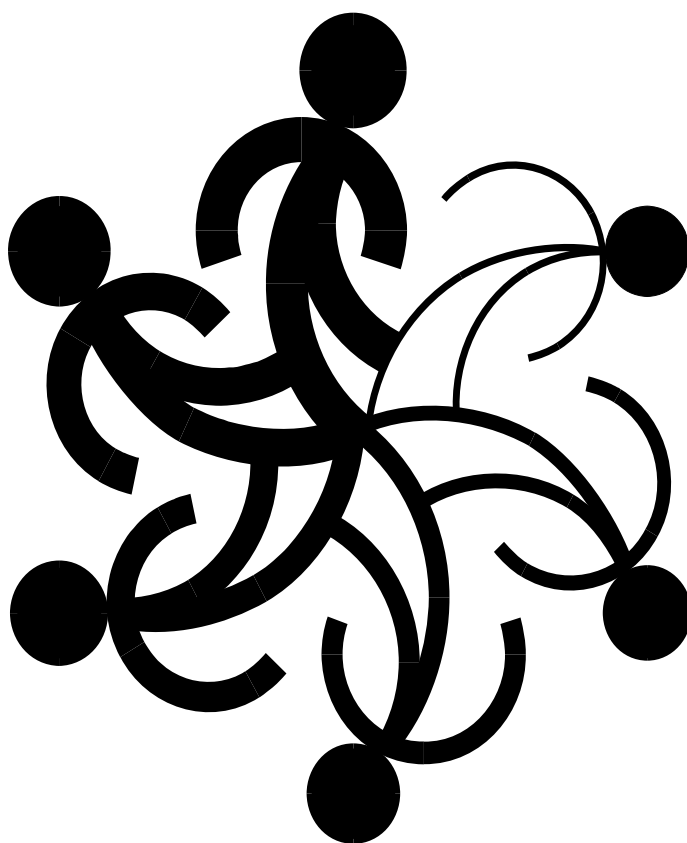


# National Longitudinal Survey of Children and Youth

Survey Overview  
for the 2008/2009 Data Collection  
Cycle 8



Statistics Canada  
Statistique Canada



Human Resources and Skills Development Canada  
Ressources humaines et Développement des compétences Canada

Canada



## Table of contents

Introduction .....	5
Acknowledgements .....	7
Background .....	9
Objectives .....	11
The National Longitudinal Survey of Children and Youth design .....	13
Content of the survey .....	19
Survey scales .....	21
Self-complete questionnaires (ages 14 to 17) .....	23
Main changes to Cycle 8 since Cycle 7 .....	25
Data collection .....	27
Direct assessments .....	31
Weighting and data analysis .....	33
How to access the National Longitudinal Survey of Children and Youth data .....	37
Detailed Cycle 8 measures .....	38



## **Introduction**

The purpose of this document is to describe the content and design of Cycle 8 (2008/2009) of the National Longitudinal Survey of Children and Youth (NLSCY). This document builds on the information presented in the previous NLSCY Overviews for Cycles 1 to 7.

The NLSCY has been conducted by Statistics Canada and is sponsored by Human Resources and Skills Development Canada (HRSDC). Statistics Canada is responsible for data collection, while HRSDC provides the overall direction to the survey. Both agencies have played a role in funding, development of survey content, research, and dissemination of findings.

In addition, HRSDC and Statistics Canada continue to benefit from the advice and contribution of a variety of interested partners. Provincial and territorial governments are kept informed of progress and their representatives provide valuable input.

HRSDC and Statistics Canada have relied on advice provided by HRSDC's Expert Advisory Group on Children and Families, a multi-disciplinary group of Canadian and international experts in child development. They are consulted on survey design, survey questions and research priorities, and are responsible for much of the original research that is conducted using the survey data. Additional experts are consulted as required.

For further information on the NLSCY, enquiries should be directed to:

### **Statistics Canada**

Client Services  
Special Surveys Division  
Telephone: 613-951-3321 or toll-free 1-800-461-9050  
Fax: 613-951-4527  
E-mail: [ssd@statcan.gc.ca](mailto:ssd@statcan.gc.ca)

This is a companion report to the documents containing the survey questionnaires, *National Longitudinal Survey of Children and Youth: Survey Instruments for the 2008/2009 Data collection, Cycle 8*, Books 1 and 2 available at the following link:

<http://www.statcan.gc.ca/cgi-bin/imdb/p2sv.pl?Function=getInstrumentLink&SurvId=4450&SurvVer=1&Instald=16044&InstaVer=8&lang=en&db=imdb&adm=8&dis=2>



## ***Acknowledgements***

The National Longitudinal Survey of Children and Youth (NLSCY) is the responsibility of a joint team of staff at Human Resources and Skills Development Canada and the Special Surveys Division at Statistics Canada.

The team would also like to acknowledge the continuing cooperation of the responding children and youth and their families, and the work of the survey interviewers and the Statistics Canada Regional Office staff.





## **Background**

The National Longitudinal Survey of Children and Youth (NLSCY) is a long-term study conducted, every two years, by Statistics Canada and sponsored by Human Resources and Skills Development Canada (HRSDC). The primary objective of the NLSCY is to monitor the development and well-being of Canada's children from infancy to adulthood.

The NLSCY follows a representative sample of Canadian children from birth to early adulthood, with data collection occurring at two year intervals. The first collection of information (Cycle 1) took place in the winter and spring of 1994/1995. In addition to following the original longitudinal panel of children, now aged 14 to 25 years in Cycle 8, the survey has continued to add and follow a new sample at each cycle to monitor early childhood development. In Cycle 8, this sample was comprised of children aged 0 to 7 years.

Much of the information in the NLSCY is collected from parents on behalf of their children by means of a household interview. Children aged 14 to 17 complete a separate written questionnaire in the home. Finally, the NLSCY includes direct measures of achievement: interviewers administer a receptive vocabulary test as well as a test of early writing and numeracy skills for children aged 4 to 5 years. Children in grades 2 to 10 complete a short mathematics/computation assessment. Youth aged 16 and 17 years complete a Problem Solving Exercise. The 18- and 19-year-olds complete a Literacy assessment and the 20- and 21-year-olds complete a Numeracy assessment.

Cross-sectional data from Cycles 1 to 3 are available on a public use microdata file (PUMF); no PUMF was created for Cycles 4, 5, 6, 7 and 8. Statistics Canada retains a master microdata file from which tabulations can be requested. Other options for access to the longitudinal data include remote data access and the Statistics Canada research data centres.



## **Objectives**

The objectives of the National Longitudinal Survey of Children and Youth (NLSCY) are:

- to determine the prevalence of various risk and protective factors for children and youth;
- to understand how these factors, as well as life events, influence children's development;
- to make this information available for developing policies and programs that will help children and youth;
- to collect information on a wide variety of topics – biological, social, economic;
- to collect information about the environment in which the child is growing up – family, peers, school, community.



## The National Longitudinal Survey of Children and Youth design

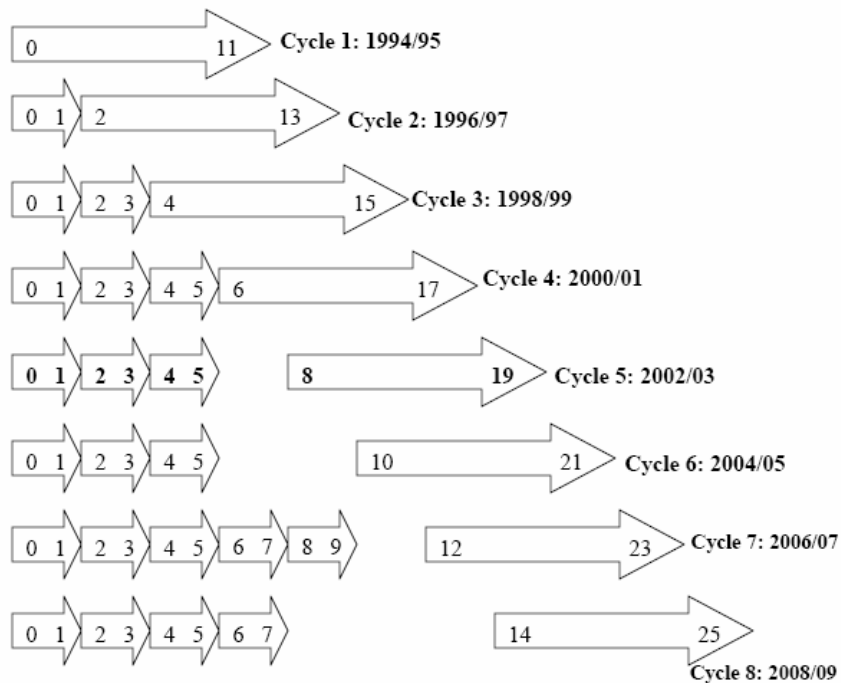
### Survey methodology – Sample

The National Longitudinal Survey of Children and Youth (NLSCY) is a probabilistic survey developed to provide information on Canadian children and youth. To produce reliable estimates that respond to the expressed needs of the clients, a representative sample of children and youth was selected. This section describes the sample selection method and size.

At Cycle 8, the NLSCY sample consists of children aged 0 to 7 years old (Early Childhood Development or ECD children) and youth and young adults aged 14 to 25 years old (original cohort). The effective age at Cycle 8 is as of December 31<sup>st</sup>, 2008. Thus, 0-year-olds are born in 2008 and 1-year-olds are born in 2007.

The diagram below illustrates the NLSCY sample. The years indicate when collection occurred. The larger arrows represent the original cohort, and the smaller arrows represent the ECD cohorts.

**Figure 1**  
**Age of children at each cycle, original cohort versus ECD cohorts**



**Notes:** Ages of children in years are shown in arrows.  
Longer arrows represent the original cohort and shorter arrows represent the Early Childhood Development (ECD) cohorts.

**Source:** Statistics Canada, National Longitudinal Survey of Children and Youth.

## Details of the National Longitudinal Survey of Children and Youth sample, Cycle 8

At Cycle 8, the NLSCY sample consists of:

- a new ECD cohort of 0- to 1-year-old children selected at Cycle 8
- a top-up sample of new 2- to 5-year-olds selected at Cycle 8.

Plus the following returning samples:

- returning 2- to 3-year-old children from the ECD cohort of 0- to 1-year-old children selected at Cycle 7
- returning 4- to 7-year old children from the top-up sample of 2- to 5-year old children selected at Cycle 7
- returning 4- to 7-year-old children from the top-up sample of 0- to 3-year old children selected at Cycle 6
- returning 6- to 7-year-old children from the ECD cohort of 0- to 1-year-old children selected at Cycle 5
- returning 14- to 25-year-old children from the original cohort of 0- to 11-year-olds selected at Cycle 1.

All children were sampled from the Labour Force Survey (LFS). Note that at Cycle 8, there are no children aged 10 to 13 years old.

This section describes which children were surveyed. Some children are sampled but not surveyed because at the previous cycle they were found to be cross-sectionally out-of-scope, e.g., they were deceased, had left the country or had too many cycles of non-response.

For details on how sampling was performed at each cycle, please refer to a cycle's Microdata User's Guide.

### Original cohort, Cycles 1 to 8

The following describes the composition of the original cohort at each cycle (see Figure). The original cohort contains a maximum of two children per household.

#### Cycle 1

The sample of children selected at Cycle 1 was designed to produce reliable—but not equally reliable—provincial estimates for children from age 0 to 11, by two-year age groupings: 0 to 1, 2 to 3, 4 to 5, 6 to 7, 8 to 9 and 10 to 11. A maximum of four children per household was selected. Households were sampled from the following sources:

- the old LFS (prior to 1994)
- the 1994 redesigned LFS
- the National Population Health Survey (NPHS), which is conducted by Statistics Canada.

At the end of Cycle 1, there were 22,831 respondent children in the NLSCY. The child-level response rate was 86.5%.

#### Cycle 2

At Cycle 2, some children were dropped from the sample for budgeting reasons: all NLSCY households belonging to the NPHS sample were dropped, and to reduce the burden on households, the maximum number of children selected per household was cut from four to two. This resulted in a sample of 16,903 children at the beginning of Cycle 2. The child-level response rate for collection was 91.5%. The cumulative, longitudinal response rate for children in the original cohort was 79.1%.

#### Cycle 3

At Cycle 3, 185 children were excluded from the sample because at the end of Cycle 2 they were either cross-sectionally out-of-scope (71) or hard refusals (114). Children who are cross-

sectionally out-of-scope include those who died, whose age was not in-scope, who had permanently left the country, or who had moved to an Indian reserve. Thus, of the 16,903

children sampled for the original cohort, collection was performed on 16,718. The child-level response rate for collection was 89.2%. The cumulative, longitudinal response rate for children in the original cohort was 76.0%.

#### **Cycle 4**

At Cycle 4, to make collection more efficient, it was decided that households with two or more consecutive cycles of non-response would be dropped from collection (along with households with one cycle of non-response followed by the status “Temporarily moved”). Consequently, 1,086 children were dropped from collection at Cycle 4 because at the end of Cycle 3 they were either cross-sectionally out-of-scope (106) or had had two or more cycles of non-response (980). A total of 15,632 children were surveyed. The child-level response rate for collection was 84.5%. The cumulative, longitudinal response rate for children in the original cohort was 67.8%.

#### **Cycle 5**

At Cycle 5, it was decided that 18- and 19-year-olds would be dropped from collection only after three consecutive cycles of non-response (versus two for younger children). The reason for this is that at age 18, the youth becomes the sole respondent, whereas before age 18 the primary respondent is the person most knowledgeable (PMK), who is typically the mother.

At Cycle 5, 469 children were dropped from collection because at the end of Cycle 4 they were either cross-sectionally out-of-scope (32) or had too many consecutive cycles of non-response (437). A total of 15,163 children were surveyed at Cycle 5. The child-level response rate for collection was 81.3%. The cumulative, longitudinal response rate for children in the original cohort was 63.1%.

#### **Cycle 6**

At Cycle 6, 1,506 children were dropped from collection because they had too many consecutive cycles of non-response by the end of Cycle 5. A total of 13,657 children were surveyed at Cycle 6. The child-level response rate for collection was 82.4%. The cumulative, longitudinal response rate for children in the original cohort was 57.6%.

#### **Cycle 7**

At Cycle 7, 613 children were dropped from collection because at the end of Cycle 6 they were either cross-sectionally out-of-scope (11) or had too many consecutive cycles of non-response (602). A total of 13,709 children were surveyed at Cycle 7. The child-level response rate for collection was 80.5%. The cumulative, longitudinal response rate for children in the original cohort was 56.6%.

Note that at Cycle 7 a new rule was applied for returning children who were 18 or older: the PMK’s history of non-response was ignored when deciding if the child should be sent to collection or not.

For more details, see the Microdata User’s Guide for Cycle 7.

#### **Cycle 8**

At Cycle 8, 56 children were dropped from collection because at the end of Cycle 7 they were either cross-sectionally out-of-scope (17) or had too many consecutive cycles of non-response (39). A total of 15,056 children were surveyed at Cycle 8. The child-level response rate for collection was 68.0%. The cumulative, longitudinal response rate for children in the original cohort was 52.7%.

Note that the rule that was applied at Cycle 7 for returning children who were 18 or older, was also applied at Cycle 8. For the 14- to 17-year-old children, the ones that last responded at Cycle 5, 6 or 7 were surveyed at Cycle 8.

For more details, see the Microdata User's Guide for Cycle 8.

### **Early childhood development cohorts present at Cycle 8**

The ECD children present at Cycle 8 were first sampled in Cycles 5, 6, 7 and 8. When the first ECD cohort of 0- to 1-year-olds was selected at Cycle 2, the rule was a maximum of one child per household, except for twins, in which case both were sampled.<sup>1</sup> At Cycle 5, the rule changed to one child per household without exception.<sup>2</sup> There were no twins in the returning ECD cohorts at Cycle 8.

Prior to Cycle 7, for the ECD samples, only respondents from the previous cycle were surveyed at subsequent cycles. At Cycle 7, this rule was modified so that non-respondents from previous cycles were surveyed, so long as they did not have 2 or more consecutive cycles of non-response. At Cycle 8, this latter condition was dropped so that all non-respondents from previous cycles were surveyed, regardless of the number of consecutive cycles of non-response.

### **Cycle 5 ECD cohort**

At Cycle 5, a sample of 0- to 1-year-olds was selected from the LFS. The total sample size was 4,492 children and households. At the end of Cycle 5 collection, there were 3,252 responding children. The response rate was 74.0%.

At the end of Cycle 5, 98 children were cross-sectionally out-of-scope and 1,142 were non-respondents. Consequently, there were 3,252 respondents at Cycle 5, returning as 2- to 3-year-olds, who were surveyed at Cycle 6. The response rate was 88.6%. The cumulative longitudinal response rate was 65.3%.

At the end of Cycle 6, 4 children were cross-sectionally out-of-scope and 33 were non-respondents. Consequently, there were 3,215 who responded to a previous cycle, returning as 4- to 5-year-olds, who were surveyed at Cycle 7. The response rate was 85.9%. The cumulative longitudinal response rate was 62.4%.

Finally, there were 3,214 respondents from a previous cycle, returning as 6- to 7-year-olds, who were surveyed at Cycle 8. The response rate was 80.5%. The cumulative longitudinal response rate was 58.8%.

### **Cycle 6 ECD cohort**

At Cycle 6, a sample of 0- to 5-year-olds was selected from the LFS. The total sample size was 5,795 children and households. At the end of Cycle 6 collection, there were 4,684 responding children. The response rate was 81.3%.

At the end of Cycle 6, 21 children were cross-sectionally out-of-scope and 143 were non-respondents. Consequently, there were 5,631 children aged 2 to 7 years old who were surveyed at Cycle 7. The response rate was 83.0%. The cumulative longitudinal response rate was 79.7%.

Finally, there were 5,039 children aged 4 to 7 years old who were surveyed at Cycle 8. The response rate was 76.9% (see Table 1). The cumulative longitudinal response rate was 75.1%.

### **Cycle 7 ECD cohort**

---

<sup>1</sup>. The ECD cohort sampled in Cycle 2 included 0- to 1-year-olds who were younger siblings of children belonging to the original cohort. This was the only cycle in which siblings from the original cohort were selected. No Cycle 2 ECD children are present in the Cycle 8 sample.

<sup>2</sup>. For the Cycle 7 sample, it was decided that one of the returning twins was dropped for returning ECD cohorts. The original cohort continues to have a maximum of two children per household.



At Cycle 7, a sample of 0- to 5-year-olds was selected from the LFS. The total sample size was 5,843 children and households. At the end of Cycle 7 collection, there were 4,691 responding children. The response rate was 80.8%.

At the end of Cycle 7, 35 children were cross-sectionally out-of-scope and 11 were non-respondents. Consequently, there were 5,797 children aged 2 to 7 years old who were surveyed at Cycle 8. The response rate was 79.1%. The cumulative longitudinal response rate was 78.2%.

**Cycle 8 ECD cohort**

At Cycle 8, a sample of 0- to 5-year-olds was selected from the LFS. The total sample size was 6,685 children and households. At the end of Cycle 8 collection, there were 5,065 responding children. The response rate was 76.0%.



## **Content of the survey**

The National Longitudinal Survey of Children and Youth (NLSCY) covers a broad range of characteristics and factors affecting child growth and development. Extensive information was gathered about the child, the child's parent(s), and the characteristics of the family, the neighbourhood, and the child's school and school experiences. This was achieved by using different survey components directed at different groups of respondents such as:

- Household component collecting the basic demographic information for all household members;
- Adult component collecting information about the person most knowledgeable (PMK) and spouse;
- Child component which is completed for each selected child aged 0 to 7 years and 14 to 17 years;
- Youth component, completed by the 16- to 25-year-olds, who answer questions about themselves;
- Self-complete questionnaires, completed by the 14- to 17-year-olds, who answer questions about themselves in a paper questionnaire;
- Direct assessments which are administered directly to the children and youth.



## **Survey scales**

For some of the concepts deemed important to measure in the National Longitudinal Survey of Children and Youth (NLSCY), it was decided that the concept would most appropriately be measured through the use of a scale. A scale is simply a group of questions or items that measures a certain concept when the answers to the items are put together.

For example, it was determined that it was important to assess three parenting behaviours using a scale on the Child Questionnaire. The scale measures positive interaction, ineffective parenting and consistent parenting.

### **Parent-reported scales**

Depression scale  
Family Functioning scale  
Neighbourhood Safety scale  
Social Support scale  
Behaviour scales  
Motor and Social Development scale  
Parenting scales  
Ages and Stages scale

### **Youth-reported scales**

Depression scale  
Neighbourhood Structure scale  
General Self-image scale  
Emotional Quotient scale  
Social Support scale  
Friends scale  
My Parents and Me scales  
Conflict Resolution scale  
Behaviour scales



## Self-complete questionnaires (ages 14 to 17)

Respondents from 14 to 17 years of age completed a paper questionnaire on various aspects of their lives. They were given the questionnaire during the interview and asked to complete it. To ensure confidentiality, each respondent placed the completed questionnaire in an envelope, sealed the envelope and gave it to the interviewer.

The Self-complete questionnaires consist of a set of two booklets, one for each age group. The table below shows the topics covered by each age-group section in the booklet. The booklets are reproduced in Book 2 – Youth Questionnaires of the *National Longitudinal Survey of Children and Youth, Cycle 8 Survey Instruments 2008/2009* which can be accessed electronically at:

<http://www.statcan.gc.ca/cgi-bin/imdb/p2sv.pl?Function=getInstrumentLink&SurvId=4450&SurvVer=1&Instald=16044&InstaVer=8&lang=en&db=imdb&adm=8&dis=2>

**Table 1**  
Self-complete questionnaires, by age group and booklet

Topic	Self-complete section by age and booklet	
	14 and 15 years Booklet 22	16 and 17 years Booklet 23
Friends and family	A	A
School	B	...
About me	C	B
Feelings and behaviours	D	C
My parent(s)	G	G
Smoking, drinking and drugs	F	D
Puberty	H	...
Activities	E	...
Dating / My relationships	H	F
Health	H	E
Work	I	...
Thank you	J	H

... not applicable

**Source:** Statistics Canada, National Longitudinal Survey of Children and Youth.





## **Main changes to Cycle 8 since Cycle 7**

### **Content changes**

At each cycle, there are changes made to the content of the National Longitudinal Survey of Children and Youth (NLSCY). Any new variable or any variable that changed, e.g., wording, response categories, and eligible population, will have an “h” as the fifth character of the variable name. Variable name conventions are described in Chapter 7.0 and survey content is described in detail in Chapter 8.0 of the Cycle 8 Microdata User's Guide.

The following is a list of the main changes to the content of the survey for Cycle 8:

#### **Self-Complete Booklet**

The Self-Complete Booklet for 12- to 13-year-olds, Booklet 21, has been retired. The youngest respondents for the longitudinal cohort in Cycle 8 are the 14- to 15-year-olds.

#### **Political Engagement (ages 20 – 25)**

In addition to asking youth if they voted in a previous election, youth were asked if they believed they had a duty to vote and how often they talked about political events with others (HPOLhQ1 to HPOLhQ7).

#### **Abilities (ages 16 – 25)**

The Abilities questions were added in Cycle 8 but had previously appeared in Cycle 5 on the self-complete paper questionnaires (HABYhQ1 to HABYhQ6).

#### **Family Formation and Fertility (ages 18 – 25)**

A new section of questions was added to gather information about the factors influencing young adults' transition to marriage and parenthood (HFFYhQ1 to HFFYhQ16).

### **Methodology changes**

At Cycle 7, ECD children ranged in age from 0 to 9, while at Cycle 8, the ECD children range in age from 0 to 7.



## **Data collection**

Data for Cycle 8 of the National Longitudinal Survey of Children and Youth (NLSCY) were collected between the fall of 2008 and the summer of 2009.

### **Child Component**

A Child component was created for each selected child from 0 to 17 years of age, except for youth aged 16 or 17 years old who are living independently. The PMK answered the Child component questions. The PMK was usually the child's mother, but it could also be the father, a step-parent or an adoptive parent who lived in the same dwelling. Only the PMK or his/her spouse was permitted to answer the questions in this component.

For households in which the only child selected was 16 or 17 years old and was living with his/her parents, a shorter version of the Child component was asked. If the child was no longer living with his/her parents, the component was not created.

### **Youth Component**

This component is used for selected respondents aged 16 and above. The youth was the only person permitted to answer the questions in this component, whether he/she was living in the family home or not.

### **Adult Component**

An Adult component was created for the PMK and his/her spouse or partner, if the selected child was 17 years old or younger, except for youth aged 16 or 17 years old who are living independently. Only the PMK or his/her spouse was permitted to answer the questions in this component. Questions in the Adult component are asked once per household, even if more than one child was selected in the household.

For households in which the only child selected was 16 or 17 years old and was living with his/her parents, a shorter version of the Adult component was asked. If the child was no longer living with his/her parents, the component was not created.

### **Sample sizes at Cycle 8**

The number of children and youth sampled in Cycle 8 is shown by age and province in the following tables.

**Table 2**  
**Number of sampled children and response rate, by age at Cycle 8**

Age as of December 31, 2008  (Years)	Sampled	In-scope	Respondents	Cycle 8 response rate  (%)
	Number			
0	2,400	2,389	1,808	75.7
1	3,082	3,074	2,298	74.8
2	2,064	2,060	1,643	79.8
3	3,516	3,495	2,729	78.1
4	2,291	2,278	1,761	77.3
5	3,113	3,094	2,369	76.6
6	1,992	1,984	1,636	82.5
7	2,279	2,272	1,814	79.8
8	1	0	0	n/a
11	1	0	0	n/a
14	1,528	1,527	1,227	80.4
15	1,606	1,602	1,274	79.5
16	1,136	1,134	890	78.5
17	1,102	1,101	880	79.9
18	1,304	1,302	850	65.3
19	1,219	1,213	785	64.7
20	1,178	1,172	689	58.8
21	1,183	1,182	677	57.3
22	1,232	1,221	757	62.0
23	1,186	1,185	713	60.2
24	1,250	1,246	755	60.6
25	1,132	1,122	711	63.4
<b>Total</b>	<b>35,795</b>	<b>35,653</b>	<b>26,266</b>	<b>73.7</b>

Source: Statistics Canada, National Longitudinal Survey of Children and Youth.

**Table 3**  
**Number of sampled children and response rate, by province of residence at  
Cycle 8**

Province	Sampled	In-scope	Respondents	Cycle 8 response rate (%)
	Number			
Newfoundland and Labrador	2,104	2,098	1,704	81.2
Prince Edward Island	1,513	1,511	1,167	77.2
Nova Scotia	2,347	2,339	1,744	74.6
New Brunswick	2,318	2,312	1,778	76.9
Quebec	5,658	5,642	4,344	77.0
Ontario	8,579	8,559	6,059	70.8
Manitoba	2,710	2,697	1,940	71.9
Saskatchewan	2,868	2,857	2,219	77.7
Alberta	4,244	4,227	2,988	70.7
British Columbia	3,375	3,360	2,272	67.6
Outside the 10 provinces	79	51	51	100.0
<b>Total</b>	<b>35,795</b>	<b>35,653</b>	<b>26,266</b>	<b>73.7</b>

Source: Statistics Canada, National Longitudinal Survey of Children and Youth.

### **Collection personnel (training, supervision and control)**

The NLSCY is conducted by Statistics Canada interviewers. A number of them have worked on one or more cycles of the NLSCY.

At each cycle, senior interviewers are responsible for ensuring that NLSCY interviewers are familiar with the survey's concepts and procedures. The senior interviewers ensure that prompt follow-up action is taken in the case of refusal and other kinds of non-response. If necessary, the non-response cases are transferred to a senior interviewer and reassigned. The senior interviewers, in turn, report to the program managers in Statistics Canada's regional offices.

For the NLSCY, a combination of classroom training and self-study materials is used to ensure that interviewers and supervisors have a proper understanding of the survey concepts. In the self-study portion, which precedes the classroom training, the program managers, senior interviewers and interviewers read the interviewer's manual for the survey. The classroom training is initially given by Head Office to the senior interviewers who subsequently train all the interviewers in their respective areas.

### **Interviewing in non-official languages**

The NLSCY questionnaires are only available in English and French. If a respondent wishes to be interviewed in another language, the case will be given to an interviewer who speaks the respondent's language, if possible. In Cycle 8, fewer than eighty cases were not completed because of a language barrier.



## Direct assessments

Research on early childhood and youth development plays a significant role in the formulation of policy for young children and youth. Using various assessment tools in the National Longitudinal Survey of Children and Youth (NLSCY) will help to enhance the knowledge about developmental processes in early childhood and youth and provide relevant data on which to base policy directions for these stages.

Choices about the assessment tools to be included in the NLSCY were made on the basis of an extended literature review, development of a research framework on child development and learning, consultations with many experts in Canada and internationally, review of material on many different possible instruments and field testing of the most likely possibilities. The instruments selected for consideration were also reviewed using a number of criteria. The criteria included reliability and validity of the instrument, coverage of domains in the research framework, ability of the instrument to indicate normal development and developmental delays, the ease of administration by lay interviewers and the availability of the instrument in English or French (or ease of translation to French or English). The final decision was strongly influenced by key experts who had a history of providing advice to the NLSCY Team.

Below is a table summarizing the many direct assessments administered to the selected respondents. For detailed information about the assessments, see Chapter 14.0 of the Cycle 8 Microdata User Guide.

**Table 4**  
**Direct assessments, by age group and method of administration**

<b>Name of assessment</b>	<b>Age group</b>	<b>Method of administration</b>
Peabody Picture Vocabulary Test – Revised (PPVT-R)	4- and 5-year-olds	Computer-assisted interview
Who Am I?	4- and 5-year-olds	Paper questionnaire
Number Knowledge assessment	4- and 5-year-olds	Computer-assisted interview
Mathematics Computation Exercise	7- to 15-year-olds in grades 2 to 10	Paper questionnaire
Problem Solving Exercise	16- and 17-year-olds	Paper questionnaire
Literacy assessment	18- and 19-year-olds	Paper questionnaire
Numeracy assessment	20- and 21-year-olds	Paper questionnaire

**Source:** Statistics Canada, National Longitudinal Survey of Children and Youth.





## **Weighting and data analysis**

The National Longitudinal Survey of Children and Youth (NLSCY) is a probability survey. As is the case with any probability survey, the sample is selected so as to be able to produce estimates for a reference population. Therefore, each unit in the sample represents a number of units in the population.

In a longitudinal survey such as the NLSCY, two types of populations are possible: longitudinal and cross-sectional. The longitudinal population is the initial population when the sample was first drawn and does not change over time; a cross-sectional population may refer to some other time period. Differences between the longitudinal and cross-sectional populations are due to births, deaths, immigration and emigration.

The NLSCY produces three sets of weights at each cycle, two longitudinal (funnel and non-funnel) and one cross-sectional. Funnel weights are assigned to longitudinal children who have responded at every cycle, while non-funnel weights are assigned to longitudinal children who responded at the most recent cycle, but not necessarily at all previous cycles.

Survey weights are calculated by taking the child's design weight and making adjustments for survey non-response and post-stratification to ensure that the final survey weights sum to known counts of children by age, sex and province. The design weight is the inverse of the probability of selection, that is, the probability that a child in the population is selected by the NLSCY sample.

Data analysis involves summarizing the data and interpreting their meaning in a way that provides clear answers to questions that initiated the survey. Sometimes the analyst simply wishes to describe the sample, but more often he or she wants to use the sample to describe some population.

When making inferences about a population that was surveyed, Statistics Canada recommends that the survey weights be used (either cross-sectional or longitudinal, depending on the analysis). Because of the complex sample design, the distribution of a characteristic of interest in the sample is probably different from its distribution in the population. Only by applying the survey weights can the population's distribution be preserved.

Stratification and clustering (both present in the NLSCY sample design) lead to unequal probabilities of selection. For example, the probability that a child in the population is sampled by the NLSCY depends on the age of the child, the child's province of residence, etc. (In the sample there is a disproportionate number of children from small provinces.) Unequal non-response rates within the population can also lead to unequal representation of children in the sample. Finally, clustering in the sample leads to the statistical non-independence of units: children belonging to the same household are not independent.

Suppose that the analyst wants the distribution of children across Canada, i.e., by province, for the original cohort. The population of inference is children aged 0 to 11 as of December 31, 1994, who were living in any province at the time of Cycle 1 collection (1994/1995). Two different sets of longitudinal weights could be used: the 'funnel' weights (for children who have responded to every cycle) or the 'non-funnel' weights (for children who responded at Cycles 1 and 8, but not necessarily all in-between). The table below illustrates the difference between weighted and unweighted estimates of the number and proportion of children in Canada, using the funnel weights.

**Table 5**  
**Distribution of children in the population, original cohort, weighted versus unweighted estimates using Cycle 8 funnel weights<sup>1</sup>**

Province	Unweighted		Weighted	
	Frequency	Percent	Frequency	Percent
Newfoundland and Labrador	491	6.57	88,986	1.91
Prince Edward Island	227	3.04	23,148	0.50
Nova Scotia	536	7.17	144,088	3.09
New Brunswick	450	6.02	115,131	2.47
Quebec	1,457	19.50	1,090,582	23.41
Ontario	1,814	24.28	1,773,616	38.08
Manitoba	569	7.62	182,869	3.93
Saskatchewan	643	8.61	173,611	3.73
Alberta	688	9.21	489,913	10.52
British Columbia	597	7.99	576,125	12.37
<b>Total</b>	<b>7,472</b>	<b>100.00</b>	<b>4,658,069</b>	<b>100.00</b>

<sup>1</sup> HWTCWd1L. See section 15.3 of the Cycle 8 Microdata User Guide.

**Source:** Statistics Canada, National Longitudinal Survey of Children and Youth.

Without the weights, the analyst would incorrectly conclude that 22.8% of children reside in the Atlantic provinces when, in fact, the true number is only 7.97%. The unweighted proportions reflect the fact that the sample has a disproportionate number of children from the smaller provinces (to ensure adequate sample size in small provinces). Without the weights, the analyst would also incorrectly conclude that there are only 7,472 children in the population when, in fact, there are over 4.6 million.

For more details on how the survey weights are calculated and what populations they refer to, see Chapter 11.0 of the Cycle 8 Microdata User Guide. For more details on analytical issues, see Chapter 15.0 of the the same guide.

### Data quality

Data quality is affected by various sources of error. Efforts are made at all steps (interviewer training, collection monitoring, processing, weighting, etc.) to reduce the potential for errors. There are two main types of error: sampling error and non-sampling errors.

#### Sampling error

The estimates derived from this survey are based on a sample of children. If we had done a census of the target population with the same questionnaires, interviewers, supervisors, processing methods and so on, we might have obtained slightly different values. The difference between the estimates produced by a sample and the estimates obtained through complete enumeration under similar conditions is known as the sampling error of the estimates.

Sampling error can be estimated using the sampling variance. For more details on calculating the estimated sampling error, see chapter 13.0 of the Cycle 8 Microdata User Guide.

#### Non-sampling errors

There are many sources of non-sampling errors in any survey. Interviewers may misunderstand survey instructions, respondents may make mistakes in answering the questions, responses may be recorded in the questionnaire incorrectly and errors may be made in processing the data. These examples of non-sampling errors are difficult to quantify. Other kinds of error, especially non-response and the coverage of the intended population, are more easily quantifiable.

Non-sampling errors can cause bias, defined as a difference between the expected survey estimated value and the true population value. As the true population values are not known, it is very difficult to measure bias.

Data users are encouraged to consider how sampling and non-sampling errors may affect the variables they are attempting to analyze.

For more details on data quality, see Chapter 12.0 of the Cycle 8 Microdata User Guide. For more details on sampling error and variance estimation, see Chapter 13.0 of the same guide.



## **How to access the National Longitudinal Survey of Children and Youth data**

### **Public use microdata file (PUMF)**

Although there is no public use microdata file (PUMF) for Cycles 4, 5, 6, 7 or 8 of the National Longitudinal Survey of Children and Youth (NLSCY), there is a PUMF available for Cycles 1 to 3. Each microdata file includes NLSCY public use data and accompanying documentation.

To ensure respondent confidentiality, a longitudinal file is not available to the public. In addition, certain variables are not available on the PUMF. Those wishing access to suppressed or longitudinal data can do so by either remote access or through custom tabulations.

### **Remote data access**

Through remote access, researchers can have access to suppressed data by submitting programs to run on the NLSCY data set at Statistics Canada. A “dummy” research file will be made available to researchers to check the logic and syntax of their programs. Researchers will transmit their programs electronically to Statistics Canada, which will then be moved into the Department’s internal, secure environment. Next, the code would be processed, the results vetted for confidentiality, and sent back to the client. It should be noted that the onus is with the user to submit retrieval programs that are correct and tested. Statistics Canada will review results only for confidentiality concerns and will not make any assessment whatsoever as to whether or not the submitted program has worked properly.

Remote data access is available on a cost recovery basis. For further information on remote data access, please refer to contact information at the beginning of this document.

### **Custom tabulations**

Statistics Canada retains a master microdata file from which specific microdata files and personalized tabulations can be requested on a cost recovery basis; please contact [ssd@statcan.gc.ca](mailto:ssd@statcan.gc.ca).

### **The Research Data Centres Program**

The Research Data Centres (RDC) provide researchers with access, in a secure university setting, to microdata from population and household surveys. The centres are staffed by Statistics Canada employees. They are operated under the provisions of the *Statistics Act* in accordance with all the confidentiality rules and are accessible only to researchers with approved projects who have been sworn in under the *Statistics Act* as 'deemed employees'.

RDCs are located throughout the country, so researchers do not need to travel to Ottawa to access Statistics Canada microdata.

Please see <http://www.statcan.gc.ca/english/rdc-cdr/index.htm> for more information.

**Detailed Cycle 8 measures**

NLSCY Cycle 8 measures – Cognitive development and language outcomes	Respondent	0-11	1	2	3	4	5	6-7	8-9	10-11	12-13	14-15	16-17	18-19	20-21	22-23	24-25	
		mths	year	years														
<b>Language</b> Receptive skills: Peabody Picture Vocabulary Test – Revised	Child					X	X											
Communication skills: <i>Ages and Stages</i> Questionnaire	Parent	X	X	X	X													
Communication skills	Parent				X	X	X											
<b>Literacy</b> Emerging literacy skills: <i>Who Am I?</i> assessment	Child					X	X											
Literacy assessment	Youth														X			
Literacy and learning activities	Parent	X	X	X	X	X	X	X										
	Youth											X	X	X	X	X	X	X
<b>Numeracy</b> Quantitative knowledge: <i>Number Knowledge</i> assessment	Child					X	X											
Math achievement: <i>Math Computation Exercise</i> Numeracy – <i>Cognitive measure</i>	Child							X**				X						
Cognitive measure	Youth												X					
Numeracy assessment	Youth														X			
School performance (math)	Parent					X	X	X				X						

Detailed Cycle 8 measures, continued

NLSCY Cycle 8 measures – Cognitive development and language e outcomes continued	Respondent	0-11	1	2	3	4	5	6-7	8-9	10-11	12-13	14-15	16-17	18-19	20-21	22-23	24-25
		mths	year	years													
Abilities (e.g., counting to 5)	Parent				X	X	X										
<b>Science</b> School performance (science)	Parent											X					
<b>Overall achievement</b> School performance	Parent					X	X	X				X					
	Youth											X	X	X	X	X	X
Educational aspirations	Parent												X				
	Youth											X					
<b>Learning processes</b> Task persistence	Parent				X	X	X										
<b>Analytical reasoning / problem solving</b> Problem solving: <i>Ages and Stages</i> Questionnaire	Parent	X	X	X	X												
<b>General knowledge</b> Computer skills	Parent				X	X	X	X									
	Youth											X	X	X	X	X	X
Labour force participation	Youth											X	X	X	X	X	X
Career aspirations	Youth													X	X		

**Detailed Cycle 8 measures, continued**

NLSCY Cycle 8 measures – Emotional development outcomes	Respondent	0-11	1	2	3	4	5	6-7	8-9	10-11	12-13	14-15	16-17	18-19	20-21	22-23	24-25	
		mths	year	years														
<b>Temperament</b> Good-natured / difficult	Parent	X	X	X														
Adaptable	Parent	X	X	X														
Self-esteem	Youth											X	X	X			X	
Emotional disorder / anxiety	Parent			X	X	X	X	X										
	Youth											X						
Depression	Youth												X	X	X	X	X	X
Emotional intelligence	Youth											X	X		X			X

NLSCY Cycle 8 measures – Social development outcomes	Respondent	0-11	1	2	3	4	5	6-7	8-9	10-11	12-13	14-15	16-17	18-19	20-21	22-23	24-25	
		mths	year	years														
Personal / social skills: <i>Ages and Stages</i> Questionnaire	Parent	X	X	X	X													
Prosocial behaviour	Parent							X										
	Youth											X						
Positive behaviour, including perseverance and independence	Parent				X	X	X											
Interpersonal relationships	Parent					X	X	X										
	Youth											X	X	X	X	X	X	X
Hyperactivity / inattention	Parent			X	X	X	X	X										
	Youth											X						
Physical aggression / opposition	Parent			X	X													



NLSCY Cycle 8 measures – Social development outcomes continued	Respondent	0-11	1	2	3	4	5	6-7	8-9	10-11	12-13	14-15	16-17	18-19	20-21	22-23	24-25	
		mths	year	years														
Physical aggression / conduct disorder	Parent					X	X	X										
	Youth											X						
Separation anxiety	Parent			X	X													
Indirect aggression	Parent					X	X	X										
	Youth											X						
Property offence	Parent																	
	Youth											X	X	X				
Delinquent behaviours	Parent																	
	Youth											X	X	X			X	
Non-sport activity participation	Parent				X	X	X	X										
	Youth											X	X					
<b>Community participation</b> Volunteering	Youth											X	X		X	X	X	
Engaged citizen	Youth											X	X		X	X	X	
Life-long learning	Youth												X	X	X			
Social support	Youth													X	X	X	X	
Family formation	Youth													X	X	X	X	

Detailed Cycle 8 measures, continued

NLSCY Cycle 8 measures – Physical health outcomes	Respondent	0-11	1	2	3	4	5	6-7	8-9	10-11	12-13	14-15	16-17	18-19	20-21	22-23	24-25	
		mths	year	years														
<b>Perinatal period</b> Delivery details	Parent	X	X	X*	X*	X*	X*	X										
Infant's health at birth	Parent	X	X	X	X	X	X	X										
Infant's birth weight	Parent	X	X	X*	X*	X*	X*	X*				X*	X*					
<b>General health</b> Current health	Parent	X	X	X	X	X	X	X				X						
	Youth											X	X	X	X	X	X	X
Height / weight	Parent	X	X	X	X	X	X	X										
	Youth											X	X	X	X	X	X	X
Health Status Index	Parent					X	X											
<b>Physical activity</b> Participation in sports and other physical activities	Parent	X	X	X	X	X	X	X										
	Youth											X	X				X	
<b>Health problems</b> Injuries	Parent	X	X	X	X	X	X	X				X						
	Youth												X		X			
Asthma	Parent	X	X	X	X	X	X	X				X						
	Youth												X	X				X
Chronic conditions	Parent	X	X	X	X	X	X	X				X						
	Youth												X	X		X	X	
Activity limitations	Parent	X	X	X	X	X	X	X				X						
	Youth												X		X		X	
Infections	Parent	X	X	X	X													

**Detailed Cycle 8 measures, concluded**

NLSCY Cycle 8 measures – Physical health outcomes continued	Respondent	0-11	1	2	3	4	5	6-7	8-9	10-11	12-13	14-15	16-17	18-19	20-21	22-23	24-25	
		mths	year	years														
<b>Development</b> Fine and gross motor skills: <i>Ages and Stages Questionnaire</i>	Parent	X	X	X	X													
<b>Milestones</b> <i>Sit up, solid food, feed self, first steps</i>	Parent	X	X	X	X													
Puberty	Youth											X	X					
<b>Exposure to risk</b> Use of alcohol, drugs, tobacco	Youth											X	X	X	X	X	X	X
Sexual health	Youth											X	X	X	X	X		
<b>Healthy lifestyles</b> Eating breakfast	Youth											X	X	X				
Dieting, weight gain / changes	Youth											X	X				X	X
Use of seat belts, bicycle helmets	Youth																	

\* Delivery details and birth weight are collected only in the first interview.

\*\* Math test administered to 7-year-olds who are in Grade 2 or above.

**Note:** The shaded columns indicate that these cohorts were not surveyed in Cycle 8. The youngest age group of the original longitudinal cohort was 14 to 15 years old.

**Source:** Statistics Canada, National Longitudinal Survey of Children and Youth.