

# Fairness of Standardized Assessments: Discrepancy between Provincial and Territorial Results

Karen Fung  
Centre for Research in Applied Measurement and Evaluation  
University of Alberta  
[Karen.Fung@ualberta.ca](mailto:Karen.Fung@ualberta.ca)

Man-Wai Chu  
Werklund School of Education  
University of Calgary  
[Manwai.Chu@ucalgary.ca](mailto:Manwai.Chu@ucalgary.ca)

## Abstract

This paper discusses the issue of unfairness due to the discrepancy in educational assessments among provinces and territories in Canada. An overview of all provincial and territorial assessments is illustrated and major assessment results are compared to provide an idea on how such discrepancy would be problematic. We argue that such non-uniformity of educational assessments is unfair to students, in which scores achieved by students of the same ability level may obtain different scores solely because of the province or territory they live in. It causes unfairness especially when students are competing for limited opportunities such as scholarships. The presence of such issue makes it difficult for students' scores to be comparable within the nation. In the final section of this paper, recommendations on solutions are provided to address the unfairness issues for the benefit of students and other stakeholders who would be affected.

## Introduction

Providing fairness, or educational equity, to every Canadian student during standardized assessments is a necessity, but it is difficult to attain (AERA, APA, & NCME, 2014). According to the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014), fairness in assessments explores many aspects of testing including the broad goal of achieving equality of opportunity in our society (Haladyna, Nolen, & Haas, 1991). The recently published *Standards for Educational and Psychological Testing* (2014) introduced a foundational chapter on fairness, which highlights the increasing need to investigate fairness issues within educational assessment programs. This new chapter emphasises accessibility and fairness and fundamental issues in testing (AERA, APA, & NCME, 2014). Additionally, the chapter better supports appropriate testing for all examinees (AERA, APA, & NCME, 2014). Fairness is an important concern to all stakeholders of any assessments. Although most assessment programs aim to achieve fairness, when considering students' performance across Canada this standard seems to be lacking. In Canada, the educational programs and curricula are governed by each province and territory individually. This non-uniformity in educational systems creates a variety of assessment programs used to assess Canadian students' learning. Although the educational assessments within Canada may share similar goals and values (Klinger, Deluca, & Miller, 2008; McEwen, 1995), students of the same grade level are evaluated differently across the country.

Klinger, Deluca, and Miller (2008) made a comparison of educational assessment programs between provinces and territories. The researchers identified differences in terms of scoring, reporting, and the use of assessment results. Such difference could also become problematic when students are assessed at different intervals and different benchmarks are used, making comparability and tracking of Canadian achievement results almost impossible. On the discussion of making comparisons among students who took different assessments, Lazer et al. (2010) illustrated that, in a scenario where assessment systems are very different, with little collaborations done on the assessment design, it would be very difficult to make comparisons of results. In such a case, linking or establishing a relationship between test scores from tests measuring different contents with possibly different levels of difficulty, can never be done when the types of exercises are different on the assessment and with no common items. It is also almost impossible to have the same students complete assessments from different provinces. Therefore, comparisons can only be made at a very surface, descriptive level. One solution is to establish empirical relationships between the assessments that are different and relate them to an independent assessment, but it would be difficult to do across nation when there are so many different assessments being used across Canada.

When comparisons are difficult to be made due to the differences in educational system, the issue of fairness is raised from such discrepancies, especially when the consequences of assessment results are high. For example, when a student is moving across provinces for university, how would his/her high school scores be compared with those from other provinces for university admission purposes? A student from Saskatchewan who received a B+ in grade 12 may receive a B- in grade 12 from Ontario due to the different assessment policies and benchmarks from each province. Additionally, many assessment programs are designed to measure whether or not students meet the provincial/territorial standards to graduate secondary school, but many post-secondary institutions use these grades as part of their admission requirements. This misuse of secondary school grades is a fairness issue that is highlighted as a problem in assessment programs (AERA, APA, & NCME, 2014). Another problem that is often asked is whether or not students of the same age from different provinces would have the same level of knowledge in particular subject area. For jurisdictions that have standardized testing, a trend of teachers focusing more on question types and contents that will be on the test was found (Herman & Golan, 1991). Whether the standardized assessment is at a high or low stakes also make a large impact. Madaus (1988) noted that "the greater the consequences attached to the test, the more likely it would be to have an impact [on teaching]" (p. 14). Therefore it is possible that students from jurisdictions with standardized assessment may learn differently than those from a jurisdiction with no standardized testing at the same grade level. These problems are extremely important when students are moving across the country and needing to enrol in a different educational system. The issue of fairness needs to be considered by educators when students' futures are at stake.

The purpose of this paper is to illustrate the lack of fairness and equality created by the current assessment programs across Canada. This paper is divided into four sections. First, a review of the different assessment programs around Canada is presented. Second, international assessments in which Canadian students have participated and results from a Canadian university highlighting students' achievement levels are presented. Third, the problem with the current assessment

policy in Canada is discussed. Fourth, recommendations on improvements to the current educational assessment system in Canada are made.

### **Assessment Policies of Canadian Provinces and Territories**

In this section, assessment programs from each province and territory are described. Each description provides an overview of the programs measuring Canadian students' achievement in each province and territory.

#### **Alberta**

Alberta administers an achievement testing program (PAT) for students in grades 1-9 and a diploma examination program for students in grade 12 (Alberta Education, 2013a). The PATs are developed to determine: if students are learning what they are expected to learn; report to Albertans how well students have achieved provincial standards at the end of grades 3, 6, and 9; assist schools, authorities, and the province in monitoring and improving student learning (Alberta Education, 2013b). Grade 3 students are assessed in English/French language arts and mathematics while grades 6 and 9 students are assessed in English/French language arts, mathematics, science, and social studies. The marks from PATs do not affect students' class grades. However, in 2013 the PATs would be phased out and replaced with the student learning assessments which will be administered at the beginning of grades 3, 6, and 9 to better inform student learning through interactive approaches to help educators and parents understand how well students can apply what they have learned (Alberta Education, 2013a). The diploma examinations were developed to: certify the level of individual student achievement in selected courses; ensure province-wide standards of achievement are maintained; report individual and group results (Alberta Education, 2013c). Diploma examinations assess grade 12 students in: biology, chemistry, English/French language arts, mathematics, physics, science, and social studies. The grade 12 diploma examinations marks account for 50% of students' final course grades. However, starting September 2015, the diploma examination will only account for 30% of students' final grades (Alberta Education, 2015).

#### **British Columbia**

British Columbia administers an annual foundation skills assessment (FSA) to students in grades 4 and 7 as well as provincial examinations (PE) in grades 10-12 (British Columbia Ministry of Education, 2013a). The main purpose of the FSAs are to help the province, school districts, schools and school planning councils evaluate how well students are achieving basic skills and make plans to improve student achievement. Grades 4 and 7 students in their second semester are assessed in reading comprehension, writing, and numeracy. The marks achieved on the FSA do not affect students' class grades. PEs are developed to ensure students meet consistent provincial standards of achievement in academic subjects; they were developed as a response to strong public desire for improved standards in education (British Columbia Ministry of Education, 2013b). Graduating students from all schools in the province are treated equitably when applying for admission to universities and other post-secondary institutions; Grade 10 students are assessed in science, English/French/First Peoples language arts, and math. Grade 11 students are assessed in social/civic/BC First Nations studies. The PE marks from grades 10 and 11 accounts for 20% of students' final course grades. Grade 12 students are assessed in

English/French/English First People/Communication language arts and that PE mark accounts for 40% of students' final course mark.

### **Manitoba**

In Manitoba, the provincial assessment program serves to measure how well students have learned the outcomes listed in the programs-of-studies (Manitoba Education, Citizenship and Youth, 2013). In grade 12, students in Manitoba write the provincial tests to assess their skills on: English Language Arts, Applied Mathematics, Essential Mathematics, Pre-Calculus Mathematics, French (for French program), and French for students in immersion programs. The tests are administered each semester to students who are enrolled in the respective classes. The languages (i.e., English and French), Applied Mathematics, and Pre-calculus Mathematics tests contribute to 30% of the students' final course grades. The Essential Mathematics test contributes to 20% of the students' final course grade. Students may rewrite a test if failed, and the higher mark will be used as the students' provincial test score (Manitoba Education, Citizenship and Youth, 2013).

### **New Brunswick**

The education system in New Brunswick is divided into the Anglophone and the Francophone sector and focuses on a *Balance Assessment* framework. This framework believes that assessment is a process instead of an individual event (New Brunswick Department of Education and Early Childhood Development, 2013), which includes ongoing formative assessment (e.g. classroom assessment), and summative assessment (i.e., The New Brunswick Assessment Program [NBAP] - provincial assessments). For the purposes of this paper, only the standardized summative assessments will be discussed. There are four main purposes of the NBAP: (a) determine if students demonstrate the skills and concepts outlined in the curriculum documents; (b) report provincially on how well students have achieved the provincial standards at given points; (c) assist schools, districts, and the province in monitoring student learning; and (d) coordinate the administration of national and international assessments. Provincial assessments are administered annually in almost all grades depending on which school program, Anglophone or Francophone, the students are enrolled in. Students in the Anglophone sector are assessed in almost every year from grades 2 to 8, except for grade 6. Assessment subjects are mainly on literacy and mathematics. Students in the Francophone sector are only assessed in grade 5 on Mathematics and Natural Sciences, and in grade 8 on Literacy.

As part of the high school graduation requirement students in both Anglophone and Francophone sectors need to complete a series of provincial assessments. In the Anglophone sector, students in grade 9 are required to write the English Language Proficiency Assessment (ELPA) which includes a reading and a writing component. However if the student obtained the category of the lowest level, or the Below Appropriate Achievement (BAA) in the ELPA, he/she may take the English Language Proficiency Reassessment (ELPR) - an assessment similar to the ELPA, but administered in grades 11 and 12. The ELPR may also be written by grade 11 and 12 students who did not have a chance to write the ELPA in grade 9. In the Francophone sector, at the end of their mathematics course, a mathematics assessment is required. In June, a French assessment is also performed. These assessments are mandatory for graduation and account for 40% of the students' final grade. In grade 10, and in grade 12 if the students did not get a chance to write in

grade 10, an assessment on Oral Proficiency Assessment is performed for French as Second Language students.

### **Newfoundland and Labrador**

Students in Newfoundland and Labrador have their provincial tests written in grade 3, 6, and 9 at the end of the school years (Newfoundland and Labrador Department of Education, 2013). These assessments are based on the curriculum of Newfoundland and Labrador to determine students' proficiency on English language arts and mathematics for grades 3 and 6. Students in grade 9 are assessed on: French, English language arts, mathematics, and science. The goals of these assessments are to inform stakeholders about students' level of performance. In addition, the Department of Education in Newfoundland and Labrador offers Public Examinations (PE) for particular subject areas in high school (i.e., English, mathematics, biology, world geography, and world history) for those who are enrolled in the associated Level III courses (the final year level of courses for senior high school), and these examinations account for 50% of the students' final mark. The PE scores are also used to determine students' provincial scholarship scores for obtaining scholarships from the province (Newfoundland and Labrador Department of Education, 2013).

### **Northwest Territories**

The student assessment, evaluation, and reporting program in the Northwest Territories (NWT) was developed to consistently evaluate and report student progress in relation to the curricular outcomes (Northwest Territories Education, Culture, & Employment, 2010). The results of the assessments provide valuable information on students' success in relation to curriculum expectations, identifies areas of strength, and challenges at the students, school, district education authority, divisional educational authority, and territory-wide level. NWT uses the same assessment program as Alberta (Northwest Territories Education, Culture, & Employment, 2011). The NWT administer the achievement testing program (ATP), which are called Alberta achievement tests (AAT) in NWT and will soon be replaced by the student learning assessments (SLA), for students in grades 3, 6, and 9 in math and language arts classes. Grade 12 students write the diploma examination in the same subjects as students from Alberta. The marks from the ATP, soon to be replaced by SLA, are not used to determine students' final grades, but the diploma exams account for 50% of students' final course grades. The NWT Minister of Education is currently considering reducing the weighting of the diploma exam to 30% starting September 2015, similar to the reduction seen in Alberta, but nothing official has been decided (J. Stuart, personal communication, May 4, 2015; please refer to the 'Alberta' section for more details of the NWT assessment program).

### **Nova Scotia**

Provincial assessments in Nova Scotia are administered in grades 3, 4, 6 and 8 (Nova Scotia Education and Early Childhood Development, 2013). These provincial assessments are part of the Program of Learning Assessment for Nova Scotia (PLANS). They are administered in the Fall for students in elementary schools (grade 3, 4, and 6), and in late Spring for junior high schools (grade 8). A reading and writing assessment is administered in grade 3, and the mathematics assessment is administered in grade 4. In grades 6 and 8, an assessment of reading,

writing, and mathematics is written. Further in high school, students of Grade 10 write the Nova Scotia Examinations (NSE) in the courses of English 10, Mathematics 10, and Mathematics 10 at Work. These examinations contribute to the students' final course mark of 20%. The goal of these assessments is to evaluate how well students have learned the curriculum.

### **Nunavut**

The student assessment and evaluation office performs research and develops assessment tools, strategies, and programs that: support the vision of a Nunavut education system built on intelligence quotient, societal values, beliefs, and perspectives; are culturally and linguistically appropriate to Nunavut; provide valid and reliable evidence of student learning and achievement; reflect current best practices in the area of educational assessment; meet the needs of Nunavut's K-12 students, parents, teachers, and the public; enhance teaching and student learning (Nunavut Department of Education, 2013). Nunavut uses Alberta's diploma examination program for students in grade 12 (D. Mearns, personal communication, November 6, 2013). The diploma examinations currently account for 50% of students' final course grades. The Nunavut Minister of Education is currently considering reducing the weighting of the diploma exam to 30% starting September 2015, similar to the reduction seen in Alberta, but nothing official has been decided (D. Mearns, personal communication, May 7, 2015; please refer to the 'Alberta' section for more details of the Nunavut assessment program).

### **Ontario**

In Ontario, students across the province write the provincial assessment developed by the Education Quality and Accountability Office (EQAO) in grades 3 and 6 (in reading, writing, and mathematics), and in grade 9 (mathematics). The main purpose of these assessments is to measure students' learning, so stakeholders can help students improve their academic achievement (Education Quality and Accountability Office, 2012). As a graduation requirement for high school, students in grade 10 are required to pass the Ontario Secondary School Literacy Test (OSSLT). The purpose of the OSSLT is to determine whether students are meeting the minimum standard for grade 9 literacy. The OSSLT is a pass/fail assessment, meaning if the student fails, the student must retake and pass the assessment in order to obtain a secondary school diploma. Retests are provided if students fail the OSSLT (Education Quality and Accountability Office, 2012). Additionally, the Ontario Secondary School Literacy Course (OSSLC) is available for students who failed the OSSLT to complete as a support course. It is restricted for those who have taken the OSSLT twice and were unable to pass. The OSSLC is a credit course intended to provide intense support on students' literacy skills required for success in high school.

### **Prince Edward Island**

Before children attend kindergarten, an Early Years Evaluation (EYE) is conducted for all children to determine the level of their early developments and to screen for possible academic or developmental delays. Results of this assessment would be provided to parents so assistance could be sought if needed. Students in Prince Edward Island (PEI) are assessed in reading, writing, and mathematics for grades 3, 6, and 9. These provincial assessments are developed by teachers and are intended to determine students' performance. These assessments do not serve as a graduation requirement for students. The PEI Department of Education and Early Childhood Development (2013) has emphasized that "provincial assessments are just one of many tools

used to monitor student learning. Parents should talk to the teacher about the full scope of their child's performance" (p.1).

### **Quebec**

In Quebec, the elementary school system is divided into three two-year cycles (Quebec Ministry of Education, 2001). In secondary education, there are two cycles (three years in cycle one, and two years in cycle two). During cycle two of secondary education (grades 10 and 11), as part of the certification of studies, provincial examinations also known as the uniform examinations are administered in January, June, and August for all students in the subjects: History and Citizenship, Mathematics, Science and Technology, English Language Arts, and French (for French as Second Language). The uniform exams contributed 50% of students' final grades. Teachers and educational consultants from both the English and French sectors develop the tests. Upon completion of secondary education, students must enter the general and vocational colleges for two years (or Collège d'enseignement général et professionnel [CEGEPs]) before entering university. The two-year CEGEPs lead students directly into university, and the three-year technical programs leads to a Diploma of College Studies (DCS). DCS is a requirement for admission to the university, and a ministerial English exit exam is required to be taken and passed in order to earn the DCS (Quebec Ministry of Education, 2001).

### **Saskatchewan**

Grade 12 students taught by non-accredited teachers will write the departmental examinations in: English language arts, biology, chemistry, physics, mathematics and pre-calculus (based on the courses students are enrolled in). These examinations are developed by accredited teachers, and are administered in January, March, May, June, August, and December of each year. Aside from these examinations, there is also an Assessment for Learning Program (AFL) implemented to improve students' learning and achievement, support the development of professional learning communities, strengthen the capacity of stakeholders to use data for decision making, and strengthen the ability of school divisions to report to the public (Saskatchewan Ministry of Education, 2012). It is not an annual assessment and its administration years are based on the subject areas and grade levels. For mathematics, it is administered for grades 5 and 8 in every odd year (i.e., 2011, 2013, 2015), for reading, it is administered for grades 4, 7, and 10 every odd year (i.e., 2011, 2013, 2015), and the writing assessment is administered for grades 5 and 8 in every even year (i.e., 2010, 2012, 2014). These assessments do not contribute to students' grades.

### **Yukon**

Yukon's assessment program is designed to allow educators, parents and support staff to address the strengths and challenges students face in learning, and decide how best to provide support in the classroom and at home (Government of Yukon Education, 2013a). Since the 2012-13 school year, Yukon's educational system started to use the British Columbia curriculum and provincial exams (Government of Yukon Education, 2013b). As such, grades 4 and 7 students write the British Columbia's foundation skills assessments, which are called the Yukon achievement tests (YAT), and grades 10-12 students write the British Columbia provincial examinations, which are called British Columbia Provincials (BCP). The YAT assessment marks do not affect students' class grades while the BCP account for 20% of students' final course grades in grades 10 and 11,

and 40% of students' final course grades in grade 12 (please refer to the 'British Columbia' section for more details of the Yukon assessment program).

This section of the paper showcased the various assessment programs used in each province and territory. It indicated that each province and territory administers different subjects, schedules, and quantities of assessments to their students. Additionally, the contributions of the assessment marks in students' final grades are different among each province and territory. For example, in Alberta the grade 12 assessment marks contribute 50% of students' final course grade; while in Prince Edward Island, there are no grade 12 assessments administered which make the final course grade comprised of only the teacher-awarded class mark. This disparity between the assessment policies around Canada makes measuring Canadian students an unfair process. The next section will present international and national assessment results that highlight Canadian students' achievements and a review of students' university entrance grades at a Canadian university.

### **Differences in Canadian Students' Achievement**

This section will highlight the different achievement levels of students on international and national assessments. The section will begin by presenting Canadian results broken down into the different provinces of two international assessments, the Programme for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS). Then the provincial and territorial results of a national assessment called the Pan-Canadian Assessment Program (PCAP) will also be presented. This section will end with an analysis of Canadian students' university entrance grades across Canada at one university.

#### **Programme for International Student Assessment (PISA)**

PISA is a large-scale, standardized test administered to 15 year-old students every three years to determine whether students approaching the end of compulsory education have acquired the essential knowledge and skills for full participation in society (PISA, 2009). Alongside PISA, background surveys (questionnaires) are also administered to students and other stakeholders to supplement and explain a range of factors that contribute to successful students, schools and education systems. The chosen skills measured in the PISA are generally recognized as key to the educational process.

The subjects assessed in PISA are reading, mathematics, and science. One of the recent data released from the Organization for Economic Co-operation and Development (OECD) are from the 2009 administration of PISA which focused on reading achievement (PISA, 2009). PISA 2009 was written by 470,000 students representing 65 countries (PISA, 2009). Approximately 23, 000 Canadian students took part in the 2009 administration of the PISA exam (Statistics Canada, 2010). The sampling procedure used in PISA is such that the collected data constitute nationally representative samples (see PISA, 2009, for details on the sampling procedure).

The results of PISA are used in conjunction with other sources of information, such as the Pan-Canadian Assessment Program (PCAP; please refer to the 'PCAP' section for more details), to inform various education-related initiatives and changes in Canada's educational systems. The 2009 PISA results are presented in Figure 1. Although the 2009 administration included math

and science, the majority of the assessment focused on reading achievement. Thus, only the reading achievement results are presented in the paper (please refer to Statistics Canada, 2010 for the Canadian math and science results). Overall, Canadian students performed well compared with students in most countries. Canadian students had a mean score of 524 on the overall reading scale which is above the OECD mean score of 496. Only four countries, Shanghai-China, Korea, Finland, and Hong Kong-China, significantly outperformed Canada in the overall reading scale. This result indicated Canadian students are performing very strong on international assessment. However, when the Canadian assessment marks are broken down into the respective provinces a different picture is present. Figure 1 breaks down the Canadian average into the different provinces and represents the respective students' reading achievement using white achievement bars. It can be seen that Ontario students are performing significantly better than the Canadian average while Alberta, British Columbia, and Quebec are performing at the Canadian average. However, Nova Scotia, Newfoundland and Labrador, Saskatchewan, New Brunswick, Manitoba, and Prince Edward Island are all performing significantly lower than the Canadian average. These results show a disparity between 15-year old Canadian students who are educated in different provinces in terms of reading achievement.

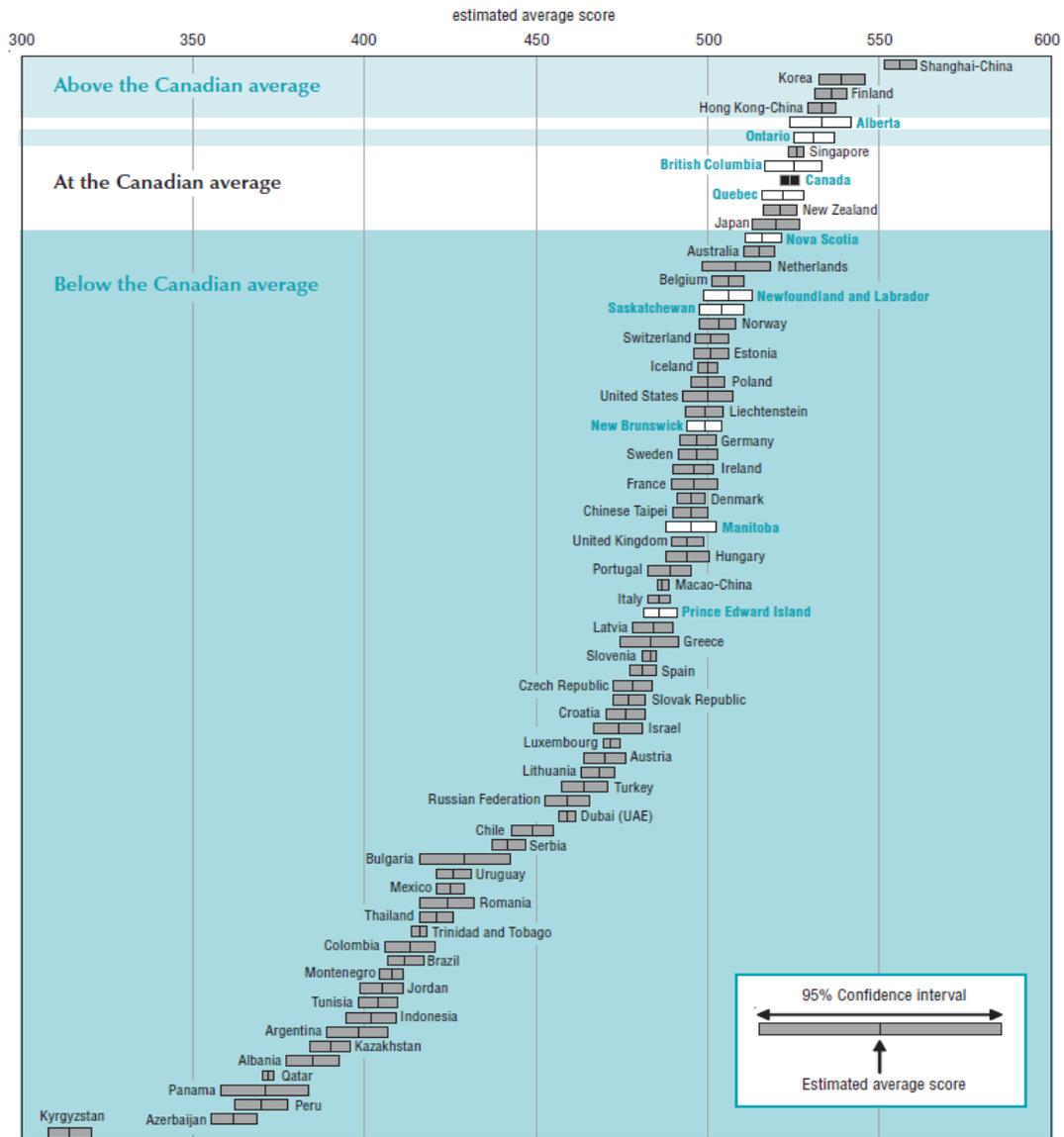


Figure 1. Average scores and confidence intervals of combined reading for provinces and countries of 2009 PISA (Statistics Canada, 2010, p. 16)

### Trends in International Mathematics and Science Study (TIMSS)

TIMSS is another international large-scale assessment dedicated to improve teaching and learning in mathematics and science (International Association for the Evaluation of Educational Achievement [IEA], 2013). TIMSS is administered every four years to students in grades four and eight. Fifty-seven countries participated in the TIMSS at the grade four level and 56 countries participated in the TIMSS at the grade eight level. Diversity among the participating countries, in terms of economic development, geographical location, and population size, are fundamental to the IEA who develops and administers the TIMSS. IEA's vision is that diversity of educational philosophies, models, and approaches that characterize the world's education systems constitute an opportunity for each country to learn from the experiences of others.

Participating TIMSS countries “share the conviction that comparing education systems in terms of their organization, curricula, and instructional practices in relation to their corresponding student achievement provides information crucial for effective education policy-making” (IEA, 2013, para. 2).

In 2011 only three provinces, Alberta, Ontario, and Quebec, participated in the TIMSS to represent Canada (Education Quality and Accountability Office [EQAO], 2012). Canadian students’ achievement on the TIMSS is higher than the TIMSS scale center point of 500. The results are presented in Figures 2 and 3. However, once Canadian achievement results were broken into the three different provinces significant differences were seen between the provinces. Although results for both grades four and eight were presented in both mathematics and science, only the mathematics results will be presented in this paper because of the similarities between the mathematics and science results in terms of indicating disparity between provinces. In the grade four mathematics results Ontario students’ marks were significantly higher than those of Alberta students, but significantly lower than those of Quebec students as shown in Figure 2. In the grade eight mathematics results Ontario and Alberta students scored similarly, but Quebec students’ were significantly higher than those of Ontario and Alberta as shown in Figure 3. These results show a disparity between grades four and eight Alberta, Ontario, and Quebec students in terms of mathematics achievement.

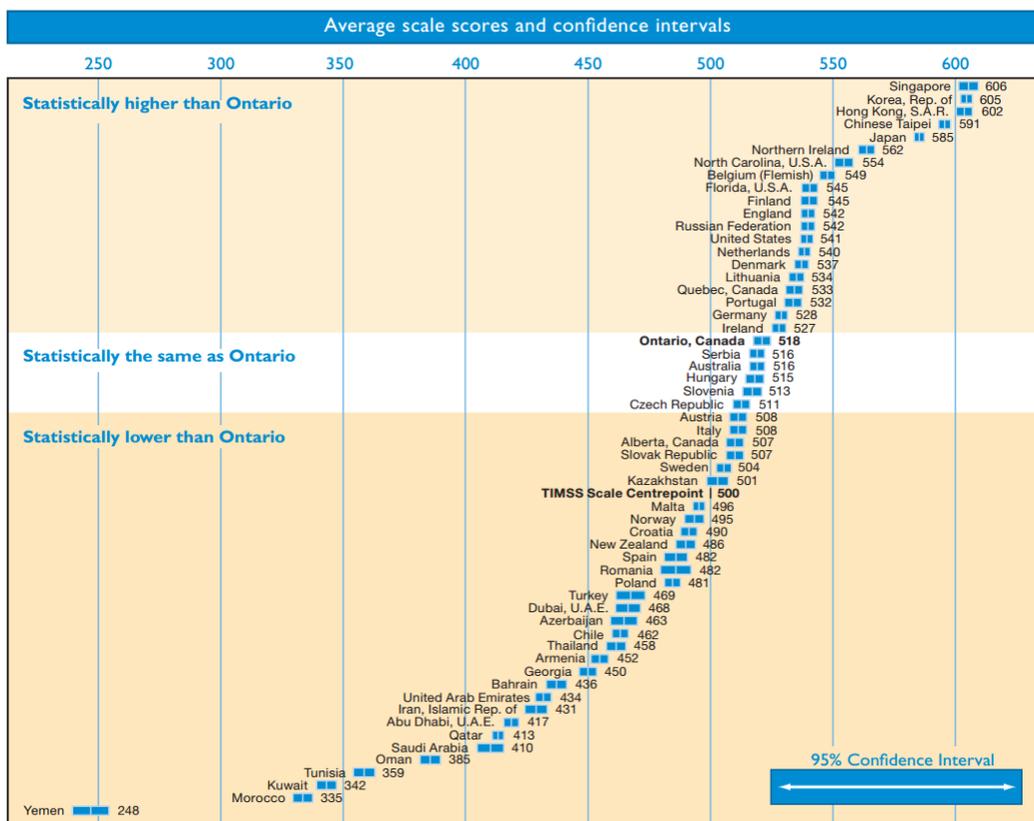


Figure 2. Grade four mathematics achievement, results by country and area of the 2011 TIMSS assessment (EQAO, 2012, p. 1)

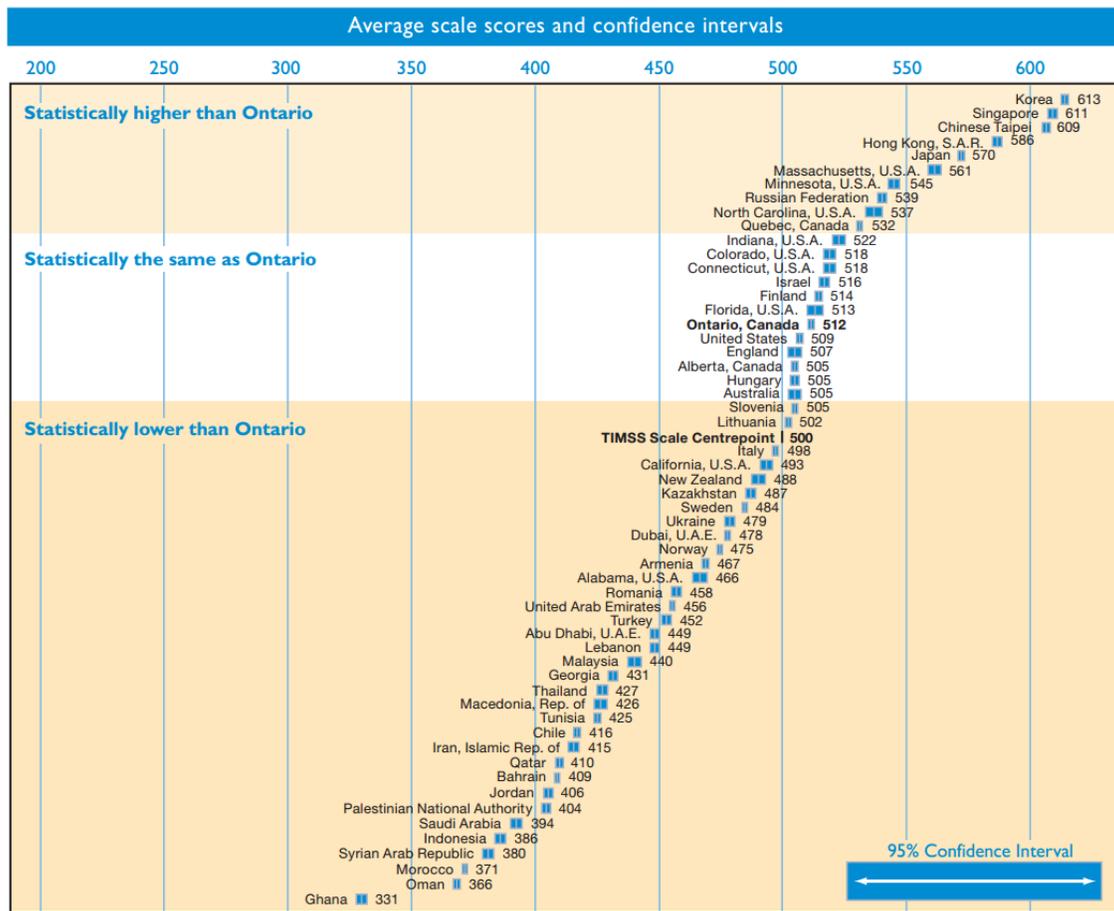


Figure 3. Grade eight mathematics achievement, results by country and area of the 2011 TIMSS assessment (EQAO, 2012, p. 3)

### Pan-Canadian Assessment Program (PCAP)

The PCAP is a Canadian national assessment that is administered every three years to 13-year-old students (i.e., Grade 8). It is administered in both official languages of Canada (i.e., English and French). Developed by the Council of Ministers of Education, Canada (CMEC), the PCAP is intended to measure whether or not students' skill levels in reading, mathematics, and science are similar across Canada, since curriculum between provinces and territories are different. The focus of the assessment varies by each administration, for example for PCAP's first administration in 2007, its focus was reading, with mathematics and science as minor domains, and in 2010, its focus was mathematics with reading and science as minor domains. Each administration the focus is rotated through. There are two main purposes of the PCAP: (a) to help ministries and departments of education make informed decisions about education, and (b) to determine how prepared 13-year-old students are for the academic challenges of high school (CMEC, n.d., p. 1). The CMEC (n.d.) stresses that results of the test should not be used to reflect students' individual performance, nor should they be used to compare between schools and

jurisdictions. All jurisdictions in Canada participate in PCAP, however not all schools are selected to write the assessment. A stratified sampling process of random sampling is used to determine the class within a school that will be writing PCAP. Results of raw scores on the PCAP are converted into standard scale scores, with a mean of 500 and standard deviation of 100.

In 2010, the main domain of assessment on the PCAP was mathematics. A comparison of average scores (for both the English and the French administrations) between provinces on the mathematics assessments can be found in Figure 4. As seen in Figure 4, Quebec and Ontario scored significantly above the national mean of 500 points. Alberta scored slightly below the mean, while the rest of the jurisdictions scored significantly below the Canadian average.

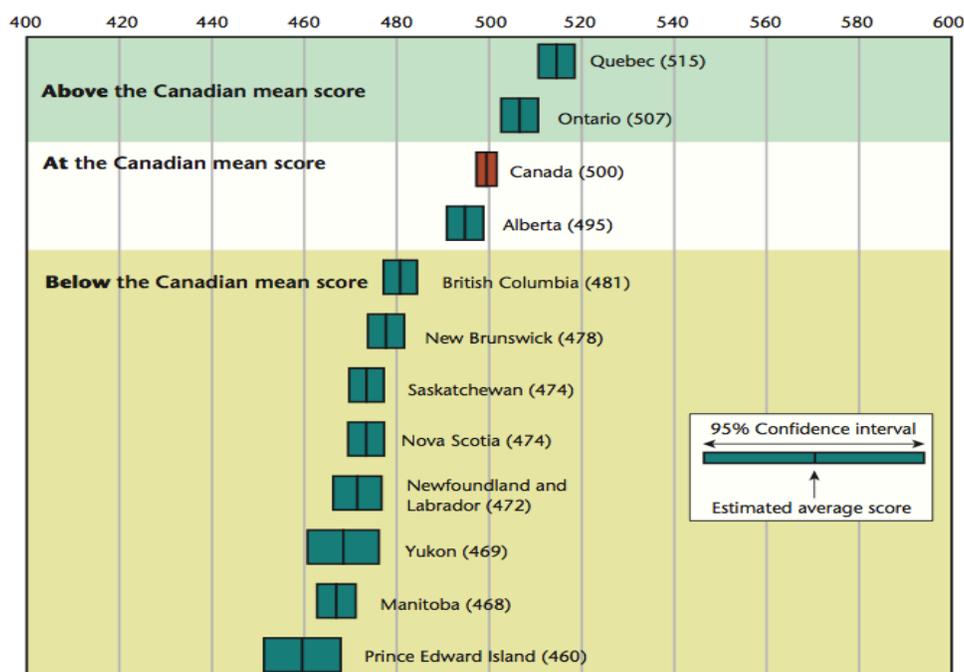


Figure 4. 2010 PCAP mathematics average scores of all jurisdictions (CMEC, 2011).

### Students' Entrance Grades into a University

Grade 12 students applying for limited post-secondary positions tend to apply to more than one post-secondary institution at once. Sometimes, these students will apply to post-secondary institutions across the country to ensure entrance into an institution that is reputable or has a specific program. When students are admitted into post-secondary institutions, their grade 12 final grades are often the determining factor in whether they are granted acceptance into a post-secondary institution or not. However, different provinces and territories have different assessment policies that may significantly affect students' final grades therefore making comparison on these students unfair.

Research from the Strategic Analysis Office from one Canadian University, which analyse statistics from undergraduate students to inform program changes, has shown that graduates from

each province perform at different levels despite equivalent secondary grades (Strategic Analysis Office, 2011). For example, a recent study of first year English students at the university who received a 'B' in their first year English class corresponded to different high school English grades when analysed according to the province where the student completed their high school (Strategic Analysis Office, personal communication August 4, 2011). Figure 5 indicates the high school English grade of each provinces' and territories' students who applied to the university and received a 'B' in English class.

Figure 5 indicates a six percent difference between Canadian students who all ended with the same first year English grade. According to this analysis, students from Alberta, the Maritimes, and the Territories may have been at a disadvantage when competing for entrance scholarships or admission with students from other provinces. This creates disparity among Canadian students when they are being compared nationally for resources such as scholarships and admission to post-secondary institutions.

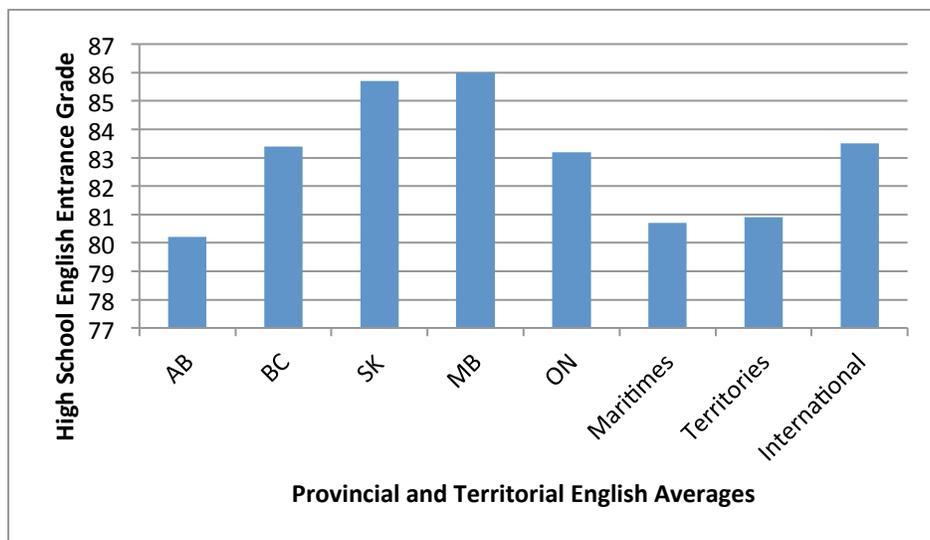


Figure 5. High school English grades corresponding to a 'B' in university English class

This section of the paper presented results from international and national assessments which all indicated Canadian students score significantly different achievement levels. This unbalance in achievement levels among students is not a problem when looking at groups of students in each province because students are measured against the same education system and assessed with the same tools. However, this disparity between students' achievement in each province and territory becomes a large issue when students are being compared against each other, such as for post-secondary admission or nationally funded scholarship opportunities. The next section will discuss the different Canadian assessment programs and international, national, and university entrance data previously presented in the paper. The problems with the current assessment policies in Canada are also addressed in the next section.

### The Issue of Unfairness

The previous two sections have outlined the differences between the assessment policies around Canada and the disparities between the academic achievements of Canadian students as measured with international assessments. The review of the different provincial assessment programs across Canada indicated that there are vast differences between provinces. This difference in provincial assessment policies is problematic because a portion of Canadian students are consistently being evaluated unfairly. This unfairness in Canadian assessment programs is best seen when students are compared with one another for limited resources. Common arenas for Canadian students to be compared with one another are post-secondary admissions and entrance scholarship applications where their final year of high school course grades are usually the determining factor in whether a student is admitted to an institution or receives scholarship funding. Current Canadian assessment programs vary so much that some students are put at a disadvantage when compared with their peers from another province or territory. For example, a student who tends to perform poorly on high stakes tests would not receive a very high final grade if they were from Alberta, NWT, or Nunavut because 50% of their grade 12 core subjects' final course grades are accounted for by the diploma exam. Despite the weighting reduction in Alberta and possible reduction in NWT and Nunavut of the diploma exam to 30% starting September 2015, the effects of this reduction on fairness of students' grades are currently unknown. However, if this student was from Nova Scotia, Ontario, or Prince Edward Island then the student's grade 12 core subjects' final course grades could be much higher because the high stakes exams administered in these provinces do not have direct effects on students' grade 12 final course grades. Depending on which province or territory this student completed their grade 12 education; this could have significant impacts on their post-secondary admissions and scholarship opportunities. This example is supported with empirical evidence from the 'Student's Entrance Grades into a University' section where one university compared students who achieved the same first year English course grade and compared their high school English grade. Their results revealed a six percent difference between students of different provinces and territories. The different assessment programs around Canada are ranking students of particular provinces and territories unfairly, with some jurisdictions administering high stakes large scale testing during critical years of education, and some without formal standardized assessment. One critical year in a students' life could be their grade 12 year where they transition between high school and post-secondary education, but it also include grades kindergarten to 11.

As students move through the Canadian education system from kindergarten to secondary graduation, the different assessment programs across Canada are causing students to be measured using different benchmarks and standards. Not all Canadian students are learning and performing equally. This is evidenced by the significant differences observed between Canadian provinces on international and national assessments such as the PISA, TIMMS, and PCAP. In all these assessments, significant differences are seen between provinces and territories indicating Canadian students are not learning and performing to a national standard. For jurisdictions that administer standardized testing, a notion known as test score pollution, or "factors affecting the truthfulness of a test score interpretation" (Haladyna, Nolen, & Haas, 1991, p. 4) may occur, especially when the assessment is high stake. These polluting factors may include the methods schools are preparing their students to take the standardized assessment, and having practices and conditions that are nonstandard. For example, matching the teaching objectives based on the test, practicing question types that are in the test, and informal alignment of the curricula based on the

test (Haladyna et al., 1991). Such occurrence of test score pollution would greatly affect the validity of the test. When such test results are used to compare with students from other jurisdictions without high stake standardized assessment, fairness becomes a concern.

Provincial and territorial differences between students' achievement levels are not a problem until students are compared against one another or they move to a different part of Canada half way through their educational program. With approximately 277,846 inter-provincial and territorial moves in 2009 (Statistics Canada, 2011), it is very likely some students moved with their parents and continued their education in the new province. This interprovincial and territorial move may make it difficult for students to adjust to the level of education in the new province. For example, if a student moves from Prince Edward Island, a province who consistently scores low on PISA, TIMMS, and PCAP, to Ontario, a province who consistently scores high on PISA, TIMME, and PCAP, then this student may have great difficulties adjusting to the new level of educational expectations. Research suggests that students who move during their primary and secondary education perform approximately one-tenth of a standard deviation lower for each move the student makes and also diminishes their prospects for graduation (Beatty, 2010; Rumberger, 2004). This decrease in student achievement could be mitigated by developing a common Canadian standard so that all students of a specific grade could perform at similar levels.

The disparity of Canadian assessment programs create an unfair disadvantage for students in some provinces and territories when they are being compared to their Canadian peers who are assessed using different measures for limited post-secondary admissions and scholarship funding opportunities. Additionally, the different Canadian assessment programs create different standards and benchmarks for students being educated in different provinces and territories. This unfairness to students is related to the different assessment programs that exist within Canada. The next section will provide recommendations and improvements to combat this unfair educational system created by different assessment programs across Canada.

### **Suggestions on Restoring Fairness**

The previous section identified the issue of educational equity, specifically the unfairness brought about by the current Canadian assessment program. Assessment programs are a provincial and territorial jurisdiction which leads to discrepant standards being set across Canada. Hence, there is a need to discuss changes to the current system that will restore the fairness to our assessment programs.

The discussion, as presented in this paper, is to present suggestions on assessment systems that would help restore the fairness by highlighting the goals of these tests. Some of the common goals among all the assessment programs across Canada are to (a) support student learning and (b) to measure how well students are meeting provincial or territorial standards (Alberta Education, 2015). Addressing these goals is not difficult considering the advances that have been made in the field of educational assessments. In recent years, various innovative assessment formats have been developed to better address these goals. Hence, an update of the traditional assessment formats (e.g., multiple-choice and written response) towards a more innovative approach is required to help combat the unfairness of the current assessment programs. One such

innovative assessment format that aligns well with these common goals is the cognitive diagnostic assessment (CDA).

CDAs aim to unify cognitive psychological theories with assessment design to advance a number of instructional and assessment concepts. For example, instructional scaffolding, which involves a temporary framework aimed at promoting learning when knowledge and skills are first being introduced to students (Sawyer, 2006), and learning progressions, a framework for developing assessments aimed at moving students from novice toward expert understanding over time (Pearson, 2010). CDA was developed to measure students' knowledge and skills in distinct content domains to help identify areas of cognitive strengths and weaknesses (Leighton & Gierl, 2007). Knowing students' cognitive weaknesses allows educators to direct students' learning towards enhancing those weaknesses to improve learning. Further, CDAs can be used to provide formative diagnostic feedback through performance reports that indicate where students are in terms of the acquisition of knowledge and skills (Jang, 2008).

The feedback reports associated with CDAs, shown in Figure 6, are developed to provide fine grain information regarding students' knowledge and skills. The feedback reports are highly detailed and supportive of student learning by indicating, explicitly, their areas of weaknesses. This will allow educators to focus their attention towards enhancing these weaknesses. In Figure 6, the CDA used to generate this feedback report focused on providing feedback on Specific Outcome 1, as listed in the program-of-study. Hence, CDAs are able to measure whether or not students are meeting provincial or territorial standards by providing detailed information of their achievement on each learning outcome.

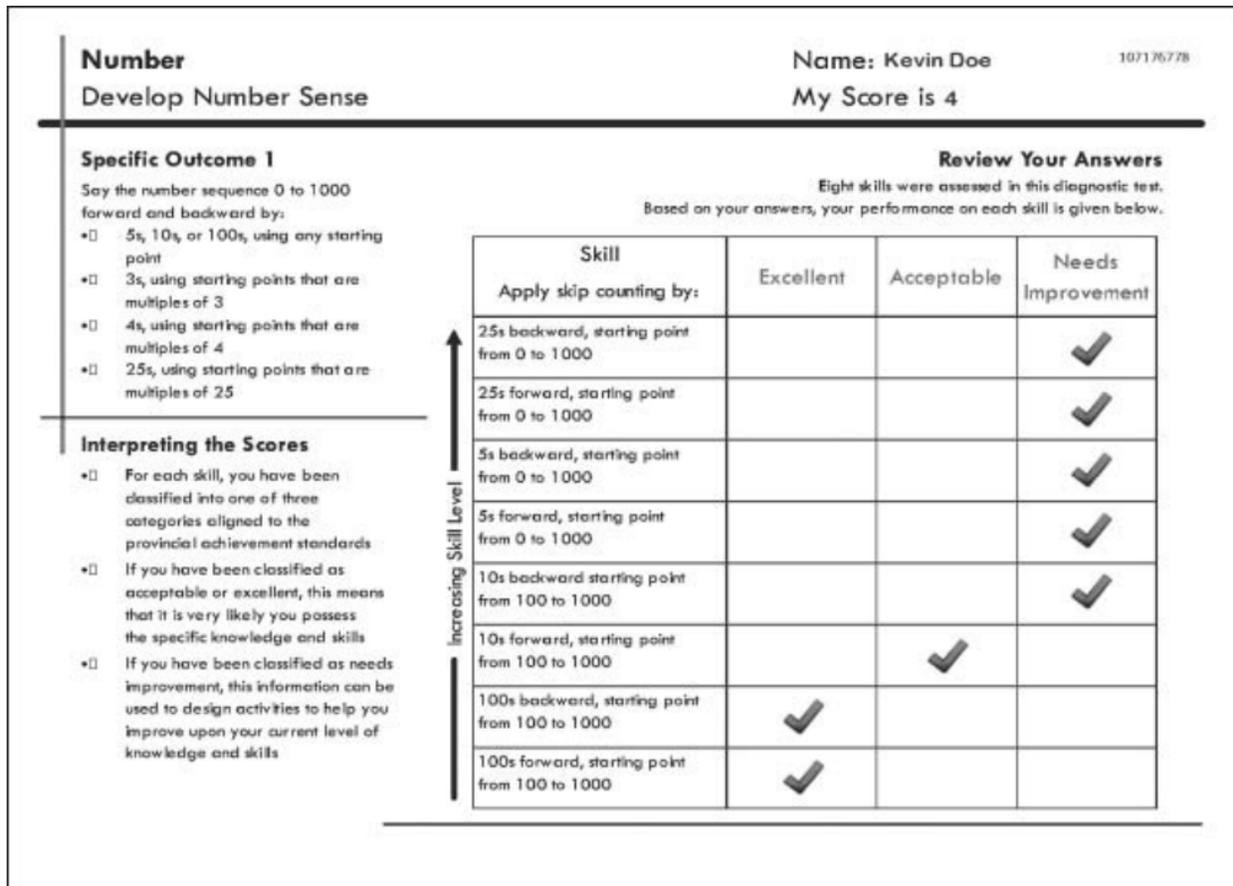


Figure 6. Example of a feedback report from a cognitive diagnostic assessment (Gierl, Alves, & Taylor-Majeau, 2010)

The two issues of unfairness discussed in this paper – (1) using assessment scores as indicators of admission and scholarship allocations and (2) ensuring all Canadian students achieve a specific level of knowledge and skills – may also be resolved by using CDAs. For the first issue, since CDAs are diagnostic in nature, the score is associated with the achievement level of a specific outcome. As such, the feedback reports would not easily allow for students to be compared to one another for admission and scholarship purposes, yet it provides information on how well students are learning the necessary subject areas. For the second issue of ensuring all Canadian students achieve a specific level of knowledge and skills, having students being evaluated on a common set of curriculum content would ensure fairness among students. For example, having students to write the same items at each grade and within similar content areas on the CDA test would allow educators to establish a link between tests and determine whether or not students are functioning at similar achievement levels across Canada. These common items would be developed by a central Canadian committee periodically so that classroom teachers could easily administer them and report the results.

This discussion recommends the use of a few common set of short CDA tests to be administered by teachers across the nation. Canadian schools may then report their students achievement

levels using a centralized system and these data may be gathered and reported by an organizational responsible for these data. Standardized assessments are a vital aspect of Canada's assessment system. However, the current programs do not allow for fair evaluations of all students. An update of our assessment programs are required to restore the fairness to the current problematic system. By using CDAs as the item format for standardized assessments, students may be provided with a more detailed description of their cognitive strengths and weaknesses. This will allow educators to make better changes to their lessons to focus on enhancing students' learning. CDAs' feedback reports may also help prevent assessment results to be used in student comparisons for admission and scholarship purposes because the reports focus on specific levels of each learning outcome instead of an overall final mark.

Aside from CDAs as a tool, it is also important to ensure that students are learning similar content of each subject within each grade in order to address the issue of disadvantages for students from some jurisdictions when they are moving to another province or territory. Therefore our second recommendation is to develop a committee of members within Canada to develop common assessment items and examine each jurisdiction's student assessments to ensure that they are measuring similar contents and at the similar levels. When jurisdictions are governing their own education policies, it is best to ensure common sets of knowledge will be learned across the country so a more uniform skill set would be developed in students. A list of criteria or a checklist should be established for members of the committee to evaluate on for each jurisdiction, including the specific contents for each subject and grade level. Although currently the PCAP developed by The Council of Ministers of Education, Canada, has collaborated with the curriculum across Canada, the PCAP is not administered to all students, therefore, whether students from difference jurisdictions are learning the same thing, or whether they have the same level of skill set, is not truly known. Having such an evaluating committee could provide a guideline for jurisdictions to follow.

There are different perspectives one may take on restoring fairness in the Canadian education system, the suggestions discussed above are only a few in many possibilities. A pandemic of unfairness of Canadian assessments is seen based on the discussion of this paper, and it is time for the education systems in all provinces and territories to change for the better by ending inequality of education policies, promoting equal learning opportunities, providing more useful learning feedback, and making fair use of assessment results.

## References

- AERA, APA, & NCME (2014). Standards for educational and psychological testing. Washington, D.C.: AERA, APA, & NCME
- Alberta Education. (2013a, November 1). *Provincial testing* [website]. Retrieved from <http://www.education.alberta.ca/admin/testing.aspx>
- Alberta Education. (2013b). *General information bulletin: Introduction to the achievement testing program*. Edmonton, AB: Alberta Education. Retrieved from [http://www.education.alberta.ca/media/6419105/02-ach-gib-2013-14\\_%20introduction%20&%20revisions\\_2013-08-02.pdf](http://www.education.alberta.ca/media/6419105/02-ach-gib-2013-14_%20introduction%20&%20revisions_2013-08-02.pdf)
- Alberta Education. (2013c). *General information bulletin: Introduction to the diploma examination program*. Edmonton, AB: Alberta Education. Retrieved from [http://www.education.alberta.ca/media/6902694/03-dip-gib-2013-14\\_intro%20revisions\\_2013-08-01.pdf](http://www.education.alberta.ca/media/6902694/03-dip-gib-2013-14_intro%20revisions_2013-08-01.pdf)
- Alberta Education (2015). *Diploma examinations* [website]. Retrieved from <https://education.alberta.ca/admin/testing/diplomaexams.aspx>
- Beatty, A. (2010). *Student mobility: Exploring the impacts of frequent moves on achievement – summary of a workshop*. Washington, DC: The National Academies Press. Retrieved from [http://www.nap.edu/openbook.php?record\\_id=12853&page=R1](http://www.nap.edu/openbook.php?record_id=12853&page=R1)
- British Columbia Ministry of Education. (2013a, November 1). Provincial examinations [website]. Retrieved from <http://www.bced.gov.bc.ca/exams/>
- British Columbia Ministry of Education. (2013b). Handbook of procedures for the graduation program. Victoria, BC: Ministry of Education. Retrieved from [http://www.bced.gov.bc.ca/exams/handbook/1314/handbook\\_of\\_procedures.pdf](http://www.bced.gov.bc.ca/exams/handbook/1314/handbook_of_procedures.pdf)
- Council of Ministers of Education, Canada. (n.d.). *Pan-Canadian assessment program (PCAP)* [website]. Retrieved from <http://www.cmec.ca/240/>
- Education Quality and Accountability Office [EQAO]. (2012). *Trends in international mathematics and science study (TIMSS), 2011: Highlights of Ontario results*. Toronto, ON: EQAO. Retrieved from [http://eqao.com/pdf\\_e/12/TIMSS\\_Highlights\\_2011.pdf](http://eqao.com/pdf_e/12/TIMSS_Highlights_2011.pdf)
- Education Quality and Accountability Office. (2012). EQAO: Ontario's provincial assessment program: Its history and influence. Retrieved from [http://www.eqao.com/Research/pdf/E/ON\\_ProvAssessmentPgm\\_9612\\_eng.pdf](http://www.eqao.com/Research/pdf/E/ON_ProvAssessmentPgm_9612_eng.pdf)
- Gierl, M. J., Alves, C., & Taylor-Majeau, R. (2010). Using the Attribute Hierarchy Method to make diagnostic inferences about examinees' knowledge and skills in mathematics: An

- operational implementation of cognitive diagnostic assessment. *International Journal of Testing*, 10, 318-341.
- Government of Yukon Education. (2013a, November 4). *Assessment and evaluation* [website]. Retrieved from <http://www.education.gov.yk.ca/kto12/261.html>
- Government of Yukon Education. (2013b, November 4). *Assessments used in school* [website]. Retrieved from <http://www.education.gov.yk.ca/kto12/280.html>
- Haladyna, T.M., Nolen, S.B., & Haas, N.S. (1991). Raising standardized achievement test scores and the origins of test score pollution. *Educational Researcher*, 20(2), 2-7.
- Herman, J. L., & Golan, S. (1991). Effects of standardized testing on teachers and learning – Another look. Retrieved from <https://www.cse.ucla.edu/products/reports/Tech334.pdf>.
- International Association for the Evaluation of Educational Achievement [IEA]. (2013, November 3). *TIMSS 2011* [website]. Retrieved from <http://timss.bc.edu/timss2011/index.html>
- Jang, E. E. (2008). A framework for cognitive diagnostic assessment. In C. A. Chapelle, Y.-R. Chung, & J. Xu (Eds.). *Towards adaptive CALL: Natural language processing for diagnostic language assessment* (pp. 117-131). Ames, IA: Iowa State University. Retrieved from [http://apling.public.iastate.edu/TSSL/5th\\_2007/proceedings2007/papers/TSSL07\\_11\\_Jang.pdf](http://apling.public.iastate.edu/TSSL/5th_2007/proceedings2007/papers/TSSL07_11_Jang.pdf)
- Klinger, D., DeLuca, C., & Miller, T. (2008). The evolving culture of large-scale assessments in Canadian education. *Canadian Journal of Educational Administration and Policy*, 76, 1-34.
- Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices* (2nd ed.). New York: Springer-Verla.
- Lazer, S., Mazzeo, J., Way, W. D., Twing, J. S., Camara, W., & Sweeney, K. (2010). Thoughts of linking and comparing assessments of common core standards. Retrieved from [http://images.pearsonassessments.com/images/tmrs/tmrs\\_rg/LinkingandComparingCCAsessments.pdf](http://images.pearsonassessments.com/images/tmrs/tmrs_rg/LinkingandComparingCCAsessments.pdf).
- Leighton, J. P., & Gierl, M. J. (Eds.). (2007). *Cognitive diagnostic assessment for education: Theory and applications*. Cambridge, UK: Cambridge University Press.
- Madaus, G. (1988). The influence of testing on the curriculum. In L. Tanner (Ed.). *The politics of reforming school administration*. London: Farmer Press.

- Manitoba Education, Citizenship and Youth. (2013). *Grade 3 assessment in reading, lecture, and numeracy, and grade 4 assessment in French immersion lecture*. Retrieved from [http://www.edu.gov.mb.ca/k12/assess/docs/gr3\\_policy/full\\_doc.pdf](http://www.edu.gov.mb.ca/k12/assess/docs/gr3_policy/full_doc.pdf)
- Manitoba Education, Citizenship and Youth. (2013). *Middle years assessment of key competencies in mathematics, reading comprehension, expository writing, and student engagement*. Retrieved from [http://www.edu.gov.mb.ca/k12/assess/docs/my\\_policy/document.pdf](http://www.edu.gov.mb.ca/k12/assess/docs/my_policy/document.pdf)
- Manitoba Education, Citizenship and Youth. (2013). *Policies and procedures for provincial tests*. Retrieved from [http://www.edu.gov.mb.ca/k12/assess/docs/pol\\_proc/document.pdf](http://www.edu.gov.mb.ca/k12/assess/docs/pol_proc/document.pdf)
- McEwen, N. (1995). Accountability in education in Canada. *Canadian Journal of Education*, 20, 1-17.
- New Brunswick Department of Education and Early Childhood Development. (2013). *Framework for provincial assessments*. Retrieved from <http://www.gnb.ca/0000/results/pdf/AssessmentFrameworkDocument.pdf>
- Newfoundland and Labrador Department of Education. (2013). *Provincial assessments* [website]. Retrieved from <http://www.ed.gov.nl.ca/edu/k12/evaluation/crts/index.html>
- Northwest Territories Education, Culture, & Employment. (2010). *Educating all our children: Departmental directive on student assessment, evaluation, and reporting*. Yellowknife, NWT: Government of the Northwest Territories. Retrieved from <http://www.ece.gov.nt.ca/files/Early-Childhood/directives/2010%20SAER%20DIRECTIVE%20.pdf>
- Northwest Territories Education, Culture, & Employment. (2011). *Education all our children: Procedures, roles and responsibilities for student assessment, evaluation and reporting*. Yellowknife, NWT: Government of the Northwest Territories. Retrieved from <http://www.ece.gov.nt.ca/files/pages/596/saer-procedures-2011-2012.pdf>
- Nova Scotia Education and Early Childhood Development (2013). *Nova Scotia assessments* [website]. Retrieved from <http://plans.ednet.ns.ca/nova-scotia-assessments>
- Nunavut Department of Education. (2013). *Student records and assessment in Nunavut* [website]. Retrieved from <http://www.edu.gov.nu.ca/apps/authoring/dspPage.aspx?page=60>
- Pearson. (2010, February). *Test, measurement & research services: What is a learning progression? (Bulletin Issue No. 12)*. Pearson Education: Nichols, P. D. Retrieved from [http://www.pearsonassessments.com/NR/rdonlyres/6C8F4D6F-EFB1-47CE-9247-3712D274190F/0/Bulletin\\_12.pdf](http://www.pearsonassessments.com/NR/rdonlyres/6C8F4D6F-EFB1-47CE-9247-3712D274190F/0/Bulletin_12.pdf)

- Prince Edward Island Department of Education and Early Childhood Development. (2013). *Provincial assessment program* [website]. Retrieved from <http://www.gov.pe.ca/eecd/studentassessment>
- Programme for International Student Assessment (PISA) 2009. (2009). *PISA 2009 at a Glance*. Retrieved from <http://www.oecd.org/dataoecd/31/28/46660259.pdf>
- Quebec Ministry of Education (2001). *Quebec education program – approved version*. Retrieved from <http://www.learnquebec.ca/toolkit/documents/new/qepelem.pdf>
- Rumberger, R. (2004). *Student mobility and academic achievement*. Retrieved from <http://www.ericdigests.org/2003-2/mobility.html>
- Saskatchewan Ministry of Education. (2013). *Assessment for learning program: Questions and answers*. Retrieved from <http://www.saskschoolboards.ca/scc/ALPQandA.pdf>
- Saskatchewan Education. (2013). *Assessment learning program*. Retrieved from <http://www.education.gov.sk.ca/AFL/AFLProgram>
- Sawyer, R. K. (2006). *Explaining creativity: The science of human innovation*. New York: Oxford University Press.
- Strategic Analysis Office (2011, August 4). *Comparability of high school grades versus other jurisdictions*. [Preliminary study to investigate the comparison of how predictive high school grades were in a first year English class].
- Statistics Canada. (2010). *Measuring up: Canadian results from the OECD PISA study: The performance of Canada's youth in reading, mathematics and science 2009 first results for Canadians aged 15*. Ottawa, ON: Ministry of Industry. Retrieved from <http://www.statcan.gc.ca/pub/81-590-x/81-590-x2010001-eng.htm>
- Statistics Canada. (2011). *Tables for migration: interprovincial 2008/2009*. Retrieved from <http://www.statcan.gc.ca/pub/91-209-x/2011001/article/11514/tbl/tbl-eng.htm#t1>