

## Student achievement trajectories in Ontario: Creating and validating a province-wide, multi-cohort and longitudinal database

Jeanne Sinclair<sup>1,\*</sup>, Scott Davies<sup>2</sup>, and Magdalena Janus<sup>3</sup>

Submission History	
Submitted:	17/07/2022
Accepted:	14/12/2022
Published:	02/02/2023

<sup>1</sup>Memorial University Faculty of Education, 323 Prince Philip Drive, St. John's, NL A1B 3X8, Canada

<sup>2</sup>Department of Leadership, Higher and Adult Education, Ontario Institute for Studies in Education, 252 Bloor Street West, Toronto, ON M5S 1V6, Canada

<sup>3</sup>Offord Centre for Child Studies, BAHT 132, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada

### Abstract

#### Introduction

Longitudinal data that tracks student achievement over many years are crucial for understanding children's learning and for guiding effective policies and interventions. Despite being Canada's most populous province, Ontario lacks such large-scale and longitudinal data on student learning. Linking datasets across cohorts requires rigorous linkage protocols, flexible handling of complex cohort structures, methods to validate linked datasets, and viable organizational partnerships. We linked administrative data on early child development and educational achievement and merged two datasets on characteristics of students' neighborhoods and schools. We developed a linkage protocol and validated how the resulting database could be generalized to Ontario's student population.

#### Methods and analysis

Two main individual-level data sources were linked: 1) the Early Development Instrument (EDI), a school readiness assessment of all Ontario public school kindergartners that is administered in three-year cycles, and 2) Ontario's Educational Quality and Assessment Office's (EQAO) math and reading assessments in grades 3, 6, 9, and 10. To compensate for their lack of a common personal identification number, a deterministic linkage process was developed using several administrative variables. A school-level and a neighborhood-level dataset were also later linked. We examined differences between unlinked and linked cases across several variables.

#### Results and implications

We successfully linked 50% of the EDI's 374,239 cases, 86,778 of which contained all five datapoints, creating a database tracking achievement for multiple cohorts from kindergarten through grade 10, with covariates for their development, demographics, affect, neighborhoods, and schools. Analyses revealed only negligible differences between linked and unlinked cases across several demographic measures, while small differences were detected across a neighborhood socioeconomic index and some measures of child development. In conclusion, we recommend the filling of key voids in sustainable research capacity by creating representative data through linkage protocols and data verification.

#### Keywords

longitudinal data; data linkage; student achievement data

\*Corresponding Author:

Email Address: [jeannes@mun.ca](mailto:jeannes@mun.ca) (Jeanne Sinclair)



## Introduction

Children's academic growth unfolds over many years through their school careers, forming what is known as "achievement trajectories." Successive academic assessments over several years can reveal children's learning pathways and their patterning by demographics, developmental traits, and types of schools and neighborhoods. Such data can offer tools by which educators and policymakers can anticipate students' needs and plan effective programs [1]. These tools have been particularly useful during world-wide COVID-19 school disruptions. Researchers in jurisdictions with rich achievement data have estimated learning shortfalls and helped policy-makers plan learning recoveries [2]. But not all jurisdictions, particularly those outside the USA, have suitable data on achievement trajectories. Even those that value educational research and prioritize evidence-based policies often lack longitudinal data on student achievement. Unlike several other provinces in Canada, namely Manitoba and British Columbia, so far, Ontario, despite its high rank on international assessments [3], lacks longitudinal data on student achievement. Such 'data voids' are likely due in part to the intrusiveness of longitudinal surveys in education, their potential limitations such as sample attrition after a few follow-ups, and prohibitive costs. As an extreme example, the National Children's Study (NCS) in the USA had cost over \$1 billion before it was cancelled [4].

Fortunately, recent technological developments have made it possible to create alternatives to longitudinal surveys. Linking complementary and large-scale administrative databases can permit researchers to track student achievement over multiple timepoints. Further, if those databases are annually or cyclically re-generated, successive cohorts can be continually linked, and thereby provide renewed and up-to-date sources of data that can support sustainable research. However, such linkages need to be anchored by clear protocols and data validation procedures, and they require careful, well-documented, management of complex cohort structures.

This paper describes a new dataset tracking Ontario students' achievement from kindergarten through high school, which was created through collaboration between university-based researchers and provincial administrative authorities, and which successfully filled a crucial data void.

## Research rationales

High-quality data on achievement trajectories require the following data specifications. First and foremost, they need to be longitudinal. Ideally, they should contain successive achievement scores that track students' learning over the bulk of their compulsory schooling careers from the early elementary years through the end of secondary school. Such data can reveal the longer-term dynamics of learning, revealing profiles of students whose achievement levels remain steady over time, or improve or worsen. Second, such data should optimally have initial time points from students' early primary school years, since research shows that early skills can be foundational for later academic success [5–7]. Third, data

should be large scale and jurisdiction-wide so they can be generalized to the full student population and also support analyses broken down by various sub-groups defined by student demographics and/or region. Fourth, combined data sources should offer a rich array of covariates that can be used to model trajectories with suitable predictors. Whereas surveys commonly include multitudes of demographic and attitudinal measures, single administrative data sets rarely do so, and thus benefit from being combined to other data sources. Finally, such data should be multi-level, supplementing information on individual students with measures of their neighborhood and school contexts. Multi-level data can reveal the embedding of achievement trajectories within different kinds of schools and neighborhoods; studies show that socioeconomic characteristics of schools and neighborhoods can have their own residual effects on achievement [8, 9], whether by attracting students with different levels of achievement [10] or by themselves moderating trajectories.

Creating such a dataset in our province was a challenge. Since no single data source met all of these specifications, we had to combine multiple sources. Further, relevant data sources were governed by different authorities. This fragmentation of authority necessitated that we establish data-sharing partnerships. The lead partner in our collaboration was McMaster University's Offord Centre for Child Studies (OCCS). The Centre maintains and administers the Early Development Instrument (EDI), a school-readiness assessment for kindergarteners used internationally to assess their early development in physical, social, emotional, cognitive and linguistic, and communicative and general knowledge domains [11]. Educational and government authorities use EDI indicators to identify variations and trends in early childhood development which guide various programs and initiatives. In Canada, over 1.5 million children have been assessed with the EDI since its inception in the year 2000 [12]. The other major data-sharing partner was Ontario's Educational Quality Assessment Office (EQAO), a semi-autonomous agency of the provincial government and the province's main provider of standardized academic tests for its public schools. Since 1996, EQAO annually tests all students in the province, although assessments were cancelled in 2015 due to teacher labour action and in 2020-2021 due to the COVID-19 pandemic.

These partners provided the main data to be merged. The EDI dataset provided its key baseline: child development indicators measured in kindergarten. EQAO provided its main indicators of achievement trajectories, combining measures from grades 3, 6, 9 and 10. Both datasets contained measures of student demographics, while EQAO also offered measures of student affect and variables pertaining to students' home environments and routines outside of school (i.e., time spent working, reading, watching television, etc.). After establishing this main partnership, the research team was broadened to include academics with access to key school-level and neighborhood-level data, described further below. A team of researchers at the University of Toronto created the Ontario School Census, while a team from McMaster University, the University of British Columbia, the University of Manitoba, and the University of Saskatchewan created

the CanNECD index [13, 14] for neighborhood socioeconomic status.

Our aim was to create a dataset – and validate the linkage – that contains measures of student development collected when children were in kindergarten, their literacy and math scores, attitudes, and various statuses across elementary and high school<sup>1</sup>, a socioeconomic indicator at the neighborhood level, and various school measures at the school level. We included kindergarten cohorts from 2004 through 2012 to maximize the number of data points. Earlier, a subset of EDI and EQAO had been linked ( $N \approx 49,000$ ) [15]. All five EDI domains were shown to reliably predict later academic outcomes in Ontario children, both among those attending English-language [15] and French-language schools [16], replicating patterns found in Manitoba [17], British Columbia [13], Australia [18] and United States [19], jurisdictions that also routinely collect the EDI at a population level. The dataset linked earlier has also been used to examine challenges for children with health disorders, taking advantage of the census-style nature of both assessments. In an investigation linking EDI records on speech and language difficulties in kindergarten with Grade 3 EQAO databases, Janus et al. [20] demonstrated that having such difficulties conveyed higher likelihood of having a special education needs and lower academic achievement levels in Grade 3. Children who were diagnosed with autism spectrum disorder (ASD) by Grade 3 showed quite different behavioural profiles in kindergarten than children who were diagnosed with non-ASD developmental disorder or neurotypical children [21]; these findings have important implications for early identification of ASD.

To build on this work and create a multicohort longitudinal database on student learning in our province, a strong partnership was needed, because authority over relevant data sources was fragmented and uncoordinated. Our partnership allowed us to merge multiple administrative data sources to create longitudinal, province-wide and multi-level data on achievement trajectories for multiple cohorts of students. Its rich data set contains multidimensional measures of children's early development, achievement scores from elementary grades into high-school, demographics and attitudes, and characteristics of their schools and neighborhoods. However, because the linked datasets lacked common personal identification number, they required a viable process for linking cases, which in turn resulted in inability to link all cases. Therefore, we conducted comparisons of linked and unlinked cases in order to validate the linked database's representativeness and thus its capacity to generate inferences about Ontario's student population.

## Methods

Four major databases were merged. Two contained data at the individual-level: EDI and EQAO; and two contained data at the area-level: the Canadian Neighbourhoods Early Child Development (CanNECD) database, and the Ontario School Census.

<sup>1</sup>Literacy is assessed at grades 3, 6, and 10; mathematics is assessed at grades 3, 6 and 9.

## Individual-level databases

### Child development in kindergarten: Early Development Instrument (EDI)

#### Focus and outcome measures

The EDI is a kindergarten-teacher completed checklist for children age 3.5 to 6.5 years. The EDI measures five major domains of early childhood development: physical health and well-being, social competence, emotional maturity, language and cognitive development, and communication and general knowledge. Items are further divided into 16 subdomains within the five major domains. EDI domain scores can range from 0 to 10, with 10 indicating high-level skills. The main outcome measures based on the EDI are domain mean scores, domain vulnerability and overall vulnerability. Vulnerability is assigned using the first baseline dataset: children who score at or below the 10<sup>th</sup> percentile in a given domain are classified as vulnerable on this domain. Children who are vulnerable in one or more domains are classified as vulnerable overall.

#### Additional variables

In addition to child development constructs, the EDI data include child-level demographics (such as date of birth, sex at birth, postal code, first language, English/French language learner status, age at EDI administration), information on child's special needs and functioning status, and class type (e.g., split grade).

#### Mode and periodicity of administration

In Ontario, children can be enrolled in non-mandatory, publicly funded kindergarten in September of the year they turn 4 years of age (junior kindergarten), and in September of the year they turn 5 years they can be enrolled in senior kindergarten. Publicly funded compulsory schooling begins at age 6. Approximately 97% of students attending public school in Ontario entered school in kindergarten [22]. Senior kindergarten teachers complete the EDI for each child in their classroom, based on their observation of the child for at least 4 to 6 months (median age of administration in the present dataset is 5.68 years). The present Ontario-based dataset consists of the EDI data collected over nine years in three-year cycles [23]: Cycle 1: 2004–2006, Cycle 2: 2007–2009, Cycle 3: 2010–2012. In each Cycle, the EDI was completed for all senior kindergarten students in approximately a third of all publicly funded school board in Ontario; thus by the end of the 3-year cycle the EDI assessment was implemented just once in every school district in Ontario. This roll-out strategy allowed communities to utilize longitudinal data by which they could examine changes in patterns of child vulnerabilities and thereby guide new policy and funding directions, without the cost of administering the EDI every single year.

#### Personal identification numbers

A unique numeric indicator called the 'EDI ID' is created by the OCCS for each student, which is based on files supplied by school districts that include the school and school district identification numbers.

## Psychometric and research evidence

The EDI's internal structure and concurrent validity have been shown to be similar across Canada, Australia, the US, and Jamaica [24–26], as well as the Philippines and Indonesia [27]. Numerous studies confirmed the EDI's validity and reliability [28–31]. Teacher ratings gathered through the EDI have predicted child's academic progress in Grade 1 as accurately as a direct assessment in kindergarten [32], and further, the EDI has been shown to reliably predict later academic, social and mental health trajectories [33].

The EDI has anchored a substantial research base on associations between early childhood development and mental health [34], special needs [20, 31], early education and wrap-around programs [35], social determinants of health [33, 36–39], geographic units at various levels [40–42] and a range of indicators of socioeconomic status [43–47]. Bibliography of published papers pertaining to the EDI is available at <https://edi.offordcentre.com/resources/bibliography-of-the-edi/>.

## Student achievement from elementary through high school: EQAO scores

### Focus and outcome measures

EQAO's standardized provincial achievement tests are curriculum-based academic measures of Ontario students' reading, writing, and math achievement, which are recalibrated every year. Literacy skills are assessed at grades 3, 6, and 10, and mathematics at grades 3, 6, and 9. Ontario places children in the appropriate grade based on the calendar year; they must turn a given age by December 31. Thus, Ontario students enter Grade 3 in September at age 7–8, Grade 6 at age 10–11, Grade 9 at age 13–14, and Grade 10 at age 14–15. The EQAO assessment is taken in spring of the academic year.

The EQAO assessments are criterion-referenced measures mapped to respective provincial grade-level curricula in each subject, and they are offered in both English and French. Both the mathematics and literacy assessments include constructed response and multiple-choice items weighted approximately equally. The Grade 9 mathematics assessment is offered at two levels: applied and academic. For all the EQAO assessments except Grade 10 literacy, scores are reported on a scale of 1 to 4, with level 3 meeting the provincial standard. Unlike the other grades' assessments, which inform decisions at the classroom level and higher, passing the Grade 10 literacy assessment (which is scaled from 200–400 with a passing score of 300) is a graduation requirement; students who do not pass may retake the assessment the following year or enroll in a literacy course that they must pass to graduate.

### Mode and periodicity of administration

The EQAO assessments are administered to every eligible student in all Ontario publicly funded schools every year (except for unusual circumstances, such as the COVID-19 pandemic disruptions and union-related job action such as partial strikes; see Table 1) and are recalibrated for every implementation. Students complete the tests at school.

## Personal identification number

The EQAO assigns a student-level person identification number at the time of their first assessment based on the child's unique Ontario Education Number, and the identification number is associated with the same student from primary to secondary school.

## Additional variables

The EQAO dataset contains approximately 450 variables per student, Grades 3–10 inclusive; beyond gender, date of birth, and academic achievement in literacy and mathematics, the EQAO dataset also includes variables representing students' perceptions of learning, self-efficacy, learning strategies, activities in and outside of school, home language environments, and parental engagement and support; these variables are derived from students' responses to a questionnaire administered each year after completing the EQAO assessment. In addition, EQAO asks teachers and principals to add information about students' language, immigration, special education status, and learning and assessment accommodations.

## Psychometric and research evidence

Lower bounds of reliability for the EQAO academic achievement assessments in terms of Cronbach's alpha are provided in Appendix A. All values are within acceptable range [48].

EQAO assessments are primarily used for system-level accountability purposes rather than measuring individual achievement [49]. EQAO data have also been used for a rich variety of research purposes, including investigations into literacy profiles across middle childhood into adolescence [50], relationships between academic achievement and health [51, 52], and the beneficial impact of outdoor greenspaces on student achievement [53]. Ample research using EQAO has focused on mathematics; for example, Larsen and Jang [54] utilized responses to the EQAO student questionnaire about self-efficacy, and teachers' responses to questions on instructional practices to develop a multi-level path analysis articulating the relationship between Grade 6 math achievement and IEP status. Wickstrom et al. [55] examined the association between special learning needs status and students' EQAO math trajectories from grades 3 to 6. EQAO data have also been employed to investigate literacy development across different home language environments [50, 56].

## Area-level databases

### Neighborhood socioeconomic status: Canadian Neighbourhoods and Early Child Development Data Index (CanNECD)

#### Focus

CanNECD SES Index is a neighbourhood-level measure that reflects a combination of socioeconomic data at the neighbourhood level, validated to maximize prediction of

children's outcomes (such as those measured by the EDI) in contrast to existing indicators that were developed and validated with focus on adults (Forer et al. 2020).

### Index development

The index was developed as a part of the Canadian Neighbourhoods and Early Child Development (CanNECD) study [13]. This study created 2,058 neighbourhoods covering Canada (with the exception of Nunavut), each with a minimum of 50 and a maximum of 600 valid EDI records. Over 2,000 socioeconomic status (SES) variables were derived from the Canadian Census and Canadian Income Taxfiler databases in both 2005/2006 and 2010/2011, custom-aggregated to the neighbourhood level [57]. Those variables broadly included income, employment, levels of education, immigration, compositions of family and dwelling, residential transience/stability, language, ethnicity, and age. Using exploratory factor analysis and stepwise multiple regression to predict EDI outcomes, the number of SES variables was reduced to ten, which were then standardized and summarized to form z-score socioeconomic indicators for each neighborhood [57]<sup>2</sup>.

### Database

The CanNECD index contains 277 variables including item- and domain-level data, at risk status, missingness, neighbourhood postal code, SES indices, and derived metrics. For the purpose of the current project, only data for Ontario were included (797 unique neighborhoods).

## School-Level data: the Ontario School Census

### Focus

The Ontario School Census is composed of school-level variables that we will include in our statistical models as potential predictors of student achievement: age of school, number of students enrolled, average standardized test scores in math and literacy (EQAO), school specialties, and the density of public and independent schools in each school's surrounding area.

### Index development

A University of Toronto research team led by Scott Davies assembled a dataset that combined school-level measures for every public and independent school in Ontario over a ten-year span. The researchers first obtained official lists of all public schools and independent schools from Ontario's Ministry of Education. Those lists contained each school's identification number, location, level (elementary and/or secondary), year of opening, and number of students enrolled. The research team then added measures of average standardized test scores

<sup>2</sup>Those ten variables were: the percentage of low income, lone parent families with children under 6; separated or divorced individuals; incomes twice or higher than the provincial median, families with children under 6; union/association dues, families with children under 6; investment income, families with children under 6; non-migrant movers in the past year; charitable donations, families with children under 6; no high school diploma; individuals not speaking either official language at home, and the Gini Coefficient quintile for lone female families with children under 6.

in several years as available, as well as a series of codes on school specialties (e.g., unique curricula, mandates, student populations) as ascertained from their websites, and then merged those two datasets to create a full school census. Since only students in public schools were included in the EDI-EQAO datasets, data on independent schools were used only to construct additional context measures for each public school, specifically measures of "density" – the number of other schools, public and private, within a certain radius of each school. Such measures help measure the surrounding organizational ecology for each school. Once measures of density were calculated in the census, the private schools were deleted from the version of the dataset that was then merged to the larger dataset using school identification numbers that were present in all datasets.

### Database

The school census has dozens of variables on a wide range of school attributes, including aspects of school location, type (e.g., religious, secular, international), level (elementary or secondary), curricular focus, advertising practices, language (French or English), school-average standardized test scores, and density of surrounding schooling options.

### Merging strategy

The merging process began with the EDI and Grade 3 EQAO datasets. Since the two components did not share common personal identification numbers, the team had to develop a different merging protocol. Both datasets contained a deidentified school board (district) identification number, gender, and birth date (with month and year) for each student. As a first step in the linkage process, only students with unique and exact links on these three variables were included, and duplicates were omitted. The process resulted in link rates ranging from 8.6% to 14.9%. These low rates were largely due to the size of the Toronto District School Board (TDSB), Canada's largest school board, as multiple students shared birthdays and gender within the TDSB and these duplicates had to be omitted. Subsequently, the team acquired TDSB school-level data, which allowed school identification numbers to be used in the linking process. This reduced the number of duplicates and added 51,000 new links to the 2005, 2008, and 2011 cohorts, representing a 68% increase in linked students across those three cohorts. If a Grade 3 personal identification number was not available, the Grade 6 personal identification number for the same student could be used instead, as happened with the 2007 EQAO Grade 3 dataset, which was missing birthdate information.

To further improve linking rates for duplicate entries, the team incorporated two additional linkage variables that were present in both datasets: English learner status and French Immersion program participation. If two students had the same information, data regarding English or French as a Second Language (EFSL) and French Immersion were used to determine the correct link. If a student's data did not link on all variables, they were excluded.

We merged the subsequent grade 6, 9, and 10 EQAO scores to the EDI-Grade 3 dataset simply by isolating each EDI administration year, linking the datasets on the EQAO

student personal identification number, and then re-merging all administration years together. Three versions were created using a stepwise inner join merging strategy that restricted each cohort's linked data to the EQAO administration with the least number of participants present for each EQAO administration that was available (see Figure 1). The least restrictive version of the dataset captures all students who were present for at least one EQAO administration; the second least restrictive restricts to students present for all four EQAO administrations (3, 6, 9, and 10) except missingness due to job action and COVID-19, and the most restrictive restricts to students who were present for all four EQAO administrations with no missingness. The fewer assessment timepoints included, the greater the size of the dataset, since each additional assessment adds the constraint that the participant be present for the linked dataset with no missingness. For researchers interested in pursuing analyses about specific academic outcomes (i.e., just literacy or numeracy, or only academic outcomes through Grade 6), the size of the dataset will be larger, as literacy is not assessed in Grade 9, and numeracy is not assessed in Grade 10. The rationale for creating a dataset with no missingness is described later in the section on validation of the linkage and discussion below.

CanNECD data were subsequently linked using the EDI student personal identification number, while Ontario School Census Data were linked using the school identification number.

## Results

### Multiple cohort structure

The merging process created a complex cohort structure which stems from the EDI's cyclical administration until 2014, wherein school boards would administer the EDI once every three years. The complexity is furthered by Ontario teacher unions' job actions (i.e., partial strikes enacting "work-to-rule" policies, in which teachers do not administer the provincial standardized assessments) and pandemic-related cancellations that have prevented EQAO assessment in some years (i.e., 2020, 2021). Further, the sizes of successfully linked cohorts differ, as they are contingent upon which specific boards participated in which year of the first three EDI cycles. The procedures described above effectively handled that complexity, where both EDI and EQAO student personal identification numbers maximized the linked cases. Table 1 presents a summary of the available mergeable components.

### Total linked observations

The proportion of EDI participants linked to the four EQAO datapoints ranged from 42–50%, depending on the cohort year (Table 2). For cohorts from 2009–2012, which had fewer EQAO assessments to match due to job action or COVID-19, up to 57% of EDI cases were linked to at least one EQAO administration.

We began the linkage process with 374,239 total EDI cases, to which 174,685 EQAO Grade 3 cases were linked

using the process described above. EQAO grades 6, 9, and 10 administrations were then linked to that dataset using a stepwise inner-join linking approach. "Inner-join" refers to a linkage of two datasets that restricts the outcome dataset to participants who exist in both datasets. The result was 155,082 cases when including participants who are missing EQAO administrations due to job action or COVID-19. From there, a fully linked dataset ( $n = 86,778$ ) was created which includes participants who took the EDI and each of the four EQAO assessments (with no missing EQAO datapoints). Figure 1 visually depicts the linking process.

The unlinked data ( $n = 170,865$ ) are EDI cases that lacked a link to any EQAO datapoint. Table 2 summarizes the size of each data component as well as the linked and unlinked datasets. The sums are validated as follows: the total EDI cases ( $n = 374,239$ ) equal the sum of the Grade 3 available data ( $n = 174,685$ ), the Grade 6 data for the 2012 EDI cohort ( $n = 28,689$ ), as job action precluded Grade 3 administration in 2015, and the total unlinked EDI cases ( $n = 170,865$ ).

### Validation of the linkage

To validate the utility of the linked dataset, we examined similarities between the linked and unlinked data on key variables. We reasoned that if the two datasets had similar distributions of key variables, we could consider them to be functionally equivalent, and thus potentially generalize from the linked data to the Ontario student population. We selected the most restricted linked dataset (fully linked by inner join across all four EQAO datapoints with no missing administrations,  $n = 86,778$ ) to compare with the unlinked EDI dataset. To judge their similarities, we calculated effect sizes using the R packages "effectsize" [58] and "sjstats" [59] for differences across several demographic, socioeconomic, and developmental variables in the EDI (see Tables 3, 4, and 5). The phi coefficient ( $\varphi$ ) was calculated to measure associations between categorical variables. We interpreted effect sizes as small (0.10), medium (0.30), or large (0.50) [60]. The omega squared ( $\omega^2$ ) coefficient was calculated for differences in continuous variables. This coefficient has been recommended as a relatively unbiased estimate for ANOVA or t-test comparisons by Lakens [61] and can be interpreted like a  $r^2$  value: it measures proportions of variance in a continuous variable that is explained by a grouping variable. Effect sizes are interpreted as small at 0.01, medium at 0.06, and large at 0.14 or greater.

The effect sizes comparing linked and unlinked cases on EDI demographic variables (Table 3) were only negligible, less than 0.10. The largest effect sizes were for English learner status (mean difference in proportion of 0.05) and special needs status (mean difference of 0.03). None met the small effect size threshold.

Effect sizes for EDI domains and age (Table 4) also show only small or negligible differences. Children in the linked dataset had higher EDI scores on all domains, and were slightly older than those in unlinked EDI data set, but mean differences ranged from only 0.32 to 0.63 on a scale of 0 to 10.

Comparisons of neighborhood SES (Table 5) similarly found small effect sizes, with linked samples having slightly higher z-scores.

Table 1: EDI and EQAO components merged into multi-cohort time-series database

Cycle	EDI cohort	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
1	2004	G3			G6			G9	G10							
	2005		G3			G6			G9	G10						
	2006			G3			G6			G9	G10					
2	2007				G3			G6			G9	G10				
	2008					G3			G6			G9	G10			
	2009						G3			JA			G9	G10		
3	2010							G3			G6			G9	C19	
	2011								G3			G6			C19	C19
	2012									JA			G6			C19

G3 = Grade 3; G6 = Grade 6; G9 = Grade 9, G10 = Grade 10; JA = job action; C19 = COVID-19 pandemic.

Table 2: Results of EDI and EQAO linkage process

EDI year	2004	2005	2006	2007	2008	2009	2010	2011	2012	TOTAL
EDI total cases	20,185	46,689	57,992	20,494	40,742	59,066	33,305	38,728	57,038	374,239
<i>EQAO available data</i>										
Grade 3	9,267	24,562	31,851	10,679	23,188	33,377	18,758	23,003	n/a <sup>a</sup>	174,685
Grade 6	9,257	23,807	31,156	10,402	22,433	n/a <sup>a</sup>	18,176	22,163	28,689	166,083
Grade 9	8,468	21,794	27,675	9,602	20,617	29,872	16,762	n/a <sup>b</sup>	n/a <sup>b</sup>	134,790
Grade 10	8,980	23,071	30,390	10,208	21,713	31,885	n/a <sup>b</sup>	n/a <sup>b</sup>	n/a <sup>c</sup>	126,247
<i>EDI stepwise linked to eqao through...</i>										
Grade 3	9,267	24,562	31,851	10,679	23,188	33,377	18,758	23,003	n/a <sup>a</sup>	
Grade 6	9,257	23,805	31,134	10,402	22,433	n/a <sup>a</sup>	18,175	22,161	28,689	
Grade 9	8,460	21,613	27,473	9,505	20,422	29,866	16,519	n/a <sup>b</sup>	n/a <sup>b</sup>	
Grade 10	<b>8,397</b>	<b>21,451</b>	<b>27,266</b>	<b>9,428</b>	<b>20,236</b>	29,624	n/a <sup>b</sup>	n/a <sup>b</sup>	n/a <sup>c</sup>	
EDI unlinked to grade 3 EQAO	10,918	22,127	26,141	9,815	17,554	25,689	14,547	15,725	28,349	170,865
% EDI linked to highest grade level assessed	42% (G10)	46% (G10)	47% (G10)	46% (G10)	50% (G10)	50% (G10)	50% (G9)	57% (G6)	50% (G6)	

a = no assessment due to 2015 job action; b = no assessment due to COVID-19 pandemic (2020–2021); c = no assessment because administration date is in the future. Bolded values represent the total of the fully linked dataset with all five datapoints and no missingness.

Table 3: Effect sizes ( $\varphi$ ) of differences in demographic variables between unlinked ( $n = 170,865$ ) and fully linked ( $n = 86,778$ ) datasets

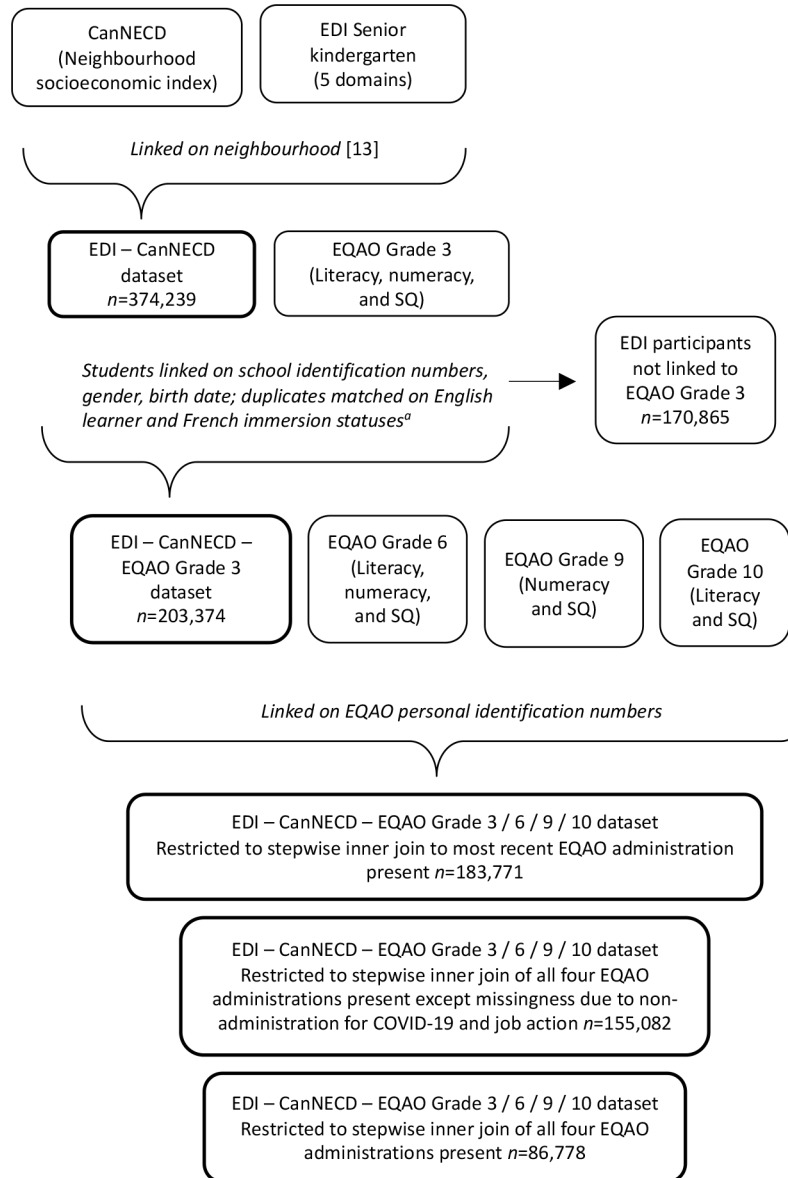
Variable	Proportion in unlinked	Proportion in linked	Effect size result ( $\varphi$ )	Interpretation
Gender (female)	.48	.50	.02	Negligible
English learner	.14	.09	.07	Negligible
Special needs	.05	.02	.07	Negligible
French or English immersion	.09	.10	<.01	Negligible

## Discussion

Overall, we accomplished the goal of merging all four data sources to create a longitudinal and multi-level dataset that is unique within Ontario. This newly created dataset contains child development variables measured when students were

in kindergarten, as well as their literacy and math scores, affective and environmental self-reports, and various statuses when students were in grades 3, 6, 9, and 10. In addition, it contains a socioeconomic indicator at the neighborhood level, and various school measures at the school level. These qualities of the data will make it possible to investigate a range of

Figure 1: Data linkage diagram



Note: <sup>a</sup>2012 EDI cohort matched to Grade 6 EQAO due to job action in 2015. “Stepwise inner join” refers to restricting each cohort’s linked data to the EQAO administration with the least number of participants. Table 2 displays specific sample sizes. Superscripted reference refers to reference list item. SQ = student questionnaire.

Table 4: Effect sizes ( $\omega^2$ ) of differences in EDI domain scores between children in unlinked ( $n = 170,865$ ) and fully linked ( $n = 86,778$ ) datasets

EDI domain	Mean in unlinked	Mean in linked	Effect size result ( $\omega^2$ )	Interpretation
Physical health and well-being	8.67	8.99	.02	Small
Social competence	8.01	8.52	.02	Small
Emotional maturity	7.85	8.21	.01	Small
Language and cognitive development	8.29	8.83	.02	Small
Communication skills and general knowledge	7.32	7.95	.02	Small
Age	5.72	5.68	<0.01	Negligible

pertinent research questions spanning the length of student learning trajectories.

Typically, validating data linkages involves examining discrepancies in linked and non-linked datasets across common

variables such as gender or birthdate [e.g., 62–64]. In our project, we used these and other variables in the linkage process. We compared linked and unlinked samples on the earliest data points (EDI scores) and neighbourhood SES,



Table 5: Effect sizes ( $\omega^2$ ) of differences in CanNECD SES Index between unlinked ( $n = 170,865$ ) and fully linked ( $n = 86,778$ ) datasets

CanNECD SES Index	Mean in unlinked	Mean in linked	Effect size result ( $\omega^2$ )	Interpretation
Time 1(2005–2006)	−0.02	0.19	0.01	Small
Time 2 (2010–2011)	−0.08.	0.14	0.01	Small

and found only negligible to small differences across several indicators of child development and their demographics. Accordingly, we believe that future analyses of our linked dataset can be generalized to the Ontario student population.

While we achieved our aims and created a comprehensive database for over 86,000 students, this dataset has some important limitations. First, it excluded children attending private schools. That exclusion, however, is an inherent characteristic of any dataset that uses EQAO data, as few students in private schools participate in those tests. While percentages vary from year to year and by grade, about 9.5% of Ontario children attended private schools in 2019 [65].

Another limitation of our merging protocol is that it excluded any public-school student who transferred between schools or left Ontario between kindergarten and Grade 3. Indeed, the differences in developmental indicators and the socioeconomic index between linked and unlinked groups, despite their small effect sizes, suggest that the linkage process created some bias by necessitating students to have attended the same school in both kindergarten and Grade 3. This can create an ‘emigration bias’ wherein students who migrate away from a given school district differ non-randomly from those who remain, along observable and non-observable variables [66]. Our linked sample excluded students who switched schools between those grades or had birthdays and genders that were identical to a schoolmate if they were also in the same language background and education group.

Since sharing a birthday with a peer is a random event, this linkage requirement should not have generated sample bias. But, switching schools can be associated with both socioeconomic and academic outcomes. For instance, Goldhaber et al. [67] found moving between schools beyond standard transitions between elementary and secondary schools lowered student performance in Grade 3. Academic performance can be associated with student mobility through the latter’s associations with financial instability, disruption, and poverty, as well as behavioral issues [68]. While relationships between early student mobility and achievement are complex and require further research, comparing the EDI cases unlinked to Grade 3 EQAO participants with EDI cases that were linked by inner join to Grades 3, 6, 9 and 10 (the most restricted sample, which demonstrates the most continuity) appeared to reveal small to negligible bias in our linked sample. In the future, we hope some of that bias can be reduced by institutionalizing common personal identification number throughout our province; doing so would capture students who switch public schools within Ontario. Reducing emigration bias among students who leave Ontario entirely, whether by leaving the province or by switching into the independent school sector, is more challenging, since it would require cooperation among multiple educational authorities to

track students and obtain their standardized tests results over repeated years.

Despite some limitations, our validation procedures suggest that our linkage protocols created data that can represent the full population of several cohorts of Ontario public school students with little bias. In a province that otherwise lacks suitable data, our project has created data with a wealth of covariates that can be used to model student achievement trajectories across the province. The data merging process is partially automated and can easily incorporate future datapoints as they become available. As such, the EDI-EQAO merging process is sustainable and can easily and continually add new cohorts, thus allowing analysis of changes in trajectory trends which can inform school leaders’ and policymakers’ decision-making.

While several provinces and at least one territory in Canada practice data linkage aiming to enable the study of children’s educational trajectory similar to our study goals (e.g., Manitoba [17], British Columbia [69]), the methods of linking and their success are not directly comparable with ours. The major limitation of our study was lack of common identifiers, which is a specific feature of government administration in Ontario – Canada’s most populous province – where separate identifiers are issued through health authorities and education authorities without an existing cross-walk between the two. In the future, our study may pave the way for connecting educational data to administrative databases (e.g., [64]), once legislative obstacles allow the combination of health and education identifiers. In the meantime, however, our study has tested a unique method that allows building useful, linked datasets despite substantial barriers, and creates a potential to catch up to the educational research conducted in other parts of the country.

## Conclusion

Effective policies require strong evidence to guide and support informed decision making. In the fields of child development and education, many jurisdictions, particularly in the USA, have public agencies that implement longitudinal surveys that track children achievement over their school careers. To be useful, such surveys need to follow student achievement over many years, have multiple measures of their demographics and developmental traits, as well as measures of their neighborhood and school characteristics. But since such surveys can be very expensive and intrusive, and can suffer from attrition, many provinces lack them, such as ours.

Our project has created a data set with many appealing features, described above, for providing key insights into crucial questions of student achievement across their

school careers. It was made possible by developing data-sharing partnerships between public educational agencies and academe. Such partnerships are important for establishing precedents. Many public bodies in education are reluctant to share their administrative data, sometimes due to wariness of potential legal ramifications. But viable and successful linkages can ease such worries among public education officials. We hope our project can showcase the potential of data-sharing partnerships, and serve as a role model for similar initiatives in other jurisdictions that lack comparable data.

## Data access statement

The data set at the centre of this study is held securely at OCCS. Data-sharing agreements prohibit OCCS from making the data set publicly available, but access may be granted to those who meet pre-specified criteria for confidential access by contacting the EDI team at OCCS at <https://edi.offordcentre.com/contact>. The full data set creation plan and underlying analytic code are available from the authors upon request; programs may rely upon coding templates or macros that are unique to OCCS.

## Acknowledgments

This paper was supported by an Insight Grant from Canada's Social Science and Humanities Research Council (435-2020-0235). The collection of data for this research was funded by the Ontario Government. The authors gratefully acknowledge the participation of school districts, schools, and teachers. The authors would like to thank leaders and staff at Ontario's Education Quality and Assessment Office for their partnership on this project.

## Statement on conflicts of interest

The Authors declare that there is no conflict of interest.

## Ethics statement

The study received approval from the Hamilton Integrated Research Ethics Board (Approval 8149) and University of Toronto Research Ethics Board (Approval 39677).

## References

1. Lovett MW, Frijters JC, Wolf M, Steinbach KA, Sevcik RA, Morris RD. Early intervention for children at risk for reading disabilities: The impact of grade at intervention and individual differences on intervention outcomes. *Journal of Educational Psychology*. 2017 Oct;109(7):889–914. <https://doi.org/10.1037/edu0000181>
2. Davies S, Aurini J. Estimates of student learning during COVID-19 school disruptions: Canada in international context. In: Vaillancourt, T. et al., editors. *Children and Schools During COVID-19 and Beyond: Engagement and Connection Through Opportunity* [Internet]. Royal Society of Canada; 2021. p. 50–69. <https://rsc-src.ca/en/covid-19-policy-briefing/children-and-schools-during-covid-19-and-beyond-engagement-and-connection>
3. OECD. *Lessons from PISA for the United States, Strong Performers and Successful Reformers in Education*. OECD Publishing; 2011. <https://doi.org/10.1787/9789264096660-en>
4. Kaiser J. NIH cancels massive U.S. children's study. *Science*. 2014 Dec. <https://www.science.org/content/article/nih-cancels-massive-us-children-s-study>.
5. Jimerson S, Egeland B, Teo A. A longitudinal study of achievement trajectories: Factors associated with change. *Journal of Educational Psychology*. 1999;91(1):116–26. <https://doi.org/10.1037/0022-0663.91.1.116>
6. Judge S, Bell SM. Reading achievement trajectories for students with learning disabilities during the elementary school years. *Reading & Writing Quarterly*. 2010 Dec 30;27(1–2):153–78. <https://doi.org/10.1080/10573569.2011.532722>
7. Newton X. End-of-high-school mathematics attainment: How did students get there?. *Teachers College Record*. 2010 Apr;112(4):1064–95. <https://doi.org/10.1177/016146811011200402>
8. Berkowitz R, Moore H, Astor RA, Benbenishty R. A research synthesis of the associations between socioeconomic background, inequality, school climate, and academic achievement. *Review of Educational Research*. 2017 Apr;87(2):425–69. <https://doi.org/10.3102/0034654316669821>
9. Nieuwenhuis J, Hooimeijer P. The association between neighbourhoods and educational achievement, a systematic review and meta-analysis. *Journal of Housing and the Built Environment*. 2016 Jun;31(2):321–47. <https://doi.org/10.1007/s10901-015-9460-7>
10. Marcotte DE, Dalane K. Socioeconomic segregation and school choice in American public schools. *Educational Researcher*. 2019 Nov;48(8):493–503. <https://doi.org/10.3102%2F0013189X19879714>
11. Janus M, Offord DR. Development and psychometric properties of the Early Development Instrument (EDI): A measure of children's school readiness. *Canadian Journal of Behavioral Science/Revue canadienne des sciences du comportement*. 2007 Jan;39(1):1
12. Offord Centre. *Monitoring children's healthy development around the world*. 2018 Oct. <https://edi.offordcentre.com/monitoring-childrens-healthy-development-around-the-world/>
13. Guhn M, Janus M, Enns J, Brownell M, Forer B, Duku E, Muhajarine N, Raos R. Examining the social determinants of children's developmental health: protocol for building a pan-Canadian population-based monitoring system for

- early childhood development. *BMJ Open*. 2016 Apr 1;6(4):e012020 <https://doi.org/10.1136/bmjopen-2016-012020>
14. Janus M, Enns J, Forer B, Raos R, Gaskin A, Webb S, Duku E, Brownell M, Muhajarine N, Guhn M. A Pan-Canadian Data Resource for Monitoring Child Developmental Health: The Canadian Neighbourhoods and Early Child Development (CanNECD) Database. *IJPDS*. 2018;3(3). <https://doi.org/10.23889/ijpds.v3i3.431>
  15. Davies S, Janus M, Duku E, Gaskin A. Using the Early Development Instrument to examine cognitive and non-cognitive school readiness and elementary student achievement. *Early Childhood Research Quarterly*. 2016 Apr 1;35:63–75. <https://doi.org/10.1016/j.ecresq.2015.10.002>
  16. Davies S, Janus M, Reid-Westoby C, Duku E, Schlanger P. Does the Early Development Instrument predict academic achievement in Ontario French schools?. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*. 2021 Jul 1. <https://doi.org/10.1037/cbs0000285>
  17. Brownell MD, Ekuma O, Nickel NC, Chartier M, Koseva I, Santos RG. A population-based analysis of factors that predict early language and cognitive development. *Early Childhood Research Quarterly*. 2016 Apr 1;35:6–18. <https://doi.org/10.1016/j.ecresq.2015.10.004>
  18. Brinkman S, Gregory T, Harris J, Hart B, Blackmore S, Janus M. Associations between the Early Development Instrument at age 5, and reading and numeracy skills at ages 8, 10 and 12: a prospective linked data study. *Child Indicators Research*. 2013 Dec;6(4):695–708. <https://doi.org/10.1007/s12187-013-9189-3>
  19. Duncan RJ, Duncan GJ, Stanley L, Aguilar E, Halfon N. The kindergarten Early Development Instrument predicts third grade academic proficiency. *Early childhood research quarterly*. 2020 Oct 1;53:287–300. <https://doi.org/10.1016/j.ecresq.2020.05.009>
  20. Janus M, Labonté C, Kirkpatrick R, Davies S, Duku E. The impact of speech and language problems in kindergarten on academic learning and special education status in grade three. *International journal of speech-language pathology*. 2019 Jan 2;21(1):75–88. <https://doi.org/10.1080/17549507.2017.1381164>
  21. Janus M, Mauti E, Horner M, Duku E, Siddiqua A, Davies S. Behavior profiles of children with Autism Spectrum Disorder in kindergarten: Comparison with other developmental disabilities and typically developing children. *Autism Research*. 2018 Mar;11(3):410–20. <https://doi.org/10.1002/aur.1904>
  22. Elementary Teachers' Federation of Ontario. Ontario's Kindergarten Program. A Success Story. Executive summary. 2021 Jan. [https://www.etfo.ca/getmedia/05d5d7d5-4253-48ca-92d5-6b0b7e261792/210204\\_ExecSummaryDrGordon.pdf](https://www.etfo.ca/getmedia/05d5d7d5-4253-48ca-92d5-6b0b7e261792/210204_ExecSummaryDrGordon.pdf)
  23. EDI. EDI in Ontario over time – Cycles 1–5 Report. (2021) <https://edi-offordcentre.s3.amazonaws.com/uploads/2021/03/ONT-C1-C5-Web-Report.pdf>
  24. Janus M. Impact of impairment on children with special needs at school entry: Comparison of school readiness outcomes in Canada, Australia, and Mexico. *Exceptionality Education International*. 2011 May 1;21(2). <https://doi.org/10.5206/eei.v21i2.7674>
  25. Janus M, Brinkman SA, Duku EK. Validity and psychometric properties of the Early Development Instrument in Canada, Australia, United States, and Jamaica. *Social Indicators Research*. 2011 Sep;103(2):283–97. <https://doi.org/10.1007/s11205-011-9846-1>
  26. Brinkman SA, Silburn S, Lawrence D, Goldfeld S, Sayers M, Oberklaid F. Investigating the validity of the Australian Early Development Index. *Early Education and Development*. 2007 Oct 11;18(3):427–51. <https://doi.org/10.1080/10409280701610812>
  27. Duku E, Janus M, Brinkman S. Investigation of the cross-national equivalence of a measurement of early child development. *Child Indicators Research*. 2015 Jun;8(2):471–89. <https://doi.org/10.1007/s12187-014-9249-3>
  28. Forer B, Zumbo BD. Validation of multilevel constructs: Validation methods and empirical findings for the EDI. *Social Indicators Research*. 2011 Sep;103(2):231–65. <https://doi.org/10.1007/s11205-011-9844-3>
  29. Hymel S, LeMare L, McKee W. The Early Development Instrument: An examination of convergent and discriminant validity. *Social Indicators Research*. 2011 Sep;103(2):267–82. <https://doi.org/10.1007/s11205-011-9845-2>
  30. Janus M, Reid-Westoby C. Monitoring the development of all children: the Early Development Instrument. *Early Childhood Matters*. 2016;125(1):40–5 <https://bernardvanleer.org/ecm-article/2016/monitoring-development-children-early-development-instrument/>
  31. Janus M, Zeraatkar D, Duku E, Bennett T. Validation of the Early Development Instrument for children with special health needs. *Journal of Paediatrics and Child Health*. 2019 Jun;55(6):659–65. <https://doi.org/10.1111/jpc.14264>
  32. Forget-Dubois N, Lemelin JP, Boivin M, Dionne G, Séguin JR, Vitaro F, Tremblay RE. Predicting early school achievement with the EDI: A longitudinal population-based study. *Early Education and Development*. 2007 Oct 11;18(3):405–26. <https://doi.org/10.1080/10409280701610796>
  33. Janus M, Reid-Westoby C, Raiter N, Forer B, Guhn M. Population-Level Data on Child Development at

- School Entry Reflecting Social Determinants of Health: A Narrative Review of Studies Using the Early Development Instrument. *International Journal of Environmental Research and Public Health*. 2021 Jan;18(7):3397. <https://doi.org/10.3390/ijerph18073397>
34. Thomson KC, Richardson CG, Gadermann AM, Emerson SD, Shoveller J, Guhn M. Association of childhood social-emotional functioning profiles at school entry with early-onset mental health conditions. *JAMA Network Open*. 2019 Jan 4;2(1):e186694. <https://doi.org/10.1001/jamanetworkopen.2018.6694>
  35. Corter C, Patel S, Pelletier J, Bertrand J. The Early Development Instrument as an evaluation and improvement tool for school-based, integrated services for young children and parents: the Toronto First Duty Project. *Early Education and Development*. 2008 Oct 2;19(5):773–94. <https://doi.org/10.1080/10409280801911888>
  36. Halfon N, Aguilar E, Stanley L, Hotez E, Block E, Janus M. Measuring equity from the start: Disparities in the health development of US kindergartners. *Health Affairs*. 2020 Oct 1;39(10):1702–9. <https://doi.org/10.1377/hlthaff.2020.00920>
  37. Hertzman C, Boyce T. How experience gets under the skin to create gradients in developmental health. *Annual Review of Public Health*. 2010 Apr 21;31:329–47. <https://doi.org/10.1146/annurev.publhealth.012809.103538>
  38. Lesaux NK, Vukovic RK, Hertzman C, Siegel LS. Context matters: The interrelatedness of early literacy skills, developmental health, and community demographics. *Early Education and Development*. 2007 Oct 11;18(3):497–518. <https://doi.org/10.1080/10409280701610861>
  39. Webb S, Janus M, Duku E, Raos R, Brownell M, Forer B, Guhn M, Muhajarine N. Neighbourhood socioeconomic status indices and early childhood development. *SSM-Population Health*. 2017 Dec 1;3:48–56. <https://doi.org/10.1016/j.ssmph.2016.11.006>
  40. Kershaw P, Forer B, Irwin LG, Hertzman C, Lapointe V. Toward a social care program of research: A population-level study of neighborhood effects on child development. *Early Education and Development*. 2007 Oct 11;18(3):535–60. <https://doi-org.qe2a-proxy.mun.ca/10.1080/10409280701610929>
  41. Kershaw P, Forer B, Lloyd JE, Hertzman C, Boyce WT, Zumbo BD, Guhn M, Milbrath C, Irwin LG, Harvey J, Hershler R. The use of population-level data to advance interdisciplinary methodology: a cell-through-society sampling framework for child development research. *International Journal of Social Research Methodology*. 2009 Dec 1;12(5):387–403. <https://doi.org/10.1080/13645570802550257>
  42. Raos R, Janus M. Examining spatial variations in the prevalence of mental health problems among 5-year-old children in Canada. *Social Science & Medicine*. 2011 Feb 1;72(3):383–8. <https://doi.org/10.1016/j.socscimed.2010.09.025>
  43. Carpiano RM, Lloyd JE, Hertzman C. Concentrated affluence, concentrated disadvantage, and children's readiness for school: a population-based, multi-level investigation. *Social Science & Medicine*. 2009 Aug 1;69(3):420–32. <https://doi.org/10.1016/j.socscimed.2009.05.028>
  44. Janus M, Duku E. The school entry gap: Socioeconomic, family, and health factors associated with children's school readiness to learn. *Early Education and Development*. 2007 Oct 11;18(3):375–403. <https://doi.org/10.1080/10409280701610796a>
  45. Lapointe VR, Ford L, Zumbo BD. Examining the relationship between neighborhood environment and school readiness for kindergarten children. *Early Education and Development*. 2007 Oct 11;18(3):473–95. <https://doi.org/10.1080/10409280701610846>
  46. Oliver LN, Dunn JR, Kohen DE, Hertzman C. Do neighbourhoods influence the readiness to learn of kindergarten children in Vancouver? A multilevel analysis of neighbourhood effects. *Environment and Planning A*. 2007 Apr;39(4):848–68. <https://doi.org/10.1068/a37126>
  47. Puchala C, Vu LT, Muhajarine N. Neighbourhood ethnic diversity buffers school readiness impact in ESL children. *Canadian Journal of Public Health/Revue Canadienne de Sante'e Publique*. 2010 Nov 1;S13–8. <https://www.jstor.org/stable/41995368>
  48. Bernardi RA. Validating research results when Cronbach's alpha is below .70: A methodological procedure. *Educational and Psychological Measurement*. 1994 Sep;54(3):766–75. <https://doi.org/10.1177/00131644940540030>
  49. Jang EE, Sinclair J. Ontario's educational assessment policy and practice: a double-edged sword?. *Assessment in Education: Principles, Policy & Practice*. 2018 Nov 2;25(6):655–77. <https://doi.org/10.1080/0969594X.2017.1329705>
  50. Sinclair J, Jang EE, Vincett M. Investigating linguistically diverse adolescents' literacy trajectories using latent transition modeling. *Reading Research Quarterly*. 2019 Jan;54(1):81–107. <https://doi.org/10.1002/rrq.220>
  51. Buajitti E, Fazio X, Lewis JA, Rosella LC. Association between lead in school drinking water systems and educational outcomes in Ontario, Canada. *Annals of Epidemiology*. 2021 Mar 1;55:50–6. <https://doi.org/10.1016/j.annepidem.2020.09.011>
  52. Ho ES, Hong B, Steen K, Stephens D, Phillips JH, Forrest CR. Academic achievement in school-aged children with single suture craniosynostosis

- over time. *Plastic Surgery*. 2021:22925503211048526. <https://doi.org/10.1177/22925503211048526>
53. Marmureanu, C. (2021). *The Effect of Vegetation on Student Achievement in the Toronto District School Board* [Ph.D., University of Toronto (Canada)]. <https://www.proquest.com/docview/2556088885/abstract/9D1C2157121D4E30PQ/1>
  54. Larsen NE, Jang EE. Instructional practices, students' self-efficacy and math achievement: A multi-level factor score path analysis. *Canadian Journal of Science, Mathematics and Technology Education*. 2022 Jan 14:1–21. <https://doi.org/10.1007/s42330-021-00181-3>
  55. Wickstrom H, Fesseha E, Jang EE. Examining the relation between IEP status, testing accommodations, and elementary students' EQAO mathematics achievement. *Canadian Journal of Science, Mathematics and Technology Education*. 2020 Jun;20(2):297–311. <https://doi.org/10.1007/s42330-020-00088-5>
  56. Kim H, Barron C, Sinclair J, Eunhee Jang E. Change in home language environment and English literacy achievement over time: A multi-group latent growth curve modeling investigation. *Language Testing*. 2020 Oct;37(4):573–99. <https://doi.org/10.1177/2F0265532220930348>
  57. Forer B, Minh A, Enns J, Webb S, Duku E, Brownell M, Muhajarine N, Janus M, Guhn M. A Canadian neighbourhood index for socioeconomic status associated with early child development. *Child Indicators Research*. 2020 Aug;13(4):1133–54. <https://doi.org/10.1007/s12187-019-09666-y>
  58. Ben-Shachar MS, Lüdtke D, Makowski D. effectsize: Estimation of effect size indices and standardized parameters. *Journal of Open Source Software*. 2020 Dec 23;5(56):2815. <https://doi.org/10.21105/joss.02815>, <https://doi.org/10.21105/joss.02815>.
  59. Lüdtke, D. *sjstats: Statistical Functions for Regression Models (Version 0.18.1)*. 2021. <https://doi.org/10.5281/zenodo.1284472>, <https://CRAN.R-project.org/package=sjstats>.
  60. Wiedmaier, B. Phi coefficient. Allen M, editor. *The SAGE encyclopedia of communication research methods*. SAGE Publications; 2017 Apr 11.
  61. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*. 2013:863. <https://doi.org/10.3389/fpsyg.2013.00863>
  62. Ark TK, Kesselring S, Hills B, McGrail KM. *Population Data BC: Supporting population data science in British Columbia*. IJPDS. 2019;4(2): 1133. <https://doi.org/10.23889/ijpds.v5i1.1133>
  63. Jarvis S, Parslow RC, Carragher P, Beresford B, Fraser LK. How many children and young people with life-limiting conditions are clinically unstable? A national data linkage study. *Archives of Disease in Childhood*. 2017 Feb 1;102(2):131–8. <https://doi.org/10.1136/archdischild-2016-310800>
  64. Saunders NR, Janus M, Porter J, Lu H, Gaskin A, Kalappa G, Guttman A. Use of administrative record linkage to measure medical and social risk factors for early developmental vulnerability in Ontario, Canada. *IJPDS*. 2021;6(1). <https://doi.org/10.23889/ijpds.v6i1.1407>
  65. Statistics Canada. Vast majority of students attended public schools prior to the pandemic. 2020. <https://www150.statcan.gc.ca/n1/daily-quotidien/201015/dq201015a-eng.htm>
  66. Nitsch D, DeStavola BL, Morton SM, Leon DA. Linkage bias in estimating the association between childhood exposures and propensity to become a mother: An example of simple sensitivity analyses. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 2006 Jul;169(3):493–505. <https://doi.org/10.1111/j.1467-985X.2006.00400.x>
  67. Goldhaber D, Koedel C, Özek U, Parsons E. Using longitudinal student mobility to identify at-risk students. *AERA Open*. 2022 Jan;8:23328584211071090. <https://doi.org/10.1177/23328584211071090>
  68. Rumberger RW. *Student Mobility: Causes, Consequences, and Solutions*. National Education Policy Center. 2015 Jun. <https://eric.ed.gov/?id=ED574695>
  69. Lloyd, JV, Hertzman, C. From Kindergarten readiness to fourth-grade assessment: Longitudinal analysis with linked population data. *Social Science & Medicine*. 2009; 68(1). <https://doi.org/10.1016/j.socscimed.2008.09.063>

## Abbreviations

ASD:	Autism spectrum disorder
CanNECD:	Canadian Neighbourhoods Early Child Development
EDI:	Early Development Instrument
EFSL:	English or French as a Second Language
EQAO:	Educational Quality and Accountability Office
NCS:	National Children's Study
OCCS:	Offord Centre for Child Studies
SES:	Socioeconomic Status
TDSB:	Toronto District School Board

## Appendix A

: Ranges of Cronbach's alpha reliability values for EQAO assessments (2007–2019)

Assessment	Minimum	Maximum
Primary Reading English	0.86	0.89
Junior Reading English	0.86	0.90
Primary Reading French	0.85	0.90
Junior Reading French	0.83	0.89
Primary Writing English	0.78	0.84
Junior Writing English	0.78	0.85
Primary Writing French	0.77	0.83
Junior Writing French	0.77	0.83
Primary Math English	0.86	0.91
Junior Mathematics English	0.89	0.92
Primary Math French	0.82	0.91
Junior Mathematics French	0.81	0.91
OSSLT English	0.87	0.90
OSSLT French	0.84	0.88
Grade 9 Mathematics (Applied, Winter, English)	0.81	0.86
Grade 9 Mathematics (Applied, Spring, English)	0.82	0.87
Grade 9 Mathematics Academic, Winter, English)	0.84	0.87
Grade 9 Mathematics Academic, Spring, English)	0.84	0.88
Grade 9 Mathematics Applied, Winter, French)	0.81	0.86
Grade 9 Mathematics Applied, Spring , French)	0.81	0.86
Grade 9 Mathematics Academic, Winter, French)	0.81	0.87
Grade 9 Mathematics Academic, Spring, French)	0.83	0.88

Note: OSSLT = Ontario Secondary School Literacy Test.

