Contrasting Cross-Sectional and Longitudinal Early School Leaver Rates in Canada

Vianne Timmons et Randy Ostridge

Résumé de l’article

L’analyse des données est essentielle à la planification des services éducatifs. En ce sens, déterminer le nombre de décrocheurs est un processus crucial pour planifier les interventions et les initiatives de support au sein des commissions scolaires. Dans le cadre de recherches sur le nombre de décrocheurs précoces à l’Île-du-Prince-Édouard, il s’est avéré que la méthode de présentation des données influence les taux de décrochage rapportés. En fait, deux considérations importantes doivent être considérées lors du compte-rendu ou de l’analyse des taux de décrochage : la définition d’un décrocheur et la manière dont les données ont été recueillies (transversale ou longitudinale). Cet article explore ces deux dimensions.
ABSTRACT. Data analysis is critical to educational planning. Determining the number of school leavers is crucial for a school board when planning for interventions and supports. In researching the number of early school leavers in the province of Prince Edward Island, the method in which the data were reported affected the rates. Two critical considerations should be looked at when reporting or analyzing school leaver rates: the definition of a school leaver, and whether the data collected are cross-sectional or longitudinal. This paper explores these two elements.

CONTRASTING CROSS-SECTIONAL AND LONGITUDINAL EARLY SCHOOL LEAVER RATES IN CANADA

VIANNE TIMMONS University of Regina
RANDY OSTRIDGE InternetWorks, Prince Edward Island

Early school leaver rates are an important indicator of the degree to which schools are achieving their goals. By monitoring early school leaver rates, the educational system can respond by implementing programs and interventions that work towards increasing the success of all students. Two elements affect the rate that is reported: the definition of an early leaver, and the methodology of data collection. Depending on data availability and research goals, various definitions of an early school leaver have been used. In addition, a cross-sectional or a longitudinal methodology may be used to collect data and
to determine the rate, resulting in two different perspectives. Considerations involving these two aspects of early school leaver rates, as well as the merits of both cross-sectional and longitudinal rates, are discussed. This article proposes the use of both cross-sectional and longitudinal rates to provide comprehensive results and understanding of the early leave rate and the process of early exit from the school system.

**INTRODUCTION**

The reporting of early school leaver rates has been considered essential to educational policy development within Canada in recent decades. Lowering the early school leaver rate is an issue of great importance for educators, students, parents, and for Canada as a whole. There are difficulties and variances in statistics regarding this issue, as will be discussed further below, but the overall rate of students not completing secondary school has been estimated at 12%. This leaves many individuals ill-equipped to find employment within Canada’s growing higher skilled labour force, and overall this lack of education limits Canada’s human capital and economic competitiveness on the international scale. These serious implications for both Canada and the early school leavers themselves requires improved data analysis of this issue.

In general, an early school leaver is a student who leaves or “drops out” of the secondary school system, resulting in them not achieving a high school diploma. An accurate understanding of the early leaver rate is necessary for judging how well the school system is serving those students who are in danger of exiting the system. Unfortunately, different early leaver data are reported by various research bodies, making it difficult to determine which information draws the most accurate picture of the educational system. To compound the problem, different organizations often use dissimilar definitions for an early school leaver, which makes direct comparisons not only problematic, but potentially meaningless. In order to compare, contrast, and validate reported early school leaver rates, it becomes increasingly important to understand the major constituents of the early school leaver definition.

Audas and Willms (2001) provide an excellent summary of the ongoing debate in educational literature with regard to the early school leaver problem. Their critique focuses on various methodologies that have been used to collect information about early school leavers. It is beyond the scope of this discussion to address all the possible variables used to define the early school leaver, so our focus will be on two specific components of current definitions. These components are the conditional definition of an early leaver and the nature of data collection. By understanding early school leaver rate definitions, comparisons between rates reported by different studies become more meaningful.

This issue transcends borders, but our focus here is on Canada and the United States, where recent studies have identified early school leaver rates as an area
Contrasting Cross-sectional Longitudinal Early School Leaver Rates in Canada

of concern for Canadian and American education. American educators have been struggling with the early school leaver rate issue over the last decade as states attempt to quantify their early school leaver population (Clements, Ligon, & Paredes, 2000; House Research Organization, 1999; Kaufman, Alt, & Chapman, 2000). In Canada, the early school leaver issue has also been pertinent for a number of years. Attention to this issue has increased recently due to the implementation of systems of accountability for educators, as well as broader systems of educational comparisons between Canada and the international community (Canadian Teachers’ Federation, 2003; Statistics Canada, 2003a). Recently, results from the Human Resources and Skills Development Canada and Statistics Canada Youth in Transition Survey caused many educational organizations to focus on early school leavers. As the cost of students dropping out of school becomes more clearly documented by research, the desire of both provincial and national groups to reduce the early leaver rates increases (Human Resources and Skills Development Canada [HRSDC], 1998, 1999, 2001a, 2000b; Lafleur, 1992). Early leaver rates in Canada have been reduced over the last decade (Bowlby & McMullen, 2002), but it is widely accepted that more could be done for students who are at risk of leaving (Black, 2003; Epp & Epp, 2001; Fallis & Opotow, 2003; MetLife Foundation, 2002; McNeal, 1997; National Dropout Prevention Center, 2002).

From a statistical perspective, the early school leaver rate is challenging to describe (Audas & Willms, 2001; Clements, et al., 2000). Examining the existing research, one can find considerable variance within early school leaver rates reported across Canada. For example, the New Brunswick Department of Education (2004) reports New Brunswick’s 1999 annual early school leaver rate to be 5.0%, whereas HRSDC and Statistics Canada reports the province’s early leaver rate to be 7.6% (Bowlby & McMullen, 2002). If both of these statistics have equally high confidence levels, the natural question is: how can different leaver rates be expressed for the same province?

This question can be answered by investigating the components that define an early school leaver rate. Perhaps the most significant component impacting the magnitude of early leaver rates is the manner in which an early school leaver is defined. In order to categorize each potential leaver, the criteria of categorization must be clearly outlined and expressed. A set of consistent conditions must be laid out which will clearly describe a person either as an early school leaver or not, as this would allow for accurate accounting. The second major influence in understanding early school leaver rates lies in the contrast between a cross-sectional leaver rate and a longitudinal leaver rate. Cross-sectional rates provide an indicator of how many individuals are leaving school at a specific point in time, whereas longitudinal rates express how many individuals in a particular time-dependent cohort are exiting. As studies continue to use a mix of definitions and methods, it becomes increasingly important for educational researchers to distinguish how those components affect the data towards early
school leaver rates (Featherman, Lerner, & Perlmutter, 1994; Magnusson, Bergman, Rudinger, & Torestad, 1991).

**Definition of an early school leaver**

The definition of an early school leaver varies depending upon the outcomes of a particular body of research, which can cause frustration when attempting to make comparisons between studies. The definition becomes very important for quantitative research when an accurate count of early leavers is necessary, but has a lesser impact in qualitative research where sociological models are being developed. For comparison between studies, it is important that the same early school leaver definition be used, or at the very least, differences between definitions are understood and clearly expressed.

A common early school leaver definition used by both Canadian and American institutions and researchers is based upon the timing and circumstances under which the student leaves school (e.g., New Brunswick Department of Education, 2004; Nova Scotia Department of Education, 2003; Kaufman, et al., 2001; HRSDC, 1996). If a student becomes absent without an approved reason at any time during a given academic year, and does not return by the end of the enrollment period of the next academic year, he or she is considered an early school leaver. The enrollment period can vary among school authorities, but it is often within the first 30 days of the new academic year. This definition excludes students who have graduated or received a graduate certificate, transferred between schools or grades (domestic or foreign), left for medical reasons, or have died. The definition includes students who have left for work-related reasons, personal reasons (such as pregnancy or caring for children), and those who were expelled. Students are normally only counted as leaving once within a particular year, regardless of how many times or under what circumstances they exit and return. This particular definition is suitable when determining the cross-sectional early school leaver rate, but not always appropriate when calculating the longitudinal early leaver rate.

When discussing a longitudinal early leaver rate, the complication arises when students exit and return to school several times. Since members of the target population could become school leavers by the cross-sectional definition, but then return two or three years later, they might not be considered early leavers within the context of a longitudinal study. In a temporal perspective, a person would be considered an early school leaver only if she/he exited the system and did not return within the life of the cohort (or the duration of study for the cohort). In that regard, longitudinal rates can be seen as somewhat self-adjusting. Aside from that difference, the cohort-based definition will normally include the same inclusions and exclusions of individuals found in a cross-sectional (or annual) leaver rate. In practice, however, many survey-based longitudinal studies have persons self-identify their educational status, which potentially lessens the precision of the early leaver definition. Strict adherence to an early
leaver definition is of greater importance for educational organizations that generate leaver rates based upon existing student records, for which careful validation is possible.

Even with a clear early school leaver definition, the actual categorization of each member of a population can be difficult. Data validation becomes extremely important depending upon the system used, and it can be complicated by the day-to-day movement of students (House Research Organization, 1999). Though most institutions now use computer databases to keep track of student records, these databases are often decentralized, making it possible for local anomalies to appear within record sets. For example, while working with the student information database system used in Prince Edward Island, we encountered such an irregularity (Timmons, 2005). A student was having difficulties within a particular school that caused him or her to leave for one month. That student was recorded as an early leaver within the school’s localized database system. At the end of the month, the student decided to return to finish his or her education, but enrolled in a different school. This new school recorded the student as an incoming transfer student. The originating school was never informed of the transfer, meaning the student would remain listed as an early leaver in one location, and as a transfer student in another location. If this type of situation is not corrected when the record sets are reconstituted at the departmental level, it is possible to have an inaccuracy within reported statistics. Improved communication within institutions is necessary in order to help avoid these situations, but to ultimately address this issue, a clear and comprehensive early school leaver definition is required.

Some organizations prefer to report the high school graduation or completion rates instead of the early leaver rate (e.g., British Columbia Ministry of Education, 2000; Ontario Ministry of Education, 2000; Statistics Canada, 2003b). A graduation rate represents how many students in a specific age group have earned their high school graduate certificate or equivalent, including a GED. These rates are parallel to the early leaver rate, indicating how many students successfully complete high school. Although a graduation rate provides a picture of how many students are graduating, it provides no information about the students who are not completers. One cannot assume any information about the early leaver population from the graduation rate because not all non-completers are early leavers (for example, some may have moved out of province, become ill or deceased). Graduation rates are attractive because it can be easier to count graduates than it can be to account for non-graduates. Focusing on graduation rates also tends to shift perspective from a deficit-based model to a success-based model, where organizations can report the number of students who have successfully negotiated the educational system. However, both graduation rates and leaver rates have value to educators, as they are not explicit opposites of each other.
Once an early school leaver definition has been established, it can be used to count the number of early leavers within a particular research paradigm. This total can be used in conjunction with enrollment or cohort numbers to calculate an early school leaver rate. The actual value of the rate will depend to a great degree upon how early leaver information is collected. The following discussion will look at how the two different data collection methodologies affect the early school leaver rates, focusing on how they contrast with each other. As mentioned, the two types of early school leaver rates are cross-sectional and longitudinal.

**CHARACTERISTICS OF A CROSS-SECTIONAL EARLY SCHOOL LEAVER RATE**

According to the common definition of an early school leaver, a cross-sectional (or annual) leaver rate represents the number of students who exited school within a specific academic year. Cross-sectional rates provide a snapshot of the target population within any given year, usually reporting based upon age or grade level.

A cross-sectional leaver rate is convenient because it inherits the classical features of cross-sectional analysis. The key quality of cross-sectional research is that it measures characteristics of a target population at a single point in time or over a very short period of time (Diggle, Liang, & Zeger, 1994). Cross-sectional studies are good at describing information that has few independent variables, with the general understanding that many assumptions about dependent variables are needed to correlate different characteristics within a population. Although a small number of independent variables can be somewhat limiting, one positive aspect is that uncertainty within results can be better quantified and controlled. It is very difficult to address temporal effects in cross-sectional analysis, as there are typically not multiple time-dependent variables to consider for comparison. For example, early leaver data taken cross-sectionally will provide no information about how many members of that population will ultimately return to the educational system in successive years. The scope of the data collection is simply too narrow. Cross-sectional data can be collected prospectively (from the present and following individuals forward over time) and retrospectively (from historical or pre-existing sources). In either case, the information collection is completed for a single time frame within the life of the sample.

A cross-sectional early school leaver rate is typically calculated for a single academic year. The mechanics of the calculation will vary based upon the manner in which the data have been collected, but normally only the total enrollment numbers and the total leaver numbers are required. Cross-sectional annual rates provide organizations with descriptive baseline data from year to year. As long as consistent data collection and analysis procedures are used, these values can be a convenient method for comparison of leaver rates over time. Given
Contrasting Cross-sectional Longitudinal Early School Leaver Rates in Canada

the cumulative nature of longitudinal rates, cross-sectional rates tend to be the smaller. This often makes them more attractive to self-reporting educational institutions, such as school boards or departments of education.

The ability to inspect school records retrospectively by year is another reason why educational institutions find cross-sectional rates attractive. Logistically, annual studies are typically the most manageable method of data collection for research purposes. This quality remains true when calculating leaver rates. Cross-sectional data collection also tends to be less expensive than a longitudinal study because the information and personnel management issues are reduced, and participant attrition is a much lower concern. The major limitation of a cross-sectional annual leaver rate is that it only captures those students who have left school within a given academic year; it does not capture information about those who have left in any other year nor those who will later return to complete their education.

In Canada, cross-sectional early school leaver rates will most commonly be reported by provincial ministries and departments of education (e.g., New Brunswick Department of Education, 2004; Nova Scotia Department of Education, 2003). Enrollment information must be processed as part of educational accountability for these organizations, so the leaver rates are calculated as an extension of this mandate. Cross-sectional rates presented by the provinces normally have low uncertainty because they are generated based upon accounting for each student in the province, not by a random or stratified sampling methodology.

CHARACTERISTICS OF A LONGITUDINAL EARLY SCHOOL LEAVER RATE

A longitudinal early school leaver rate measures the number of students who have exited school over the lifetime of a given population. Typically, the population will be defined by age, since cohorts created by grade level can have internal fluctuations because of how students are promoted or repeated. A longitudinal leaver rate provides not only useful statistical information for a single year, but can provide a picture of what is happening within the population from year to year. This sets the longitudinal approach apart from the cross-sectional approach, adding new layers of depth to the information collected and the analysis options available.

The defining quality of longitudinal research is that individuals are measured repeatedly throughout time. This contrasts with cross-sectional studies, where the data collection occurs at a single point in time. As described in Diggle, et al., (1994), longitudinal studies are very good for describing cohort and age effects, which are more difficult to assess with a cross-sectional approach. The cohort effect considers differences within baseline information about a population. The age effect considers changes within a population over time. For example, within a population of early school leavers, a cohort effect would
include how many males and females exist within the population. An age effect for the same population might consider the change in income rates for early school leavers over time. When working with longitudinal information, it is very important that the research design is precise and appropriate for the analysis being conducted, accounting for both the age and cohort effect. A flaw in methodology is very difficult to correct from one stage to the next without potential contamination of the results.

Most longitudinal studies collect information prospectively, since historically collected data may be of lower quality. Prospective collection also provides researchers with a built-in opportunity to manage variability factors over the life of data collection. Logistically, longitudinal studies are often the most difficult type of study to manage. Continued observation of the target population requires a large investment in resources, and careful monitoring of the data collection process is also necessary. In the case of at-risk groups like early school leavers, participant attrition becomes an issue over longer studies, affecting the quality and consistency of collected data.

The longitudinal early school leaver rate is calculated for a specific age cohort by counting all members of the population that have exited school up to that point in the lifetime of the cohort. The total number of leavers is then used in conjunction with the total population of the cohort to calculate an exit rate. On the surface, the longitudinal rate would appear to be a cumulative total of each year’s cross-sectional leaver rate, but this is not the case. As mentioned when discussing the cross-sectional leaver rate, a student may be counted as a leaver within a given academic year, but then return some time later. This type of re-entry causes the longitudinal rate to be slightly less than the sum of a cross-sectional rate. Some jurisdictions in the United States use a formula to convert the sum of cross-sectional rates into a longitudinal rate. A scaling factor would be applied to reflect a re-entry rate of a jurisdiction, thereby approximating the longitudinal rate. These rates are not true longitudinal rates, however, and that distinction is sometimes not clear.

In North America, there have been several large longitudinal projects focusing on education. In Canada, two large-scale national longitudinal surveys include the Statistics Canada-administered National Longitudinal Survey of Children and Youth (NLSCY) and Youth in Transition Survey (YITS). In the United States, the U.S. Department of Education National Educational Longitudinal Study and the U.S. Department of Labor National Longitudinal Survey of Youth are two important longitudinal studies investigating school-aged youth. Studies of this nature have done a great deal to collect information relating to the overall school experience, including components specifically connected to early school leavers.
CONTRASTING CROSS-SECTIONAL AND LONGITUDINAL EARLY SCHOOL LEAVER RATES

The purpose of identifying early school leaver rates is to provide a snapshot of what is occurring within the educational system, in order to better evaluate current practices and help formulate more effective educational services. Early school leaver rates provide important data for this analysis; however, both the cross-sectional and the longitudinal rates provide slightly different views of the leaver situation. While it is not possible to directly compare cross-sectional and longitudinal leaver rates, there are some situations in which it may be desirable to contrast the two types of calculations. As discussed when defining a cross-sectional leaver rate, this type of rate is used most often by educational institutions who are required to report leaver rates on an annual basis. Longitudinal rates are more often used by large national studies or academic research groups. The clearest picture of the early leaver situation can be found using the cross-sectional and longitudinal approaches to validate each other. Since direct comparison of cross-sectional and longitudinal rates is not possible, indirect comparisons must be made.

As noted, early school leaver rates reported by educational institutions are often cross-sectional, presenting a picture of the leaver population year by year. Many departments and ministries of education in Canada fall into this category, and their annual reports will include cross-sectional leaver rates. Nationally, the Government of Canada has sponsored several studies that have focused on building a profile of the educational system in Canada. These national studies are generally completed on a larger scale, capturing longitudinal information that produces longitudinal leaver rates. As these national studies are conducted less frequently than the annual reporting provided by provincial agencies, the two types of studies cannot be compared on a year-by-year basis. Also, because of the differences in both early school leaver definitions and data collection methodologies, there are some challenges in attempting to contrast the rates against each other when they do coincide for a given year. It should also be noted that most of the national longitudinal surveys use a random or stratified sampling methodology to collect their information, which results in large uncertainties in provinces with smaller populations.

Let us return to the example provided in the introduction of this discussion. The New Brunswick Department of Education (2004) reports the 1999 annual early school leaver rate to be 5.0%. The YITS results report that the early school leaver rate for New Brunswick in 1999 is 7.6% (Bowlby & McMullen, 2002). Which provides the most accurate picture of what is happening in the school system during that particular year? The New Brunswick Department of Education rate is cross-sectional, and is drawn from enrollment numbers for students that are in grades 10 to 12 during 1999. Therefore, the New Brunswick Department of Education rate is providing information on students who were enrolled in high school and did not return for the 2000 school year. The
YITS rate is a longitudinal rate based upon a stratified sample of 20-year-olds who self-identified their school status. Therefore, the YITS rate infers the total number of 20-year-olds that do not have a high school diploma or equivalent certification. The two rates are both called early leaver and dropout rates, but they are clearly not expressing the same information. Both rates do have value to policy makers, however, and one rate cannot be viewed as necessarily superior to the other. Under what circumstances would an organization choose a cross-sectional rate as opposed to a longitudinal rate, or conversely, does this division even have to exist?

The most accurate picture of early school leavers across Canada can ultimately be gained by utilizing the cross-sectional information generated by provincial organizations in concert with the longitudinal data produced by national surveys. By the nature of its definition, a longitudinal study is capable of expressing all of the same information that a cross-sectional study can collect, but continuous longitudinal study of youth nationwide is not a realistic goal. To date, instances of large Canadian longitudinal studies on youth have been separated by five to ten years. It would be beneficial to policy makers to have more frequent information on early leaver rates, which is where a cross-sectional approach can assist in completing the picture. The cross-sectional leaver rate can provide information from year to year, highlighting possible immediate effects of policy changes. If all provinces in Canada were to report cross-sectional leaver rates with a similar baseline definition, then a summative national rate could be produced within any given year. All limitations of the cross-sectional methodology would apply, but this information could be used to guide policy between national surveys. Since the longitudinal approach typically represents a sampled rate, a cross-sectional rate based upon school records could reduce the uncertainty for jurisdictions that have smaller populations.

**CONCLUSION**

Various indicators of early school leaving are reported by different organizations and researchers. An understanding of these indicators requires that both the definition of an early school leaver and the specific methodology be considered. The nature of the reported rates varies by the methodology used – be it cross-sectional or longitudinal. Both methodologies result in meaningful measures. However, using cross-sectional data to determine annual early leaver rates provides year-to-year measures that may be more reliable than longitudinal data from samples, whereas a richer understanding of educational status over time, as well as factors related to school success, can be gained from longitudinal data. By using data from both cross-sectional and longitudinal studies, more complete knowledge about early school leavers can be gained. Given the costs of early school leaving, both to society and the individual, taking advantage of both cross-sectional and longitudinal data is recommended.
REFERENCES


VIANNETIMMONS is President and Vice-Chancellor of the University of Regina. Her research interests include student retention, family literacy, and the relationship between inclusion and health.

RANDYOSTRIDGE has worked with technology in an educational capacity for nearly a decade. Starting in analysis and interpretation, he has since moved on to educational software development, focusing on best practices in multimedia integration and usability design.

VIANNETIMMONS est présidente et vice-rectrice à l’Université de Regina. Ses intérêts de recherche portent sur la rétention des élèves, la littératie au sein des familles ainsi que les relations existant entre l’inclusion et la santé.

RANDYOSTRIDGE a travaillé pendant presqu’une décennie dans un contexte technologique. Débutant sa carrière en analyse et en interprétation, il a depuis développé une expertise en développement de logiciels éducatifs, privilégiant des pratiques éprouvées en intégration multimédia et en utilisabilité du design.