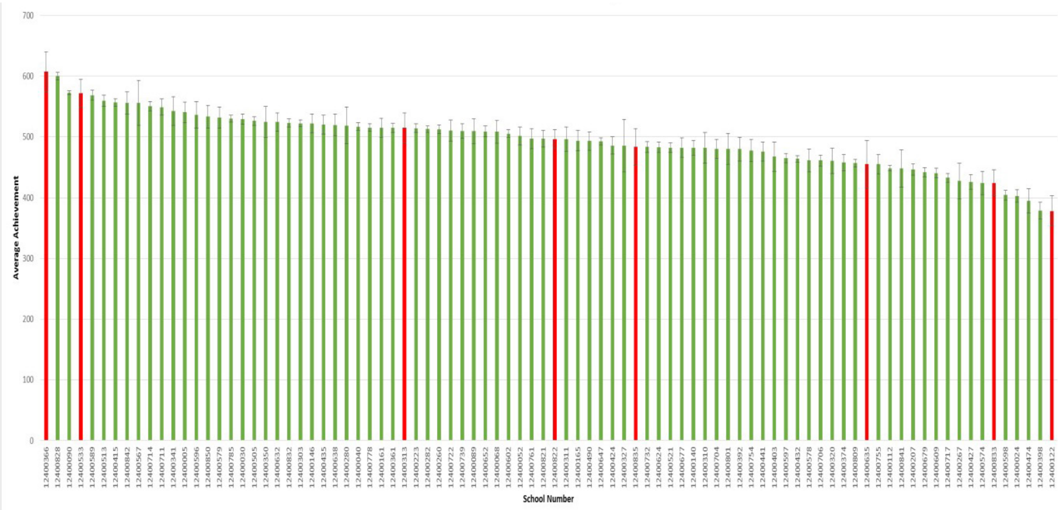
In mathematics, 16 schools scored more than 20 points above the provincial average of 489, while over 25 schools scored more than 20 points below the average (Figure 3). Figure 4 shows the results in reading in a different format, and confirms that there are many more low-performing than high-performing schools in the province. These results suggest that systemic interventions may be more appropriate than interventions at the individual school level to improve low-performing schools.

Figure 3 **PISA 2015 mathematics average score distribution by school**



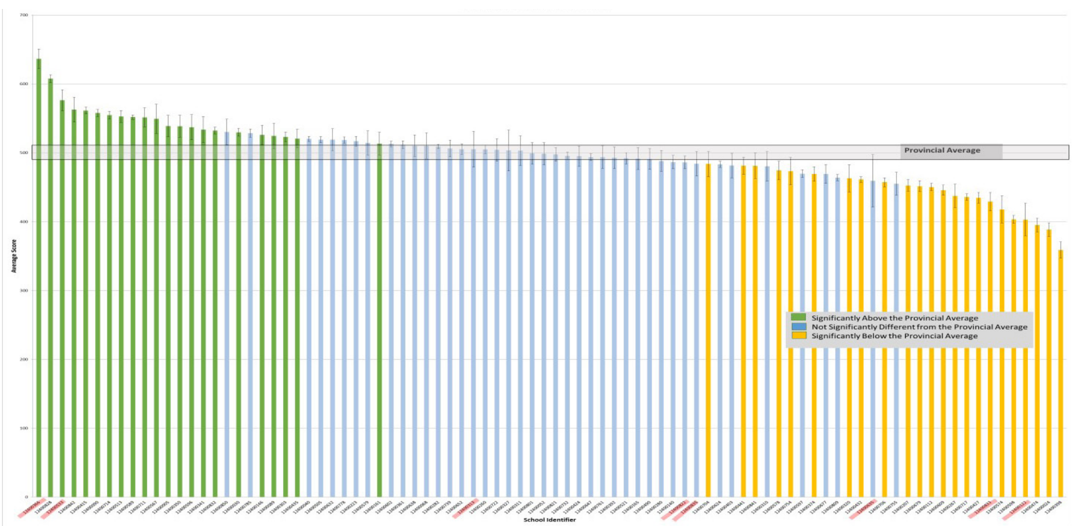
Note: Some 63 per cent of schools have an average score between 485 and 515.

FIGURE 4 PISA 2015 school results in reading



In science, with a provincial average of 499, almost 300 points separate high- and low-achieving schools in the province (Figure 5). This is a huge discrepancy considering that the difference between the highest-performing country in PISA 2015 (Singapore) and the lowest-performing country (Dominican Republic) is 224 points (OECD, 2016).

FIGURE 5 PISA 2015 school results in science



Another way to look at school results is by language of the school system. There were significant differences between school systems in Manitoba with respect to student performance in science and reading (Table 1). Figure 6 shows school average by language of the school system. As confirmed by Table 1, there are more lower-achieving schools in the French system than in the English one.

TABLE 1 PISA 2015 results by language of the school system

	Science		Reading		Mathematics	
	Average	Standard error	Average	Standard error	Average	Standard error
English	501	5.0	501	5.3	489	4.5
French	473	6.9	461	8.1	482	8.9

FIGURE 6 **PISA 2015 school results in science by language of school system**

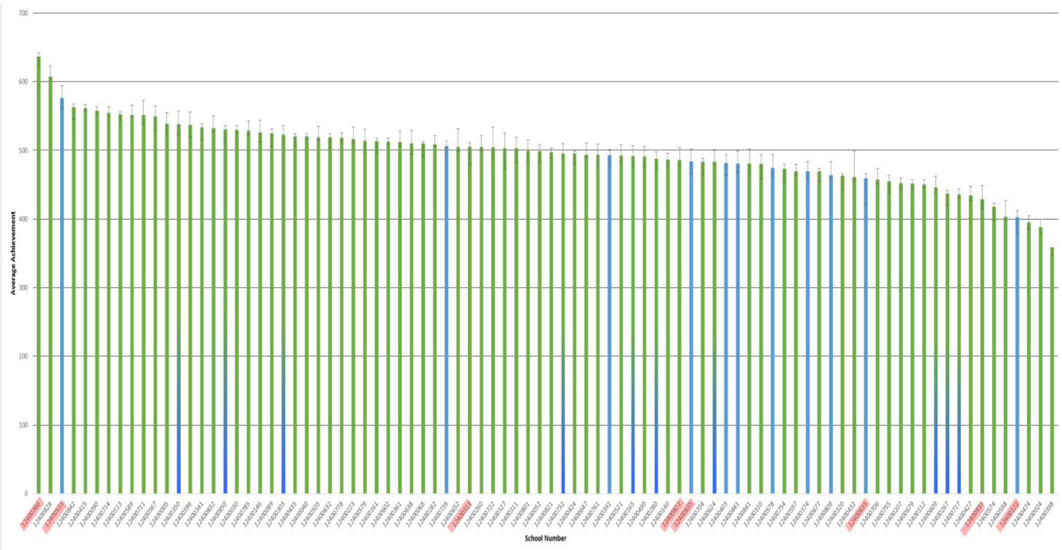


Figure 6 also shows another finding of particular interest: in 10 Manitoba schools, some students wrote the PISA test in English and others wrote the test in French. Three of these schools showed low performance—between 436 and 446—whereas three others showed high performance, with averages between 523 and 538. More analysis of the results from these schools is required, as the language of the test and the language of the school system are not easily tracked in Manitoba.

An issue with confidentiality?

In Canada, participation in PISA is confidential. As such, it should not be possible to identify either participating students or schools. Nevertheless, in some cases maintaining school anonymity can be problematic. As an example, a quick look at Figure 1, above, shows that one school has a sample size of 80 students. Based on Figure 6, this is a school in the French program. We also know that this school had average performance of 469 in science, 457 in reading, and 477 in mathematics. Further, it is clear that the principal did not complete the school questionnaire, as all data categories for this school are blank. Analysts who are familiar with the school demographics in the province would be able to identify the school with a fairly high degree of certainty. In provinces with a small number of schools in a given stratum (e.g., French schools), such situations should be treated with caution, regardless of whether individual school results are ever generated.

Learning from high- and low-performing schools

Using Figure 5 for illustrative purposes, the next section focuses on schools at the extremes of the science score distribution (either very high or very low performers) based on some characteristics of interest. Although the number of cases is relatively limited, we can look at the 10 Manitoba schools with an average score of 550 points or more (high performance) and at the 10 schools with a score of 440 or less (low performance) in science.

Given the multiple dimensions covered by the PISA study and their intricate relationships, bivariate analyses may be misleading. Multivariate analyses are therefore often best suited when analyzing the complex relationships between variables in the PISA study (Fuchs and Wöβmann, 2005). Given the illustrative purpose of this paper, however, we have conducted a series of bivariate analyses to identify possible factors that may help differentiate Manitoba schools based on their average performance in PISA science. Table 2 presents characteristics of high-performing schools that differed significantly from other schools and characteristics for which there was no significant difference.

TABLE 2 PISA 2015 selected factors and their statistical significance in high-performing schools

Statistically significant difference between high-performing schools and all other schools		No statistically significant difference between high-performing schools and all other schools
Variable	**Relationship**	**Variable**
Index of economic, social, and cultural status (ESCS)	Positive	Index of science-specific resources
Index of shortage of educational materials	Negative	Index of science self-efficacy
Index of educational leadership	Negative	
Index of school responsibility for curriculum and assessment	Positive	
Index of school responsibility for resources	Positive	
Index of school autonomy	Positive	
School type	Negative	
Index of science enjoyment	Positive	

p =0.95

These analyses reveal that high-performing schools in Manitoba differ from non-high-performing schools with respect to a number of factors:

- Their students have a higher socioeconomic status.
- Their principals believe to a lesser extent that a shortage of educational resources hinders the capacity to provide instruction.
- Principals report providing less active leadership in their schools.
- Principals believe that they and the teachers in their schools have greater responsibility for the curriculum and assessment than the school boards and provincial/national authority do.
- Principals believe that they and the teachers in their schools have greater responsibility for resources than the school boards and provincial/national authority do.
- Principals believe that they and the teachers in their schools have more autonomy in the governance of their schools.
- Although the proportion of private schools (independent and government-dependent) is relatively small in the province, there are proportionally more private schools among high-performing schools.[3]
- Students in high-performing schools report more enjoyment of science.

High-performing and non-high-performing schools in Manitoba do not differ with respect to a number of other factors:

- Principals in high- and non-high-performing schools reported similar levels of science-specific resources in their schools.
- Students in high- and non-high-performing schools report having similar levels of science self-efficacy.

Similarly, it is possible to compare low-performing schools in Manitoba with non-low-performing schools (Table 3).[4]

[3] This is where multivariate analysis would be most useful as socioeconomic status is likely to be related to this variable.

[4] Although it would have been technically possible to compare high-performing and low-performing schools directly, this was not done in this study because of the small number of schools in each group.

TABLE 3 PISA 2015 selected factors and their statistical significance in low-performing schools

Statistically significant difference between low-performing schools and all other schools		No statistically significant difference between low-performing schools and all other schools
Variable	**Relationship**[a]	**Variable**
Student factors hindering learning	Positive	Test anxiety
Grade repetition	Positive	Teacher factors hindering learning
Number of books in the home	Negative	Student eating breakfast before going to school
		Student with an immigration background
		Student–teacher ratio

[a] In this context, a positive relationship means that the factor affects low-performing schools more.
$p = 0.95$

These analyses reveal that low-performing schools in Manitoba differ from non-low-performing schools with respect to certain factors:

- Principals are more likely to believe that student factors hinder learning more in their schools.
- Proportionally more students in low-performing schools have repeated a grade.
- Students in low-performing schools have fewer books in the home.

Low-performing and non-low-performing schools do not differ on a number of factors:

- Students report having similar levels of test anxiety.
- Principals believe that teacher factors hinder learning in their schools to a similar degree.
- Similar proportions of students report eating breakfast before going to school.
- Similar proportions of students report having an immigration background.
- Principals report similar student–teacher ratios in their schools.

Based on these analyses, several characteristics related to policy considerations, teaching practices, and parental matters may help explain why some schools are high-performing while others are low-performing. There are also a number of characteristics that do not differentiate between schools. This information may be useful not only to the Ministry of Education but also to schools in Manitoba as they consider what can be done to improve student achievement.

School reporting

As explained previously, school-level reports are not reported publicly in Canada. The National Centre for PISA has determined that school participation must remain confidential and as such, that no school-level reporting can be done. However, it can be argued that confidentiality can be maintained if a school report is shared only with the actual school and its school board. Given the way PISA operates in Canada, any province wishing to share school reports with some of its schools would need to obtain the agreement of the Canadian PISA Steering Committee. One could also expect a ripple effect if some jurisdictions disseminate school reports based on participation in PISA while others don't. Furthermore, the OECD would probably need to agree to such an approach, although this should not be a serious issue since other countries are sharing school reports with their schools. Also, in some jurisdictions, partner organizations such as teacher unions might need to be reassured about the intent of any school-level reporting. Provinces would also have to determine whether they wished to make results available to all participating schools or only to those requesting their results. Relatedly, the issue of whether participating schools that received their results decided to make them available publicly would need to be addressed explicitly with schools.

In addition to these political considerations, some technical matters would have to be taken into account if school-level reporting were to be conducted:

1. **Sample size** The TCS for PISA 2015 was set at 42 students. Based on an 80 per cent response rate, most schools should have tested approximately 34 students. However, when exemptions and absences were taken into account, the modal number of students participating was 31 in Manitoba. At the lower end of the spectrum, in 11 schools fewer than 12 students participated. In these schools, student confidentiality could be

compromised when disaggregating results by gender with fewer than 6 students per cell. When considering both sample size and standard error, it would therefore be advisable not to generate aggregated results for any group with fewer than six students.

2. **School identification** Once schools are selected to participate in PISA, their school identifier is kept in a database at the National Centre A unique PISA identifier is assigned to each school and is used throughout the administration process. For the purposes of producing school-level reports, the Canadian National Centre would need to maintain and update the file linking the original school identifier used by the province and the school identifier in the final data set issued by the international contractor.
3. **Reporting average performance** In PISA, student achievement can be reported by either mean score or level of performance. A graphical representation including confidence intervals should be used to report mean scores, rather than an exact mean score, in order to avoid possible ranking interpretations. To avoid confidentiality issues associated with having very few students at the extremes of the distribution when generating results by performance level, only the proportion of students meeting and not meeting the PISA baseline (Level 2) or of those in the high-achieving group (Levels 5 and 6) should be reported. Also, reports of PISA scores and some components of the ESCS should include a brief, non-technical explanation indicating that the values are based on imputation rather than on true scores (OECD, 2017b).
4. **Reporting on questionnaire variables** It is commonly understood that the reliability of a scale is higher than the reliability of the individual items composing the scale (Cronbach, 1951). As such, to the extent possible, reporting based on PISA composite indices should be preferred to reporting on individual questionnaire items.
5. **Comparisons** When comparisons are deemed useful, school results could be compared to the following groups: provincial average, Canadian average, and OECD average. Direct comparisons should not be made with other schools by naming them.
6. **Data suppression** When reporting the number of students, results should be suppressed when the size of the cell is six students or fewer in any given category.

Sample school report

The following section illustrates what a school report might look like using PISA 2015 data from an actual school in Manitoba.

TABLE 4 Demographic information

School name	Sample high school
PISA identifier	12400850
Manitoba Education and Training School Number	XXXX
District/local authority	XXXX
School type	Public
Total number of students enrolled	667 boys 669 girls
Number of 15-year-old students in the school	159
Average class size for 15-year-olds	26–30 students
Proportion of PISA students with an immigration background (first or second generation)	24.2%
Number of students tested	31 students in English 9 students in French

The key indicators could be selected by the province. A brief generic description of each indicator would be provided, as shown in Table 5.

TABLE 5 **Key PISA indicators for your school**

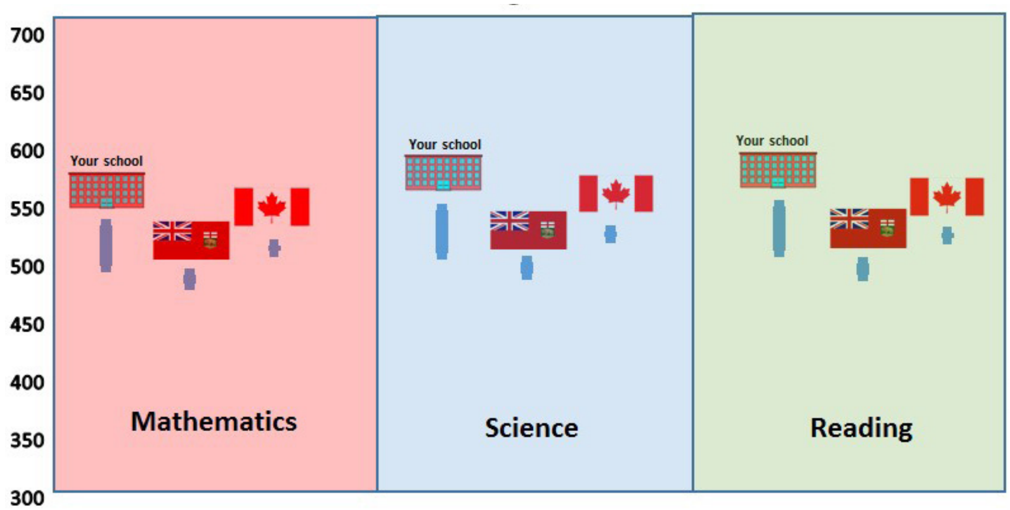
	Manitoba	Canada
Economic, social, and cultural status (ESCS)	↑	↑
Disciplinary climate in science classes	↑	↑
Inquiry-based science teaching and learning practices	*	*
Environmental awareness	*	*
Enjoyment of science	*	*
Science self-efficacy	*	*

* The average for your school is not measurably different from the comparison group.
↑ The average for your school is higher than for the comparison group.
↓ The average for your school is lower than for the comparison group.

Your school's achievement in the PISA domains

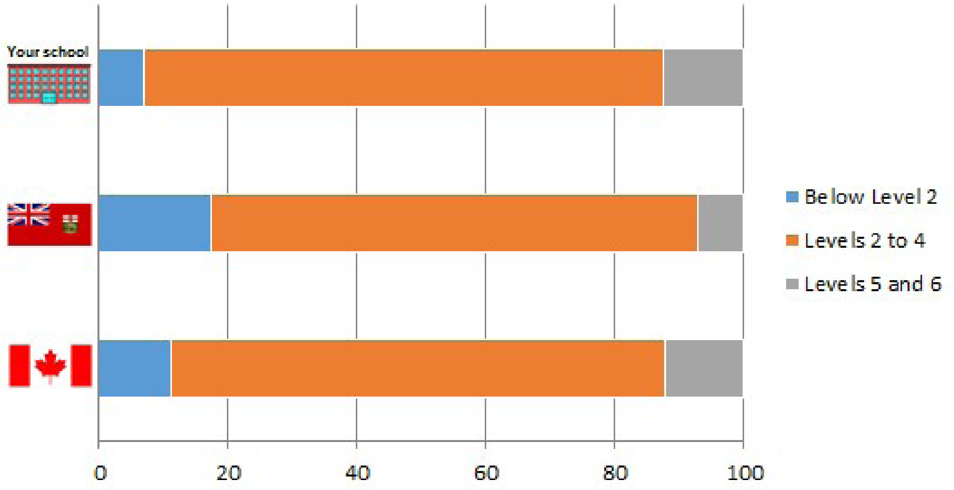
The average school achievement would be provided for the PISA core domains (Figure 7). Results on the PISA additional domains (collaborative problem solving, financial literacy) could also be provided, depending on the quality of the data. Provided that they had sufficient discriminatory power, results for the sub-scales (competency, knowledge, and system) in the major domain of science could also be compared with provincial and Canadian results. Instead of showing exact scores, the graph could represent the confidence intervals of the scores with bars of differing width.

FIGURE 7 **PISA 2015 average achievement**



Owing to the school sample size, results by level could be provided for each domain (as in Figure 8) and, as explained previously, aggregated to reduce the margin of error. Again, no exact percentage would be given.

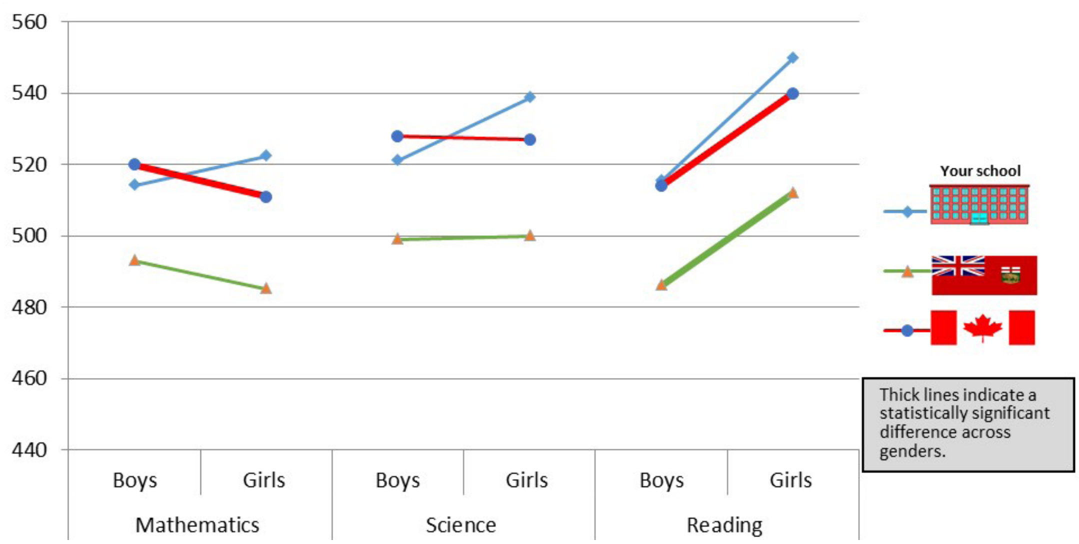
FIGURE 8 Proportion of low- and high-achieving students in science



Gender differences in the PISA domains

A number of graphical representations could be selected to illustrate the gender gap in each subject area (as in Figure 9). Depending on the specific needs of a province, other group differences could be selected. For instance, results by immigration status or by language group might be especially pertinent in some jurisdictions. The important point is that the selected groups should be meaningful for participating schools.

FIGURE 9 Selected contextual factors from PISA at your school



Note: Thick lines indicate statistically significant difference across genders.

Based on responses from the student and school questionnaires, a number of variables of interest have been selected to illustrate the type of analysis that could be provided to participating schools. The variables could be chosen according to the specific needs of provinces or on idiosyncratic patterns of responses from participating schools. In each case, a brief generic explanation of the variables would be provided, as well as its relationship with student achievement. Depending on the level of sophistication required, results could be provided graphically (Figures 10 and 11) or in a tabular format. As previously mentioned, no exact percentages would be provided.

FIGURE 10 Selected student-related factors

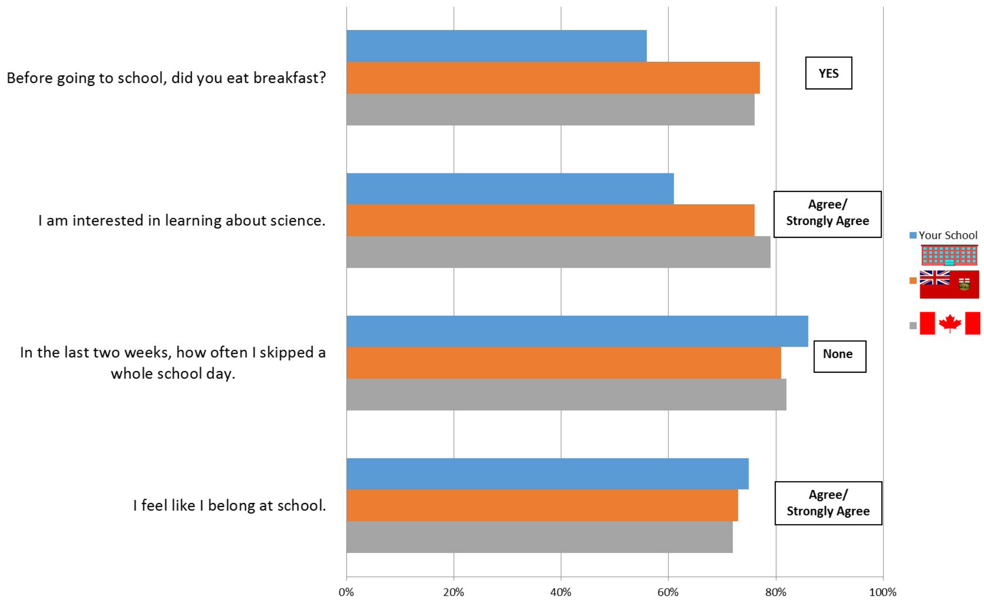
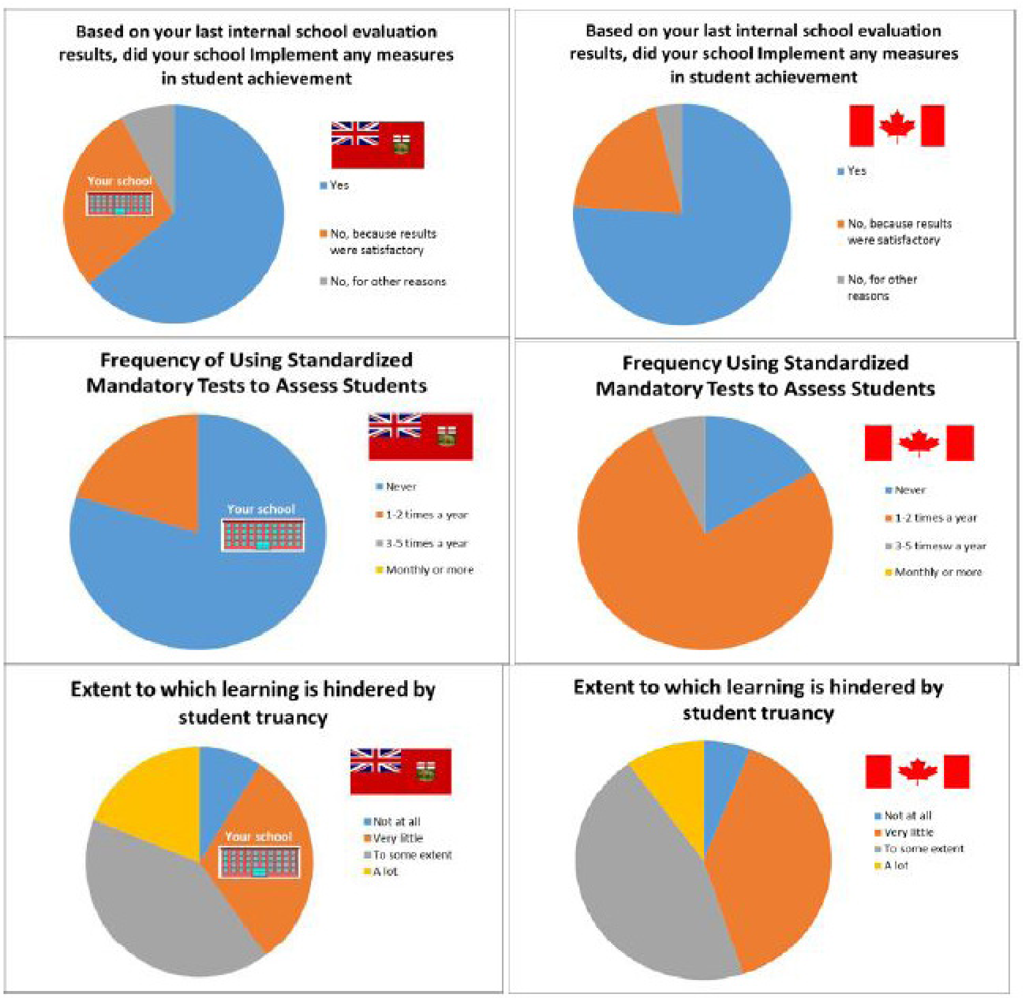


FIGURE 11 Selected school-related factors



Generating school reports

The more standardized school reports are, the more efficient their preparation will be. The selection of the right contextual variables will be key to ensuring that reports respond to the needs of provinces and schools. It is also advisable to consider providing only electronic versions of reports in order to save both time and money for provinces. Smaller provinces would be well advised to make use of templates developed for larger provinces to the greatest extent possible, as it would significantly reduce the per school cost. Using a similar template across assessments and across cycles could also greatly reduce time and costs. Finally, it would be prudent to ask schools to confirm the accuracy of the demographic information table before generating a detailed report.

Conclusion

Participation in national or international assessments places a considerable burden on schools, and as long as Canadian schools receive no reports of their participation they see no tangible benefit in participating. Producing school-level reports using the approach described in this paper would therefore probably be well received as long as certain parameters were clearly understood at the outset:

- Reports should remain confidential to the participating schools and their districts. Provinces would need to determine whether to allow schools to make their results more widely available.
- The reports should be used for information and goal-setting purposes rather than as a way to reward or sanction schools.
- Results would need to be provided in a reasonably timely manner (within a few months of the release of the international results) and be accompanied by appropriate interpretation and in-servicing.

An item analysis of the released PISA items might also provide provinces with a tool to illustrate concretely examples of concepts on which students showed relative strength and areas requiring improvement. This type of reporting would be suitable at the provincial level and could prove very valuable for curriculum review purposes. Because of the PISA test design, however, item analysis would lack reliability at the school level.

Schools participating in international assessments such as PIRLS and TIMSS would probably benefit from school reports even more than from PISA, because of the parent questionnaire they include. There is no doubt that school staff would be greatly interested in knowing more about how parents view their school, and these assessments have traditionally had high parental response rates in Canada. Although international education assessments also include a teacher questionnaire, including data from these in school-level reports would need to be carefully considered because of the small number of teachers responding in each case and the related confidentiality risks.

References

Campbell, J.R., Kelly, D.L., Mullis, I.V.S., Martin, M.O., & Sainsbury, M. (2001). *Framework and specifications for PIRLS assessment 2001*. Chestnut Hill, ME: Boston College.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334.

Fuchs, T., & Wößmann, L. (2005). *Computers and student learning: Bivariate and multivariate evidence on the availability and use of computers at home and at school* (IFO Working Paper No. 8). Munich: Institute for Economic Research at the University of Munich.

International Association for the Evaluation of Educational Achievement (2016). IDB Analyzer 4.0 [Computer software]. Retrieved from https://www.iea.nl/data

Mullis, I.V.S., Martin, M.O., & Loveless, T. (2016). *20 Years of TIMSS: International trends in mathematics and science.* Chestnut Hill, ME: Boston College.

O'Grady, K., Deussing, M-A., Scerbina, T., Fung, K., & Muhe, N. (2016). *Measuring up, Canadian results of the OECD PISA study: The performance of Canada's youth in science, reading and mathematics—2015 first results for Canadians aged 15.* Toronto: Council of Ministers of Education, Canada.

Organisation for Economic Co-operation and Development. (1999). *Measuring knowledge and skills: A new framework for assessment*. Paris: OECD Publishing.

Organisation for Economic Co-operation and Development. (2016). *PISA 2015 results: Vol. 1. Excellence and equity in education.* Paris: OECD Publishing.

Organisation for Economic Co-operation and Development. (2017a). *How your school compares internationally: OECD test for schools (based on PISA) – Sample report high school*. Paris: OECD Publishing.

Organisation for Economic Co-operation and Development. (2017b). *PISA 2015 technical report*. Paris: OECD Publishing.

Westat. (2016a). *PISA 2015 school report: Pleasantville High School* (O.M.B. No. 1850-0755). Prepared under contract to the National Center for Education Statistics, U.S. Department of Education, Rockville, MD.

Westat. (2016b, March). *Sampling in PISA* (Document CY7_NPM(1603)AB_SMP_SamplinginPISA_1.docx). Report produced for the First Meeting of the PISA 2018 National Project Managers, Prague, Czech Republic. Retrieved from https://www.oecd.org/pisa/pisaproducts/SAMPLING-IN-PISA.pdf

A Perspective of Ontario Pre-University Teachers on Global Competencies

Muriel Péguret
Associate Professor, York University

Liam Bekirsky, OCT
Candidate for Master of Public Policy/Master of Global Affairs, University of Toronto

Dominique Scheffel-Dunand
Associate Professor, York University

Abstract

The Organisation for Economic Co-operation and Development (OECD) and the Council of Ministers of Education, Canada (CMEC) have developed reference frameworks and conducted studies of global competencies as part of a strategy to promote the creation of meaningful (professional) development activities for teachers and learners. Nonetheless, given the differing definitions of global competencies across frameworks, our study sought to better understand what the perceptions and pedagogical practices of Ontario pre-university teachers with regard to global competencies can tell us about their preparedness for integrating these competencies into the classroom. Our survey of Ontario pre-university teachers, based on York University's Globally Networked Learning (GNL) approach to integrating global competencies into the curriculum, revealed some of the difficulties teachers have in this area. The survey attempted to assess whether variables such as geographic location, teacher experience, subject of instruction, level of instruction and language of instruction are significant in the evaluation of perceptions related to the integration of global competencies in the curriculum in language classrooms. This initial study–whose objective aligns with and complements that of the Programme for International Student Assessment (PISA)–describes and illustrates the questions raised by the data we collected in the survey and our analysis of the results. The study sought to measure Ontario language teachers' perceptions of their redefined role: as an activator of, and collaborator in, learning in an environment that integrates global competencies and has as a goal the personal and economic success of future citizens who are civically and socially engaged at the local and global levels.

Key terms: global competence; global competencies; interdisciplinary learning; professional development; curriculum internationalization; language learning; evaluation of perceptions and representations.

Introduction

According to the Organisation for Economic Co-operation and Development (OECD), teacher education and professional development are vital to the integration of global competencies into curricula as measured by the Programme for International Student Assessment (PISA). Education systems must support teachers' global competency development if students are to truly succeed in this area (Ramos & Schleicher, 2016, p. 19). This being said, we encountered a certain lack of clarity around this concept, mainly due to the varying terminology used in different frameworks. The concept of global competence itself has a variety of formulations and conceptions: global competence/competencies (Mansilla & Jackson, 2011; Ramos & Schleicher, 2016); 21st-century competencies (Council of Ministers of Education, Canada [CMEC], 2017; Ontario Ministry of Education [OME], 2016a); deep learning competencies (Fullan, Quinn, & McEachen, 2017); and essential competencies (Ipperciel & ElAtia, 2014).

With a diversity of interpretations of these various theoretical frameworks, what can the perceptions and pedagogical practices of Ontario pre-university teachers with regard to global competencies tell us about their preparedness for integrating these competencies into the classroom? We postulate that this question needs to be answered in order to supplement and better understand the PISA 2018 results on global competencies (teacher beliefs and practices, student competencies), as defined by the OECD, in a Canadian context. It is also necessary to do so in order permit professional development activities to be designed to meet the needs of Canadian teachers.

Theoretical framework

At York University, the initial focus of the Globally Networked Learning (GNL) project was curriculum internationalisation, internationalisation at home and civic and social engagement among university students. The GNL project proposes an institutionalisation of pedagogical practices and resources to guide and support the experiences of two educators, in different countries or regions, in classroom partnership and course/module/task co-development (Starke-Meyerring & Wilson, 2008). These co-development activities involve students and instructors in authentic and meaningful tasks with the aim of acquiring global competencies (CMEC, 2017; OME, 2016b). We are in the process of developing professional development activities with this pedagogical focus for university and pre-university teachers that will also allow them to develop their own global competencies.

In this context of internationalising pedagogical practices, we inventoried a diverse range of theoretical frameworks and thereby enriched the objectives of the GNL project. Our definition of global competencies is based on that of CMEC: "Global competencies are overarching sets of attitudes, skills, and knowledge that can be interdependent, interdisciplinary, and leveraged in a variety of situations both locally and globally" (n.d.). CMEC organises competencies into six groups, to which the OME makes minor changes (2016a):

- critical thinking and problem solving
- innovation, creativity, and entrepreneurship
- learning to learn/self-awareness and self-direction (per the OME: a growth mindset [metacognition/learning to learn, perseverance, and resilience])
- collaboration
- communication
- global citizenship and sustainability (per the OME: local, global, and digital citizenship)

We also took inspiration from the rethinking of the role of the teacher as activator, culture builder, and collaborator outlined by the Ontario Ministry of Education (2016b) and Fullan et al. (2017). Within these frameworks, "deep learning" describes the development of global competencies needed for the 21st centuries (Fullan et al., 2017, p. 16).

Data sources

To answer our question about the perceptions and pedagogical practices of Ontario pre-university teachers, we began by creating a survey in English for pre-university modern language teachers in Ontario. The Ontario Modern Language Teachers' Association (OMLTA), which has roughly 900 active members, gave us permission to use their electronic distribution list to send out the survey. Respondents included teachers of modern languages and other disciplines in Ontario who work in urban or rural areas, whose first language is generally French or English, and who have a diverse range of specialisations and years of experience. We intend to broaden the scope of the survey to teachers of all specialisations in French- and English-language school boards across Ontario. Taking into account the role of the teacher as activator, culture builder, and collaborator, as defined by Fullan et al. (2017), we focused on the following areas to determine how teachers perceived their level of professional preparedness for integrating global competencies: their awareness and understanding of the concepts, their training, and the resources available to them.

Our survey received ethics approval from York University on May 4, 2018. It was sent out by email on May 11, 2018, and stayed open until June 4, 2018, with a reminder issued on May 31, 2018. In total, the survey attracted 103 respondents.

Analysis

Variables

Our independent variables include the following eight indicators:

1. Subjects currently taught
2. Subjects specialised to teach
3. Years of teaching experience
4. Grade levels currently taught
5. Grade levels taught in the past
6. Geographical location of the school
7. First language
8. Principal language of instruction

Dependent variables include 40 indicators on the three roles of the teacher: activator, culture builder, and collaborator. We also asked respondents their opinions about how competencies are reflected in the curriculum and about the quality of their resources and training with the respect to the following objectives:

- integrating the competencies
- creating an environment that is conducive to developing/integrating them
- collaborating with partners at school and in the broader community

The research questions were as follows:

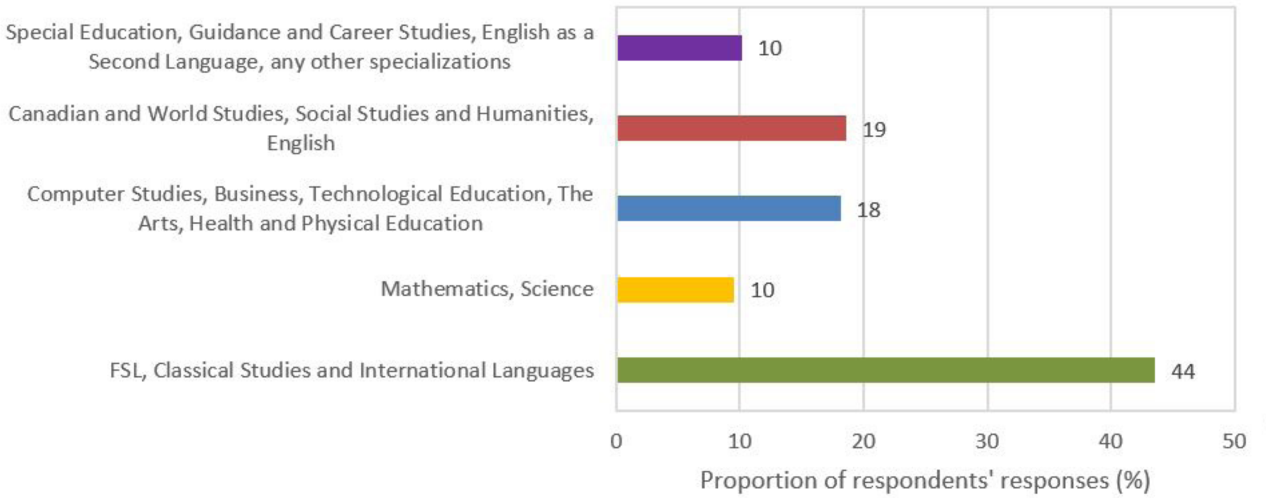
1. What are the main trends in the responses (across all respondents)?
2. Do our independent variables (geographic location, first language, years of teaching experience, etc.) have statistically significant relationships with our dependent variables (integration of global competencies into teaching, access to pedagogical resources, etc.)?
3. What types of professional development related to 21st-century learning have our respondents completed (if any), and what types are they interested in?

Given the relatively small sample size, the subjects taught by respondents were divided into five groups (Figure 1) using the following categories:

- special education
- mathematics and science
- subjects traditionally perceived as applied
- social sciences and humanities
- languages

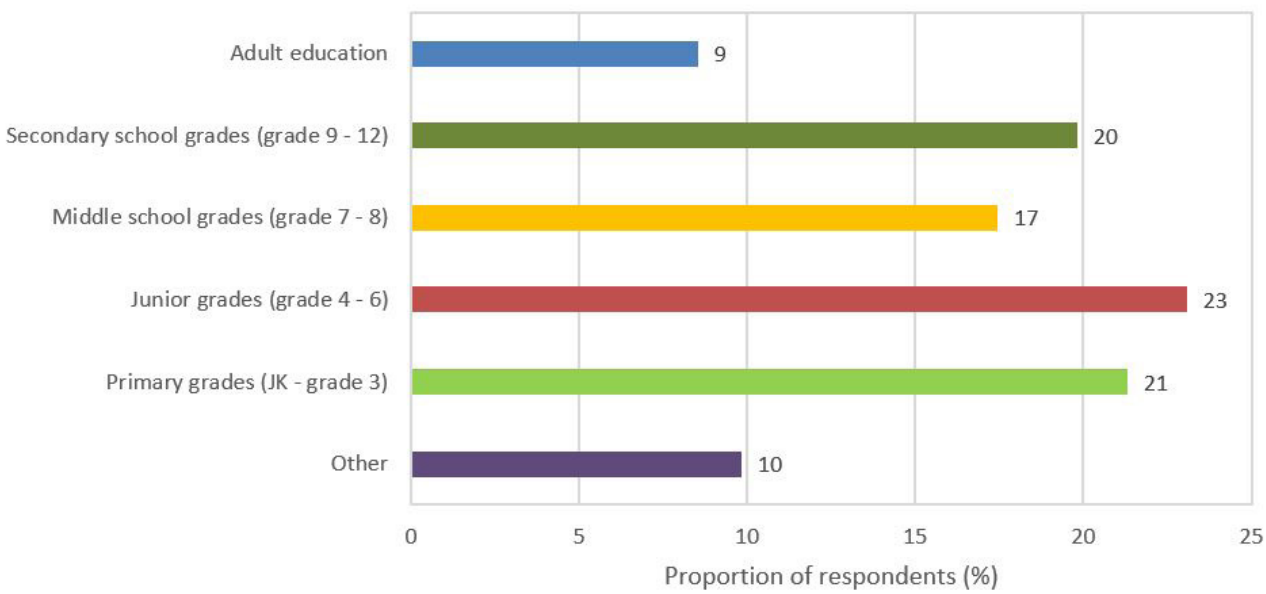
It should be noted, however, that since most of our respondents were French as a second language teachers, they probably teach the other subjects as part of an immersion program.

FIGURE 1 Subject area(s) in which respondents currently teach



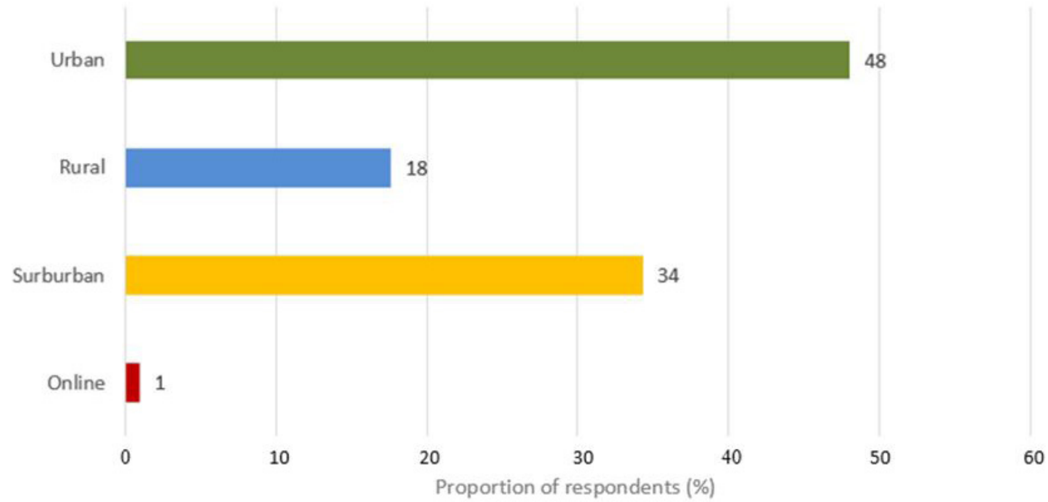
Our respondents included teachers at all grade levels, from primary to adult education (Figure 2).

FIGURE 2 Level at which respondents currently teach



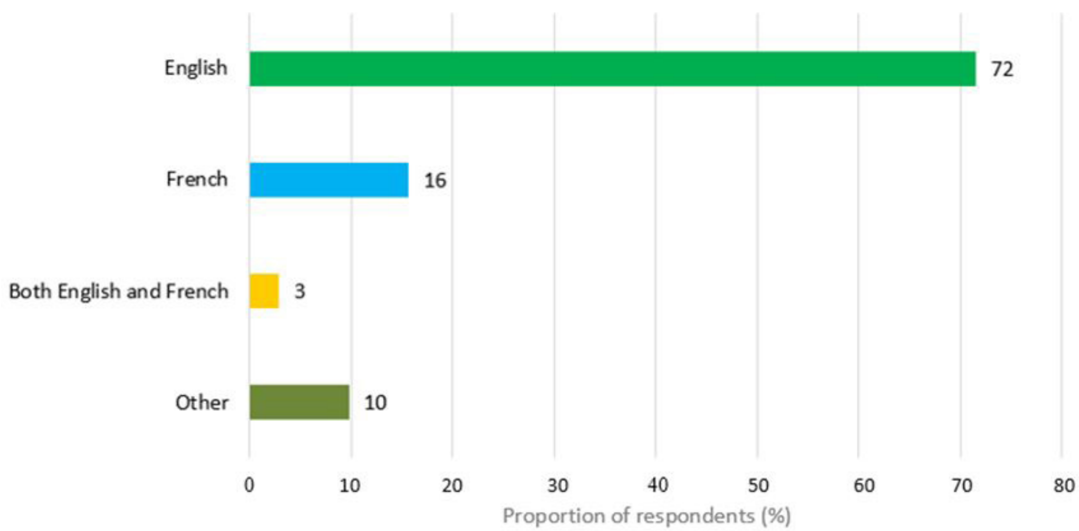
Most respondents were from urban or suburban areas, but 18 per cent taught in rural areas (Figure 3).

FIGURE 3 Location of school



Most respondents reported English as their first language (72 per cent), 16 per cent reported French, 3 per cent are native bilinguals (French and English), and 10 per cent were allophones (Figure 4).

FIGURE 4 First language(s) of respondents



Data analysis methods

We used descriptive statistics to answer our first research question, enabling us to spot general trends in respondents' answers. To answer our second research question, we used chi-squared tests and multivariate analysis of variance to test the relationships between 7 independent variables and 40 dependent variables. Given the large number of tests performed, we used the Bonferroni correction for multiple testing.

The dependent variables are all rated on either three-level Likert scales (not at all or not very, somewhat, very) or five-level Likert scales (never, rarely, sometimes, often, always). To ensure a practical and interpretable statistical analysis, the Likert scales were converted into numerical scales, with the highest number representing "very" or "always." This enabled us to compare the relationships with the independent variables more easily.

We found significant relationships between almost all independent and dependent variables in our data. As a result, we needed to identify the most significant relationships on a practical level. We were able to do so using phi coefficient tests to determine effect size. We chose to illustrate relationships that have at least one full point between high and low scores. The advantage of the statistical models we used is that they are currently considered within the field to be practical. They are particularly robust for the purpose of analysing categorical independent variables against dependent variables on a Likert scale.

To answer our third research question, our survey employed two open-ended questions. We used a word-cloud generator on the responses to give us an immediate overview of both the variety of answers and the most frequently mentioned ideas.

Results

The following section discusses responses to our three research questions.

What are the main trends in survey responses?

Integration of global competencies (presence in the curriculum, integration training, resources) The respondents perceive "communication" to be the most commonly reflected global competency in the curricula of the subjects they teach (Figure 5). Respondents believe they have adequate training to teach this competency (Figure 6) and adequate resources to develop it in the classroom (Figure 7). To a lesser degree, they identify "collaboration" and "growth mindset" as present in their classrooms (Figure 5).

For respondents, "Innovation, creativity, and entrepreneurship" is seen as the global competency least reflected their curricula (Figure 5) and the one they feel the least adequately trained to teach (Figure 6). Additionally, they do not think they have adequate resources to develop this competency in the classroom (Figure 7). To a lesser degree, respondents identified "local, global, and digital citizenship" and "critical thinking and problem solving" as relatively absent from their classrooms (Figure 5).

FIGURE 5 **To what extent are the following competencies reflected in the curriculum for the subjects you teach?**

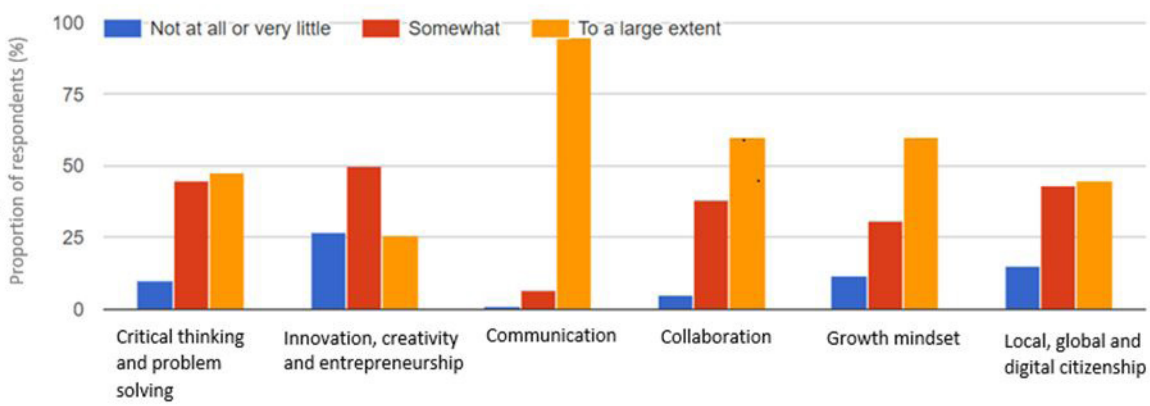


FIGURE 6 **To what extent do you feel sufficiently trained to integrate the following competencies in the learning goals and success criteria for the subjects you teach?**

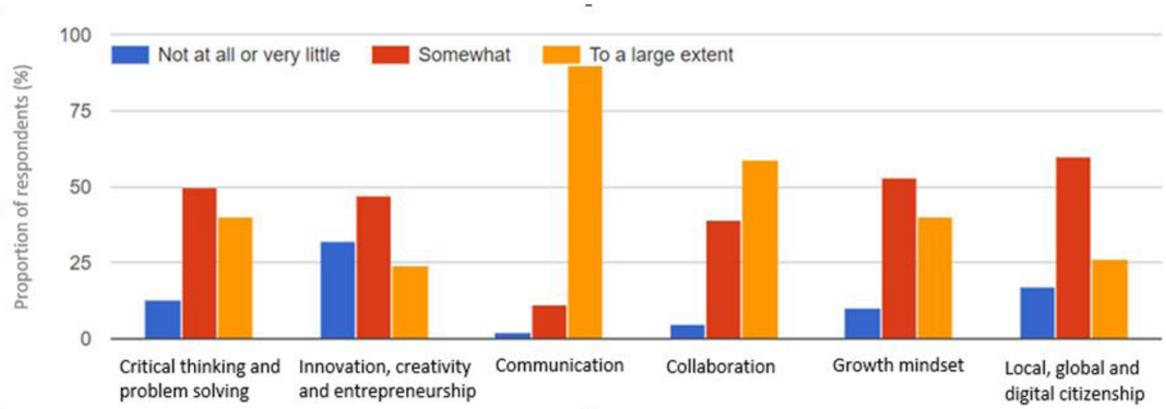
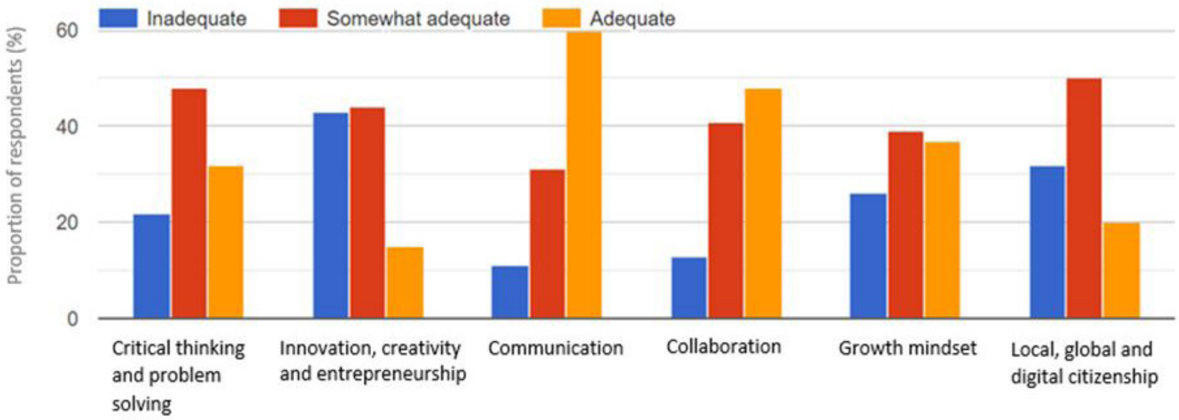
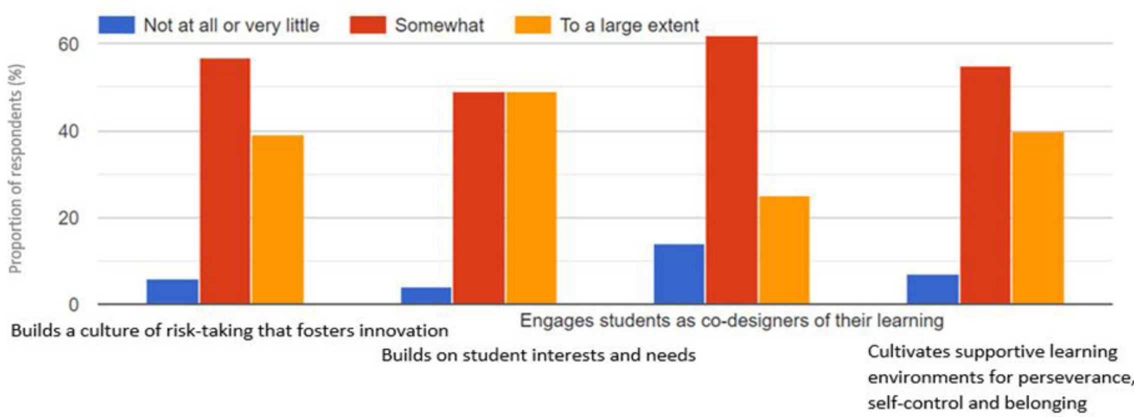


FIGURE 7 **Do you have the adequate resources to integrate the following competencies in the learning goals and success criteria for the subjects you teach?**



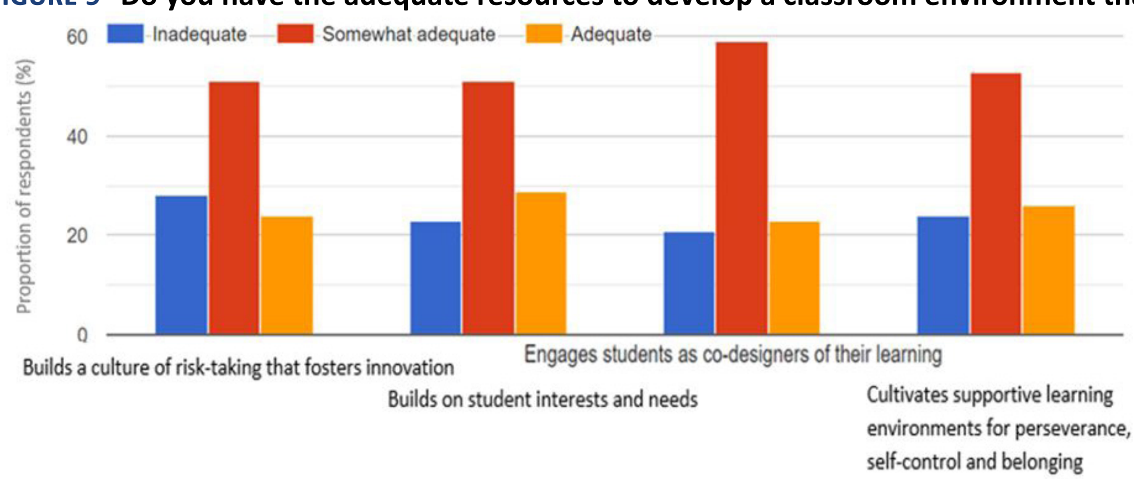
Teachers' roles in developing global competencies Our respondents don't generally assess their training as entirely lacking with respect to the three roles identified by Fullan et al. (2017), but they don't feel very well trained either. The area in which they consider their training the strongest is building on student interests and needs (Figure 8). Generally, however, they feel somewhat comfortable engaging students as co-designers of their learning, building a culture of risk taking that fosters innovation, or cultivating supportive learning environments for perseverance, self-control, and a sense of belonging (Figure 8).

FIGURE 8 **To what extent do you feel sufficiently trained to develop a classroom environment that…?**



Our survey revealed a clear lack of resources. Across all respondents, there was a sense that resources were inadequate to permit them to execute all their responsibilities as teachers, per Fullan et al. (2017) (Figure 9).

FIGURE 9 **Do you have the adequate resources to develop a classroom environment that…?**



Collaboration Between colleagues who teach the same subject, collaborating to design the learning process was generally scored either as "rarely" or "sometimes" (Figure 10). As for respondents' own habits, collaboration is most frequent (but still occasional) with colleagues in the same subject area and with students themselves. It generally scored "rarely" or "sometimes" with students' families or parents, with colleagues in other subject areas, with school administrators, and with teachers from other schools (Figure 11).

FIGURE 10 How often do your colleagues…?

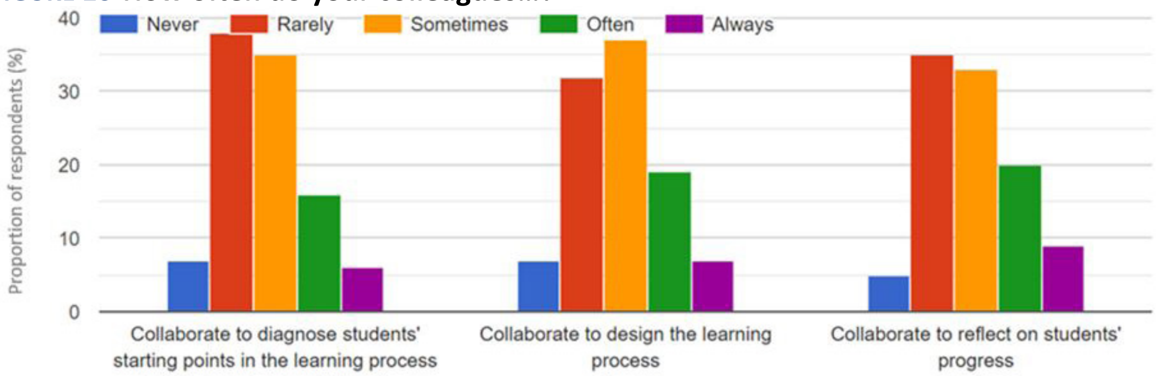
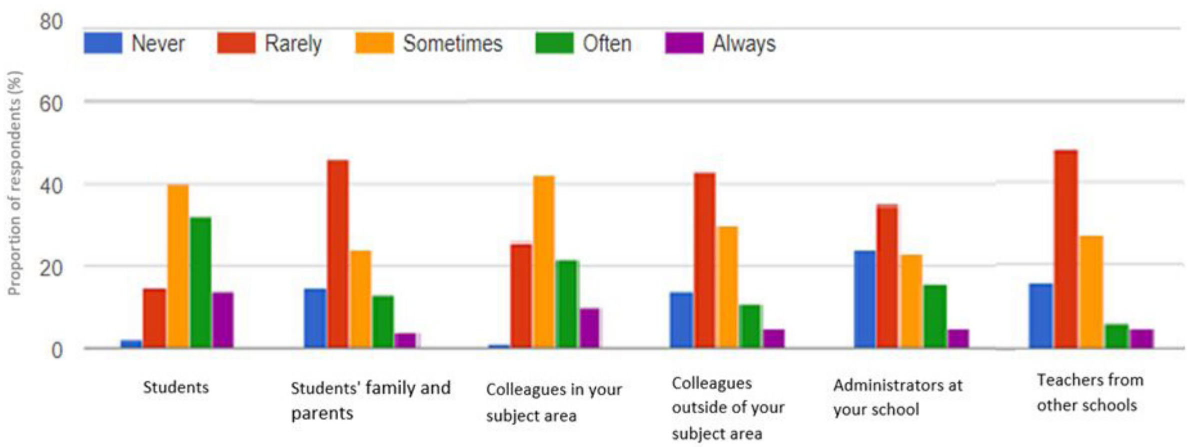
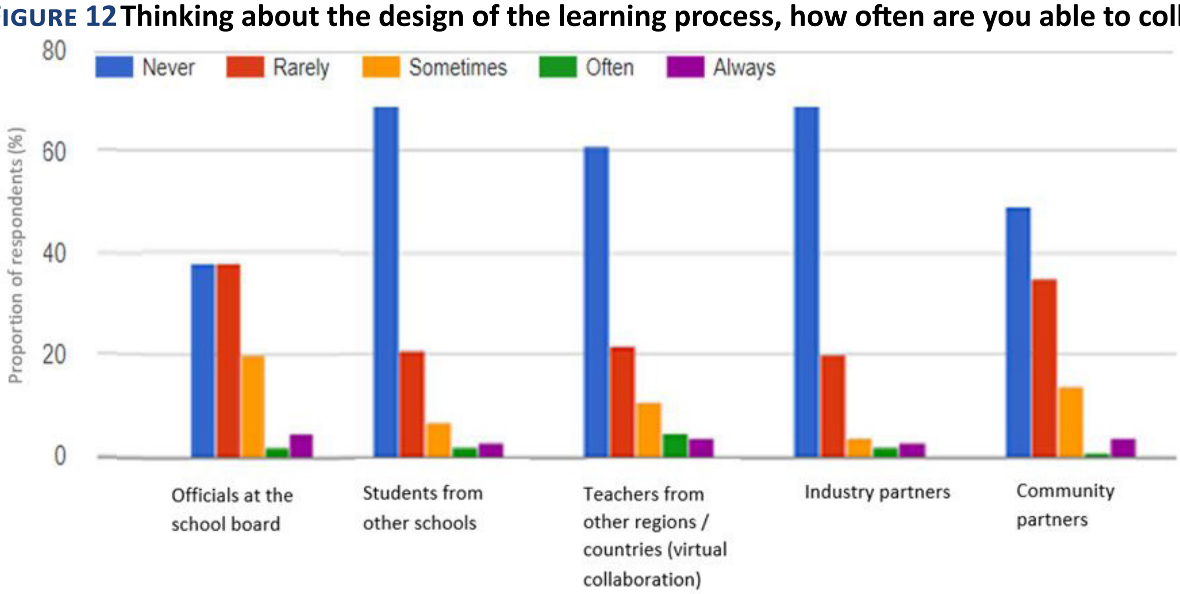


FIGURE 11 Thinking about the design of the learning process, how often are you able to collaborate with…?



Collaboration is rare or non-existent with school board officials, students from other schools, teachers from other regions/countries (via technology), industry partners, and community partners (Figure 12).

FIGURE 12 Thinking about the design of the learning process, how often are you able to collaborate with…?



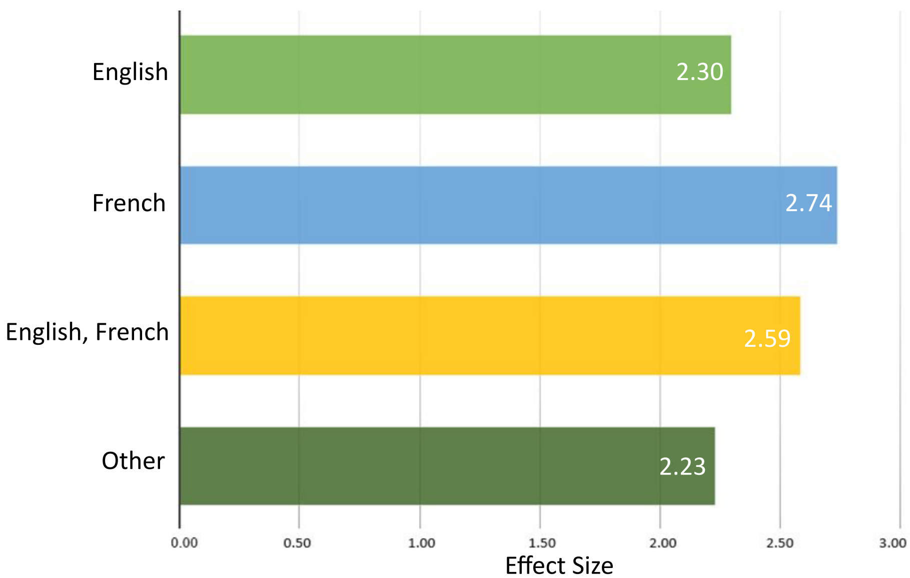
Do our independent variables have statistically significant relationships with our dependent variables?

Overall, the trends that emerged from the responses seem to indicate potential gaps in the training and preparedness of modern language teachers to integrate global competencies. Our statistical analysis also reveals significant differences for all indicators (dependent variables) based on respondent profile (independent variable). Through a deeper analysis, we were able to identify which profiles seem generally to be the most and least advanced in terms of comprehension and integration of global competencies.

The two global competencies with the largest variance based on respondent profile are "local, global, and digital citizenship" and "innovation, creativity, and entrepreneurship." There is also a large variance for "collaboration," which, according to Fullan et al. (2017), is necessary for designing the learning process to develop global competencies among learners.

We first observed that French-speaking respondents whose mother tongue was French (most likely teaching in the English-language system), feel much more adequately trained to integrate local, global, and digital citizenship into their teaching than English as a first language , native bilingual (English-French) , and especially allophone teachers do, even though they say that it is much less present in the curriculum than other competencies (Figure 13). As for language of instruction, those who teach subjects in English are more likely to feel adequately trained to integrate this competency than those who teach subjects in French.

FIGURE 13 Adequate training to integrate local, global, & digital citizenship by first language



Other trends regarding the competencies of "innovation, creativity, and entrepreneurship" (Figure 14) and "local, global, and digital citizenship" are tied to teachers' years of experience and to location. Respondents with several (more than 11) years of experience reported having considerably more training, and teachers in rural schools reported having more training in these two competencies than suburban and urban teachers did. The latter two groups reported having much less training specifically with regard to integrating innovation, creativity, and entrepreneurship (Figure 15).

FIGURE 14 Adequate training to integrate innovation, creativity and entrepreneurship by years of teaching experience

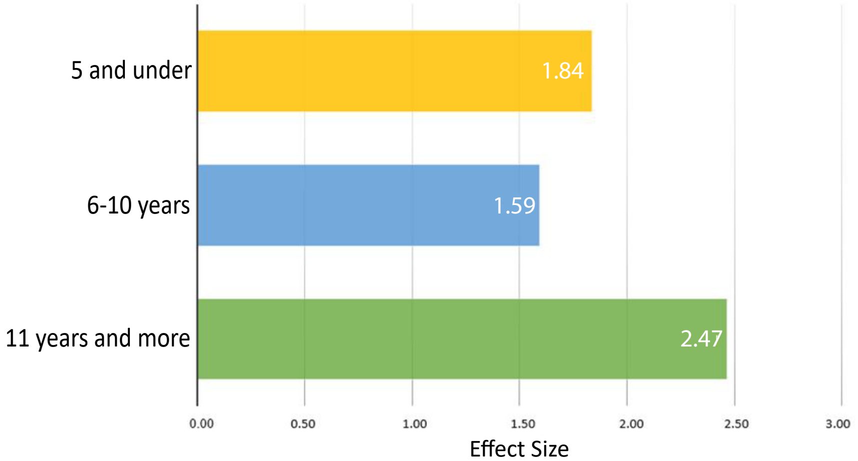
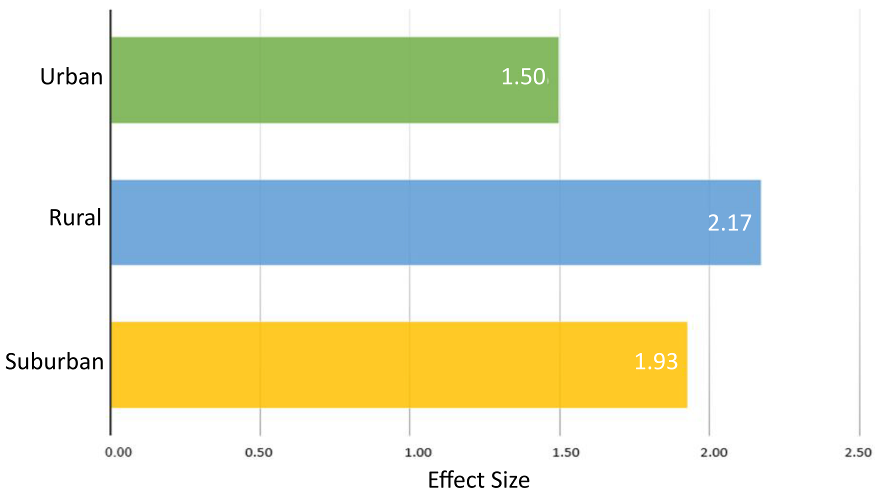
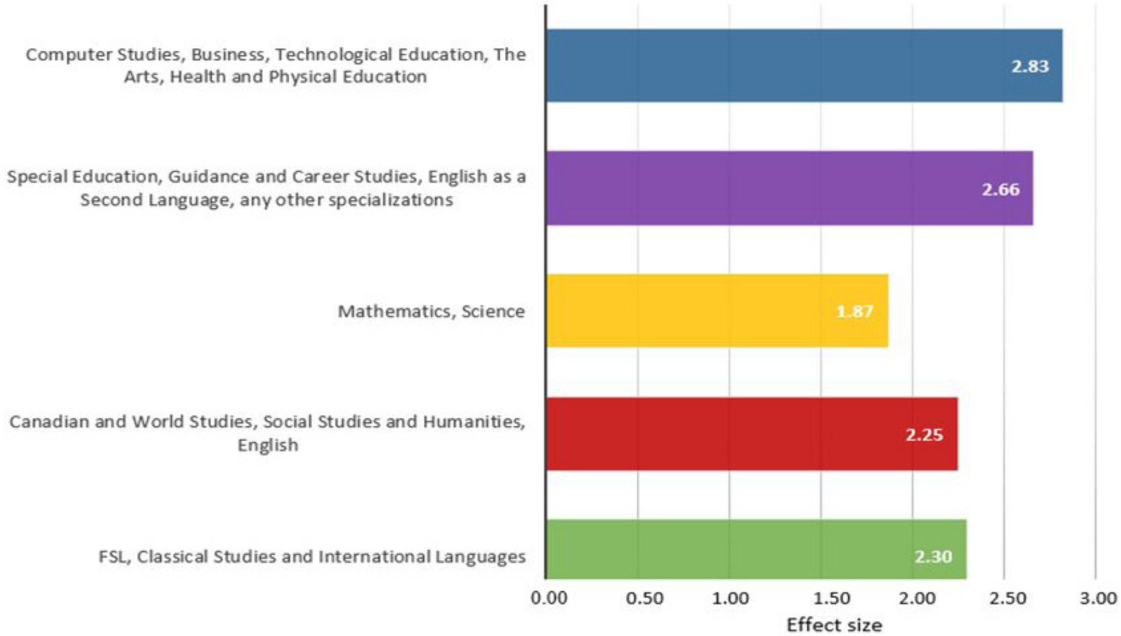


FIGURE 15 Adequate training to integrate innovation, creativity and entrepreneurship by geography of school



We also observed that teachers who teach mathematics and science seem less prepared than others to integrate innovation, creativity, and entrepreneurship into their classrooms. They also feel less well prepared with regard to local, global, and digital citizenship, which they nevertheless identified as being present in their curriculum (Figure 16). However, teachers who teach science collaborate more with school administrators than their colleagues in other subject areas do (Figure 18).

FIGURE 16 **Adequate training to integrate local, global, & digital citizenship by subject currently taught**



Teachers who teach applied subjects (computer studies, etc.) feel the most prepared to develop their students' competency in local, global, and digital citizenship (Figure 16). They also collaborate to design the learning process more than their colleagues who teach social sciences and humanities (Canadian and world studies, etc.) (Figure 17). However, they collaborate less with school administrators than their colleagues in other subject areas (Figure 18).

FIGURE 17 **Collaboration with colleagues to diagnose students' starting points in the learning process by subject currently taught**

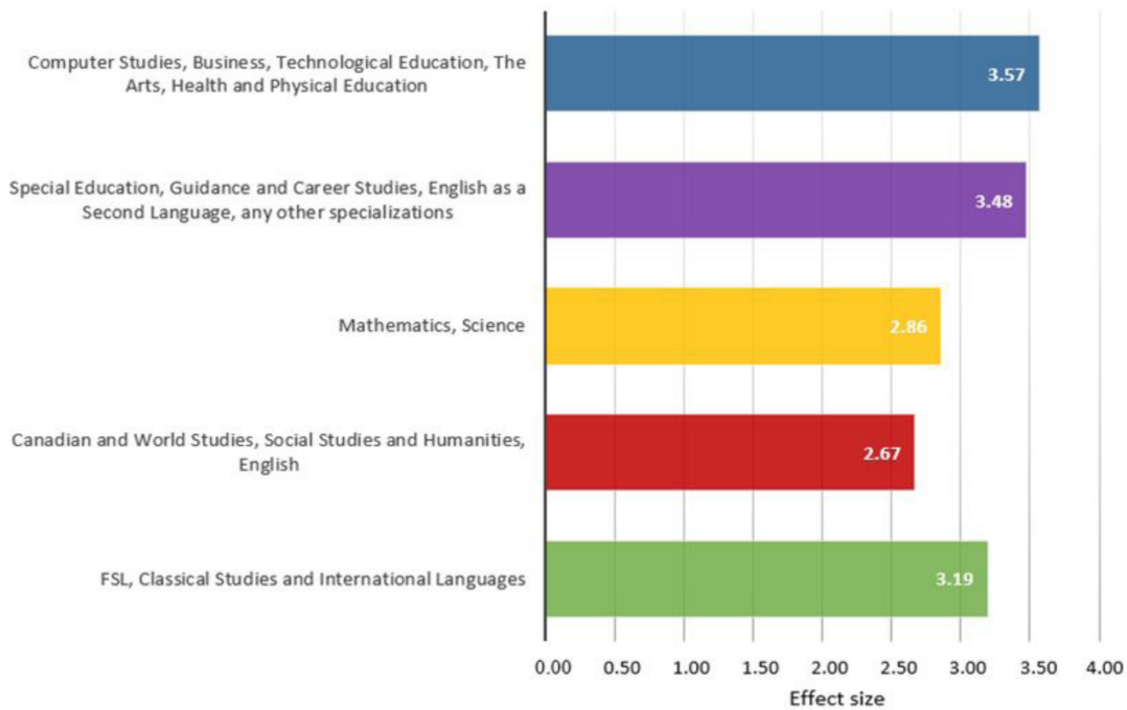
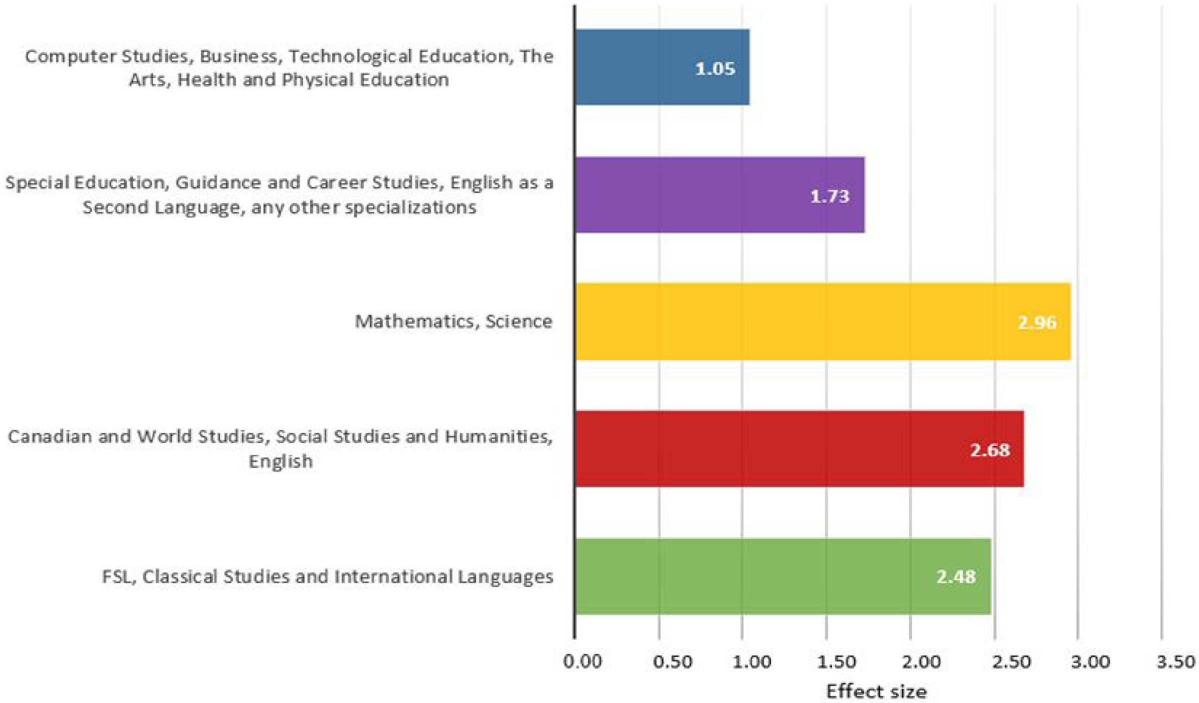


FIGURE 18 **Collaboration with administration at school by subject currently taught**



Primary-school teachers were more likely than their secondary-level colleagues to assess their facility with local, global, and digital citizenship as lacking (Figure 19). They also don't tend to collaborate with teachers in other regions or countries via technology (Figure 20). However, they do tend to collaborate more with school administrators than their colleagues who teach at other grade levels. Teachers at intermediate levels are more likely than their colleagues at other levels to collaborate with teachers in other regions or countries via technology. Finally, secondary-school teachers are more likely to collaborate with community partners than their colleagues who teach at other levels are (Figure 21).

FIGURE 19 **Adequate training to integrate local, global, & digital citizenship by level at which respondent currently teaches**

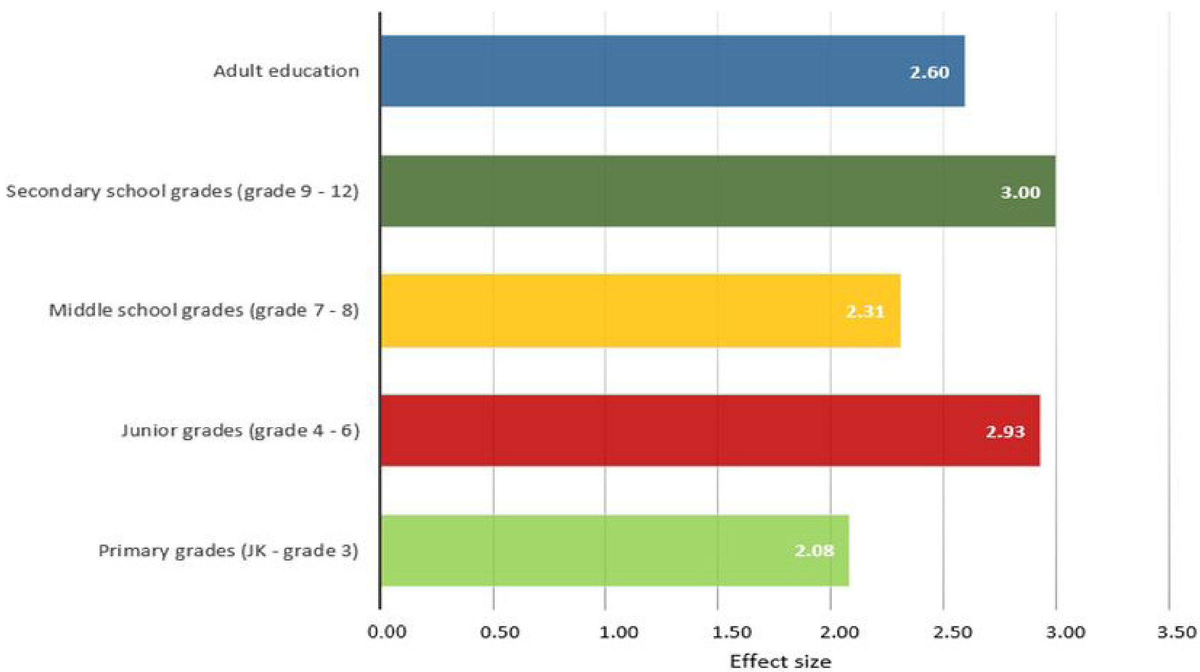


FIGURE 20 **Collaboration with teachers in other countries/regions (virtual collaboration) by level currently taught**

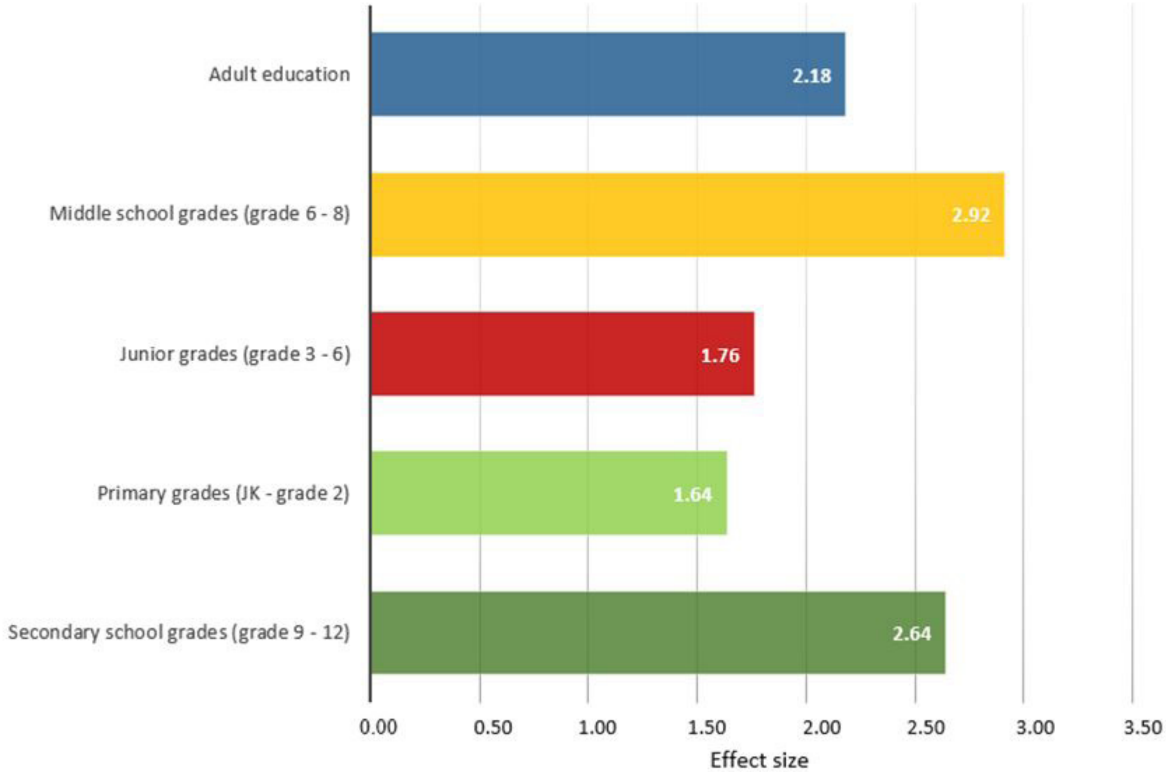
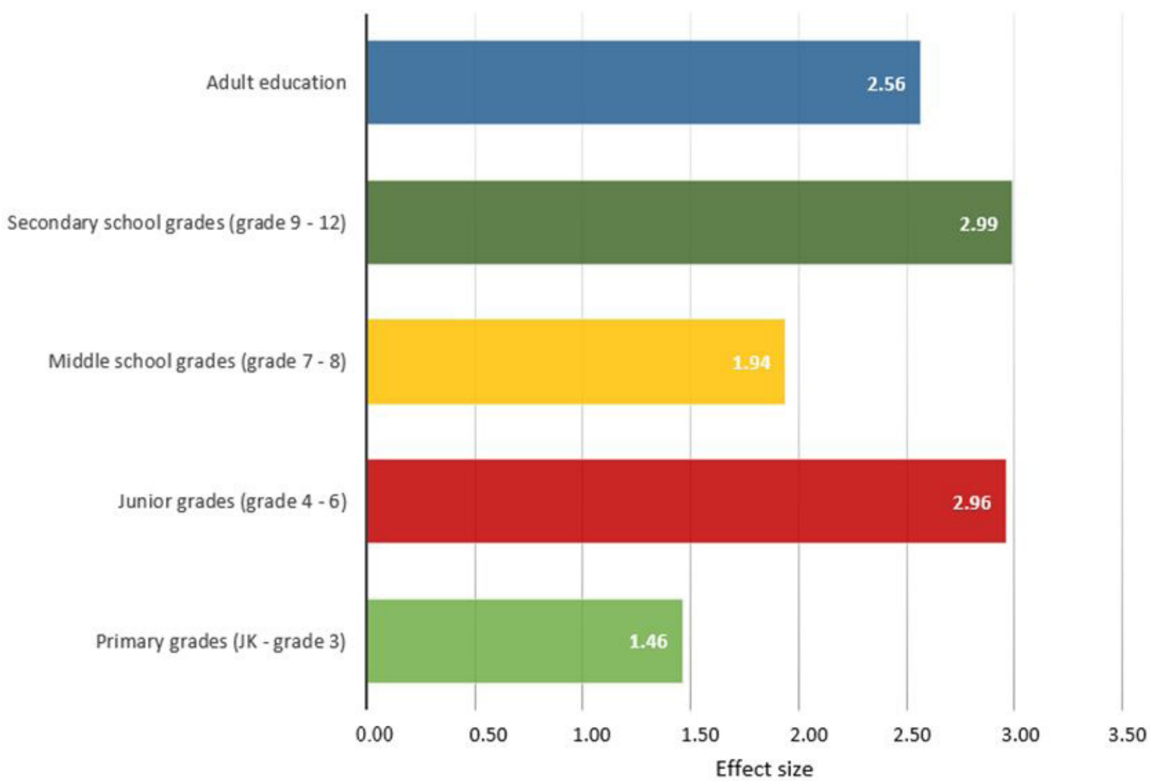


FIGURE 21 **Collaboration with community partners by level currently taught**



What types of professional development related to 21st-century learning have respondents completed (if applicable) and what types are they interested in?

The survey question on completed professional development was open ended, and a word cloud was created based on word frequency (Figure 22). A few references to professional development relate directly to new pedagogies for deep learning, but generally the activities mentioned are only somewhat related to what is meant by 21st-century learning.

The most frequently mentioned professional development activity is the use of multimedia: Google tools, technology-assisted pedagogy, and additional qualifications on classroom computer use, Twitter, and so forth. To a lesser degree, respondents mentioned growth mindset, metacognition, intercultural competence, and critical thinking. It is important to note that many reported not having completed any professional development related to 21st-century learning.

FIGURE 22 Types of professional development completed

The question about the form of professional development respondents are interested in provided a list of choices and an "other" category. Judging by respondents' most frequently used key words, the most popular forms of training are online resources and modules and print teaching resources (Figure 23).

FIGURE 23 Types of professional development desired (created using WordItOut)

Discussion and conclusions

Answering the study question

Our question was the following: What can the perceptions and pedagogical practices of Ontario pre-university teachers with regard to global competencies tell us about their preparedness for integrating these competencies into the classroom? We can now propose some answers and interpretations and formulate implications for next steps.

Our survey reveals that teachers are having to deal with a transition between educational paradigms. The data suggests two things: that respondents are no strangers to global competencies; and that there is still a lot of work to be done to help them better integrate these competencies into the classroom. Modern language teachers described a lack with regard to all global competencies other than "communication," as well as a need for relevant training and resources for deep learning.

The type of collaboration advised by Fullan et al. (2017)—increasingly outside one's classroom and one's school—is happening very little or not at all. And yet that is the type of collaboration that would engage students in the actual concerns of our society and the wider world and would support their education as young 21st-century citizens. The lack of professional development for teachers, guidance on best practices, relevant resources, and appropriate technology (remote collaboration software not being very reliable yet in school settings) are most likely to blame.

Although on the whole the teachers we surveyed felt a certain degree of preparedness for integrating global competencies into their teaching, the results of our statistical analysis seem to indicate that preparedness is not uniform. Results vary significantly depending on location, teaching experience, subjects of instruction, teaching level, language of instruction, and first language. We are well aware that these trends are significant only within our sample, but the results have convinced us that they should be studied on a larger scale. To do so would confirm whether the variables are significant across Ontario or Canada, and would support the design of professional development activities based on teachers' identified needs and abilities in pedagogical methods for building these competencies in learners, and even in themselves.

We would like to continue the study by conducting qualitative investigations of teachers who self-identify as well trained, potentially confirming the trends that have emerged and providing explanations from a more targeted sample. What happens in rural areas, where collaboration between colleagues appears to occur effectively? Why do fewer teachers in urban areas collaborate with school administrators?

Our data suggest that it is important to educate future teachers with respect to deep learning *during* their teacher training, instead of waiting for experience alone to turn them into experts. It also seems important to attempt to integrate all global competencies into all school programs (i.e., not just communication in modern language teaching), as suggested by Fullan et al. (2017). Teachers need support, especially at the primary and secondary levels, if they are to collaborate via technology.

This initial study also suggests a lack of training in global competencies in modern language teachers in Ontario. Half the respondents either chose not to answer the open-ended question about training they had completed (perhaps indicating a lack of it) or answered "none." Their answers also seem to reveal confusion about the concept itself, rooted in a strong association between global competencies and technology. (Integrating technology into the classroom doesn't in itself constitute developing global competencies.)

Lastly, it should be noted that teachers seem to have more interest in resources than in training, and that they prefer things they can do at home to activities conducted elsewhere. In other words, teachers have motivation and interest but little time for professional development. Again, this shows us that numerous professional development activities should be done at the level of teacher education.

To conclude, it seems that the York University GNL project, originally designed for the postsecondary level, could be adapted for the pre-university level and used as a pedagogical approach in the education of modern language teachers, or more generally for anyone interested in project-based pedagogy. This pedagogy would help to identify solutions to the challenges revealed by our study. It would also help to fill the gaps revealed by the study in terms of preparedness for integrating global competencies into the curriculum, especially because internationalisation of the curriculum supports the development of all global competencies, rather than predominantly the "communication" competency.

The limitations of our approach

One of the limitations of the questionnaire as a tool is that it is impossible to know how respondents interpret the concepts and questions. For example, we cannot be sure what our respondents understand the global competency of "communication" to mean, and whether their understanding matches the definition adopted by CMEC.

Another limitation of our analysis is the relatively small sample size and the large number of tests conducted. To partially compensate for this problem, we adjusted the levels of significance (p values). We also regret the limitations of our sample breakdowns (some groups were made of very few respondents). That said, the research works well as a descriptive study.

Future research and applicability

We believe it is essential to continue this type of investigation into the perceptions and needs of teachers with regard to integrating global competencies. The pilot project opened up a number of interesting avenues, shining light on the situation of modern language teachers in particular. What about teachers in other disciplines? Our next study will be conducted on a larger scale. We also believe that integrating qualitative data is essential to better understand how teachers interpret our questions and how they understand global competencies.

Variations in interpretations of and educational policies surrounding global competencies (OECD, CMEC) are particularly apparent in local adaptations. Our initial analysis has helped to define how Ontario modern language teachers understand the concept of global competencies in the Canadian context and within CMEC's theoretical framework. Having tasked ourselves with measuring the quality of the reception and comprehension of these new pedagogical frameworks, we have a responsibility to engage future generations of learners and teachers in integrating the inherent traits of these frameworks into education, work, and social and civic engagement.

Our analysis has helped us identify the needs of teachers in terms of professional preparedness for integrating global competencies into their practice. Professional development activities that aim to integrate global competencies into the curriculum in a structured way will ensure the engagement both of learners who want to be 21st century citizens and of teachers who want to activate the analytical capacities at the heart of global competencies. The objective of any such professional development should be to support teachers through this paradigm shift.

References

Council of Ministers of Education, Canada. (2017). *Cadre pancanadien du CMEC pour les compétences globales*. Retrieved from http://www.ontariodirectors.ca/CODE-TLF/docs/tel/Cadre_pancanadien_competences_globales_descripteu-3.PDF

Council of Ministers of Education, Canada. (n.d.). *Global competencies*. Retrieved from https://www.cmec.ca/682/Global_Competencies.html

Fullan, M., Quinn, J., & McEachen, J. (2017). *Deep learning: Engage the world, change the world*. Thousand Oaks, CA: Corwin.

Ipperciel, D., & ElAtia, S. (2014). Assessing graduate attributes: Building a criteria-based competency model. *International Journal of Higher Education*, *3*(3), 27–38.

Mansilla, V. B., & Jackson, A. (2011). *Educating for global competency: Preparing our youth to engage the world*. New York: Asia Society.

Ontario Ministry of Education. (2016a). *Compétences du 21e siècle : document de réflexion.* Retrieved from http://pedagogienumeriqueenaction.cforp.ca/wp-content/uploads/2016/02/Ontario-21st-century-competencies-foundation-FINAL-FR_AODA_EDUGAINS_Feb-19_16.pdf

Ontario Ministry of Education. (2016b). *Pédagogie numérique en action : document de fondements.* Retrieved from https://pedagogienumeriqueenaction.cforp.ca/wp-content/uploads/2016/03/Document-de-fondements-Fevrier-2016-version-r%C3%A9visee.pdf

Ramos, G., & Schleicher, A. (2016). *Global competency for an inclusive world*. Paris: OECD Publishing.

Starke-Meyerring, D., & Wilson, M. (Eds). (2008). *Designing globally networked learning environments: Visionary partnerships, policies, and pedagogies*. Rotterdam: Sense.

Tracking Resources and Infrastructure in Schools: A Proxy Measure of System Capacity to Support All Students in Creativity, Social-Emotional Development, Citizenship, and Health

David Hagen Cameron
Former Research Director, People for Education

Annie Kidder
Executive Director, People for Education

Christine Schandl
Research Manager, People for Education

Eloise Tan
Research Director, People for Education

Abstract

In line with a current global movement in public education (Organisation for Economic Co-operation and Development [OECD], 2018), this paper discusses the relevance of specific competencies in creativity, social-emotional learning, citizenship, and health in relation to the core purposes of public education: its pursuit of equity, and its role in preparing the next generation for success. It draws from two separate but related studies. The first study, Measuring What Matters, was a large partnership project that included a case study of schools and teachers using these competencies over a two-year period. The second study, the Annual Ontario School Survey, uses large-scale school survey data focused on infrastructure and resources as a proxy measure of the system's capacity to provide learning opportunities that support the development of competencies in creativity, social-emotional learning, citizenship, and health. This paper uses Ozga's understanding of education policy to unpack both the challenges and the possibilities faced by jurisdictions in supporting these areas of learning (Ozga, 2005). In Ozga's theory, policy is more than the text from which it is created. The policy experience—the ways in which school boards, schools, students, and educators interpret and act in response to policy—is an essential part of the policy itself.

Introduction

This paper draws on the findings of two research projects conducted by People for Education—Measuring What Matters (MWM) and the Annual Ontario School Survey—in order to consider the implications of a shift in Ontario public education to embed and support skills referred to as "transferable" or "21st-century" skills within the curriculum. The paper uses these two sources of information to investigate the following:

1. What themes emerge from the pedagogic experiences of educators as they integrate competencies in creativity, social-emotional learning, citizenship, and health into their school and classroom learning processes with students?
2. What are the jurisdictional implications of optimizing opportunities for students to develop these competencies?
3. Given these implications, does Ontario currently have sufficient resources and infrastructure to support learning and development of these competencies for all students?

Theoretical framework

There has been a general global trend in public education toward supporting competency-based learning. In Canada, British Columbia and Quebec have followed this trend by moving towards a competency-based curriculum (British Columbia Ministry of Education, 2018; Gouvernement du Québec, 2007; OECD, 2016).

In line with this movement, People for Education conducted MWM, a province-wide four-year project in which education scholars, policy-makers, and practitioners have identified, designed, and trialed key competencies within areas of creativity, social-emotional learning, citizenship, and health in order both to highlight their importance to students and society and to understand how they come to life in classrooms and schools (Bascia, 2014; Cameron, Watkins, and Kidder, 2015; Ferguson & Power, 2014; Sears, 2014; Shanker, 2014; Upitis, 2014). For the past 22 years, People for Education has also conducted the Annual Ontario School Survey, a province-wide survey of principals of publicly funded schools that tracks the impact of policy and funding changes on school resources and programs.

This paper follows Ozga's understanding of education policy as being more than the text from which it is created. Ozga (2005) asserts that policy experience—the ways in which school boards, schools, students, and educators interpret and act within any particular policy frame (e.g., student report cards or grading)—is part of the policy itself. This invites a more careful understanding of how policies are experienced in the school on a daily basis within both structured and unstructured time for educators and students.

Methods

MWM is a five-year project designed to work with the education sector and beyond to identify, define, and test various areas of learning that are deemed important for school and life success but are not necessarily thought of as direct "academic capacities." The Annual Ontario School Survey tracks the impact of policy and funding changes on school resources and programs.

MWM offers evidence to support increasing the commitment to these areas of learning in relation to their value to students and society. The survey findings provide insight into the quality of the existing resource commitment.

Case study methods

From the winter of 2016 to the summer of 2017, researchers from People for Education partnered with educators in schools and school boards to explore (1) the use of specific competencies within the domains of creativity, social-emotional learning, citizenship, and health, and (2) the connections between the competencies and the learning conditions that may support their development.[1]

This case study marked a critical shift in the MWM project, which went from being largely in the purview of scholarship and research communities to being infused and integrated with practitioner knowledge and expertise. The information elicited helps deepen our understanding of the implications of system support for areas of school success beyond direct knowledge acquisition and traditionally defined academic achievement.

Participating in the case study were 100 educators in 26 publicly funded schools (6 secondary and 20 elementary) and seven school boards. The schools were located in south, north, central, and west Ontario: 8 in rural areas and 18 in urban areas. Of participating schools, 3 were Catholic, 2 were French public, and 21 were English public. The educators who took part varied in terms of their roles and the depth to which they embedded the MWM project into existing programs or initiatives.

The MWM researchers identified and recruited potential participants through existing contacts, looking for educators who wanted to take part and who were already involved in collaborative environments. The educators thus constituted a purposive sample of convenience, as opposed to a strictly randomized or representative sample. The case study was not focused on generating results that could be generalized to a whole population, seeking instead to describe what occurred when educators collaborated to evaluate select groups of MWM competencies in their schools and classrooms.

This is an emergent case study. It presents a research narrative that is drawn from educator experiences in addressing the competencies within the classroom or schools. (See Appendix A for examples.) It does not provide a specific set of expectations in how teachers might address the competencies with students. This presents challenges to the methodology

[1] See: https://peopleforeducation.ca/mwm-defining-the-competencies/ for further details.

because the wide diversity of approaches makes building any common understanding across case participants difficult. The study addresses this challenge by providing a consistent set of collaborative processes or practices for the educator teams involved. In this way, the content and approaches educators used to address the competencies varied, but the way in which they addressed the learning and worked together was consistent across the study (See Appendix B for a detailed description of the educator team-based collaborative approaches.)

Data from the field trials were collected through four means: observation, interviews, focus groups, and other artifacts (e.g., digital pictures and student work). Focus groups and interviews were recorded and transcribed, and observations were recorded in field notes and journals. All participant names are confidential, as are their schools and school boards. All recordings were transcribed and coded using inductive analysis so that themes are grounded in the data (Hatch, 2002).

Annual School Survey methods

This report also relies on data from People for Education's 2017 and 2018 Annual Ontario School Surveys (People for Education, 2017, 2018a). Surveys were mailed and emailed to principals in every publicly funded school in Ontario and could be completed in English or French. In 2018 a total of 1,244 elementary and secondary schools, in 70 of Ontario's 72 publicly funded school boards responded to the survey, representing 22 per cent of the province's publicly funded schools. Survey responses are disaggregated to examine representation across provincial regions. Regional representation in both the 2017 and 2018 surveys corresponds relatively well with the regional distribution of schools.

Data collected from the survey are also matched with school-level data provided by the Education Quality and Accountability Office (EQAO) and Statistics Canada via the Ontario Ministry of Education (OME) School Information Finder. Data provided by EQAO are derived from student self-report questionnaires completed during the Grade 3, 6, 9, and 10 provincial assessments in spring 2017. Data from Statistics Canada include the estimated percentage of children attending each school whose families' after-tax income is below the low-income measure for their family type and size; and the estimated percentage of children attending each school who have at least one parent whose highest certificate, degree, or diploma is from a university. These data are based on the 2011 National Household Survey. School-level data from external sources are matched to People for Education's principal responses using MIDENT (school identification number), for the purpose of analysis.

The quantitative analyses in this report are based on both descriptive and inferential statistics. The chief objective of the descriptive analyses is to present numerical information in an illuminating format that is accessible to a broad public readership. All data were analyzed using SPSS statistical software.

For geographic comparisons, schools were classified as either rural or urban (including both urban and suburban areas) using postal codes. Rural schools are located in jurisdictions with under 75,000 people and not contiguous to an urban centre greater than 75,000 people. All other schools were classified as urban schools. Based on scholarly literature and governmental sources, it was determined that a population of 75,000 provided the most accurate dividing line between small town/rural and urban/suburban areas in the context of Ontario.

Where significant shifts were found in year-over-year comparisons, the trends were confirmed by a comparison with the sample of repeating schools. The average student-to-staff ratio was calculated for schools that reported both the total number of students and the full-time equivalents for staff positions. All survey responses and data are kept confidential and stored in conjunction with recommendations for safeguarding data issued by the Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, & Social Sciences and Humanities Research Council of Canada (2014, December).

Analysis

The analysis in this paper is divided into three parts, each of which addresses the three questions stated above. The sections cover the following areas:

1. school-based themes arising from the case study of educators working to support competency development in creativity, social-emotional learning, citizenship, and health
2. systemic implications for supporting student development in these areas
3. tracking school resources and infrastructure as a proxy measure of system capacity for supporting student development in these areas

School-based themes

A set of themes was established to provide markers in the ongoing narrative of educators in the case study as they worked with students on particular competencies in the domains of creativity, social-emotional learning, citizenship, and health. The themes are as follows:

- importance to student learning
- multi-varied approaches to teaching, learning, and assessment
- the quality of the learning opportunity
- interrelationship of competencies in socio-emotional learning, creativity, citizenship, and health
- common competency lexicon, or language of learning
- broadened perspectives on where learning occurs in schools

The themes emerged through the work with educators in an iterative process; new themes grounded focus groups and observations that then served to refine and modify these themes (Morgan, 2008). They provide insights into potential jurisdictional approaches to measuring broad areas of success, which in Ontario are defined under policy umbrellas such as "student well-being" or "21st-century competencies" (OME, 2016a, 2016b).

Importance to student learning

Participating educators found that the competencies resonated with work they were already focused on, or wanted to focus on, as core components of their teaching. Some expressed concern about the notion of "measuring" students within the domains articulated in MWM. However, the work itself—and positioning the work in relation to classroom and school assessment by establishing specific opportunities for students to demonstrate particular competencies—was well received. The following excerpts are typical of the impressions that educators expressed:

> [One teacher] commented last week [that] the competencies actually validated what she knew to be important in her heart. She works with kids who are identified [as having exceptionalities] and for the one or two students she focused on particularly, she was really looking to support their wellness … It made her feel that "I always knew this is important, and now something is telling me that it's okay to work this way; it's not just about x, y, and z."
> – Program coordinator

> For me it's always, "How do I make this practical?" I believe there is a sense of urgency in education. I believe it's so important to bring those ideas to the centre of learning in our classrooms, but also in our schools, our boards and across boards … That's why I love being involved in this group, because we do have the capacity to share with other boards.
> – School principal

> We have lots of frameworks, but [despite] the title, Measuring What Matters, I found it easy to put the "measuring" aside. I am not quantifying, but trying to say, "If these are my goals that underlie what I do as a teacher, it isn't about what [students] understand about seasonal changes in my science

> curriculum, it's how they're thinking critically and asking questions around those ideas within Science." I see it as a framework that gives greater purpose to what we are doing. And values the things we know are intrinsically important.
> – Science teacher

The ideas expressed in relation to learning within the MWM project are not new to education in Ontario, nor are the competencies completely absent from Ontario curriculum and policy. What is new is the formal and systematic act of focusing on specific, concrete areas within the domains of creativity, social-emotional learning, citizenship, and/or physical and mental health.

The work brought energy to teachers, resonating with their sense of elements that are central to learning but that often do not get the same attention as academic achievement. In short, conducting relevant and integrated assessment of these areas, often with students as co-assessors, was congruent with many participants' professional values as educators. Sergiovanni (2007) speaks of educators as stewards, focusing on the responsibility they carry for the development of children. Participants found that the experience of working in the case study rang true to this sense of stewardship.

Multi-varied approaches to teaching, learning, and assessment

Educators took a range of approaches in their use of the MWM competency framework. Some had a narrow focus, addressing one or two competencies within a single domain; others viewed their inquiry through combinations of competencies from more than one domain.

The critical moment for all came in translating a set of competencies into actual classroom and student experiences. While this "translation moment" was similar for all the participants—and the initial team-based discussions were helpful for diversifying, refining, and broadening their perspectives—the methods the participants used varied widely. That individuality of focus and approach underscores how personalized this work is. It also signals how important it is to protect non-standardized learning contexts, so that educators and their students can find relevance to their own learning experiences and processes within the sets of competencies:

> I think [there are] multiple ways to engage in a relationship with these competencies: they could be my jumping off point and I start with them, or they might just come out at the very end and I lay them on [top]. It depends on … the questions I'm playing with [and how] my wondering, my curiosity, leads me to connect with them.
> – School district coordinator

> One of the things that came out of this year is the different types of evidence that were used [by teachers]. [I] wouldn't want a limiting structure; [maybe a] show of options versus structure. [For example,] "You need to bring something back—here's what it could look like: narrative, survey pre- to post-, measurable 'look fors,' frequency perspective, qualitative perspective." It depends on what aspect of program and competencies you are focusing on.
> – School principal

> I determined an entry point based on student emotional needs—big glaring challenges that were not addressed—elephant-in-the-room pieces no one wanted to address [such as the student's] emotional needs. [I think] we are uncomfortable as a system in addressing [student] identity "issues."
> – Elementary school teacher

The quality of the learning opportunity

There appears to be an inextricable and dynamic link between learning conditions and the specific competencies that students express. Learning conditions frame and support the expression of specific competencies and, conversely, a focus on particular competencies in relation to teaching, learning, and assessment supports teachers in exploring a greater range of possible learning conditions and/or opportunities. In the course of field trials, the educators in the case study established classroom conditions to work on a specific competency or set of competencies. Then, in reaction to student responses, they often reframed the conditions. This opened the door to new possibilities in the learning experience.

For example, educators found that the domain of creativity provided a launching point for collaborative discussion about how to teach and scaffold creativity, and for self-reflection about learning conditions that allowed students to express specific competencies within creativity. Others puzzled out the problematic nature of risk taking and its connection to students' identity. Students had a wide array of feelings about the productiveness of taking risks in classrooms, and this diversity of opinion was often associated with their identity as students. This prompted teachers to examine the conditions in their classrooms that allowed students to express and understand their identities, thus encouraging risk-taking behaviour.

> I guess our observations have proved … that we both think there is a need for some scaffolding in order for students to begin their creative process. We wondered how much scaffolding is too much, because sometimes if you provide too much, then that shuts down the creative process. I think one big thing that we learned was that students need the vocabulary to be able to negotiate and show, or publish, their creativity and their ideas.
> – Secondary school English teacher

> In the Grade 1 and 2 classrooms, teachers were really curious about the [competencies in the] domain of creativity – specifically "students think flexibly" and "students reflect on their own thinking process," and "students are primarily driven by intrinsic motivation". That intrinsic motivation [condition] was of huge interest to the teachers. In our first conversation [they were] saying, "Is there anything I can do about that? Where does it come from? Can we actually facilitate that in the classroom? Can we change it? Does it have something to do with environment?"
> – School district coordinator

> I think [using the MWM lexicon] helped us learn about the [creativity] domain, because we were forcing ourselves to notice and name those [competencies] … Why aren't we seeing and hearing them? Why might that be? Are there conditions that we need to change or reset so that we do see more of that? When didn't we see that? Did we see examples of creativity? What was it about that moment—that lesson—that pulled that out? How can we get more of it?
> – School district coordinator

Interrelationship of competencies

From the earliest stages of MWM, it was clear that competencies related to socio-emotional learning, health, citizenship, and creativity are strongly linked. This interrelationship was also evident across the work of the educators in the study.

One striking example involved the work of two elementary educators to teach social-emotional competencies related to self-awareness and to develop the creative competency "apply metaphorical thinking" at the same time. The class used the metaphor of an iceberg to explore personal identity and relate it to the historical perspectives of Indigenous peoples in Canada. As they moved from diverse peer and personal notions of identity to their study of Indigenous identities and perspectives, given First Nations, Métis, and Inuit historical, cultural, and political realities in Canada, students developed competencies in the citizenship and social-emotional learning domains:

> It's interesting that as we explored First Nations history and current issues, the students began to talk about residential schools and foster children who are Aboriginal, and then make sense of them as individual children and their loss of language … interweaving notions between their own understanding of identity and the book [they had read for class]. They wondered how the main characters were also [like them] struggling with identity and what [the characters'] "Iceberg" might look like and what their

> own "Iceberg" looked [like]. That was amazing. The intra[-personal] connections really deepened their thinking.
> – Elementary school teacher
>
> This is a quote from one student talking about identity [reads from observation notes], "I think we've been talking about identity, because we all have identity. We're learning that some people have parts of themselves that they want to keep hidden and some people have the exact same parts that they want to reveal; reveal it all at the top of the Iceberg. The things that they want to reveal could be something like talent. Some people have a talent that they might hide all the way out at the bottom of their Iceberg."
> – Program coordinator quoting a student

The way that the competencies blended into one another within the learning experience was typical in the experience of educators in the case study. It illustrates the problematic nature of translating discrete, rational frameworks, such as MWM, into complex experiences. It also points to the importance of creating policy and institutional space to support professional autonomy, reflection, and exploration.

Common competency lexicon, or language of learning

Participants in the study noted how helpful it was to have a specific, common language with which to bind these complex areas of learning. The specific language of the competencies created opportunities for educators to communicate with each other, to generate new conditions, and to articulate different meanings out of everyday classroom experience. One program coordinator shared this idea as they reflected on the use of the competencies:

> I sensed that they [teachers on this team] felt that creativity was sort of this nebulous concept and [they thought,] "What does it look like in areas other than art?" Then I was at their school with them a couple of weeks ago and it was interesting to hear how their thinking had evolved; that actually it's not just this nebulous thing, it lives in so many different places and spaces and [the specific competencies] helped them name and unpack what it looked like and actually make connections to the curriculum.
> – Program coordinator

The use of a common language to articulate sometimes broad areas of learning, allows teachers working with different age groups and in different learning contexts to communicate with one another in more meaningful and specific ways. In one school board, a kindergarten teacher was able to engage in a discussion with two teams of secondary school teachers about her efforts to infuse intentional learning within her students' imaginative play. This led to a discussion about the tricky balance between the progressive learning experiences often defined in the curriculum and the creative competency "work without an end goal in mind," and also about whether some sort of scaffolding is required to do productive work using this and other competencies that support imaginative thinking.

These examples illustrate the value of having specific language available to support educators who are working together within these broad areas of learning, in much the same way as they would discuss specific skills when working on mathematics or reading. The language permitted more accessible entry points into discussions about pedagogy and learning theory. It allowed for cross-contextual understanding and dissonance, and provided a concrete way to discuss, plan, document, and analyze classroom micro-adjustments.

Broadened perspectives on where learning occurs in schools

A number of schools in the study explored student experiences outside the classroom. Here, perspectives on where learning occurs and what constitutes a learning experience broadened from situated moments within scheduled classroom times to student experiences throughout the school day.

One principal led her staff in shifting the use of school hallways and stairwells as age-segregated corridors to places where students from all grades represented their learning in relation to the creativity, social-emotional learning, citizenship, and health domains:

> What we've done is we've been working towards designating and labeling [different] school spaces [with each MWM domain]. Our decision was to take each area and link them. We're going to use our staircases to link them. The primary [students] took on creativity. The intermediates really wanted social-emotional. The juniors wanted citizenship, and then right in front of the gym we have our health component.
>
> We're trying to break down barriers to create a sense of community in our school. When you go down to creativity we will have examples from the intermediate [students' work] in the primary spaces. We're breaking down those paradigms of "This is kindergarten alley and this [is] kindergarten work." We want the kids to connect, but we also want the teachers to connect.
> – School principal

The interaction of school space and function, and the experience of students, staff, and community members as they used these spaces, was central to how educators approached this aspect of the student experience in schools. In these ways, exploring how all members of the school community used the school's various organizational and functional structures engendered new ways of thinking about the qualities of learning moments, and about where and how they occur.

System implications

The MWM project represents a system-wide collaboration of scholars, policy-makers, and educators (People for Education, 2016). While each partner engaged in the project in ways that were relevant to their particular interests and roles within public education, all had a hand in working through the various aspects of the process. The intention was to use research and evidence to provide a proof of concept for the argument that the creativity, social-emotional learning, citizenship, and health domains are essential to the purpose(s) of public education and, as such, need to be supported—with focus and resources—in ways that allow schools to provide the best opportunities for students to develop these critical capacities.

The work within the project also offered a deeper understanding of how a jurisdiction might approach supporting these areas in relation to ongoing school-based work and system-level measurement and accountability. The project elicited several core insights, that can be grouped into four main topics:

- assessment and measurement
- equity
- common provincial competency framework
- infrastructure and resource demands

Assessment and measurement

Assessment of these areas of learning is possible and important, but it is also complex. Based on our work with experts, educators, and policy-makers, we believe that assessment policies, resources, and infrastructure should be developed from the expertise found in schools, with an understanding that educators may require some capacity-building in their use of assessments. Assessment approaches to areas of learning beyond literacy and numeracy, at both the local and jurisdictional level, face a critical tension between a desire for systematically collected data, and the potentially negative impact of its collection (Kempf, 2016). This tension needs to be dealt with explicitly and productively in order to realize the potential that assessment may have to influence the learning experiences of the next generation.

For that reason, it is vital to first value and understand the school and classroom conditions that have been shown to enhance and deepen learning of these essential competencies, and then begin to explore jurisdictional approaches to assessment that would support or catalyze classroom and school learning, rather than inhibit it. People for Education believes there is a path through which central and local forms of assessment can be harmoniously connected, but the work is challenging. It will require sector-wide inter- and intra-organizational collaboration, coordination, resources, capacity building, trust, and the right kind of political will.

Equity

It makes good sense to design, implement, and assess the impact of any equity strategy both centrally and locally by measuring specific competencies in creativity, social-emotional learning, citizenship, and health alongside general achievement indicators. This approach:

- opens up new and different ways of thinking about learning;
- privileges the variety of ways in which students experience content and tasks;
- supports an inclusive and diverse definition of school success;
- creates links between individual assessment and the learning opportunities from which the assessment is drawn; and
- supports the recognition of student identity within the definition of success.

Common provincial competency framework

As detailed earlier, many jurisdictions are using multiple strategies, programs, policies, and initiatives for student learning that have similar goals but different names and details. A specific lexicon that defines areas of competence in creativity, social-emotional learning, citizenship, and health increases educators' ability to communicate across the system and to the general public. System-wide use of such a lexicon would build coherence of interpretation and the capacity to personalize approaches within a common understanding. Although early learning programs, postsecondary institutions, and employer organizations focus on similar areas (e.g., critical thinking, persistence, well-being, problem solving, collaboration, resilience, innovation, learning through play), by and large they operate independently of one another. Using a common lexicon would facilitate more connected operations in these policy and program environments, generating policy cross-fertilization and interaction.

When communicating with the general public about the broad purpose of education, the education sector often refers to developing citizenship or creativity. However, those big ideas are rarely described in terms of explicit competencies. A lexicon would help the public understand more concretely how the education system is building these competencies in our students and enable deeper, more meaningful communication between the education sector, the public, and other sectors concerned with the development of creativity, social-emotional learning, citizenship, and health.

Infrastructure and resource demands

System performance in public education is currently defined largely by measurable literacy and numeracy outcomes and by graduation rates, whereas education policies define a far wider range of areas as key to student success in schools (e.g., healthy peer relations, career education through portfolios and self-assessment). This tension can confound work in schools when policy mandates and suggestions expand to include a wide range of goals while the allocation of resources and infrastructure remains focused on areas of learning designed to boost student achievement on large-scale assessments (People for Education, 2018a). Such an environment is not conducive to the kind of learning conditions explored in the MWM project.

Tracking school resources and infrastructure

We have already argued that the development of competencies within the creativity, social-emotional learning, citizenship, and health domains is essential for our next generation. This paper has also emphasized the value of providing a range of learning opportunities, both in and out of school, to support balanced and deep development of these competencies. Following from these two points, any central or local policy attempt to address the first—changing, broadening, or increasing the expectations placed on schools to address these areas as outcomes for students—will not be fully successful unless it also includes careful consideration of the resources and infrastructure available to do so.

Do schools in Ontario currently have the resource and infrastructure capacity to deliver on the imperatives about student learning articulated by the MWM project? This is the central issue of the third part of our analysis, which is based on the Annual Ontario School Survey distributed to all elementary and secondary schools in the fall of each year.

For 22 years, People for Education has conducted the Annual Ontario School Survey to keep track of the impact of policy and funding changes on resources and programs in schools. In 2018, 22 per cent of Ontario's nearly 5,000 schools responded to the survey. The schools in this sample are an adequate representation of the regional diversity of Ontario. The survey serves as the basis for an annual report in which People for Education describes the infrastructure and resources that Ontario schools have in a wide range of areas: the arts, libraries, special education, health, careers, and Indigenous education, to name a few (e.g., People for Education, 2018a).

As detailed earlier, competencies within the MWM project are inextricably linked to the conditions in which students access and use them while learning. Equally, teachers approach the competencies in diverse ways in different subjects and grades. These factors underscore that schools need the capacity to offer varied and diverse learning experiences and opportunities, supported by infrastructure. For example, arts education can significantly support the development of competencies in creativity, social-emotional learning, and citizenship (Deasy, 2002; Shanker, 2014; Upitis, 2014). A well-resourced library, access to guidance counsellors, and school fundraising to support students' experiential learning might offer different but equally important learning opportunities (e.g., Alberta Education, 2009). Individual schools' capacity to offer these opportunities are one way to examine a system's capacity to support the development of the competencies focused on above. Systematic tracking of these opportunities and resources may offer a viable proxy measure of system capacity. For the purposes of this study, we examine four areas:

- the school principal
- guidance counsellors
- education in the arts
- school fundraising

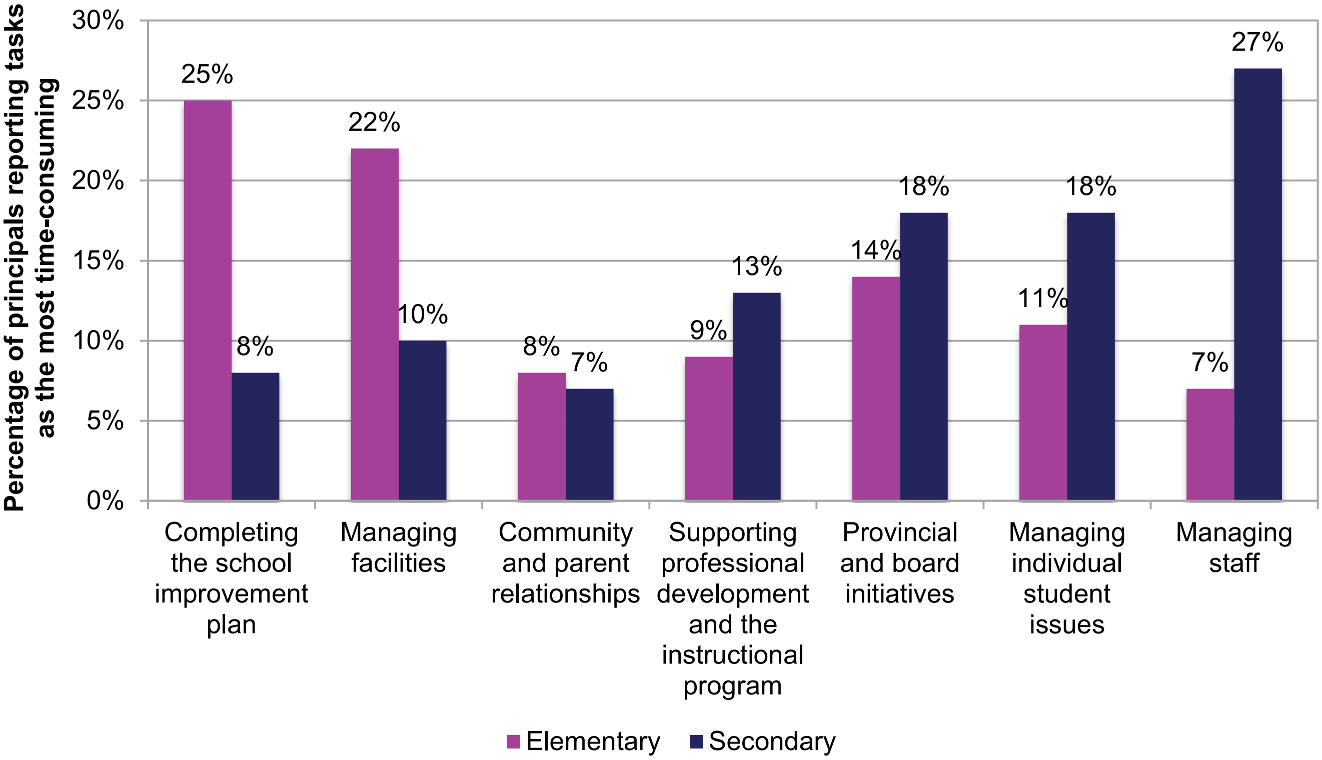
The school principal

Out of all school-related factors beyond teachers, school principals have the highest impact on students' education (Leithwood, Seashore Louis, Anderson, & Wahlstrom, 2004). They are expected to lead school improvement and instruction efforts, attend to individual student needs, support implementation of Ministry and school board initiatives, support special education programs and students, supervise staff and human resources, attend to school facilities, and communicate with parents and the community (Pollock, 2014; McCarthy, 2016). However, in Ontario, principals' role has expanded far beyond leading a school's instruction and learning program.

As a secondary school principal in the Limestone District School Board put it, "We spend so much time managing significant mental health issues with students and managing staff and facility issues that I never feel we can move ahead with the student learning agenda." This principal's perspective seems relevant when mapping what all elementary school principals report as their most time-consuming tasks (Figure 1).

Only 9 per cent of elementary and 13 per cent of secondary school principals report that their most time-consuming task is supporting professional learning and improving the instructional program. This is consistent with recent research showing an overload of administrative work for principals in schools across Canada (Alberta Teachers' Association, 2014; Leithwood & Azah, 2014).

FIGURE 1 Most time-consuming tasks for Ontario school principals by grade level



These findings are especially relevant when considering the importance of leadership and support in a school in relation to the competencies under discussion in this paper. Teachers participating in the MWM project were supported with the time and space to work on integrating the competencies they had selected into existing curriculum. Allowing teachers the licence to direct their own professional development in these areas and supporting their collaboration within a whole school strategy requires leadership attention and the capacity to focus on the school's instructional program. Leadership by the principal can go a long way in authorizing teachers' judgment and decision making, as well as in supporting the kind of ongoing peer collaboration that optimizes work on these areas of student learning (Schmoker, 2006).

Principal support is thus a centrally important organizational capacity. However, in 2018, more than 20 per cent of Ontario elementary school principals reported that managing school facilities took up the largest proportion of their time. Many principals reported feeling stretched too thin and not having the capacity to attend to the variety of student needs within their building. This coincides with studies on the role of the Canadian principal that report they work an average of 59 hours per week, and that they are feeling increasingly overwhelmed by the breadth of their responsibilities (e.g., Alberta Teachers' Association, 2014).

Guidance counsellors

Guidance counsellors can play a key role with respect to a variety of competencies within social-emotional learning and health. They support students in planning transitions, seeking help for mental health issues, learning to act as advocates for themselves in connection with special education needs, and developing educational pathways that will lead to long-term success (People for Education, 2018a). All of these competencies are critical for student development in social-emotional learning and health. The work of guidance counsellors to help students learn to become active agents in their educational careers and beyond is also central to developing their civic engagement and agency (Sears, 2014).

Principals report that the role of guidance counsellors is multifaceted and varies across the province. Generally, guidance counsellors support students with career and life planning, although there is no set job description in Ontario. In 2018, 26% of Ontario secondary schools reported that the most time-consuming task for guidance counsellors was providing one-on-one counselling to students for mental health needs. Another 4% reported that providing behaviour-related interventions took the most time for guidance counsellors.

Principals also report that they do not have enough guidance staff to support their students, especially in the elementary panel. In 2017, only 14 per cent of Ontario elementary schools had guidance counsellors and the majority were part time. In elementary schools with Grades 7 and 8—where students are preparing to make the transition to secondary school—only 20 per cent had guidance counsellors and the majority were part time. Among elementary schools with guidance staff, counsellors were scheduled for an average of 1.5 days per week. In secondary schools, the average ratio per school of students to guidance teachers was 396:1, but in 10 per cent of schools the average ratio was 826:1.

Education in the arts

Students can develop a number of the MWM competencies in all four domains through arts education. Its importance to student learning is documented in fields related to the MWM competencies, such as improving spatial reasoning (Hetland & Winner, 2001) and deepening motivation for learning (Deasy, 2002). Most significantly, arts education has the potential to enrich students' creativity and social development (Hunter, 2005). These two qualities are closely associated with two competency domains in the MWM project and included in the Ministry of Education's 21st-century competencies (Government of Ontario, 2015).

Despite the widely recognized importance of arts education, equitable access to arts programs and resources is an ongoing challenge in Ontario. While some schools offer many extracurricular arts activities, students in small and rural schools, in schools with higher levels of poverty, and in schools with lower levels of parental education are less likely to have access to learning opportunities in the arts (Figures 2 and 3). A high parental education indicates at least one parent has a university education, low parental education indicates no university education among parent(s). Budgets for the arts in Ontario schools ranged from less than $500 to almost $5,000 in 2018 and there was a relationship between schools' arts budgets and their fundraising capacity.

Collectively, these statistics suggest significant differences in the capacity of elementary schools to deliver learning opportunities in the arts, based on the context in which each school finds itself. This is especially problematic when considered alongside the varying capacity of schools to raise money in order to support arts programs.

FIGURE 2 Elementary school access to specialist art teachers, by school location

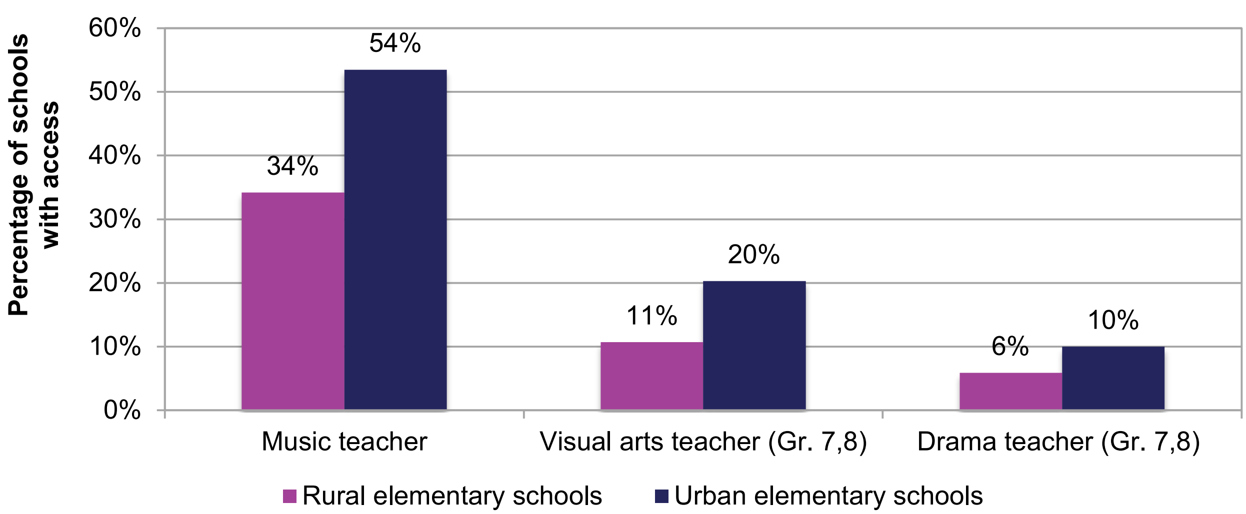
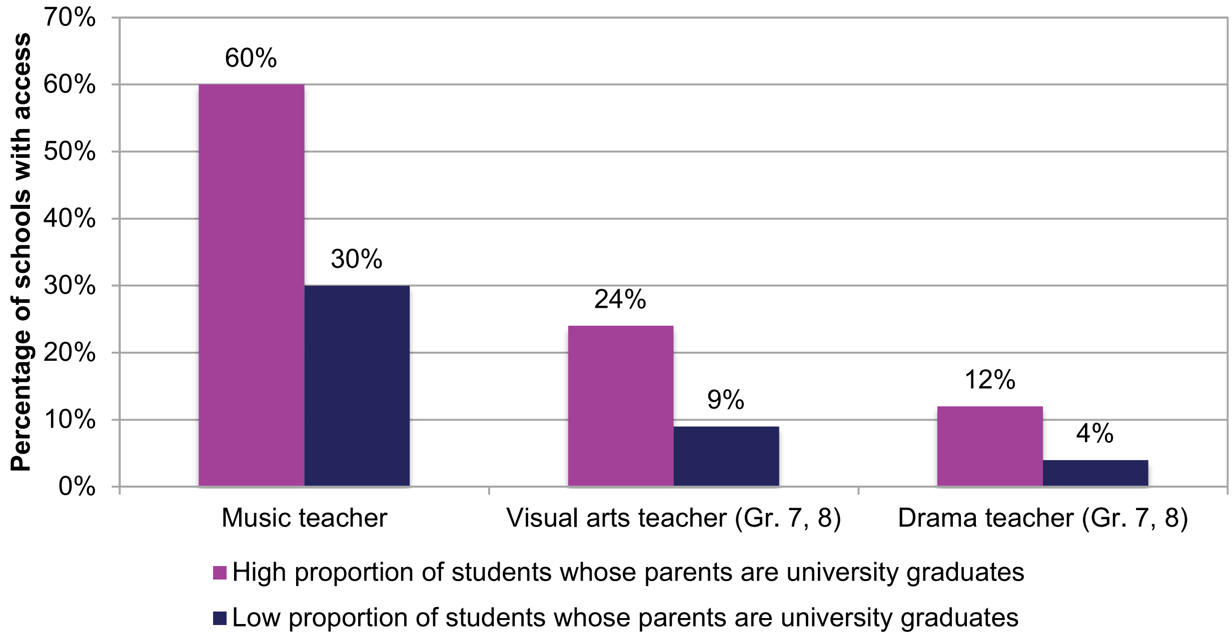


FIGURE 3 Elementary school access to specialist art teachers, by students' parental education



School fundraising

In 2018, 99 per cent of elementary and 87 per cent of secondary schools reported that they fundraised (People for Education, 2018b). While fundraising is not often thought of as a vital resource for core learning opportunities within public education, the money schools raise helps to meet needs related to the arts, classroom materials, experiential learning out of the school, technology, libraries, and sports and recreation (Figure 3). In this way, school fundraising can support the diverse range of learning opportunities that the MWM project has shown to be critical in optimizing the ability of students from kindergarten to Grade 12 to develop creative, social-emotional, health, and citizenship capacity.

However, inequities related to fundraising have been a persistent issue in Ontario (Winton, 2016), and these inequities appear to be growing. Currently, the top 10 per cent of fundraising elementary schools raise 37 times the amount raised by the bottom 10 per cent, and some schools report raising as much as $123,000. Among secondary schools, the top 5 per cent of fundraising schools raise as much as the bottom 81 per cent combined, and some schools report raising $150,000 (People of Education, 2018b).

Conclusion

It is widely accepted that it is vital for students to develop creative, social-emotional, citizenship, and health competencies alongside their academic outcomes (e.g., OECD, 2016). To support further exploration of the implications of this view, this paper draws on two studies to describe how both educators and school systems can support students' development of these competencies.

The paper highlights the various approaches educators can take in addressing similar competencies; the importance of respecting educators' judgment while also providing resources for collaborative peer support; and the need to be judicious about what educators can do in addition to their existing teaching and leadership demands.

The themes outlined in the paper have a number of implications, including challenges to some of our current approaches to assessment; a requirement for evolution in how we think about equity and system coherence; and suggestions of ways we could track the provision of learning opportunities that are connected to the development of the competencies articulated through the MWM project.

Analogous to the use of large-scale student performance data as a proxy measure of a public education system's overall effectiveness (which can create its own problematic consequences), the proposition here is that counting specifically selected opportunities and resources in schools may be an effective way to measure whether or not the school system is able to offer the learning opportunities required to optimize the development of these competencies for long-term success.

The study uses Ontario as an example to describe the disparities that often exist between schools in relation to infrastructure and resources. These disparities mean that not all schools can offer a diverse range of opportunities for students to develop competencies in creativity, social-emotional learning, citizenship, and health, regardless of social or economic background. These competencies are all teachable and learnable skills that can be developed given the right opportunities. Providing those opportunities may prove to be the real challenge.

Appendix A: Selected examples of activities and their relevant domains

Examples from schools	Relevant domains from project
In a secondary-school applied program, a mathematics and English teacher team explored the roles of conditions and constraints on idea generation in mathematics and English. The teachers set a variety of conditions through which students generated new ideas related to the topics at hand. They observed each other's classes and collected information from the students' responses.	Creativity Social-emotional learning
One elementary school principal in northern Ontario brought a school-wide focus on health to her community. Every morning began with an "active start," when students kicked the day off by participating in facilitated games and activities in the gym rather than by listening to school announcements in their classrooms.	Health Social-emotional learning
An elementary-school principal worked with staff and students to co-write a school code of conduct that centred on three themes: respect, engagement, and responsibility. From there, teachers worked in professional learning community groups to develop units and learning environments that showcased social-emotional learning competencies.	Social-emotional learning
A teacher stopped providing letter grades in her science class, instead focusing on rich descriptive feedback based on the competencies.	Social-emotional learning
A school coach and a teacher focused on student self-awareness and diversity of perspectives through an exploration of residential school experiences for First Nations students.	Citizenship Social-emotional learning

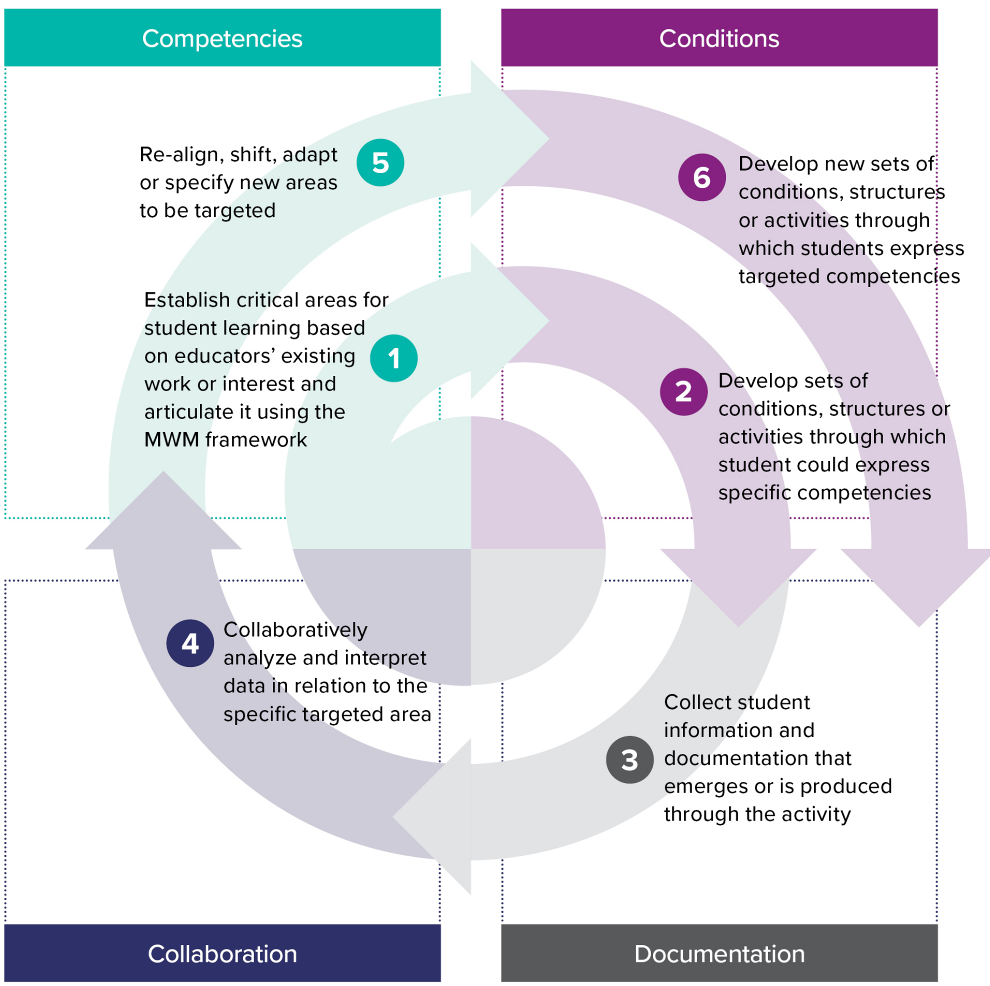
Appendix B: Process of educator teams in using MWM competencies

The process

While the composition of each educator team in the research was different in terms of grades and subjects taught, the process they undertook followed fairly consistent phases, as represented in Figure B1.

Step 1 Establish essential areas for student learning based on educators' existing work or interest, and articulate it using the MWM framework.
Step 2 Develop sets of conditions, structures, or activities through which students can express the specific competencies.
Step 3 Collect the student information and documentation that emerged or was produced through the activity.
Step 4 Collaboratively analyze and interpret the data in relation to the targeted area.
Step 5 Re-align, shift, adapt, or specify new areas to be targeted.
Step 6 Develop new sets of conditions, structures, or activities through which students express targeted competencies.

FIGURE B1 Diagrammatic representation of process



During this process, the authors met with the educators for at least two individual interviews or focus groups. Using these meetings, artifacts, and classroom observations (for a limited number of educators), the researchers attempted to answer the research questions.

Method: The activities

Each school and field trial team designed and implemented a varied and personalized set of activities, and these were integrated into their ongoing work rather than supplementing it. Activities ranged from math and drama collaborations to school-wide learning walks.

While the variety of activities demonstrates the broad application of the competencies to learning experiences for students in schools, the diversity of contexts in which the competencies were used also presents a challenge. The study addressed this issue by using a common process of school and classroom assessment that is outlined in Growing success: Assessment, evaluation, and reporting in Ontario schools (OME, 2010).

Specifically, educators focused on formative assessment, gathering information in diverse ways in order to provide feedback and re-establish conditions through which students could progress in specific learning tasks, competencies, or concepts (Black, Harrison, Lee, Marshall, & William 2004). Formative assessment of these competencies occurred both at the individual level and in collaborative teams, depending on the school or school board site. The process did not take the form of a discrete moment of assessment but instead blended into educators' planning and teaching. Unlike assessment in which teachers analyze a set of data or information gathered at one time, this type of assessment involves adaptations in classroom or school conditions and activities in response to student learning.

The type of information gathered by participating educators was diverse but can be categorized into three large areas, also outlined in Growing Success: observations, conversations, and student products (p. 39). The following is a sample of sources of evidence used by participating educators:

- notes taken while observing student interactions
- student work products
- student self-reports
- video recordings
- formal and informal conversations between teachers and students
- small-group discussion between and among teachers and students

References

Alberta Education. (2009). *Career and technology studies: Program philosophy and rationale.* Edmonton, AB: Government of Alberta. Retrieved from https://education.alberta.ca/media/482254/philosophy.pdf

Alberta Teachers' Association. (2014). *The future of the principalship in Canada*. Edmonton: Alberta Teacher Association. Retrieved from https://www.teachers.ab.ca/SiteCollectionDocuments/ATA/Publications/Research/The%20Future%20of%20the%20Principalship%20in%20Canada.pdf

Bascia, N. (2014). *The school context model: How school environments shape students' opportunities to learn*. Report prepared for Measuring What Matters, People for Education, November 8. Toronto.

Black, P., Harrison, C., Lee, C., Marshall, B., & William, D (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, *86*(1), 8–21. doi:10.1177/003172170408600105.

British Columbia Ministry of Education. (2018). *BC's new curriculum*. Victoria, BC: Queen's Printer. Retrieved from https://curriculum.gov.bc.ca

Cameron, D., Watkins, E., & Kidder, A. (2015). *Measuring what matters 2014–15: Moving from theory to practice*. Report prepared for Measuring What Matters, People for Education, November 3. Toronto. Retrieved from http://peopleforeducation.ca/measuring-what-matters/wp-content/uploads/2015/12/P4E-MWM-From-Theory-to-Practice-2014-15.pdf

Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, & Social Sciences and Humanities Research Council of Canada. (2014, December). *Tri-council policy statement: Ethical conduct for research involving humans*. Retrieved from Government of Canada Web site: http://www.pre.ethics.gc.ca/eng/policy-politique/initiatives/tcps2-eptc2/Default

Deasy, R.J. (2002). *Critical links: Learning in the arts and student academic and social development.* Washington, DC: Arts Education Partnership.

Ferguson, B., & Power, K. (2014). *Broader measures of success: Physical and mental health in schools*. Report prepared for Measuring What Matters, People for Education, November 8. Toronto.

Gouvernement du Québec. (2007). *Québec education program (secondary cycle two). Chapter 3: Cross-curricular competencies.* Québec City, QC: Gouvernement du Québec.

Government of Ontario (2015). 21st Century Competencies. Toronto, ON: Queen's Printer for Ontario.

Hatch, J.A. (2002). *Doing qualitative research in education settings.* Albany, New York: State University of New York Press.

Hetland, L., & Winner, E. (2001). The arts and academic achievement: What the evidence shows. *Arts Education Policy Review*, *102*(5), 3–6.

Hunter, M. A. (2005). *Education and the arts research overview*. Sydney: Australia Council for the Arts.

Kempf, A. (2016). *The pedagogy of standardized testing: The radical impacts of educational standardization in the US and Canada.* London, UK: Palgrave Macmillan.

Leithwood, K, Seashore Louis, K., Anderson, S. & Wahlstrom, K. (2004). *How leadership influences student learning.* Toronto: The Center for Applied Research and Educational Improvement (CAREI).

Leithwood, K., & Azah, V. N. (2014). *Elementary principals' and vice-principals' workload study: Final report*. Toronto: Ontario Ministry of Education. Retrieved from http://www.edu.gov.on.ca/eng/policyfunding/memos/nov2014/FullElementaryReportOctober7_EN.pdf

McCarthy, T. (2016). *The role of the secondary vice-principal and its relationship with the enactment of the Ontario leadership framework* (Doctoral thesis). The University of Western Ontario, London, ON.

Morgan, D. L. (2008). Emergent design. In L. M. Given (Ed.), *The SAGE encyclopedia of qualitative research methods.* Retrieved from http://sk.sagepub.com/reference/research/n128.xml

Ontario Ministry of Education. (2010). *Growing success: Assessment, evaluation, and reporting in Ontario schools*. Retrieved from the Ministry of Education Web site: http://www.edu.gov.on.ca

Ontario Ministry of Education. (2016a). *Ontario's well-being strategy for education: Fact sheet for parents*. Retrieved from the Ministry of Education Web site: http://www.edu.gov.on.ca/eng/about/WBFactSheet.html

Ontario Ministry of Education. (2016b). *21st century competencies: Foundation document for discussion.* Retrieved from http://www.edugains.ca/resources21CL/About21stCentury/21CL_21stCenturyCompetencies.pdf

Organisation for Economic Co-operation and Development. (2016). *Global competency for an inclusive world.* Paris: OECD Publishing.

Organisation for Economic Co-operation and Development. (2018). *The future of education and skills. Education 2030. The future we want.* Paris: OECD Publishing.

Ozga, J. (2005). *Models of policy-making and policy learning: ESRC research project on education and youth transitions in England, Wales and Scotland, 1984–2002*. Paper presented at the Seminar on Policy Learning in 14–19 Education, a joint seminar of the Education and Youth Transitions Project and Nuffield Review of 14–19 Education, 15 March. Retrieved from http://www.ces.ed.ac.uk/PDF%20Files/EYT_SP05_JO.pdf

People for Education. (2016). *Measuring what matters: Phase 3 progress report.* Toronto, ON: People for Education.

People for Education. (2017). *Competing Priorities: Annual Report on Ontario's Publicly Funded Schools 2017.* Toronto, ON: People for Education.

People for Education. (2018a). *The new basics for public education.* Toronto, ON: People for Education.

People for Education. (2018b). *Fundraising and fees in Ontario's schools.* Toronto, ON: People for Education.

Pollock, K. (2014). *The changing nature of principals' work: Final report*. Toronto, ON: Ontario Principals' Council.

Schmoker, M. (2006). *Results now: How we can achieve unprecedented improvements in teaching and learning*. Alexandria, VA: Association for Supervision and Curriculum Development.

Sears, A. (2014). *Measuring what matters: Citizenship domain*. Report prepared for Measuring what matters, People for Education, November 8. Toronto. Retrieved from http://peopleforeducation.ca/measuring-what-matters/domain/democracy

Sergiovanni, T. J. (2007). Leadership as stewardship. In *Rethinking leadership: A collection of articles*, 49–60.

Shanker, S. (2014). *Broader measures for success: Social/emotional learning*. Report prepared for Measuring What Matters, People for Education, 8 November. Toronto.

Upitis, R. (2014). *Creativity: The state of the domain*. Report prepared for Measuring What Matters, People for Education, 8 November. Toronto.

Winton, S. (2016). The normalization of school fundraising in Ontario: An argumentative discourse analysis. *Canadian Journal of Educational Administration and Policy*, *180*, 1–32.

PISA 2015: Do Canadian Students Have the Right Skills and Attitudes toward Science?

Marie-Anne Deussing
Analyst, Employment and Social Development Canada

Tanya Scerbina
Manager, International Assessments, Council of Ministers of Education, Canada

Yitian Tao
Analyst, Council of Ministers of Education, Canada

Abstract

This paper uses the 2015 Programme for International Student Assessment (PISA) to explore gender differences in students' attitudes toward science, beliefs in their scientific abilities, and scientific-related career expectations in Canada. In general, boys are more likely than girls to enjoy learning about science, are more interested than girls in physics and chemistry, and are more confident in their ability to perform science-related tasks. On the other hand, girls are more interested than boys in health-related topics and are more likely to perceive learning science as being useful to them and to their future studies and careers. As well, Canadian girls are generally more likely to expect to work in a science-related field than boys, although this is true most particularly among girls who are considered to be low or moderate performers in science. Top-performing girls are less likely than top-performing boys to expect a career in a science-related field. Girls are more likely than boys to expect to be working as health professionals or technicians or associate professionals in a science-related field, while boys are more likely than girls to expect to work as science and engineering professionals or information and communication technology (ICT) professionals.

Introduction

Women and men have historically worked in different occupations, and this is especially true when it comes to science, technology, engineering, and mathematics (STEM) occupations. Despite a gradual shift in recent decades, gender disparities in STEM fields persist. Given that the Canadian girls perform as well as Canadian boys in science at age 15, based on the Programme for International Student Assessment (PISA) 2015 science results, it is worth examining whether this gender disparity may be linked to differences between girls and boys in their feelings about science.

This paper examines whether disparities exist early on in the attitudes of boys and girls toward science, their beliefs about their abilities in science, and their expectations of future study and work in science- and technology-related fields. PISA is ideal for this purpose, as it not only provides data on the science skills of 15-year-old students but also measures a range of factors related to their attitudes toward science through questions about enjoyment of science, interest in broad science topics, and instrumental motivation to learn science. PISA also measures science self-efficacy—the extent to which students believe in their own ability to handle science tasks effectively and overcome difficulties—as well as student expectations of having a science-related career.

Attitudes towards science

Nurturing motivation and interest in science at the critical age when students begin to think about their future careers is believed to be important in influencing their choice of occupations that involve science or science-based technology (Organisation for Economic Co-operation and Development [OECD], 2008). In PISA, two forms of motivation to learn science are measured: students may learn science because they enjoy it, and/or because they perceive learning science to be useful for their plans.

Enjoyment of science

PISA measures students' enjoyment of science through their responses ("strongly agree," "agree," "disagree," or "strongly disagree") to statements affirming that they generally have fun when learning science topics; that they like reading about science; that they are happy working on science topics; that they enjoy acquiring new knowledge in science; and that they are interested in learning about science. The index of enjoyment of science was constructed to summarize students' answers, with higher values on the index reflecting greater levels of agreement with these statements.[1]

Canada places among the top in terms of student enjoyment of science when compared to other OECD countries.[2] More specifically, 75 per cent of students in Canada report having fun when learning science topics, 63 per cent enjoy reading about science, 69 per cent are happy working on science topics, 79 per cent enjoy acquiring new knowledge in science, and 79 per cent are interested in learning about science. All these shares are higher than the OECD average. At the provincial level, students' enjoyment of science is highest in Alberta and lowest in Saskatchewan and Manitoba, with every province reporting a higher enjoyment of science than the OECD average.

Between 2006 and 2015, students' enjoyment of science improved in 16 of the 35 OECD countries. In Canada, the index of enjoyment of science increased, with more students in 2015 than in 2006 reporting that they enjoy reading about science, enjoy acquiring knowledge in science, and are interested in learning about science. Almost all provinces observed an improvement in their students' enjoyment of science over the period, with the largest improvement observed in Prince Edward Island and the smallest improvement observed in New Brunswick. Newfoundland and Labrador is the only exception, with no statistically significant change.

On average across OECD countries, boys are more likely than girls to agree with each of the statements that make up the index of enjoyment of science. In Canada, a small but statistically significant difference exists between the shares of boys and girls who have fun when learning science topics, who enjoy acquiring new knowledge in science, and who are interested in learning about science. A more sizable difference exists between the shares of boys and girls who report reading about science, as well as those who are happy working on science topics. More specifically, relative to girls, boys are 7 percentage points more likely to like reading about science (66 per cent versus 59 per cent) and 6 percentage points more likely to be happy working on science topics (72 per cent versus 66 per cent). Both boys and girls in Canada observed an improvement in their enjoyment of science between 2006 and 2015.

Interest in broad science topics

Interest in science may be one of the reasons why students enjoy learning about science. PISA measures the extent to which students are interested in five broad science topics:

- the biosphere (e.g., ecosystem services, sustainability)
- motion and forces (e.g., velocity, friction, magnetic and gravitational forces)
- energy and its transformation (e.g., conservation, chemical reactions)
- the universe and its history
- how science can help us prevent disease.

Possible student responses are "not interested," "hardly interested," "interested," or "highly interested."

Across most countries and economies, including Canada, students prefer topics related to disease prevention and astronomy. More specifically, 74 per cent of students in Canada are interested or highly interested in disease prevention

---

[1] Indices were constructed on which the average OECD student (e.g., the student with an average level of interest) was given an index value of 0 and on which about two-thirds of the OECD student population were between the values of –1 and 1 (i.e., the index has a standard deviation of 1). Therefore, students in countries with negative mean index values did not necessarily respond negatively to the underlying questions. Rather, they responded less positively than students on average across OECD countries. Likewise, in countries with positive mean index values, students responded more positively than the OECD average. For indices that were based on trend questions from 2006, the derived variable was equated to the corresponding scale in the PISA 2006 database. As a result, the OECD mean may not be exactly 0.

[2] Some care must be taken when comparing index scores across countries, as students in different countries may not always mean the same thing when answering questions about matters such as interest in science.

and 69 per cent are interested in the universe and its history, whereas 63 per cent are interested in energy and its transformation, 55 per cent are interested in motion and forces, and 53 per cent are interested in topics related to the biosphere. Across OECD countries, students on average have a lower interest than students in Canada in all five topics. Two out of three students (66 per cent) on average across the OECD countries are interested in how science can prevent disease and in the universe and its history, while fewer than half are interested in energy and its transformation, motion and forces, and topics related to the biosphere.

An overall measure of interest in science was also constructed from students' reported interest in the five broad science topics, with higher values on this index reflecting greater interest. Relative to other OECD countries, the interest of students in Canada in broad science topics is surpassed only by that of students in Mexico. At the provincial level, the index of interest in broad science topics is highest in British Columbia and Alberta and lowest in Saskatchewan, Manitoba, and New Brunswick. The share of students in Alberta interested or highly interested in all five specific science topics is the highest among all provinces.

Across OECD countries, boys are more interested than girls in physics and chemistry (motion and forces, energy and its transformation), while girls tend to be more interested in health-related topics (how science can help us prevent disease). In Canada, 66 per cent of boys are interested in motion and forces and 69 per cent are interested in energy and its transformation, compared to 45 per cent and 58 per cent of girls, respectively. On the other hand, 69 per cent of Canadian boys are interested in how science can help us prevent disease, compared to 78 per cent of girls. Although still statistically significant, gender differences are small with respect to topics related to the biosphere or to the universe and its history.

Instrumental motivation to learn science

Instrumental motivation refers to the drive to learn science because students perceive it to be useful to them and to their future studies and careers. PISA measures the extent to which students think that science is relevant to their study and career prospects through responses ("strongly agree," "agree," "disagree," or "strongly disagree") to the following statements:

- Making an effort in my school science subject is worth it because this will help me in the work I want to do later on.
- What I learn in my school science subject is important for me because I need this for what I want to do later on.
- Studying science at school is worthwhile for me because what I learn will improve my career prospects.
- Many things I learn in my school science subject will help me to get a job.

In general, most Canadian students recognize the instrumental value of studying science as a way to improve their career prospects and work in their desired field. At least three out of four students agree or strongly agree with all four statements affirming the importance of school science for their future. Furthermore, based on the index of instrumental motivation to learn science that was constructed to summarize responses to these four statements, students in Canada have a higher instrumental motivation to learn science than students in most other OECD countries, with the mean index of instrumental motivation higher only in Mexico. At the provincial level, students in Newfoundland and Labrador and in Saskatchewan have the highest instrumental motivation to learn science, while students in Manitoba have the lowest.

Two of the four items used in PISA 2015 to measure instrumental motivation to learn science are identical to those in PISA 2006. As a result, it is possible to examine changes over the period in students' perceptions of the usefulness of high-school science courses for their future studies and careers. Students in most OECD countries, including Canada, were more likely to perceive science as useful in 2015 than in 2006, with only Chile, Germany, and Poland observing a decline in their students' instrumental motivation. Provincially, the instrumental motivation to learn science increased among students in Quebec, Saskatchewan, Ontario, and British Columbia, was unchanged for students in Nova Scotia, Newfoundland and Labrador, Alberta, and New Brunswick, and decreased among students in Prince Edward Island.

On average across OECD countries, the index of instrumental motivation to learn science is significantly higher among boys than among girls. In Canada, however, the opposite is the case, as girls are more likely than boys to strongly agree to all four statements that make up the index. Furthermore, in Canada comparing 2015 results to 2006 results, both more boys and more girls agreed or strongly agreed that making an effort in science subjects at school is worth it because it will help them in the work they want to do and improve their career prospects.

Science self-efficacy

Science self-efficacy refers to future-oriented judgments about one's competency in accomplishing particular goals in a context that requires scientific abilities (OECD, 2016). To measure science self-efficacy, PISA 2015 asked students to report how easy it would be for them to do the following:

- Recognize the science question that underlies a newspaper report on a health issue.
- Explain why earthquakes occur more frequently in some areas than in others.
- Describe the role of antibiotics in the treatment of disease.
- Identify the science question associated with the disposal of garbage.
- Predict how changes to an environment will affect the survival of certain species.
- Interpret the scientific information provided on the labelling of food items.
- Discuss how new evidence can lead them to change their understanding of the possibility of life on Mars.
- Identify the better of two explanations for the formation of acid rain.

For each of these, students could report that they "could do this easily," "could do this with a bit of effort," "would struggle to do this on my own," or "couldn't do this."

In Canada, fewer than one in four students can easily describe the role of antibiotics, identify the science question associated with garbage disposal, discuss how new evidence can change understanding of the possibility of life on Mars, or identify the better of two explanations for acid rain. Across the OECD countries, even fewer students can easily perform these tasks. Explaining why earthquakes occur more frequently in some areas and predicting how changes to an environment will affect the survival of certain species are the tasks considered easy by the highest proportion of students in Canada (36 per cent), while across the OECD countries, 34 per cent of students on average could easily explain earthquake frequency and 23 per cent could easily explain how environmental changes will affect the survival of certain species.

Students' responses were used to create the index of self-efficacy. Responses were reverse coded, so that the higher values of the index correspond to higher levels of science self-efficacy. Based on this index, the science self-efficacy of students in Canada is the highest of all OECD countries. Provincially, students in Alberta, British Columbia, and Ontario are the most confident in their science abilities, whereas students in New Brunswick are the least confident. In every province, students are more confident than students on average across OECD countries.

Between 2006 and 2015, the average science self-efficacy index remained stable across OECD countries, although students were more likely to report that they could easily describe the role of antibiotics and less likely to report that they could easily interpret the scientific information on food labels. In Canada, students' science self-efficacy improved over the period, with students in 2015 more likely than students in 2006 to report that they can easily describe the role of antibiotics in the treatment of disease, explain why earthquakes occur more frequently in some areas, recognize the science question that underlies a newspaper report on a health issue, and discuss how new evidence could lead them to change their understanding of the possibility of life on Mars. At the same time, students in 2015 are less likely than students in 2006 to report that they can easily predict how environmental changes will affect the survival of certain species or interpret the scientific information on food labels.

In Canada, girls are more likely than boys to have low science self-efficacy, and the gap is larger than the average observed across OECD countries. Interestingly, the gender gap in self-confidence depends on the type of problem students encounter. In Canada, boys are more likely than girls to report that they can easily discuss how new evidence can change their understanding of the possibility of life on Mars, identify the better of two explanations for acid rain, explain why earthquakes occur more frequently in some areas, recognize the science question that underlies a newspaper report on a health issue, identify the science question associated with the disposal of garbage, and interpret the scientific information on food labels. On the other hand, girls are as likely as boys to report that they can easily describe the role of antibiotics in the treatment of disease and predict how changes in the environment will affect the survival of certain species.

In general, both Canadian boys and Canadian girls in 2015 have more confidence about performing certain scientific tasks than their peers in 2006. In 2015, a higher share of boys can easily recognize the science question that underlies a newspaper report on a health issue, describe the role of antibiotics, identify the science question associated with garbage disposal, and discuss how new evidence could change their understanding of the possibility of life on Mars. A higher share of girls in 2015 can easily explain why earthquakes occur more frequently in some areas and describe the role of antibiotics in the treatment of disease, although lower shares are able to predict how environmental changes will affect the survival of certain species and interpret the scientific information on food labels.

Career expectations

PISA 2015 asked students what occupation they expect to have when they will be 30 years old. Students could enter any job title or description in an open-entry field, and their answers were classified according to the 2008 edition of the International Standard Classification of Occupations, or ISCO-08 (International Labour Office, 2012). The coded answers were used to create an indicator of science-related career expectations. These are defined as career expectations that require the study of science beyond compulsory education, typically in formal tertiary education. Within science-related occupations, four major groups are identified: science and engineering professionals, health professionals, science technicians and associate professionals, and information and communication technology (ICT) professionals.

Canada has a relatively high share of students expecting a science-related career (34 per cent) compared to the OECD average (24 per cent). In fact, among OECD countries, only Mexico (41 per cent), the United States (38 per cent), and Chile (38 per cent) have a higher percentage of students expecting a science-related career. At the provincial level, not much variation is observed; the share of students planning careers in science-related fields ranges from 31 per cent in Manitoba to 37 per cent in Newfoundland and Labrador.

In terms of specific science-related career categories, 12 per cent of students in Canada expect to work as science and engineering professionals compared to an OECD average of 9 per cent, with only Mexico (18 per cent), Chile (18 per cent), and Turkey (17 per cent) observing a higher share of students with this expectation. At the provincial level, the share varies from less than 10 per cent in Prince Edward Island, Nova Scotia, Manitoba, and Saskatchewan to 14 per cent in Alberta.

Likewise, Canada has a higher share of students expecting to become health professionals (19 per cent) than the OECD average (12 per cent), surpassed only by the United States (21 per cent). At the provincial level, Newfoundland and Labrador has the highest proportion of students expecting careers as health professionals (22 per cent), while Alberta has the lowest (16 per cent).

Compared to the average across OECD countries (3 per cent), a similar share of students in Canada (2 per cent) expect careers as ICT professionals. At the provincial level, the proportion varies from 1 per cent of students in Newfoundland and Labrador, Nova Scotia, New Brunswick, and Manitoba to 3 per cent in Ontario.

A career as a science-related technician or associate professional is the least popular science-related occupation, with less than 1 per cent of students in Canada and around 2 per cent of students on average across OECD countries expecting to work within this category.

PISA 2015 marks the second time that the question about career expectations was asked of all students, making it possible to examine changes in students' expectations of a science-related career between 2006 and 2015. On average across OECD countries, the share of students who expect to be working in a science-related occupation at age 30 increased from 21 per cent to 24 per cent between 2006 and 2015, mainly because the share of students planning careers as health professionals rose 3 percentage points. In Canada, the share of students expecting to pursue science-related occupations increased by 4 percentage points, from 30 per cent in 2006 to 34 per cent in 2015.[3] More specifically, the share of Canadian students expecting careers as science and engineering professionals increased by 1 percentage point, and the share of those expecting careers as health professionals rose by 4 percentage points. No statistically significant change is observed over the period in the share of students expecting careers in ICT or as science-related technicians and associate professionals.

On average across OECD countries, there is a small but statistically significant difference between boys and girls expecting to work in a science-related field: some 25 per cent of boys and 24 per cent of girls. In Canada, however, a higher proportion of girls (37 per cent) than boys (31 per cent) expect a science-related career. Interestingly, when looking at career expectations by student proficiency in science, among low (below Level 2) and moderate (Level 2 or 3) performers, a higher share of girls than boys expect to work in a science-related career (see OECD, 2016, Table I.3.10c). More specifically, 24 per cent of low-performing girls expect a science-related career, compared to 12 per cent of boys. Among moderate performers, the proportions are 34 per cent and 25 per cent, respectively. On the other hand, among students considered top performers in science (Level 5 or above), 53 per cent of boys expect to work in science-related occupations, compared to 45 per cent of girls.

Provincially, statistically significant differences in career expectations between boys and girls are observed in all provinces except Quebec, Alberta, and Newfoundland and Labrador. The difference is especially pronounced in Prince Edward Island, where girls are almost twice as likely as boys to expect a science-related career (47 per cent of girls versus 24 per cent of boys).

Girls and boys in Canada have very different views of their potential science careers. In particular, girls are almost three times as likely as boys to envisage themselves as health professionals (29 per cent of girls versus 10 per cent of boys). This is also true across OECD countries; on average, 17 per cent of girls expect to work as health professionals, compared to 6 per cent of boys. The difference of 19 percentage points between the shares of boys and girls in Canada expecting careers as health professionals is the second highest of all OECD countries, surpassed only by the United States (26 percentage points). Nevertheless, the proportion of Canadian students, both boys and girls, expecting careers as health professionals is one of the highest of all OECD countries. Girls in Canada are also close to two times more likely than boys to expect to work as science-related technicians or associate professionals, whereas across the OECD countries, boys are close to three times as likely as girls to expect to work in such occupations.

By contrast, a higher proportion of boys than girls see themselves becoming science and engineering professionals. More specifically, in Canada, 18 per cent of boys expect to work as science and engineering professionals, compared to just 7 per cent of girls. Across OECD countries on average, boys are approximately twice as likely as girls to expect to work as science and engineering professionals. It is worth noting that the shares of Canadian boys and girls expecting careers as science and engineering professionals are among the highest of all OECD countries.

Boys are also more likely than girls to see themselves as ICT professionals not only in Canada (4 per cent of boys versus 0.3 per cent of girls) but also across OECD countries (4 per cent of boys versus 0.4 per cent of girls).

Conclusion

Concerns have been raised that the proportion of students in Canada who choose careers in STEM is insufficient, particularly among girls. Nurturing motivation and interest in science at a critical age may be part of the solution. Using data from PISA 2015, this paper has shown that in general, Canadian students enjoy learning about science, with more than two-thirds reporting that they have fun when learning science topics, are happy working on science topics, enjoy acquiring new knowledge in science, and are interested in learning about science. Students in Canada are also

[3] In 2006, the question was administered in paper format, while in 2015, most countries/economies administered the question by computer. As well, in 2006 responses were coded using the 1988 International Standard Classification of Occupations (ISCO), while in 2015 they were coded using ISCO-08 (International Labour Organization, 2012). These contextual and methodological differences must be borne in mind when comparing student responses across the two cycles.

interested in a broad array of science topics, with a preference for topics related to disease prevention and astronomy. As well, most Canadian students recognize the instrumental value of studying science as a way to improve their career prospects, and are confident of their ability to perform science-related tasks. Furthermore, students' attitudes toward science and confidence in their science-related ability improved from 2006 to 2015. By 2015, enjoyment of science, instrumental motivation to learn science, and confidence in performing science-related tasks had improved for both boys and girls.

However, gender gaps are present in terms of both attitudes and confidence related to science. On the one hand, boys are more likely than girls to enjoy learning about science, more interested in physics and chemistry, and more confident in their ability to perform science-related tasks. On the other hand, girls are more likely than boys to be interested in health-related topics and are more likely to perceive learning science as useful to them and to their future studies and careers.

Based on PISA 2015 results, about a third of Canadian 15-year-old students expect to be working in a science-related occupation when they are 30, compared to approximately a quarter of students on average across the OECD countries. In fact, the share of students in Canada expecting a science-related career is surpassed by only three other OECD countries. Furthermore, this share increased between 2006 and 2015, largely because the proportion of students who expect to be working as health professionals rose. Canadian girls are in general more likely than boys to expect to work in a science-related field, although this is true most particularly among girls considered to be low or moderate performers in science. Top-performing girls are less likely than top-performing boys to expect a career in a science-related field. Boys and girls also expect careers in different science sub-fields. In particular, girls are almost three times more likely than boys to expect to be working as health professionals, and close to twice as likely to expect to work as science-related technicians or associate professionals. On the other hand, boys are close to three times as likely as girls to expect to work as science and engineering professionals and are also more likely to see themselves as ICT professionals.

Overall, results from PISA 2015 show encouraging developments in terms of the motivation and interest in science among 15-year-old students in Canada, as a relatively higher share of Canadian students report that they expect to work in science-related careers than the proportion of students in other participating countries and economies. At the same time, gender gaps still exist in attitudes toward science and with respect to student expectations of working in certain science-related careers. Of particular concern is that girls have less confidence than boys in their ability to perform certain science-related tasks.

References

International Labour Office. (2012). *International standard classification of occupations (ISCO-08): Vol. 1. Structure, group definitions, and correspondence tables.* Geneva: International Labour Office.

Organisation for Economic Co-operation and Development. (2008). *Encouraging student interest in science and technology studies*. Paris: OECD Publishing. Retrieved from www.oecd-ilibrary.org/content/book/9789264040892-en

Organisation for Economic Co-operation and Development. (2016). *PISA 2015 results: Vol. 1. Excellence and equity in education*. Paris: PISA and OECD Publishing. doi:10.1787/9789264266490-en.