An Essay on the Challenges of Doing Education Research in Canada

Karen Robson

Abstract
In this essay, I discuss the challenges faced by Canadian researchers in trying to undertake research, particularly in the area of education. I begin by focusing on the issue of data availability (with focus on the lack of race data in Canada) and the extreme limitations that these issues place on the potential for research on important Canadian education issues and then discuss what I regard as hypervigilant data access protocols for Canadian data sets. I then turn to practical issues that arise when comparing education data across cities and countries and the process of “harmonizing” the data. I address the compromises that must be made when attempting to make data comparable across different sites. I conclude by discussing how the larger context in which education occurs must be considered when understanding observed comparative differences between educational outcomes.

Keywords
education, data access, harmonization, comparative research, race

Introduction
In this essay, I discuss data issues in Canada as they pertain to education research. I begin with the issue of the general lack of data. Much of the discourse on this topic—with a few notable exceptions—exists in informal conversations between researchers, and more recently—criticisms from the media. I wish to contribute to the small body of academic dialogue that is documenting this withering of Canadian education data infrastructure despite the very real need for it to create meaningful policy. It is of great importance that we are aware of the data limitations in Canada that severely restrict our ability to produce solid evidence-based education policy on the educational achievement and attainment of equity-seeking groups. One major issue associated with this lack of data is the absence of race being measured. I discuss the implications of this omission throughout this essay. My second goal of this essay is to discuss the difficulties that accompany cross-national research in education. These differences highlight how context is important for understanding the differences that may exist between different sites of data collection. These issues of comparability are inextricably linked to the data limitations that exist in education data in Canada.

1McMaster University, Hamilton, Ontario, Canada

Corresponding Author:
Karen Robson, Department of Sociology, McMaster University, 1280 Main Street, Hamilton, Ontario, Canada L8S 4L8.
Email: kirobson@mcmaster.ca
Data Issues in Canada

There are a number of data issues that I wish to address. The first is the overall availability of data in Canada. Then I move on to critique the existing data that we have in terms of its quality and its access. I also address how the potential for better data exist, but that the will of those who can make this happen does not appear to be apparent for various reasons. In data that do exist, the issue of a key missing demographic indicator—race—is discussed at length.

Availability of Canadian Data

While there are several sources of education data in Canada, their quality pale in comparison to those available in other countries. For researchers such as myself who have had the experience of working in data-rich countries (the United Kingdom and the United States, for example), being restricted by data quality and access is tremendously frustrating.

Data can be made available by various sources, including provincial and federal governments, international organizations, and individual researchers. There are also many kinds of data that are useful for the study of education data and they are generally divisible into two different kinds of types: cross-sectional, or those that collect data at one point in time, or cohort studies, which are longitudinal and collect data from the same individuals at more than one time. There are advantages to both approaches, with the former being cheaper and often resulting in large one-time studies of several thousands of individuals. Longitudinal studies, however, are necessary to demonstrate causality, or to establish how factors can impact on outcomes over the life course. Gathering data from individuals at multiple points in time, however, is a time-consuming and costly process.

Cross-Sectional Large-Scale Assessment Data in Canada

Canadian nationally representative data sources, particularly on issues of education, are extremely limited. Large-scale assessment studies such as PISA (Programme for International Student Assessment), TIMSS (Trends in International Mathematics and Science Study), and PIRLS (Progress in International Reading Literacy) are conducted in Canada and scores of other countries, but their shortcomings for research questions beyond blanket comparisons with other countries are extensive. I will focus on PISA because it is the most known international education assessment, undertaken every three years on 15-year-olds across around 90 OECD countries (Organisation for Economic Co-operation and Development 2018).

Every three years, the results of PISA are released, usually to a media frenzy. In Canada, the fixation is usually on how much better or worse reading or math scores are, compared to other countries. The last round of PISA from 2018 had the Canadian media occupied with reading scores (see, for example, Baxter 2019; CTV News Staff 2019; Journal Pioneer 2019) in their jurisdictions. Invariably, the media also engages in the production of ranking tables, which demonstrate which countries are doing the “best” and “worst” (Anderson and Shendruk 2019). We learned from such recent accounts (by ranking the countries from best to worst scores on the reading standardized test, for instance) that Estonia and Macau were doing “better” than Canada. However, such comparisons, I argue, are not useful beyond alluring headlines. Given the focus on comparing, say, Canada to Estonia or Macao, just how relevant are these comparisons? I would say that comparisons are potentially especially useful if—and only if—they can provide light on each society. However, they must be very carefully done, and it is not entirely clear that the massive, yet quick multinational PISA tables are really the best format for this.

There are two important omissions from data like PISA and similar such studies. One is the absence of contextual characteristics that would allow researchers to thoughtfully explain why
country differences exist. The second is the demonstration of why standardized testing at 15 is an important indicator of anything significant.

While public-release data files of PISA are possible to examine, they suppress or exclude “sensitive” variables. It is unclear if Canada collects race data in their PISA survey, like other countries such as the United States (National Centre for Education Statistics 2020) or the United Kingdom (Department for Education 2019). While the findings on low-income and immigrant students are reported, glossing over this important factor associated with educational attainment and achievement in Canada severely compromises our understanding (and recognition) of systemic racism in the Canadian education system and suggests race does not matter, when an increasing body of evidence is strongly indicating that it does. While Canadian PISA data results often celebrate that the children of immigrants are outstanding performers (Cardoza 2018), “immigrant status” as an indicator does nothing to examine the systemic racism experienced by Canadian-born Black and Indigenous People of Color. This point has been further highlighted by Jones (2014:25) who argued that because Canada does not have national education data . . . it [is] impossible to provide the sort of nuanced analysis that is needed in order to actually understand who is being left behind in terms of access, . . . Observers of Canadian higher education recognize that equity is a problem, but the lack of data on, for example, race, disability, or sexual orientation is a significant barrier to advancing the conversation.

My second criticism of such forms of data is that these are snapshots at one point in time and that PISA and similar tests on 15-year-olds are important, but only insofar as they predict future post-secondary achievement. The PISA process, however, has no ability to do this. It is strongly assumed that these test scores are predictive of later-life success, but these tests have no ability to do so as they are not linked to any future outcome measures. It is also extremely important to note that University of Chicago research has demonstrated that ninth-grade GPA is a stronger prediction of college graduation than standardized tests (Easton, Johnson, and Sartain 2017). While a U.S.-based study, it arguably has a strong relevance to diverse populations, like those in Toronto, with a high proportion of “at-risk” students.

**Longitudinal Studies of Youth in Canada**

While Statistics Canada previously collected longitudinal cohort studies, namely the Youth in Transitions Survey (YITS) (Statistics Canada 2011) and the National Longitudinal Survey of Children and Youth (NLSCY) (Statistics Canada 2010), both these studies were discontinued over a decade ago. Such studies were the only nationally representative way to study the transition of youth through the life course in Canada. There are no (known) plans to restart these studies or to begin similar cohort studies in Canada. This is a great loss to the Canadian data landscape, and as noted by Jones (2014), the absence of such data makes it difficult to advance the conversation of access in Canadian education in any significant way.

There is promise with the new Education and Labour Market Longitudinal Platform offered by Statistics Canada, which links the Postsecondary Student Information System (PSIS), the Registered Apprenticeship Information System (RAIS), and the T1Family File tax records (T1FF). To the best of my knowledge, however, these files are still lacking in rich demographic characteristics, limiting key student characteristics to age, gender, and whether or not the student self identified as an Aboriginal person (Statistics Canada 2020). While it is possible to follow these students’ incomes in great detail, there is no ability to control for parental characteristics or race. For those interested, Finnie et al. (2020) provide a thorough overview of the findings of studies that have employed these data, as well as those that have connected these data with other databases.
We only need to look to other similar countries to see their wealth of longitudinal youth studies. The U.S. Bureau of Labor Statistics has three active longitudinal studies (National Longitudinal Survey of Youth that began in 1979, a similar study that began in 1997, and a study of the children of the 1979 cohort) (Bureau of Labor Statistics, N.d.). The National Longitudinal Study of Adolescent to Adult Health (Add Health) is an active longitudinal study run by the University of North Carolina. It is a nationally representative sample of over 20,000 adolescents that first began in the 1994–1995 school year and have been interviewed five times since.

The United Kingdom has a long history of birth cohort studies, which include all babies born in specific week in the United Kingdom. The first of these studies occurred in 1946, followed by similar birth cohort studies in 1958, 1970, and 2000. All of these studies are housed at University College London. Other longitudinal studies of youth include the Avon Longitudinal Study of Parents and Children, which was started in 1991 and the Longitudinal Study of Young People in England (renamed to Next Steps), which began in 2004 and studies youth born in 1990. All these studies are considered active and have had recent data collection attempts. A systematic review of cohort studies in Australia and New Zealand recently revealed that these two countries have a combined 23 active studies (Townsend et al. 2016). My point is that other similar nations (English-speaking, immigrant-receiving) have embraced and invested in youth cohort studies to drive evidence-based policy around health and education.

Administrative Data

Administrative data is information that is collected by bureaucracies in the course of daily business. In the case of education, provincial governments collect administrative data on all students, from the time they enter public education until the time they leave. Grades, attendance, suspensions, and special education needs are recorded as a matter of record. Such records offer a wealth of information for potential researchers interested in matters of education. They are also longitudinal, meaning we can theoretically examine change over time. Such data, however, are largely inaccessible to (academic) researchers in Canada. Except for British Columbia, provinces do not make such data available to researchers. The case of British Columbia is unique insofar as the BC Ministry of Education and Student Transitions Project administration files are available to qualified researchers in British Columbia who apply for the data. Data are linked by the use of the personal education number that students are assigned in kindergarten. Data from K-12 as well as any postsecondary education done within the province are linked. Linkages in BC are now being beyond postsecondary education and into the labor force.

As mentioned above, administrative data exist de facto in all jurisdictions. Access to it does not. In Ontario, the Toronto District School Board (TDSB) has been the outlier in providing longitudinal data for the study of transitions from high school to postsecondary. Most—if not all—of this work was undertaken by now-retired Research Coordinator by the name of Robert S. Brown who had a deep interest in data infrastructure. This paired with the board’s student census (a self-completion questionnaire), which occurs approximately every five years, has led to a wealth of data, when connected to the administrative data files and postsecondary application data. The TDSB is not the only board in the province to collect such data—but they are the only board to allow access to these data to approve outside researchers. The reasons for the “why” of the lack of data infrastructure in Ontario are unclear, but from an equity perspective Gallagher-Mackay (2017) found that the collection of data on students, particularly around equity-seeking groups, was essential to ensure that Ontario Human Rights around access and student success were being met, yet individuals in positions of power fail to champion such ideas, or incorrectly believe that asking students questions is “illegal,” or that linking student information by way of the Ontario Education Number is also in violation of the law (it is not).
**Individual Researcher Data Construction**

The resources available to federal and provincial bodies are, of course, obviously greater than those that are afforded to individual researchers. There have, however, been notable cohort studies that have been undertaken by academic researchers and their teams to answer research questions around the transitions through education to employment. Notable cohort studies have been undertaken by a handful of Canadian social science researchers, such as Professor Harvey Krahn at University of Alberta (Edmonton Transitions Study), Professor Lesley Andres at UBC (Paths on Life’s Way), and Paul Anisef at York University (Class of ’73). Krahn’s project has recently been in the field collecting data on the high school graduating class of 1985, rendering the study a 32-year span, while Andres’ project began in 1988 and recently went into the field to reconnect with cohort members, making this now a 28-year study. Both aforementioned studies had contact with cohort members most recently in 2010 and 2012, respectively.

**The Canadian Data Access Situation**

Unlike most western countries, there is no long-standing tradition of longitudinal data collection in Canada, particularly in education. This is highly problematic because, as noted by Jones (2014:1332) “governments are perfectly capable of making decisions in the complete absence of relevant evidence, research, or data”—and we see this time and time again with regard to decisions being made around changes to postsecondary education in Ontario. Even in contexts where government or district agencies collect data for their own purposes, making it available for researchers to analyze is difficult and shrouded in fear of the potential loss of privacy and legal ramifications. It is not possible to thoroughly document this phenomenon because much of the discussions about data access are limited to requests that are not answered and conversations that are off the record.

There is a real problem with finding an individual in a government position who is willing to be a champion of opening up data access. Gallagher-MacKay’s (2017:29) analysis of the data access situation in Ontario is the only published accounts of these issues, documenting concerns of the stakeholders that range from (unfounded) legal concerns to political management of the perceived image of the government (i.e. not wanting to be perceived as “too Big-Brotherish”). The underlying concern that education-related data could possibly be used for nefarious purposes is again unfounded by evidence in other countries where the collection and analyses of these sort of data are normalized.

The government, of course, is in the best position to undertake these large-scale data construction tasks. Data collected by the federal government is done so under the auspices of the Statistics Act (Government of Canada 2020), which was written in its first forms in 1918 and last amended in 2017. It details the roles of Statistics Canada, whose role it is to

(a) to collect, compile, analyse, abstract and publish statistical information relating to the commercial, industrial, financial, social, economic and general activities and condition of the people.

Access to data collected by Statistics Canada is limited to individuals who work for Statistics Canada or “deemed employees.” Individuals who are granted the “deemed employee” status can examine data in Research Data Centres that are located across Canada’s postsecondary institutions. The stipulations for being granted access are very stringent and require meeting the criteria of the security policy.

The umbrella Policy on Government Security indicates that:
“researchers must obtain Reliability Status from the STC Departmental Security before having access to the data in the RDC. Security checks are conducted by the RCMP for each researcher accessing data in the RDC.

- As required by the Statistics Act, each researcher accessing data in the RDCs has deemed employee status and swears a legally binding oath to protect the confidentiality of Statistics Canada data utilized in the RDC. This oath is binding for life.
- Each researcher is required to attend an orientation session during which a RDC Analyst explains the researchers’ legal responsibilities to protect the confidentiality and all the security measures in place within the RDC.
- There is a Statistics Canada employee on site to ensure the above measures are clearly understood and adhered to by all researchers participating in the RDC program.”

In order to become a deemed employee, applicants must be background-checked by the RCMP (Royal Canadian Mounted Police), give fingerprints, provide a credit check, and swear an oath to the Queen.

These criteria are required in order to adhere to government security and also to protect the privacy of the individuals from whom data were collected, although the “privacy” of the individual researcher is entirely disregarded.

If a person manages to pass all of these criteria and attend the orientation session, he or she is bound by the data center rules. Only data that have been screened and approved by an RDC analyst can be taken out of the center.

Similar processes exist in other Western countries to access federal data. However, in many of these other countries, alternative non-federal data sources also exist. These are social constructed choices that the government has made to limit access under the guise of “protecting privacy,” by violating the privacy of the individuals who wish to study social trends in Canada and by making such access very inconvenient. My point is that other countries also protect the identities of their research respondents as well as operate under far less draconian data access rules. Canada is not special (or heightened) in terms of its risks for data breaches by social researchers.

**Doing comparative research in Canada.** The previous sections of this paper have identified how the issue of data itself is problematic in Canada. The availability of data—current data—to examine issues around education is unlike that encountered in other comparable democracies. And where it does exist, it can be tedious to access. I now turn to the issues a researcher will encounter when undertaking comparative research with Canadian data, whether it be within-Canada comparative research or international comparative research. Because much of my research examines the associations that self-identified race has with education-related outcomes, much of the focus of my discussion is around the lack of race data in Canada and its important implications.

In terms of understanding the term comparative analysis, such an approach starts with the assumption that some sort of comparison is necessary. Comparisons can take various forms, but generally the comparison is of a time period or with a different population. In other words, you can compare the same type of population (e.g., Canadian students) at different points in time in order to be a “comparative study,” but you can also compare Canadian students with students elsewhere. When I use the term comparative, it is most usually to indicate that I am comparing students in one city (usually Toronto) with students elsewhere—often outside of Canada.

The TDSB is a major exception to the trend to not collect data. The (unfortunate) reason for this student census stems from a human rights case launched by parents in the early 2000s around the disproportionate numbers of racialized students being expelled and sanctioned by “zero tolerance” policies (Ontario Human Rights Commission 2017). Combining these data with the administrative data collected on student enrollments and also connecting these data to the college and university application centers in Ontario provides a rare opportunity for longitudinal analysis in Canada.
Given that Canada is the only OECD country without a federal education system (all issues of education are mandated to the individual provinces and territories), a first logical comparison of Toronto would be with other large Canadian cities. However, no other cities collect comparable data that are accessible to academic researchers. Other school boards in Ontario have recently been administering a similar student census (Peel, Ottawa, Durham), but these data are not available to researchers outside of their respective boards.\textsuperscript{13}

**Comparing Toronto Students to Students in Other Canadian Cities**

As mentioned above, British Columbia has extensive administrative data that documents students from kindergarten to postsecondary, and now even into the labor force. As Vancouver and Toronto are both major cities in Canada, it seemed obvious to compare student outcomes in Toronto with Vancouver. The BC administrative data had information on all students in British Columbia and it could easily be analyzed at the subsample level of Metro-Vancouver.

In order to undertake comparative analysis, however, one must engage in a process that is often called “data harmonization.” This term refers to the process of making various items in data sets comparable with each other. Researchers must decide on the general concepts that are the most important for them and look for how such concepts are measured in each data set. Sometimes this is very straightforward. For example, measuring grade point average on a 4-point scale is usually the same from jurisdiction to jurisdiction (just make sure to note if it is standardized or not). My team and I were able to find comparable indicators on many characteristics, like grades, special education needs, and courses taken.

There is, however, an important missing variable in the Vancouver data: race. The closest measures to race in the BC data are “language spoken at home” and whether or not the student was enrolled in an ESL (English as a Second Language) program which are, at best, weak proxies. But the lack of this simple demographic variable, however, severely limited the extent to which it was useful for comparing to Toronto youth. Black youth—those that experienced the greatest systemic barriers in Toronto—most likely spoke English as their native language. In the end, we decided such comparisons were simply not viable. Race as a demographic measure is something lacking in the vast majority of Canadian education data sets, steeped in a long history of whitewashing racial differences and masking racial disparities in the language of multiculturalism.

Surprisingly, the BC also did not measure for immigrant status, a strategy commonly used in Statistics Canada, along with “visible minority status.” As stated earlier, the topic about an absence of data is not one found in many journal articles, although its prominence is arising more and more in the popular press due to the lack of race based in Canada around issues of education, public health, and arrests (see, for example, Asiedu 2020; Balkissoon 2020; Tunney 2020). Journalists have arguably made more headway in demanding race-based data than have academics.

But why is this important? It is of particular concern that Toronto is one of the few education data sites that contains any information on race—a factor that our research has repeatedly demonstrated to have a strong impact on life chances in Toronto. Using these data, race has been found to be a determinant of various outcomes, including being identified as having a special education need (Parekh and Brown 2020), being deemed “gifted” (Parekh, Brown and Robson 2018), teachers’ perceptions of learning (Parekh, Brown and Zheng 2018), food security (Robson et al. 2021), university choices (Davies, Maldonado, and Zarifa 2014), various elementary school risk factors associated with postsecondary access (Brown, Gallagher-Mackay, and Parekh 2020), socioeconomic disadvantage (Livingstone and Weinfeld 2017), and postsecondary access (Robson et al. 2014). And although recent studies have indicated a closing gap between black students and others in terms of postsecondary access (Livingstone and Weinfeld 2017; Robson et al. 2018), careful examination of results continue to indicate that black students are still very much more likely to
experience socioeconomic disadvantage and be disproportionately affected by academic streaming. Toronto is the only Canadian educational jurisdiction where one can perform this type of analysis. The lack of data from other major cities does not mean there are no systemic racism problems—it means they are not being measured, recognized, and thus addressed.

Glen Jones (2018:24) has similarly argued that

Canada’s national data on higher education is inadequate and in decline. There is no national and little provincial data on higher education participation by groups, other than by gender and socioeconomic status (SES; frequently defined in terms of parental income), and so while equity receives considerable attention in public discussions of higher education, there is relatively little empirical data to support a nuanced analysis of participation.

As mentioned above, replacing “race” with “immigrant status” or “visible minority” is simply not an acceptable alternative to measuring race. There is considerable evidence that children of immigrants have higher educational achievement and attainment than native-born Canadians (see, for example, Abada, Hou, and Ram 2009; Boyd 2002), a fact that is celebrated triannually when PISA results are announced. However, given the highly selective points-based system of Canadian immigration, economic migrants to Canada tend to be highly educated and financially stable (Entorf and Minoiu 2005)—two parental characteristics that account for a great deal of student upward mobility in Canada. In other words, recent immigrants to Canada tend to possess many of the characteristics that are positively associated with their children’s achievement and attainment. In terms of “visible minority”—which essentially is a white/non-white demarcation—not all “non-white” students are the same in terms of educational outcomes. There is considerable heterogeneity in that category which masks the real differences experienced by different racialized groups in Canada. Lumping all immigrants into a single category hides the differences that exist between various immigrant groups. The children of Asian immigrants (with the exception of Filipinos) tend to have higher achievement across generations than most European groups. In contrast, the children of black immigrants do not experience such generalized upward mobility (Abada et al. 2009). However, the differences among immigrant groups can only be discovered in data where race and immigrant status are simultaneously collected. If race is absent and all immigrants are treated as a homogeneous group, the findings likely reveal that the “average immigrant performance” is better or at-par with children of Canadian-born parents—a conclusion that clouds the outcomes experienced by different racial immigrant groups.

**Comparing Data across National Borders**

American education researchers generally have access to school district administrative data that is combined with National Student Clearinghouse data, which is a centralized database that enables school districts to track students link their students through the U.S. postsecondary system. In contrast, we are almost entirely unable to link secondary and postsecondary education data in Ontario.

When harmonizing data between Toronto and Chicago (Robson et al. 2019), we were limited in the demographic variables that they could provide for analysis, but had tremendous envy on the detailed nature of the university and college enrollment information they had access to. Because all their data were of the administrative variety, they did not have many indicators of demographic factors that would typically be reported in our Toronto schools census (e.g. family size, parental education, country of origin—variables that are typically measured through self-report questionnaires).

Our American colleagues, however, had race data—a standard demographic feature in U.S. data sets. The standard way of measuring race in these data, however, consisted of just five
categories: white, black, Asian, Hispanic, and Other. In comparison, the TDSB census collected race data by self-reports from students. As a result, there were dozens of racial identities reported—far more than the six categories that were provided by the U.S. data sources. Because race is a key correlate of education-related outcomes in the research that I do, compromises had to be made to “collapse” the multicategorized race variable in Toronto to the categories provided in the U.S. data.

What became immediately obvious was that the racial categories between Toronto and Chicago were notably different. Black (51 percent) and Latino (33 percent) students were the vast majority of the student population in Chicago Public Schools, while “Asian” (44 percent) and white (35 percent) students comprised the majority of the Toronto Board. In Toronto, collapsing several individual categories into “Asian” meant that all students from “Asian” including South Asian, Southeast Asian, and East Asian were recoded to be grouped together. This was problematic for several reasons, but particularly because the educational outcomes of these different subgroups often differ—particularly for “southeast Asians” (e.g., Filipinos and Vietnamese), who often have lower levels of university attendance. Additionally, South Asian and other Asian groups are obviously very culturally different, sharing only a very broad continent of origin. “Other” students in the Toronto data comprised any student that identified as having a “mixed” origin, although similar students in the United States would have been likely categorized differently. And although “black” students were represented in both data, being “black” meant fundamentally different things for Chicago and Toronto students. In Chicago, these students were more than likely to be third or higher generation Americans, while in Toronto these groups were split between Caribbean and African countries of origin at the first and second generations—both with very different educational outcome patterns (Toronto District School Board, N.d.).

The patterns of white student attendance in both cities were also somewhat problematic, as Toronto percentages generally mirrored the percentages of white youth in the population, but this was certainly not true in Chicago where about 45 percent of Chicago is white, but only 10 percent of Chicago students in the public school system were white, a finding that has been reported elsewhere (Moore 2014; Sander 2006). The reason for this discrepancy between city populations and public-school populations is simply because white families are sending their children to private school. This is an extremely important difference between the U.S. and Canadian contexts where private schooling does exist, but its uptake rate is extremely low (around 6 percent in all of Ontario; Toronto-specific figures are not available). This has important implications for how we understand racial differences—if the “white majority” is by and large admonishing the public-school system in one city context and not the other.

Discussion and Conclusion

In this essay, I have focused on the challenges of doing education research in Canada. Many of the issues I focused on were around the lack of data in general, particularly nationally representative longitudinal data. And where data exist, I have argued that the processes of accessing them are overly laborious, with an underlying assumption that a researcher has nefarious intentions. I also highlighted that how even where data exist, they often fail to contain a key demographic correlate of educational outcomes—race. The lack of data and the failure to include race as an indicator in existing data are not insignificant problems when it comes to creating evidence-based policy. As noted by McKenzie (2020), the collection of these data itself is not enough—it must be linked to action and evidence-based policy. By not measuring it, the problems do not go away, but instead get buried and ignored. McKenzie (2020) argued that measuring race in health data is so critically important during our COVID-19 response because not doing so obfuscates the fact that many racialized persons in Canada are at a higher risk of becoming infected with the virus. An effective pandemic response, for example, must be able to design effective
interventions for different populations. Similarly, addressing inequities in education must be able to identify the subgroups that are most likely barriers due to systemic racism in order to create effective solutions.

Comparing education outcomes across contexts also presents a number of challenges, which was discussed in terms of making sense of racial distributions of public-school students in some U.S. cities. And often, data precision must be compromised in order to allow for comparability. Additionally, all comparative research must account for context—that is, what are the larger external forces driving any differences that are observed? For example, when comparing student postsecondary transition in Ontario to students elsewhere, we absolutely must account for processes of “tracking” or “streaming.” In Ontario, students must pick a majority “applied” or “academic” set of courses in grade 8, which ostensibly determine if they will be eligible for university applications when they graduate.\(^\text{16}\) In contrast, streaming is not practiced (officially or unofficially) in the Chicago public school systems, with the “de-tracking” reforms initiated in 1997 (Nomi and Allensworth 2014). This has important implications for interpreting where jurisdictional differences in postsecondary uptake are found. For example, research in Toronto has found that students with special education needs were disproportionately put into “applied” courses, which made it impossible to be eligible for university (Robson et al. 2014). Thus, while students with special education needs might be accounted for in all data sets being considered, understanding why this characteristic was so strongly associated with college attendance and so negatively associated with university attendance in Toronto (compared to other cities) required the interpretation of these larger structural processes at play.

A final general point about comparing contexts in educational spheres is that unlike other countries, Canada also does not have “selective” universities. The vast majority of colleges and universities are publicly funded and those that are private are not considered “elite”—they are usually religiously based. Although there is somewhat of a stratification between research-intensive and teaching-intensive universities, that is, the “U15 Group of Canadian Research Universities,”\(^\text{17}\) and some media ranking consensus that University of Toronto and McGill are “the best”—the stratification and status of postsecondary institutions in Canada is nothing like that which is observed in the United States and the United Kingdom. Canadian students generally pick undergraduate institutions based on availability of scholarships and proximity to the familial home (Drewes and Michael 2006). Thus, when doing comparative research on university enrollments in Canada with other countries, it is difficult—if not impossible—to account for the selectivity of institutions that is so prevalent elsewhere.

Furthermore, Ontario (and to some extent, Canada more generally) has a college system that is unlike college systems elsewhere insofar as they are not like American “junior colleges” but more like job training institutions where students go to learn specific industry skills. Skolnik (2016) has argued that the Ontario college system lies somewhere between the American and European models of colleges, with a continued focus on skills and job training and only marginal emphasis on baccalaureate credentialism. There are some bridging programs to four-year degrees, but they are fairly new. In terms of practicality, it is very difficult to argue that the college system in Ontario is comparable to the two-year junior college system in the United States.

To conclude, the role that understanding broader context plays in explaining differences by location cannot be understated. It is not enough to say that differences merely exist—the responsible researcher must explain the larger features of a society (e.g. history, political influences, and education policy) that will help explain why those differences exist. The researcher must also consider that differences that appear to exist may also be artifacts of the data compromises they have made to attempt comparability. And most importantly, researchers and policymakers must recognize that failure to find differences do not mean they do not exist—one must ask if the appropriate items have been asked and if data on the issue have actually been collected.
Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD
Karen Robson https://orcid.org/0000-0002-1041-1546

Notes
1. In my previous attempts to access the Statistics Canada version of PISA, the province of student was considered sensitive and not available in the data files.
2. See https://cls.ucl.ac.uk/ for more information on the various birth cohort studies.
3. For additional information, see http://www.bristol.ac.uk/alspac/.
4. See https://esrc.ukri.org/research/our-research/next-steps-formerly-longitudinal-study-of-young-people-lsype/.
5. For information on how to access these data, see https://www2.gov.bc.ca/gov/content/education-training/post-secondary-education/data-research/student-transitions-project.
6. See https://www.tdsb.on.ca/research/Research/Parent-and-Student-Census for more information about the Student Census.
7. Krahm has three cohort studies. See https://edmontontransitionstudy.weebly.com/.
8. Information about Andres’ study can be found here: http://blogs.ubc.ca/paths/.
9. See the project Web site at https://classof73.yorku.ca/.
10. The most recent example of this is removing “access” to postsecondary as a provincial postsecondary funding policy in Ontario and instead replacing it with “key performance indicators” in 2019. The rationale was not discussed in terms of evidence, but in terms of the perceived employability of university graduates (see Friesen 2019).
11. I realize that these criticisms upset people who are champions of Canada’s Research Data Centres. I, however, maintain that these centers dissuade people from doing research. There are other models where the cited concerns over privacy and the potential nefarious behavior of academic researchers have not been seen as the benchmark for designing data access rules. See the UK Data Archive https://www.data-archive.ac.uk/ for example, where it is far easier for Canadian researchers to download decades of longitudinal U.K. research than to access Canadian data records.
12. Collecting ongoing data on race and suspensions allows policymakers to observe if there is a correlation between the two factors.
13. My requests to do so have been entirely ignored.
14. See https://www.studentclearinghouse.org/about/.
15. I say “almost entirely” because there are some pilot projects linking Toronto District School Board (TDSB) data to University of Toronto and York University, but beyond that, no such linkages have been made. Universities are not obligated to make their data available to researchers.
16. It should be noted that streaming is being ended in Ontario—at least for mathematics—in 2020. This is largely a result of the decades of research that has demonstrated its disproportionately harmful effects on black students. See Alphonso (2020).
17. See http://u15.ca/.

References
Abada, Teresa, Feng Hou, and Bali Ram. 2009. “Ethnic Differences in Educational Attainment among the Children of Canadian Immigrants.” Canadian Journal of Sociology 34(1):1–30.
Alphonso, Caroline. 2020. “Ontario to End Academic Streaming that Splits Students into Applied, Academic Tracks.” The Globe and Mail, July 6. Retrieved April 1, 2021. (https://www.theglobeandmail.com/canada/article-ontario-to-end-academic-streaming-that-splits-students-into-applied/).
Anderson, Jenny and Amanda Shendruk. 2019. “The Best Students in the World, Charted.” Quartz, December 3. Retrieved April 1, 2021. (https://qz.com/1759506/pisa-2018-results-the-best-and-worst-students-in-the-world/).

Asiedu, Evelyn. 2020. “Canadian Universities Must Collect Race-based Data.” Maclean’s, July 29. Retrieved April 1, 2021. (https://www.macleans.ca/opinion/canadian-universities-must-collect-race-based-data/).

Balkissoon, Denise. 2020. “Canada’s Dire Need for Better Race-based Data.” Maclean’s, June 8. Retrieved April 1, 2021. (https://www.macleans.ca/opinion/canadas-dire-need-better-race-based-data/).

Baxter, David. 2019. “Sask. Reading, Math, Science Scores Below Canadian Average but there are Positive Trends: Study.” Global News, December 10. Retrieved April 1, 2021. (https://globalnews.ca/news/6280318/sask-reading-math-science-scores-below-canadian-average-but-there-are-positive-trends-study/).

Boyd, Monica. 2002. “Educational Attainments of Immigrant Offspring: Success or Segmented Assimilation?” International Migration Review 36(4):1037–60.

Brown, Robert S., Kelly Gallagher-Mackay, and Gillian Parekh. 2020. “Redefining Risk: Human Rights and Elementary School Factors Predicting Post-secondary Access.” Education Policy Analysis Archives 28:21.

Bureau of Labor Statistics. N.d. “Cohorts.” National Longitudinal Surveys. Retrieved April 1, 2021. (https://www.nlsinfo.org/content/cohorts).

Cardoza, Kavitha. 2018. “In Canada’s Public Schools, Immigrant Students Are Thriving.” Education Week, February 28. Retrieved April 1, 2021. (https://www.edweek.org/ew/articles/2018/02/28/in-canadas-public-schools-immigrant-students-are.html).

CTV News Staff. 2019. “Alberta Students Best in Canada at Reading, Science, International Test Results Show.” CTV News, December 3. Retrieved April 1, 2021. (https://edmonton.cbc.ca/alberta-students-best-in-canada-at-reading-science-international-test-results-show-1.4713229).

Davies, Scott, Vicky Maldonado, and David Zarifa. 2014. “Effectively Maintaining Inequality in Toronto: Predicting Student Destinations in Ontario Universities.” Canadian Review of Sociology 51(1):22–53.

Department for Education. 2019. “Achievement of 15- Year-olds in England: PISA 2018 Results.” Retrieved (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/904420/PISA_2018_England_national_report_accessible.pdf).

Drewes, Torben and Christopher Michael. 2006. “How do Students Choose a University? An Analysis of Applications to Universities in Ontario, Canada.” Research in Higher Education 47(7):781–800.

Easton, John Q., Esperanza Johnson, and Lauren Sartain. 2017. “The Predictive Power of Ninth-grade GPA.” Chicago, IL: The University of Chicago Consortium on School Research.” Retrieved April 1, 2021. (https://consortium.uchicago.edu/sites/default/files/2018-10/Predictive%20Power%20of%20Ninth-Grade-Sep%202017-Consortium.pdf).

Entorf, Horst and Nicoleta Minoiu. 2005. “What a Difference Immigration Policy Makes: A Comparison of PISA Scores in Europe and Traditional Countries of Immigration.” German Economic Review 6(3):355–76.

Finnie, Ross, Masashi Miyairi, and Mich Dubois. 2020. “How Much Do They Make? New Evidence on the Early Career Earnings of Canadian Post-secondary Education Graduates by Credential and Field of Study.” Ottawa, Ontario, Canada: Education Policy Research Initiative and Labour Market Information Council. Retrieved April 1, 2021. (https://lmic-cimt.ca/wp-content/uploads/2020/02/LMIC_How-Much-Do-They-Make-Research-Report.pdf).

Friesen, Joe. 2019. “New Metrics for Ontario University and College Funding Include Employment and Graduation Rates.” The Globe and Mail, April 16. Retrieved March 1, 2021. (https://www.theglobeandmail.com/canada/article-new-metrics-for-ontario-university-and-college-funding-include/).

Gallagher-Mackay, Kelly. 2017. “Data Infrastructure for Studying Equity of Access to Postsecondary Education in Ontario.” Higher Education Quality Council of Ontario. Retrieved February 1, 2021. (https://heqco.ca/pub/data-infrastructure-for-studying-equity-of-access-to-postsecondary-education-in-ontario/).

Government of Canada. 2020. “Statistics Act.” Retrieved April 1, 2021. (https://laws-lois.justice.gc.ca/eng/acts/s-19/fulltext.html).

Jones, Glenn A. 2014. “Building and Strengthening Policy Research Capacity: Key Issues in Canadian Higher Education.” Studies in Higher Education 39(8):1332–42.
Jones, Glenn A. 2018. “Chapter 8: Decentralization, Provincial Systems, and the Challenge of Equity: High Participation Higher Education in Canada.” Pp. 203 in High Participation Systems of Higher Education, edited by B. Cantwell, S. Marginson, and A. Smolentsev. Oxford, England: Oxford University Press. doi:10.1093/oso/9780198828877.003.0008.

Journal Pioneer. 2019. “Province Pleased with P.E.I. Students’ Performance in International Test.” Journal Pioneer, December 6. Retrieved February 15, 2021. (https://www.journalpioneer.com/news/local/province-pleased-with-pei-students-performance-in-international-test-385038/).

Livingstone, Ann-Marie and Morten Weinfeld. 2017. “Black Students and High School Completion in Quebec and Ontario: A Multivariate Analysis.” Canadian Review of Sociology 54(2):174–97.

McKenzie, Kwame. 2020. “Race and Ethnicity Data Collection during COVID-19 in Canada: If You are Not Counted You Cannot Count on the Pandemic Response.” The Royal Society of Canada, November 12. Retrieved November 15, 2020. (https://rsc-src.ca/en/race-and-ethnicity-data-collection-during-covid-19-in-canada-if-you-are-not-counted-you-cannot-count).

Moore, Natalie. 2014. “Why So Few White Kids Land in CPS—And Why it Matters.” WBEZ Chicago Radio, November 14. Retrieved November 30, 2020. (https://www.wbez.org/stories/why-so-few-white-kids-land-in-cps-and-why-it-matters/9b19752f-94cd-47d1-b313-e9e32ebfe187).

National Centre for Education Statistics. 2020. “Program for International Student Assessment (PISA).” Retrieved July 15, 2020. (https://nces.ed.gov/surveys/pisa/pisa2018/#/reading/achievement).

Nomi, Takako and Elaine Allensworth. 2014. “Skill-based Sorting in the Era of College Prep for All: Costs and Benefits,” Research Brief. Chicago, IL: The University of Chicago Consortium on School Research.

Ontario Human Rights Commission. 2017. “Under Suspicion: Concerns about Racial Profiling in Education.” Retrieved March 13, 2021. (http://www.ohrc.on.ca/en/under-suspicion-concerns-about-racial-profiling-education).

Organisation for Economic Co-operation and Development. 2018. “Programme for International Student Assessment.” Retrieved September 16, 2020. (https://www.oecd.org/pisa/).

Parekh, Gillian and Robert S. Brown. 2020. “Naming and Claiming: The Tension between Institutional and Self-identification of Disability.” Canadian Journal of Disability Studies 9(5):347–79.

Parekh, Gillian, Robert S. Brown, and Karen Robson. 2018. “The Social Construction of Giftedness.” Canadian Journal of Disability Studies 7(2):1–32.

Parekh, Gillian, Robert S. Brown, and Samuel Zheng. 2018. “Learning Skills, System Equity, and Implicit Bias within Ontario, Canada.” Educational Policy. Published electronically December 2. doi:10.1177/0895904818813303.

Robson, Karen L., Paul Anisef, Robert S. Brown, and Rhonda C. George. 2018. “Underrepresented Students and the Transition to Postsecondary Education: Comparing two Toronto cohorts.” Canadian Journal of Higher Education 48(1):39–59.

Robson, Karen L., Paul Anisef, Robert S. Brown, and Jenny Nagaoka. 2019. “A Comparison of Factors Determining the Transition to Postsecondary Education in Toronto and Chicago.” Research in Comparative and International Education 14(3):338–56.

Robson, Karen L., Paul Anisef, Robert S. Brown, and Gillian Parekh. 2014. “The Intersectionality of Postsecondary Pathways: The Case of High School Students with Special Education Needs.” Canadian Review of Sociology 51(3):193–215.

Robson, Karen L., Ashley Pullman, Paul Anisef, Robert S. Brown, and Reana Maier. 2021. “A Multi-level Approach to Examining Food Security, Race, Academic Performance, and Post-secondary Confirmation in a Toronto High School Cohort, 2011-2014.” Race Ethnicity and Education 24(1):56–75.

Sander, William. 2006. “Private Schools and School Enrollment in Chicago.” Chicago Fed Letter No. 231. Chicago, IL: Federal Reserve Bank of Chicago.

Skolnik, Michael L. 2016. “Situating Ontario’s Colleges between the American and European Models for Providing Opportunity for the Attainment of Baccalaureate Degrees in Applied Fields of Study.” Canadian Journal of Higher Education 46(1):38–56.

Statistics Canada. 2010. “National Longitudinal Survey of Children and Youth (NLSCY).” Retrieved June 14, 2020. (https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDDS=4450).

Statistics Canada. 2011. “Youth in Transition Survey (YITS).” Retrieved (https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDDS=4435).
Statistics Canada. 2020. “Postsecondary Student Information System (PSIS) Record Layout, Files, and Data Element Descriptions 2018-2019.” Retrieved January 23, 2021. (https://www.statcan.gc.ca/eng/statistical-programs/instrument/5017_Q1_V5).

Toronto District School Board. N.d. “TREND DATA Fact Sheet No. 1.” Retrieved January 23, 2021. (https://www.tdsb.on.ca/research/Research/Research-Reports).

Townsend, M. L., Riepsamen, A., Georgiou, C., Flood, V. M., Caputi, P., Wright, I. M., Davis, W. S., Jones, A., Larkin, T. A., Williamson, M. J., & Grenyer, B. F. (2016). Longitudinal Intergenerational Birth Cohort Designs: A Systematic Review of Australian and New Zealand studies. *PloS one, 11*(3), e0150491.

Tunney, Katherine. 2020. “Federal Agency Teaming Up with the Canadian Association of Chiefs of Police on New Project.” *CBC*, July 20. Retrieved January 23, 2021. (https://www.cbc.ca/news/politics/statistics-canada-race-data-police-1.5650273).

**Author Biography**

**Karen Robson** is Ontario Research Chair in Academic Achievement and At-Risk Youth and Associate Professor of Sociology at McMaster University. Over the last decade, she has regularly engaged in applied research with the Toronto District School Board, the Higher Education Quality Council of Ontario, the Ontario Council on Articulation and Transfer, and the Ontario Ministry of Colleges and Universities.