Special Surveys Division

National Longitudinal Survey of Children and Youth (NLSCY)

Cycle 3 North

DATA USER'S GUIDE

July, 2002

Statistics Canada Statistique Canada

Canada

Table of Contents

1.0 Introduction

The National Longitudinal Survey of Children and Youth (NLSCY) is a long-term survey designed to measure child development and well being. The first cycle of the survey was conducted by Statistics Canada in 1994-1995 on behalf of Human Resources Development Canada. The second cycle of the survey took place in 1996-1997 and the third cycle was conducted in 1998-1999.

Statistics Canada carried out the Territories Component in conjunction with the statistical bureaus from the Yukon and the Northwest Territories. Similar data were collected in the provinces and the territories, however different collection methods were used. The content differs slightly as well. This manual has been produced to facilitate the manipulation of the Territories Component master file and to document data quality and other analytical issues regarding the NLSCY.

Any questions about the data set or its use should be directed to:

Client Services
Special Surveys Division
Telephone: (613) 951-7355 OR 1-888-297-7355
Fax: (613) 951-3012
Email: ssd@statcan.ca

2.0 Background

Before the NLSCY was undertaken there were few statistical studies that described a broad range of characteristics of children in Canada. Measures of health, well-being and life opportunities are needed, however, if governments and researchers hope to learn more about the ongoing life conditions of Canadian children and youth, and their developmental experiences. Longitudinal data are central to discovering developmental changes occurring in children over time, and studying the impacts of the social environment of the child and various family-related factors.

Data on the prevalence of, and interaction among various characteristics and conditions will assist policy makers in understanding the processes that modify risk and protect and encourage the healthy development of children. Such information will enhance the capacity of the various partners in society to develop effective strategies, policies and programs to help children succeed in our changing society.

3.0
Objectives

The primary objective of the NLSCY is to develop a national database on the characteristics and life experiences of children and youth in Canada as they grow from infancy to adulthood. The more specific objectives of the NLSCY are:

- to determine the prevalence of various biological, social and economic characteristics and risk factors of children and youth in Canada;
- to monitor the impact of such risk factors, life events and protective factors on the development of these children; and
- to provide this information to policy and program officials for use in developing effective policies and strategies to help young people live healthy, active and rewarding lives.

Underlying these objectives is the need to:

- fill an existing information gap regarding the characteristics and experiences of children in Canada, particularly during their early years;
- focus on all aspects of the child in a holistic manner (i.e., the child, his/her family, school, and community);
- provide national, and as far as possible, provincial and territorial data; and
- explore subject areas that are amenable to policy intervention and which affect a significant segment of the population.

4.0 Survey Methodology

The requirement for the NLSCY design was to select a representative sample of children in Canada and to follow and monitor these children over time into adulthood.

4.1 Definition of the NLSCY Population

The target population of the NLSCY for Cycle 1 consisted of Canadian children aged newborn to 11 years of age. In Cycle 2, the sample was topped up for newborns and one year-olds, increasing the target population from newborns to 13 years of age.

In Cycle 3 there was a top-up of children aged up to 15 years, not just newborns and one year-olds as in Cycle 2. This increased the target population from newborns to 15 years of age.

4.2 NLSCY Sample Design

In terms of sampling, the starting point for the NLSCY design in Cycle 1 was the household. Sampled households actually came from three possible sources, which have been labeled as the Main Component, the Integrated Component and the Territories Component.

The Territories Component

The initial sample design of the NLSCY in the North was integrated with the National Population Health Survey (NPHS) in order to help alleviate response burden. The target population of the integrated sample included household residents living in private occupied dwellings located in the two territories, with the exclusion of populations on Indian reserves, Canadian Forces Bases and in institutions. Also, persons living in unorganized areas were excluded from the target population

The split of the Northwest Territories into two new territories, Nunavut and Western, which took place in 1999, also had to be considered.

4.2.1 Sample Selection of Households

A simple random sample of dwellings was selected from a list frame of addresses in each community with the exception of two strata in the Yukon where random digit dialing was used and provided essentially a simple random sample of residential telephone lines.

A selected household was included in the NLSCY if there was at least one child aged 0 to 11 in the household for Cycle 1.

For Cycle 2 of the survey, new households were added if there was a child between the ages of 0 and 1 living in the selected household. The collection of both the longitudinal panel and the top-up was delegated to the statistical agencies in both the Yukon and Northwest Territories.

In Cycle 3 local statistical agencies in each territory were given the responsibility of interviewing the longitudinal cohorts and for sampling the children for the top-up. In Cycle 3 new households were added again with the goal of representing new children aged 0 and 1, but top-ups were also added in the other age cohorts between 2 to 15. These households were selected using a non-probabilistic method such as

quota or snowball sampling. It is not certain what sampling method was used and if and whether the Yukon and the North West Territories used the same method. The absence of this information poses a problem for estimating population totals or averages using the common Horowitz-Thompson estimator since it requires knowledge of inclusion probabilities. By choosing the alternate ratio estimator, the values of the inclusion probabilities are not essential. What is required is auxiliary information known for all population units. The census and demographic update of population estimates are used as this auxiliary information.

4.2.2 Sample Selection of Families and Children

Where a household contained more than one economic family with children aged 0 to 15, only one family was chosen for inclusion in the NLSCY.

Within each selected family, the survey was administered for all children aged 0 to 15 up to a maximum of three children. Where there were four or more eligible children, three were selected randomly for inclusion in the survey.

4.3 Sample Allocation

The NLSCY sample for Cycle 1 was constructed taking two important requirements into consideration. A sufficient sample was required in each of the 10 provinces and 2 territories to allow for the production of reliable estimates for all children 0 to 11 years of age. The sample allocation was derived such that the smaller provinces and the territories had sufficient sample to meet this requirement.

Unlike the main survey in the provinces it was not necessary to have a large enough sample to produce estimates at the territorial level by seven key age groupings or cohorts. The allocation of the sample in the territories was done in such a way that it would only be possible to produce estimates for aggregated age groups.

4.3.1 Yukon Territory

4.3.1.L Longitudinal Sample Selection Cycle 1

For the longitudinal cohort introduced in Cycle 1, the Yukon Territory was divided into five strata as shown in Table 1. These same strata were used in subsequent cycles for post-stratified ratio weight adjustments for estimation purposes. The first two strata used Random Digit Dialing to select the sample for the interview. Random samples were selected from lists of dwellings in the other strata. The sample of 1,500 households was allocated proportional to the population size in each stratum. The expected number of households is given in Table 1. Only some of the selected households would contain children in the required age range and be usable in the NLSCY.

Table 1 - Allocated Sample Sizes by Stratum for the Yukon, for cycle 1

NPHS/NLSCY Stratum	Household Sample Size
Whitehorse	1,084
Medium Sized Communities	177
Aboriginal Communities	69
Mixed Communities	82
Non-Aboriginal Communities	88
TOTAL	1,500

4.3.1.C Sample Selection Cycles 2 and 3

The longitudinal sample for the Yukon was reduced for Cycle 2 by eliminating a number of households that were shared with NPHS. Of the remaining sample of households the number of children selected was also reduced to a maximum of 2 per household. In total, 170 households were dropped from the sample and a further 135 households had the number of children interviewed reduced.

A new sample of children aged 0 and 1 were introduced for the longitudinal cohort in Cycles 2 and 3. A cross-sectional top-up of children aged 2 to 15 was also introduced in Cycle 3. The Yukon Territory was divided into five strata as shown in Table 1 and a target number of children by age were prescribed for the local agency to interview. In most cases Random Digit Dialing was likely used to select the sample for the interview. Since this is a non-probabilistic method of sample selection, information about the number of contacts, refusals and other non-response in ratio to the number of interviews was not provided to Statistics Canada by the Territorial agency. There were 64 new households added to the response file in Cycle 2. In Cycle 3, 257 new households were added and new members from 41 existing longitudinal households were also added.

4.3.2 Northwest Territories and Nunavut

4.3.2.L Longitudinal Sample Selection Cycle 1

At the time of sample selection in Cycle 1 both territories were part of the Northwest Territories, both of the new territories needed to be represented well in the sample. They were treated as strata in the sample design and the sample of 1,500 households was allocated to these two strata to meet the requirements of both surveys to release estimates for both new territories. This was done using a well-known allocation method that balances the reliability requirements at the territory and stratum levels (Kish, 1988). Using this method, the sample was allocated proportional to (Wh2 + 1/22), where Wh is the 1991 Census proportion of households in stratum, h, h=1,2. This allocation produced sample sizes of 652 households for Nunavut and 848 for the Western Territory.

Within Nunavut and Western, the sample was allocated to each in-scope community proportional to its population size. Any community with a resulting sample size less than 10 households had its sample size increased to 10. A sample of 10 households was deemed the minimum to justify the cost of surveying in a community. Table 2 shows the expected sample sizes aggregated by six regions.

Table 2 - Allocated Sample Sizes by Region for the NWT, for cycle 1

REGION		Household Sample Size
Nunavut	Baffin	322
	Keewatin	204
	Kitikmeot	126
	Total	652
Western	Inuvik	205
	Fort Smith	282
	Yellowknife	361
	Total	848
TOTAL		**1,500**

4.3.2.C Sample Selection Cycles 2 and 3

Children aged 0 and 1 were introduced for the longitudinal cohort in Cycles 2 and 3. A cross-sectional top-up of children aged 2 to 15 was also introduced in Cycle 3. The Northwest Territory and Nunavut Territory were divided into three strata each as shown in Table 2 and a target number of children by age were prescribed for the local agencies to interview. In most cases Random Digit Dialing or other non-probabilistic method was likely used to select the sample for the interview. Since this is a non-probabilistic method of sample selection, information about the number of contacts, refusals and other non-response in ratio to the number of interviews was not provided to Statistics Canada by the Territorial agencies. There were 17 new households added to the response file in Cycle 2, and new members were also added from 130 existing longitudinal households. In Cycle 3, 208 new households were added and new members from 4 existing longitudinal households were added as well.

4.4 Sample Size

For Cycle 1, in both the Yukon and Northwest Territories, the initial NLSCY sample requirement was 1,500 households in order to obtain a sufficient sample of children 0 to 11 years old. As the NPHS required only 1,200 households to achieve its required sample of persons 12 years old and over, the NLSCY requirement determined the overall sample size. For subsequent cycles – 2 and 3 - the sample requirements were not defined by a prescribed number of households, but instead by a required number of respondents.

Actual Sample Sizes

Table 3 shows the responding sample sizes obtained by territory, for cycles 1, 2 and 3. Table 4 provides the sample distribution by age, for cycles 1, 2 and 3.

Table 3 - Responding Longitudinal Sample Sizes by Territory, cycles 1, 2 and 3.

	Cycle 1	Cycle 1	Cycle 2	Cycle 2	Cycle 3	Cycle 3
Territory / Region	Number of Households	Number of Children	Number of Households	Number of Children	Number of Households	Number of Children
Yukon						
Whitehorse	481	765	274	430	260	404
Medium Sized Communities	82	126	42	69	38	55
Native Communities	19	36	3	6	3	5
Mixed Communities	38	65	15	25	16	26
Non-Native Communities	37	67	13	24	11	21
Total	**657**	**1,059**	**347**	**554**	**328**	**511**
Nunavut						
Baffin	181	358	154	303	139	263
Keewatin	127	244	116	227	106	185
Kitikmeot	91	170	82	152	72	124
Total	**399**	**772**	**352**	**682**	**317**	**572**
Western						
Inuvik	74	139	59	105	48	83
Fort Smith	123	224	100	175	84	146
Yellowknife	131	210	98	158	83	132
Total	**331**	**573**	**257**	**438**	**215**	**361**
Northwest Territories						
Total	**724**	**1,345**	**609**	**1,120**	**532**	**933**
Total	**1,384**	**2,404**	**956**	**1,674**	**860**	**1444**

Table 4 Responding Sample Sizes by Age, cycle 1.

Age	Yukon Territory	Western	Nunavut
0	72	43	90
1	79	42	61
2	78	53	73
3	81	64	61
4	90	54	67
5	93	46	70
6	95	54	65
7	84	45	51
8	82	48	66
9	96	43	54
10	111	41	51
11	98	53	50
Total	1,059	586	759

Table 4 Responding Sample Sizes by Age, cycle 2.

For cycle 2, 214 cross-sectional children between 0 and 1 year old were introduced into the sample for the Yukon and the Northwest Territories in order to be able to produce cross-sectional estimates of children aged 0 to 13 years. This includes 61 children in the Yukon and 152 children in the North West Territories (77 in Western, 75 in Nunavut).

Age	Yukon Territory	Western	Nunavut
0	18	38	33
1	43	38	44
2	40	43	56
3	39	40	58
4	31	54	50
5	50	50	48
6	48	58	49
7	56	50	57
8	55	54	54
9	45	38	40
10	53	57	44
11	53	51	27
12	41	39	24
13	46	34	40
Total	618	644	624

Table 5 Responding Sample Sizes by Age, cycle 3.

For cycle 3, 134 cross-sectional 0 and 1 year old children have been introduced in the sample for Yukon and 69 for the North West Territories (32 in Western, 37 in Nunavut) to be able to produce cross-sectional estimates of children aged 0 to 15 years. There were also 250 cross-sectional children ages 2-15 introduced into the sample from the Yukon and 288 from the North West Territories (263 in Western, 25 in Nunavut) that helped to offset the declining sample for the older age cohorts due to attrition.

Age	Yukon Territory	Western	Nunavut
	Total	Total	Total
0	79	13	14
1	55	19	23
2	54	40	46
3	76	44	53
4	48	43	55
5	57	42	64
6	53	67	49
7	58	63	53
8	64	65	51
9	54	48	61
10	64	48	46
11	59	44	51
12	54	53	41
13	57	36	27
14	55	47	26
15	56	42	38
Total	943	714	698

5.0 Data Collection

Data collection for Cycle 1 of the NLSCY took place between 1994 and 1995. Cycle 2 data were collected between 1996 and 1997, and Cycle 3 collection took place between the fall of 1998 and the spring of 1999.

In cycle 1, every selected household with children newborn to 11 years of age had information collected on up to three of those children in a family. In cycle 2, these respondents were re-contacted and additions of children aged 0 to 2 years were selected in order that each age group (from 0 to 13 years) would be represented. All these respondents (now aged 2 to 15 in cycle 3) were re-contacted in cycle 3. At the same time, additional children from all age groups (aged 0 to 15) were added to the sample in this cycle.

5.1 The Household Collection

In all three cycles, for the household collection, data were collected for a variety of respondents using paper and pencil questionnaires. Collection in the Yukon and the Northwest Territories used a shortened version of the NLSCY and NPHS survey instruments. Below is a description of each type of questionnaire used in the household.

The Household Roster

The household questionnaire for the territories lists each member of the household and asks for basic demographic information for each household member as well as some questions on dwelling conditions and household income. A PMK was assigned in a similar manner as the Main and Integrated Components. A question was asked so as to determine which household member was the Person Most Knowledgeable about the child. This person was labeled as the PMK for the household. In most cases the PMK was the mother of the child.

In the Main and Integrated Components, it was possible to construct a relationship grid showing the relationship of everyone in the household to everyone else in the household. Since the Territories Component did not use computer assisted interviewing techniques, only the relationship of household members to the PMK was recorded. The PMK was then asked to complete a set of two additional questionnaires: the Parent questionnaire and the Child questionnaire.

The Parent Questionnaire

The purpose of the Parent Questionnaire was to gather general health information about the PMK and his/her spouse/partner and to obtain some general information about the child's social environment. Data were also collected for the PMK and spouse/partner. Topic areas included education, literacy, mental health of the PMK, labour force and income.

The Child Questionnaire

The Child Questionnaire was completed for selected children in the household, aged newborn to 15 years (up to a maximum of three). These maximums were used to reduce respondent burden. Topic areas on the Child Questionnaire included: health, behaviour, education, literacy, parenting, and custody history.

Cognitive Measures and Self-Complete Questionnaires

The collection in the territories does not include the Peabody Picture Vocabulary Test, the self-complete questionnaire for 10-11,12-13 and 14-15 year olds, the Teachers' and Principal's Questionnaires, the reading and mathematical aptitude indicator or the Reading Comprehension and Mathematics Computation Exercise.

Interviewing

Collection operations took place from November 1998 to March 1999. Interviews were conducted by interviewers hired and trained by the Yukon and NWT Bureaus of Statistics. Households were contacted either in person or by phone, depending on their location. Typically, households in more urban areas were contacted by phone while rural and remote communities required visits. Due to the total length of the NPHS/NLSCY interview, some questionnaires were completed at a later time because the selected respondent was not available at the time of the initial visit or because the long interview time prevented the completion of the interview in one contact.

For all responding households, a household roster (name, age, sex, marital status, relationship to selected person) was completed by a knowledgeable household member - usually the person at home at the time of the interviewer visit. Questions were also asked about the dwelling itself. A household member at least 12 years old was then randomly selected to complete the health questionnaire for NPHS.

Supervision and Control

All interviewers were under the supervision of a staff of senior interviewers. The seniors were responsible for ensuring that interviewers were familiar with the concepts and procedures of the NLSCY and the NPHS. They periodically monitored their interviewers and reviewed completed questionnaires.

5.2 Non-Response

Interviewers were instructed to make all reasonable attempts to obtain NLSCY and NPHS interviews with members of eligible households. Individuals who at first refused to participate in the NPHS were re-contacted by a senior interviewer who stressed the importance of the survey and the household's co-operation. This was followed by a second call (or visit) from the interviewer. For cases where the timing of the interviewer's call (or visit) was inconvenient, a more convenient time was arranged to call back. For cases where there was no one at home, numerous attempts were made to contact the respondent. Despite the best efforts of interviewers, a small number of non-responding households typically remain at the end of a survey collection period.

5.3 Non-response follow-up

Many strategies were put in place to reduce the number of non-response cases. Before interviews started, a maximum recommended assignment size by interviewer was calculated based on test results to allow efficient follow-up of non contact cases (i.e., to avoid over burdening interviewers).

Interviewer procedures included methods of reducing the number of non-contacts by making telephone calls or visits at various times of the day, talking to neighbours or landlords to determine who lived in the dwelling and to obtain telephone numbers, etc.

Refusals were followed-up by senior interviewers or the project supervisor who tried to convince respondents to participate in the survey.

In addition, interviewers were empowered to hire, where necessary, translators in remote northern communities to try to reduce the number of non-interviews due to language problems.

6.0 Data Processing

The main output of the NLSCY is a clean data file. This section presents a brief summary of some of the processing steps involved in producing this file.

6.1 Data Capture

As discussed in Section 5.1, all of the information was collected in face-to-face or telephone interviews and recorded on paper questionnaires. The territorial statistical bureaus were responsible for capturing the data for these questionnaires. Some preliminary edits were performed at this stage, including range checks, verification of the flow patterns (relating to the age of the child or the response given to particular questions) as well as some consistency edits. The data were then forwarded to Statistics Canada's Head Office.

6.2 Editing

6.2.1 Pre-edit

The purpose of the Pre-edit was to carry out some basic formatting and preliminary editing. The following are some of the procedures that were carried out:

Data base files were created for each questionnaire. For the child questionnaire, two files were required to accommodate the number of questions.

In some sections, different options were offered for different age groups. For example, in Cycle 3, in the long-term conditions section, all children were asked about 8 different health problems but children aged 6 years and older were also asked about two other health problems. Initially, the questions for 0 - 5 year olds were stored as separate variables, apart from the questions for 6 - 11 year olds. As part of the pre-edit, the variables were collapsed into one group of variables (so that all questions pertaining to a particular condition are stored as the same variable, regardless of the age of the child.)

The formats of several variables were altered to facilitate editing. For example, the date formats were changed from DDMMYY to DDMMYYYY to create a four-digit year.

In some cases a value was captured that was determined to be out of range for a particular item. These invalid entries were set to missing values.

6.2.2 Consistency Editing

After the pre-edit, consistency editing was carried out. The goal of consistency editing is to verify the relationship between two or more variables. For example, in the Socio-Demographic Section, for children who were not born in Canada, a question was asked to obtain the year the child first immigrated to Canada (CSDCQ3). A consistency edit compared this question to the year of birth of the child. If the year of immigration was before year of birth then it was set to not-stated in the edit. Consistency editing was done for the various sections of the questionnaire and any data quality concerns that were noted as a result of this editing are discussed in detail in Section 9 of this document.

6.3 Naming Conventions and Coding Structure for NLSCY Variables

The NLSCY master file documentation system has employed certain standards to label variable names and values. The intent is to make interpretation of the data more straightforward for the user. These standards are described in this section.

6.3.1 Naming Conventions for Variables

In the NLSCY file a naming convention has been used for each variable in order to give users specific information about the variable. All variable names are at most eight characters long so that these names can easily be used with analytical software packages such as SAS or SPSS.

The variable names are of the following format:

C SE C Q xxx

where:

C: refers to the NLSCY cycle. "A" means the first cycle, "B" the second, "C" the third etc. Obviously for this release all variable names will start with an "C".

SE: refers to the section of the questionnaire where the question was asked or the section from which the variable was derived. The table in Section 6.3.2 gives the acronyms which are used for the sections included in this release of Cycle 3 data. More information about the content for each of these sections can be found in Section 9.

C: refers to the collection unit or the unit to which the variable refers. There are four possibilities:

- **C** means the variable refers to the child
- **P** means the variable refers to the PMK
- **S** means the variable refers to the spouse/partner
- **H** means the variable refers to the household

b: refers to the NLSCY cycle in which the variable first appeared on the file. "b" indicates the variable was new in cycle 2. In subsequent cycles, new variables will also be identified using the lowercase letter representing the cycle. For example, new variables in cycle 3 will contain a "c", in cycle 4 a "d" , etc . Some revisions were made to the content of the questionnaire between cycles. If the revision resulted in a change to the meaning or the values of a question, the variable was treated as new and contains a "b".

Q: refers to the variable type. There are four possibilities:

Q means the variable refers to a question that was asked directly on one of the NLSCY questionnaires
S means that the variable refers to a score calculated for one of the scales used on the questionnaire (See Section 9.1)
D means the variable was derived from other questions that were asked on the questionnaire (See Section 6.5)

xxx: refers to the question or variable identification. Generally, xxx is a series of alpha-numeric characters that reflect, as closely as possible, the question identification on the actual questionnaire.

6.3.2 Acronym Names for Questionnaire Sections

The following table gives the acronym names that were used for each section of the various NLSCY questionnaires (in alphabetical order). As explained in Section 6.3.1 this acronym is embedded in the variable name for all variables on the NLSCY master file. The acronym is the second and third characters of the variable name.

ACRONYM	SECTION
AC	Activities variables: -Asked for children 0 to 15 on the Child's Questionnaire.
BE	Behaviour variables: - asked for children 0 to 15 years, on the Child's Questionnaire.
CH	Adult Chronic Conditions Variables: -Asked of the PMK and Spouse on the Parents Questionnaire.
CS	Custody variables: - asked for children 0 to 15 on the Child's Questionnaire
DM	Demographic variables derived to explain the living arrangements of the child: - derived from information of the household roster and relationship grid.
DP	Depression scale variables: - this scale was administered to the PMK, on the Parent Questionnaire.
ED	Education variables. - asked for children 4 to 15 years old on the Child's Questionnaire and about the PMK and spouse/partner on the Parent Questionnaire.
GE	Geographic Variables: - derived from sample information.
HH	Household variables: -These questions relate to the dwelling characteristics
HL	Health variables: - asked for children 0-15 years old, on the Child's Questionnaire and about the PMK and spouse/partner on the Parent Questionnaire.
IN	Income variables: - household income and personal income of the PMK, collected on the Parent and Household Questionnaires.

LF	Labour force variables: - collected for both the PMK and spouse/partner on the Parent Questionnaire.
LT	Literacy variables: - asked for children 0 to 9 years old, on the Child's Questionnaire.
MD	Medical/biological variables: - asked for children 0 to 3 years of age on the Child's Questionnaire.
MM	Variables collected as part of the household roster. Basic demographic variables were collected for each household member. These variables are included on the NLSCY master file for the child, the PMK and the spouse/partner.
MS	Motor and social development variables: - asked for children 0 to 3 years old, on the Child's Questionnaire.
PR	Parenting style variables: - asked for children 0 to 11, on the Child's Questionnaire.
RL	Social relationship variables: - asked for children 4 to 9 years old, on the Child's Questionnaire.
RS	Restriction of Activities Variables: -Asked of the PMK and Spouse on the Parents Questionnaire.
SD	Socio-demographic variables: - collected for the child on the Child's Questionnaire and for the PMK and spouse/partner on the Parent Questionnaire.
WB	Work after birth variables: Asked for children 0 to 15 years old on the Child's Questionnaire.

6.3.3 Examples of Variable Names

In order to illustrate the naming convention used for variables included on the Cycle 3 NLSCY master file the following examples are given.

CLFSQ2 This refers to Q2 in the Labour Force Section for the spouse/partner.
The "C" indicates it is a Cycle 3 variable.
The "LF" indicates the Labour Force Section.
The "S" indicates it refers to the spouse/partner.
The "Q" indicates it was an item asked directly on the questionnaire.
The "2" is the ID of the item.

CPRCS03 This is a positive interaction score on the parenting scale for a 2 to 11 year-old child.
The "C" indicates it is a Cycle 3 variable.
The "PR" indicates the Parenting Section.
The "C" indicates it refers to the child.
The "S" indicates the variable refers to a score.
The "03" is the ID of the variable.

6.3.4 Coding Structure for NLSCY Variables

Some standards have been developed for the coding structure of NLSCY variables in order to explain certain situations in a consistent fashion across all variables. The following describes these various situations and the codes used to describe the situations.

Refusal: During an interview, the respondent may choose to refuse to provide an answer for a particular item.

On the NLSCY file an item which was refused is indicated by a code "8". For a variable that is one digit long the code will be "8", for a 2 digit variable "98", for a three digit variable "998", etc.

Don't Know: In other cases the respondent may not know the answer to a particular item.

On the NLSCY file, the code used to indicate that the respondent did not know the answer to an item is "7". For a variable that is one digit long the code will be "7", for a two-digit variable "97", for a three-digit variable "997", etc.

Not Applicable:

In some cases a question was not applicable to the survey respondent. A code "6", "96" "996" ... has been used on the file to indicate that a question or derived variable is not applicable.

1 In some cases a single question or series of questions was not applicable. For example, the number of days the child has been away from school (CEDCcQ13) is only applicable for children who are actually in school (CEDCD01). Otherwise there will be a code 996 for this question.

2 In other cases an entire section of the questionnaire was not applicable or even an entire questionnaire. For example, the Motor and Social Development Section was applicable only to children 0 to 3 years old. For all children outside of this age group (i.e., 4 years and older) the motor and social development variables on the master file have been set to not-applicable ("6", "96", "996" etc.). For cases where the PMK did not have a spouse or common-law partner residing in the household, all "spouse" variables (e.g., the Labour Force Section and the Education Section for the spouse) have been set to not applicable.

Not-Stated: In some cases, as part of processing the answer to an item has been set to not-stated. The not-stated code indicates that the answer to the question is unknown. Not-stated codes were assigned for three main reasons:

1 As part of the interview, the interviewer was permitted to record a refusal or don't know response. When this happened the interviewer often skipped out of this particular section of the questionnaire. In the case of refusal, it was assumed that the line of questioning was sensitive and it was likely that the respondent would not answer any more questions on this particular topic area. In the case of a don't know it was assumed that the respondent was not well enough informed to answer further questions. As part of the NLSCY processing, it was decided that all of these subsequent questions should be assigned a not-stated code. A not-stated code means that the question was not asked to the respondent. In some cases it is not even known if the question was applicable to the respondent.

2 In some cases a specific questionnaire was not started or it was started but ended prematurely. For example, there may have been some kind of an interruption, or the respondent decided that she/he wished to terminate the interview. If there was enough information collected to establish this household as a responding household, then all remaining items on the questionnaire (and on

questionnaires that had not yet been started) were set to not-stated. The one exception was that if it was known that a certain section or a certain questionnaire was not applicable, then these questions were set to not applicable.

3 The third situation in which not-stated codes were used was as a result of consistency edits. When the relationship between groups of variables was checked for consistency, if there was an error, often one or more of the variables was set to not-stated. (See Section 6.2.2 for more information about consistency editing.)

For derived variables, if one or more of the input variables to the derived variable had a refusal, don't know or not-stated code, then the derived variable was set to not-stated.

6.4 Coding of Open-ended Questions

A few data items on the NLSCY questionnaire were recorded by interviewers in an open-ended format. For example, in the Labour Force Section, a PMK who had worked in the previous 12 months was asked to identify a main job. There was then a series of open-ended questions asked about this main job:

Thinking about this main job, what kind of business, service or industry is this?

Again, thinking about this main job, what kind of work are you doing?

In this work, what are your most important duties or activities?

The interviewer recorded in words the answer provided by the PMK. At Head Office, these written descriptions were coded into industry and occupation codes to describe the nature of the work of the PMK. Similar information was collected for the spouse/partner and codes assigned to describe the nature of his/her work.

The coding systems used were the 1980 Standard Occupational Classification codes (SOC) and the 1980 Standard Industrial Classification codes (SIC). Grouped versions of these codes are available on the master file (CLFPD07 and CLFPD08 for the PMK, and CLFSD07 and CLFSD08 for the spouse/partner).

6.5 Creation of Derived Variables

A number of data items on the file have been derived by combining items on the questionnaire in order to facilitate data analysis. For example, in the section on socio-demographic characteristics, the PMK was asked a series of questions about immigration including the date of immigration. Using this information and the date of birth, a variable was formed to indicate the age at the time of immigration, and another to indicate the number of years since immigration.

All derived variables on the NLSCY master file have a "D" as the fifth character of the variable name.

6.6 Territories and Provinces Release Name Concordance Table

Release Name in the Territories	Release Name in the Provinces	Comments
CGEHD01		
CGEHD02		
CGEHD03	CGEHD03	
	CGEHD04	
	CGEHD06	
	CGEHD07	
CMMPQ4/ CMMSQ4/ CMMCQ4	CMMPQ01/ CMMSQ01/ CMMCQ01	
-----	CMMCbQ1A	
-----	CMMCbD01	
CMMPQ5/ CMMSQ5/ CMMCQ5	CMMPQ02/ CMMSQ02/ CMMCQ02	
CMMPQ4A/ CMMSQ4A/ CMMCQ4A	CMMPQ03A/ CMMSQ03A/ CMMCQ03A	
CMMPQ4B/ CMMSQ4B/ CMMCQ4B	CMMPQ03B/ CMMSQ03B/ CMMCQ03B	
CMMPQ4C/ CMMSQ4C/ CMMCQ4C	CMMPQ03C/ CMMSQ03C/ CMMCQ03C	
CMMPQ6/ CMMSQ6/ CMMCQ6	CMMPQ04/ CMMSQ04/ ---	Territories ask for “current marital status” whereas Provinces ask for “marital status”
CDMCD01	CDMCD01	
CDMHD02	CDMHD02	
CDMCD03	CDMCD03	Since the Territories component provides only the relationship of household members to the PMK, it was not possible to determine “parent status” for children with a PMK other than a parent or a sibling. In these cases, if there were no other household members over the age of 15, the parent status was assumed to be “does not live with a parent”
CDMCD04	CDMCD04	
CDMCD05	CDMCD05	

CDMCD06	CDMCD06	
CDMPD06A	CDMPD06A	
CDMCD06B	CMDCD06B	For value of '50' - Provinces specify 'PMK does not have a spouse' Territories specify 'PMK does not have a spouse *living in the household*''
-----	CDMCD06C	
CDMPD06D/ CDMSD06D	CDMPD06D/ CDMSD06E	
CDMHD06F	CDMHD06F	
CDMHD07	CDMHD07	
CDMCD08	CDMCD08	Since the Territories Component provides only the relationship of household members to the PMK, it was not possible to identify siblings for children with a PMK other than a parent.
CDMCD09	CDMCD09	
CDMCD10	CDMCD10	
CDMCD11	CDMCD11	
CDMCD12	CDMCD12	
CDMCD13	CDMCD13	Since the Territories Component provides only the relationship of household members to the PMK, it was not possible to determine 'parent status' for children with a PMK other than a parent.
CDMCD14	CDMCD14	
CDMCD15	CDMCD15	
CDMCD16	CDMCD16	
-----	CDMCD17	Since the Territories Component provides only the relationship of household members to the PMK, it was not possible to identify biological parents for children with a PMK other than a biological parent.
CDMCD18	CDMCD18	
CDMCD18B	CDMCD18B	
CDMCD19	CDMCD19	
CDMCD19B	CDMCD19B	
CDMCD20	CDMCD20	
-----	CDMCbD21	
-----	CDMCbD22	
-----	CDMCbD23	
-----	CDMCbD24	

-----	CDMCbD25	
-----	CDMCbD26	
-----	CDMCbD27	
-----	CDMCbD28	
CDMHPC	CDMHPC	Territories provides only the first 5 digits of the postal code; Provinces provides the full postal code.
CEDPQ1/ CEDSQ1	CEDPQ01/ CEDSQ01	Territories include an additional option: '11 - G.E.D.'
CEDPQ2/ CEDSQ2	CEDPQ02/ CEDSQ02	
-----	CEDPcQ2A/CEDScQ2A	
CEDPQ3/ CEDSQ3	CEDPQ03/ CEDSQ03	
-----	CEDPbQ3A/CEDSc03	
CEDPQ4/ CEDSQ4	CEDPQ04/ CEDSQ04	
CEDPQ5/ CEDSQ5	CEDPQ05/ CEDSQ05	
CEDPQ6/ CEDSQ6	CEDPQ06/ CEDSQ06	
CEDPD01/ CEDSD01	CEDPD01/CEDSD01	The Territories Component groups G.E.D. with Secondary School graduation.
CEDPD02/ CEDSD02	CEDPD02/ CEDSD02	
CEDPD04/ CEDSD04	CEDPD04/ CEDSD04	
CEDPcQ7A/ CEDScQ7A	CEDPcQ7A/ CEDScQ7A	
CEDPcQ7B/ CEDScQ7B	CEDPcQ7B/ CEDScQ7B	
CEDPcQ7C/ CEDScQ7C	CEDPcQ7C/ CEDScQ7C	
CEDPcQ7D/ CEDScQ7D	CEDPcQ7D/ CEDScQ7D	
CEDHcQ8A	CEDHcQ8A	
CEDHcQ8B	CEDHcQ8B	
CEDHcQ8C	CEDHcQ8C	
CEDHcQ8D	CEDHcQ8D	
CEDHcQ8E	CEDHcQ8E	
CLFPQ12A/ CLFSQ12A	CLFPbQ04/CLFSbQ04	
CLFPQ12B/ CLFSQ12B	CLFPbQ5A-CLFPbQ5H /CLFSbQ5A-CLFSbQ5H	The Provinces breaks into 8 yes/no questions (mark all that apply). The Territories is mark one only.
CLFPQ12C/ CLFSQ12C	CLFPbQ06/CLFSbQ06	
CBLFPD05/ CLFSD05	------	
CLFPD05B/ CLFSD05B	CLFPcD5A/ CLFSD5A	

CLFPD06/CLFSD06	**-----**	
CLFPD06B/ CLFSD06B	CLFPD06A/ CLFSD06A	
CLFPD07/ CLFSD07	CLFPD07/ CLFSD07	
	CLFPcD7A/CLFScD7A	
CLFPD08/ CLFSD08	CLFPD08/ CLFSD08	
	CLFPcD8A/CLFScD8A	
CLFPD09/ CLFSD09	CLFPD09/ CLFSD09	
-----	CLFPD12/ CLFSD12	
-----	CLFPD25/ CLFSD25	
-----	CLFPD34/CLFSD34	
-----	CLFPbD38/CLFSbD38	
-----	CLFPD51/CLFSD51	
-----	CLFHD49B	
------	CLFHD50	
CLFPb17B/ CLFSb17B	CLFPb17A/ CLFSb17A	
-----	CLFPb17B/CLFSb17B	
CINHD01A	CINHD01A	
CINHD01B	CINHD01B	
CINPD02/CINSD02	CINPD02/CINSD02	
-----	CINHD03A	
-----	CINHD04A	
-----	CINHD05A	
CINHD07	CINHD07	
-----	CINHD08B	In the Territories Component, household income was only reported as an income range. Therefore, it was not possible to create an SES variable as in the Provinces. Instead, the midpoint of the income range was used as a substitute for a continuous income value and an “adjusted SES” variable was created for the Territories. The adjusted SES variable for the Territories should not be considered equivalent to the SES variable for the Provinces.
-----	CINHbD8L	
CINHD09	-----	

CDPPQ12A-L	CDPPQ12A-L	
CDPPS01	CDPPS01	
-----	CFNHQ01A-M	
-----	CFNHQ02	
-----	CFNHS01	
CSDPD01/ CSDSD01/ CSDCD01	CSDPD01/ CSDSD01/ CSDCD01	
CSDPD02/ CSDSD02/ CSDCD02	CSDPD02/ CSDSD02/ CSDCD02	
CSDPD02B/ CSDSD02B/ CSDCD02B	CSDPD02B/ CSDSD02B/ CSDCD02B	
CSDPD03/ CSDSD03/ CSDCD03	CSDPD03/ CSDSD03/ CSDCD03	
CSDPD04/ CSDSD04/ CSDSD04	CSDPD04/ CSDSD04/ CSDSD04	
CSDPD05/ CSDSD05/ CSDCD05	CSDPD05/ CSDSD05/ CSDCD05	
CSDPD05B/ CSDSD05B/ CSDCD05B	CSDPD05B/ CSDSD05B/ CSDCD05B	
CSDPD06/ CSDSD06/ CSDCD06	CSDPD06/ CSDSD06/ CSDCD06	
CSDPD06B/ CSDSD06B/ CSDCD06B	CSDPD06B/ CSDSD06B/ CSDCD06B	
CLTCcQ1	CLTCcQ1	
CLTCcQ2	CLTCcQ2	
CLTCcQ3	CLTCcQ3	
CLTCcQ4	CLTCcQ4	
CLTCcQ5	CLTCcQ5	
CLTCcQ6A	CLTCcQ6A	
CLTCcQ6B	CLTCcQ6B	
CLTCcQ7	CLTCcQ7	
CLTCcQ8	CLTCcQ8	
CLTCcQ12	CLTCcQ12	

CLTCcQ13	CLTCcQ13	
CLTCcQ14	CLTCcQ14	
CMDCQ1A	CMDCQ01A	
CMDCQ1B	CMDCQ01B	
CMDCQ1C	CMDCQ01C	
-----	CMDCbQ2A-MDCbQ2E	
CMDCQ3	CMDCQ03	
CMDCQ4	CMDCQ04	
CMDCQ5A	CMDCQ05A	
CMDCQ5B	CMDCQ05B	
CMDCQ5C	CMDCQ05C	
CMDCQ6	CMDCQ06	
CMDCQ7	CMDCQ07	
CMDCQ8A	CMDCQ08A	
CMDCQ8B	CMDCQ08B	
CMDCQ8C	CMDCQ08C	
CMDCQ9A	CMDCQ09A	
CMDCQ9B1	CMDCQ09B	
CMDCQ9B2	CMDCQ09C	
CMDCQ9B3	CMDCQ09D	
CMDCQ10A	CMDCQ10A	
CMDC10B1	CMDCQ10B	
CMDC10B2	CMDCQ10C	
CMDC10B3	CMDCQ10D	
CMDCQ12A	CMDCQ12A	
CMDCQ12B	CMDCQ12C	
CMDCQ13A	CMDCQ13B	
CMDCQ14A	CMDCQ14B	
-----	CMDCQ15	
CMDCQ16	CMDCQ16	

CMDCQ17	CMDCQ17	
CMDCb18B	CMDCQ18	The Territories has a yes/no question first, then asks what type of birthing aids used. The Provinces is one question.
CMDCb18C	-----	
CMDCb18D	-----	
CMDCb18E	-----	
CMDCQ21A	CMDCQ21A	
CMDC21B1	CMDCQ21B	
CMDC21B2	CMDCQ21C	
CMDC21B3	CMDCQ21D	
CMDC21B4	CMDCQ21E	
CMDCQ21C	CMDCQ21F	
CMDCQ22	CMDCQ22	
-----	CMDCQ23A-F	
CMDCQ24A	CMDCQ24A	
-----	CMDCQ24B	
CMDCQ25	CMDCQ25	
CMDCQ26	CMDCQ26	
CMDCQ27	CMDCQ27	
CMDCQ28A	CMDCQ28A	
CMDCQ28B	CMDCQ28B	
CMDCQ28C	CMDCQ28C	
CMDCQ28D	CMDCQ28D	
CMDCQ28E	CMDCQ28E	
CMDCQ28F	CMDCQ28F	
CMDCQ28G	CMDCQ28G	
CMDCQ28H	CMDCQ28H	
CMDCQ28I	CMDCQ28I	
CMDCQ28J	CMDCQ28J	
CMDCQ28K	CMDCQ28K	
CMDCQ28L	CMDCQ28L	

CMDCQ28M	CMDCQ28M	
-----	CWBCcQ1B	
-----	CWBCcQ1C	
-----	CWBCcQ02	
-----	CWBCcQ2B	
CMDCbQ29	CMDCbQ29	
CMDCbc29A	-----	
CMDCb29B	CMDCbQ31	
CWBCc29C	CWBCcQ4A	
CWBCc29D	CWBCcQ4B	
-----	CWBCcQ4C	
CMDCD01	CMDCD01	
CMDCD02	CMDCD02	
CMDCD03	CMDCD03	
CMDCD04	CMDCD04	
CMDCD05	CMDCD05	
CMDCD06	CMDCD06	
CMDCD07	CMDCD07	
CMDCD08	CMDCD08	
-----	CMDCD09	
-----	CMDCD10	
-----	CTMCQ01-33	
-----	CEDCbQ0	
-----	CEDCbQ0A	
CEDCD01	CEDCD01	
-----	CEDCQ02	
-----	CEDCQ03	
CEDCQ6	CEDCQ06	
CEDCD03	CEDCD03	
-----	CEDCQ09A	

-----	CEDCQ09B	
-----	CEDCQ10	
CEDCQ12A	CEDCQ12A	
	CEDCQ12B	
	CEDCQ12C	
CEDCQ13	CEDCbQ13	
CEDCQ14A	CEDCQ14A	
CEDCb14A	CEDCb14A	
CEDCQ14B	CEDCQ14B	
CEDCQ14C	CEDCQ14C	
CEDCb14C	CEDCb14C	
CEDCQ14D	CEDCQ14D	
-----	CEDCb14E	
-----	CEDCb14F	
-----	CEDCb14H	
CEDCQ15	CEDCb11A	Provinces ask “Aside from school changes, how many times in --’s life has he/she moved, that is, changed his/her usual place of residence?”; Territories ask “How many times in --’s life has he/she moved, that is, in how many residences has --lived?”
-----	CEDCQ11	
-----	CEDCQ15A	
-----	CEDCQ15B	
-----	CEDCb15C	
-----	CEDCQ16	
-----	CEDCQ17	
-----	CEDCQ18A	
-----	CEDCQ18B	
	CEDCc18C	
	CEDCc18D	
-----	CEDCQ19A	
-----	CEDCQ19B	

-----	CEDCQ19C	
-----	CEDCQ19D	
-----	CEDCb21A-CEDCb21K	
-----	CEDCbQ22	
-----	CEDCbQ23	
-----	CMACS01-02	
	CMACcS03	
-----	CRECS01-02	
-----	CBECQ1	
-----	CBECQ2	
-----	CBECQ3	
-----	CBECQ4	
-----	CBECQ5	
-----	CBECQ5A	
-----	CBECQ6A	
CBECQ6B	CBECQ6B	
CBECQ6C	CBECQ6C	
CBECQ6D	CBECQ6D	
CBECQ6E	CBECQ6E	
-----	CBECQ6E1	
CBECQ6F	CBECQ6F	
CBECQ6G	CBECQ6G	
-----	CBECQ6H	
CBECQ6I	CBECQ6I	
-----	CBECQ6J	
-----	CBECQ6J1	
CBECQ6K	CBECQ6K	
CBECQ6L	CBECQ6L	
-----	CBECQ6M	
CBECQ6N	CBECQ6N	

CBECQ6O	CBECQ6O	
CBECQ6P	CBECQ6P	
CBECQ6Q	CBECQ6Q	
-----	CBECQ6R	
-----	CBECQ6R1	
CBECQ6S	CBECQ6S	
CBECQ6T	CBECQ6T	
-----	CBECQ6T1	
CBECQ6U	CBECQ6U	
CBECQ6V	CBECQ6V	
CBECQ6W	CBECQ6W	
CBECQ6X	CBECQ6X	
-----	CBECQ6Y	
-----	CBECQ6Z	
-----	CBECQ6Z1	
CBECQ6AA	CBECQ6AA	
CBECQ6BB	CBECQ6BB	
CBECQ6CC	CBECQ6CC	
CBECQ6DD	CBECQ6DD	
-----	CBEC6DD1	
CBECQ6EE	CBECQ6EE	
CBECQ6FF	CBECQ6FF	
-----	CBECQ6GG	
CBECQ6HH	CBECQ6HH	
-----	CBECQ6II	
CBECQ6JJ	CBECQ6JJ	
-----	CBECQ6LL	
-----	CBEC6LL1	
CBECQ6MM	CBECQ6MM	
CBECQ6NN	CBECQ6NN	

-----	CBECQ6OO	
CBECQ6PP	CBECQ6PP	
-----	CBEC6PP1	
CBECQ6QQ	CBECQ6QQ	
CBECQ6RR	CBECQ6RR	
CBECQ6SS	CBECQ6SS	
-----	CBECQ6TT	
-----	CBEC6TT1	
CBECQ6UU	CBECQ6UU	
CBECQ7A	CBECQ7A	
CBECQ7B	CBECQ7B	
CBECQ7C	CBECQ7C	
CBECQ7D	CBECQ7D	
CBECQ7E	CBECQ7E	
CBECQ7F	CBECQ7F	
-----	CBECbQ9A-BECbQ9O	
-----	CBECS01	
-----	CBECS02	
-----	CBECS03	
-----	CBECS04	
-----	CBECS05	
CBECS06	CBECS06	The Provinces used 8 questions to determine this factor score, for 4-15 year olds; the Territories only used 7 questions, for 4-15 year olds.
CBECS07	CBECS07	The Provinces used 10 questions to determine this factor score, for 4-15 year olds; the Territories only used 5 questions, for 4-15 year olds.
CBECS08	CBECS08	The Provinces used 8 questions to determine this factor score; the Territories only used 7 questions.
-----	CBECS09	
-----	CBECS10	
CBECS11	CBECS11	The Provinces determined this factor score for 4-11 year olds; the Territories for 8-15 year olds.
CRLCQ1	CRLCQ01	
CRLCQ2	CRLCQ02	

-----	CRLCQ03	
CRLCQ4	CRLCQ04	
-----	CRLCQ05	
-----	CRLCQ06	
-----	CRLCQ07	
-----	CRLCQ08	
CRLCQ9	CRLCQ09	
CPRCQ1	CPRCQ01	
-----	CPRCQ02	
CPRCQ3	CPRCQ03	
-----	CPRCQ04	
-----	CPRCQ05	
CPRCQ6	CPRCQ06	
CPRCbQ7A	CPRCQ07	The Territories breaks into two questions, depending on age. The Provinces uses one question with different wording, depending on age.
CPRCQ7		
-----	CPRCQ08	
-----	CPRCQ09	
CPRCQ10	CPRCQ10	
CPRCQ11	CPRCQ11	
CPRCQ12	CPRCQ12	
-----	CPRCQ13	
-----	CPRCQ14	
-----	CPRCQ15	
CPRCQ16	CPRCQ16	
-----	CPRCQ17	
-----	CPRCQ18	
-----	CPRCQ19	
-----	CPRCQ20	
-----	CPRCQ21	
-----	CPRCQ22	

-----	CPRCQ23	
-----	CPRCQ24	
-----	CPRCQ25	
-----	CPRCQ26A-K	
-----	CPRCQ27	
-----	CPRCQ28	
-----	CPRCb30A-CPRCb30J	
-----	CPRCb31A-CPRCb3H	
CPRCS01	CPRCS01	The Provinces used 5 questions to determine this factor score; the Territories used 4 questions.
-----	CPRCS02	
CPRCS03	CPRCS03	The Provinces used 5 questions to determine this factor score; the Territories used 4 questions.
-----	CPRCS04	
CPRCS05	CPRCS05	The Provinces used 5 questions to determine this factor score; the Territories used 4 questions.
-----	CPRCS06	
-----	CPRCbS09	
-----	CPRCbS10	
-----	CCRCQ1A-bQ09	
-----	CCRCbD02-06	
-----	CPPCS01-02	
-----	CPPCD01-02	
-----	CA1CQ01-12	
-----	CD1CQ01A-03	
-----	CE1CQ01A-Q	
CRSPD01/ CRSSD01	CRSPD01/ CRSSD01	
CRSPQ1A/ CRSSQ1A	CRSPQ01A/ CRSSQ01A	
CRSPQ1B/ CRSSQ1B	CRSPQ01B/ CRSSQ01B	
CRSPQ1C/ CRSSQ1C	CRSPQ01C/ CRSSQ01C	
CRSPQ1D/ CRSSQ1D	CRSPQ01D/ CRSSQ01D	
CRSPQ1E/ CRSSQ1E	CRSPQ01E/ CRSSQ01E	

CCHPD01/ CCHSD01	CCHPD01/ CCHSD01	
CCHPQ1A/ CCHSQ1A	CCHPQ1A/ CCHSQ1A	
CCHPQ1B/ CCHSQ1B	CCHPQ1B/ CCHSQ1B	
CCHPQ1C/ CCHSQ1C	CCHPQ1C/ CCHSQ1C	
CCHPQ1D/ CCHSQ1D	CCHPQ1D/ CCHSQ1D	
CCHPQ1E/ CCHSQ1E	CCHPQ1E/ CCHSQ1E	
CCHPQ1F/ CCHSQ1F	CCHPQ1F/ CCHSQ1F	
CCHPQ1G/ CCHSQ1G	CCHPQ1G/ CCHSQ1G	
CCHPQ1H/ CCHSQ1H	CCHPQ1H/ CCHSQ1H	
CCHPQ1I/ CCHSQ1I	CCHPQ1I/ CCHSQ1I	
CCHPQ1J/ CCHSQ1J	CCHPQ1J/ CCHSQ1J	
CCHPQ1K/ CCHSQ1K	CCHPQ1K/ CCHSQ1K	
CCHPQ1L/ CCHSQ1L	CCHPQ1L/ CCHSQ1L	
CCHPQ1M/ CCJSQ1M	CCHPQ1M/ CCJSQ1M	
CCHPQ1N/ CCHSQ1N	CCHPQ1N/ CCHSQ1N	
CCHPQ1O/ CCHSQ1O	CCHPQ1O/ CCHSQ1O	
CCHPQ1P/ CCHSQ1P	**-----**	
CCHPQ1R/ CCHSQ1R	**-----**	
CCHPQ1S/ CCHSQ1S	**-----**	
CCHPQ1T/ CCHSQ1T	**-----**	
CCHPbQ1U/ CCHSbQ1U	CCHPQ1U/ CCHSQ1U	
CCHPbQ1V/ CCHSbQ1V	-----	
CCHPbQ1W/ CCHSbQ1W	-----	
CHLPQ1/ CHLSQ1	CHLPQ01/ CHLSQ01	
CHLPQ2/ CHLSQ2	CHLPQ02/ CHLSQ02	
CHLPQ3/ CHLSQ3	CHLPQ03/ CHLSQ03	
CHLPcQ5A/ CHLScQ5A	CHLPcQ5A/ CHLScQ5A	

CHLPQ5B/ CHLSQ5B	-----	
CHLPQ5C/ CHLSQ5C	-----	
CHLPQ5D/ CHLSQ5D	-----	
CHLPcQ5/ CHLScQ5	CHLPQ05/ CHLSQ05	
CHLPbQ6/ CHLSbQ6	CHLPQ06/ CHLSQ06	
CHLPcQ7/ CHLPcQ7	-----	
CHLMQ8	CHLMQ08	Provinces ask “how many times throughout your life...”; Territories ask “how many times...”
CHLMQ9	CHLMQ09	
CHLMQ11	CHLMQ11	
CHHHQ1	CHHHQ01	
-----	CHHHQ02	
-----	CHHHQ02B	
CHHHQ3	CHHHQ03	
CHHHQ6	CHHHQ06	Territories have additional ‘suite’ options for the first four dwelling types.
-----	CHHHQ06b	
CSFHcQ1	CSFHQ01	
CSFHcQ2	CSFHQ02	
-----	CSFHQ03	
CSFHcQ4	CSFHcQ04	
CSFHcQ5A	**-----**	
CSFHcQ5B	**-----**	
CSFHcQ5C	CSFHQ05C	
CSFHcQ6A	CSFHQ06A	
CSFHcQ6B	CSFHQ06B	
CSFHcQ6C	CSFHQ06C	
CSFHcQ6D	CSFHQ06D	
CSFHcQ6E	CSFHQ06E	
CSFHcQ7A	-----	
CSFHcQ7B	-----	

CSFHcQ7C	-----	
CSFHcQ7D	-----	
CSFHcQ7E	-----	
CSFHcQ7F	-----	
-----	CSPHQ01A-F	
-----	CSPHS01	
-----	CSFHS6	
-----	COBHQ01-09	
CHLCQ1	CHLCQ01	
CHLCQ2	CHLCQ02	
CHLCQ3B	CHLCQ03B	
CHLCQ4A	CHLCQ04A	
CHLCb4C1	CHLCb4C1	
CHLCQ5	CHLCQ05	
CHLCbQ6	CHLCQ06	For the Health Status section, the Territories asks 4-5 year olds only. The Provinces asks children>3 and uses different wording for questions CHLCQ06, CHLCQ07, CHLCQ20 AND CHLCQ22, depending on age.
CHLCbQ7	CHLCQ07	
CHLCQ8	CHLCQ08	
CHLCQ9	CHLCQ09	
CHLCQ10	CHLCQ10	
CHLCQ11	CHLCQ11	
CHLCQ12	CHLCQ12	
CHLCQ13	CHLCQ13	
CHLCQ14	CHLCQ14	
CHLCQ15	CHLCQ15	
CHLCQ16	CHLCQ16	
CHLCQ17	CHLCQ17	
CHLCQ18	CHLCQ18	
CHLCQ19	CHLCQ19	
CHLCbQ20	CHLCQ20	

CHLCQ21	CHLCQ21	
CHLCbQ22	CHLCQ22	
CHLCQ23	CHLCQ23	
CHLCQ24	CHLCQ24	
CHLCQ25	CHLCQ25	
CHLCQ26	CHLCQ26	
CHLCQ27	CHLCQ27	
CHLCQ28	CHLCQ28	
CHLCQ29	CHLCQ29	
CHLCQ30	CHLCQ30	
CHLCQ31	CHLCQ31	
CHLCQ32	CHLCQ32	
CHLCQ33	CHLCQ33	
CHLCQ34	CHLCQ34	
CHLCQ35	CHLCQ35	
CHLCQ36	CHLCQ36	
CHLCQ37	CHLCQ37	
CHLCQ38	CHLCQ38	
CHLCQ39	CHLCQ39	
CHLCQ40	CHLCQ40	
CHLCQ41	CHLCQ41	
CHLCQ42	CHLCQ42	Territories include an additional option ‘10 - In the bush/on the land’
CHLCb43A	CHLCQ43A	
CHLCQ43B	CHLCQ43B	
CHLCQ43C	CHLCQ43C	
CHLCQ44	CHLCQ44	
CHLC45AA	CHLCQ45A	
CHLC45AB	CHLCQ45B	
CHLC45AC	CHLCQ45C	
CHLC45AD	CHLCQ45D	

CHLC45AE	CHLCQ45E	
CHLC45AF	CHLCQ45F	
CHLC45AG	CHLCQ45G	
CHLC45AH	CHLCQ45H	
CHLC45AI	CHLCQ45I	
CHLC45AJ	CHLCQ45J	
CHLC45AK	-----	
CHLCQ45B	CHLCQ45L	
-----	CHLCQ46	
-----	CHLCQ47A-B	
-----	CHLCQ48A-I	
-----	CHLCQ49	
-----	CHLCQ50	
-----	CHLCQ51A-E	
-----	CHLCQ52A-O	
-----	CHLCD02	
-----	CHLCbD4C	
-----	CHLCbD4D	
-----	CHLCbD45	
-----	CHLCD51	
-----	CHLCc02A	
-----	CACCQ1	
-----	CACCQ2AA-AG	
-----	CACCQ2B	
CACCQ3A	CACCQ3A	
CACCb3AA	CACCb3AA	
CACCQ3B	CACCQ3B	
CACCQ3C	CACCQ3C	
CACCQ3D	CACCQ3D	Provinces have slightly different wordings for this question, depending on the child's age.
CACCcQ4A	CACCcQ4A	

CACCcQ4B	CACCcQ4B	
CACCbQ8	CACCbQ4C	
CACCQ9	CACCQ5	
-----	CACCQ6A-F	
-----	CACCQ7A-B	
-----	CACCQ8A-B	
-----	CEPCS15	
CCSCQ1	-----	
CCSCQ2A	-----	
CCSCQ2B	-----	
CCSCQ3A	-----	
CCSCQ3B	-----	
CCSCQ4A	-----	
CCSCQ4B	-----	
CCSCQ5	-----	
CCSCQ6A	-----	
CCSCQ6B	-----	
CCSCbQ6C	-----	
CCSCQ7	-----	
CCSCQ8	-----	
CCSCQ9A	-----	
CCSCQ9B	-----	
CCSCQ10A	-----	
CCSCQ10B	-----	
CCSCc10C	-----	
CCSCQ11	-----	
CCSCQ12	-----	
CCSCQ13	-----	
CCSCQ14	-----	
	CLFPQ01/CLFSQ01	

CLFPQ2/ CLFSQ2	CLFPQ02/ CLFSQ02	
-----	CLFPbQ03/CLFSbQ03	
	CLFPbQ3A/CLFSbQ3A	
	CLFPbQ04/CLFSbQ04	
CLFPQ4A/ CLFSQ4A	-----	
CLFPQ6A/ CLFSQ6A	-----	
-----	CLFPcQ06/ CLFScQ06	
CLFPc18A/CLFSc18A	CLFPcQ6A/CLFScQ6A	
CLFPc18B/CLFSc18B	CLFPcQ6B/CLFScQ6B	
CLFPc18C/CLFSc18C	CLFPcQ6C/CLFScQ6C	
CLFPQ11A/ CLFSQ11A	-----	
CLFPQ4B/ CLFSQ4B	-----	
CLFPQ6B/ CLFSQ6B	-----	
CLFPQ11B/ CLFSQ11B	-----	
CLFPQ4C/ CLFSQ4C	-----	
CLFPQ6C/ CLFSQ6C	-----	
CLFPQ11C/ CLFSQ11C	-----	
CLFPb8_4/CLFSb8_4	-----	
CLFPb84A/CLFSb84A	-----	
CLFPb84B/CLFSb84B	-----	
CLFPQ12/ CLFSQ12	-----	
CLFPQ16/ CLFSQ16	CLFPbQ13/ CLFSbQ13	For the Provinces, the question starts with, “At this work...”.
-----	CLFPb14A	
-----	CLFPb14B	
-----	CLFPb14C	
-----	CLFP14CC	
-----	CLFPb14D	
-----	CLFPb15A	
-----	CLFPbQ16	
CINHQ1A	**-----**	

CINHQ1B	-----	
CINHQ1C	-----	
CINHQ1D	-----	
CINHQ1E	-----	
CINHQ1F	-----	
CINHQ1G	-----	
CINHQ1H	-----	
CINHQ1I	-----	
CINHQ1J	-----	
CINHQ1K	-----	
CINHQ1L	-----	
CINHQ1M	-----	
CINHQ2	-----	
CINHD02B	-----	
CINHQ3	-----	
CINPQ4B/CINSQ4B	-----	
CSDPQ1/ CSDSQ1/ CSDCQ1	CSDPQ1/ CSDSQ1/ CSDCQ1	
CSDPQ2A1/ CSDSQ2A1/ CSDCQ2A1	CSDPQ2AA/ CSDSQ2AA/ CSDCQ2AA	
CSDPQ2A2/ CSDSQ2A2/ CSDCQ2A2	CSDPQ2AB/ CSDSQ2AB/ CSDCQ2AB	
CSDPQ2A3/ CSDSQ2A3/ CSDCQ2A3	CSDPQ2AC/ CSDSQ2AC/ CSDCQ2AC	
CSDPQ2A4/ CSDSQ2A4/ CSDCQ2A4	CSDPQ2AD/ CSDSQ2AD/ CSDCQ2AD	
CSDPQ2B/ CSDSQ2B/ CSDCQ2B	CSDPQ2B/ CSDSQ2B/ CSDCQ2B	
CSDPQ3/ CSDSQ3/ CSDCQ3	CSDPQ3/ CSDSQ3/ CSDCQ3	
CSDPQ4A/ CSDSQ4A/ CSDCQ4A	CSDPQ4A/ CSDSQ4A/ CSDCQ4A	
CSDPQ4B/ CSDSQ4B/	CSDPQ4B/ CSDSQ4B/	

CSDCQ4B	CSDCQ4B	
CSDPQ4C/ CSDSQ4C/ CSDCQ4C	CSDPQ4C/ CSDSQ4C/ CSDCQ4C	
CSDPQ4D/ CSDSQ4D/ CSDCQ4D	CSDPQ4D/ CSDSQ4D/ CSDCQ4D	
CSDPQ4E/ CSDSQ4E/ CSDCQ4E	CSDPQ4E/ CSDSQ4E/ CSDCQ4E	
CSDPQ4F/ CSDSQ4F/ CSDCQ4F	CSDPQ4F/ CSDSQ4F/ CSDCQ4F	
CSDPQ4G/ CSDSQ4G/ CSDCQ4G	CSDPQ4G/ CSDSQ4G/ CSDCQ4G	
CSDPQ4H/ CSDSQ4H/ CSDCQ4H	CSDPQ4H/ CSDSQ4H/ CSDCQ4H	
CSDPQ4I/ CSDSQ4I/ CSDCQ4I	CSDPQ4I/ CSDSQ4I/ CSDCQ4I	
CSDPQ4J/ CSDSQ4J/ CSDCQ4J	CSDPQ4J/ CSDSQ4J/ CSDCQ4J	
CSDPQ4K/ CSDSQ4K/ CSDCQ4K	CSDPQ4K/ CSDSQ4K/ CSDCQ4K	
CSDPQ4L/ CSDSQ4L/ CSDCQ4L	CSDPQ4L/ CSDSQ4L/ CSDCQ4L	
CSDPbQ4S/ CSDSbQ4S/ CSDCbQ4S	CSDPQ4M/ CSDSQ4M/ CSDCQ4M	
CSDPQ4M/ CSDSQ4M/ CSDCQ4M	CSDPQ4N/ CSDSQ4N/ CSDCQ4N	
CSDPQ4N/ CSDSQ4N/ CSDCQ4N	CSDPQ4O/ CSDSQ4O/ CSDCQ4O	
CSDPQ4O/ CSDSQ4O/ CSDCQ4O	CSDPQ4P/ CSDSQ4P/ CSDCQ4P	
CSDPQ4P/ CSDSQ4P/ CSDCQ4P	CSDPQ4Q/ CSDSQ4Q/ CSDCQ4Q	
CSDPQ4Q/ CSDSQ4Q/ CSDCQ4Q	CSDPQ4R/ CSDSQ4R/ CSDCQ4R	
CSDPQ4R/ CSDSQ4R/ CSDCQ4R	CSDPQ4S/ CSDSQ4S/ CSDCQ4S	
CSDPQ5A/ CSDSQ5A/	CSDPQ5A/ CSDSQ5A/	

CSDCQ5A	CSDCQ5A	
CSDPQ5B/ CSDSQ5B/ CSDCQ5B	CSDPQ5B/ CSDSQ5B/ CSDCQ5B	
CSDPQ5C/ CSDSQ5C/ CSDCQ5C	-----	CSDPQ5C in the Territories includes CSDPQ5C - 5S (inclusive) in the Provinces.
-----	CSDPQ5C/ CSDSQ5C/ CSDCQ5C	
-----	CSDPQ5D/ CSDSQ5D/ CSDCQ5D	
-----	CSDPQ5E/ CSDSQ5E/ CSDCQ5E	
-----	CSDPQ5F/ CSDSQ5F/ CSDCQ5F	
-----	CSDPQ5G/ CSDSQ5G/ CSDCQ5G	
-----	CSDPQ5H/ CSDSQ5H/ CSDCQ5H	
-----	CSDPQ5I/ CSDSQ5I/ CSDCQ5I	
-----	CSDPQ5J/ CSDSQ5J/ CSDCQ5J	
-----	CSDPQ5K/ CSDSQ5K/ CSDCQ5K	
-----	CSDPQ5L/ CSDSQ5L/ CSDCQ5L	
-----	CSDPQ5M/ CSDSQ5M/ CSDCQ5M	
-----	CSDPQ5N/ CSDSQ5N/ CSDCQ5N	
-----	CSDPQ5O/ CSDSQ5O/ CSDCQ5O	
-----	CSDPQ5P/ CSDSQ5P/ CSDCQ5P	
-----	CSDPQ5Q/ CSDSQ5Q/ CSDCQ5Q	
-----	CSDPQ5R/ CSDSQ5R/	

	CSDCQ5R	
-----	CSDPQ5S/ CSDSQ5S/ CSDCQ5S	
CSDPQ6A1/ CSDSQ6A1/ CSDCQ6A1	CSDPQ6/ CSDSQ6/ CSDCQ6	
CSDPQ6A2/ CSDSQ6A2/ CSDCQ6A2	CSDPQ6B/ CSDSQ6B/ CSDCQ6B	
CSDPQ6A3/ CSDSQ6A3/ CSDCQ6A3	-----	CSDPQ6A in the Territories includes CSDPQ6C - 6S (inclusive) in the Provinces
-----	CSDPQ6C/ CSDSQ6C/ CSDCQ6C	
-----	CSDPQ6D/ CSDSQ6D/ CSDCQ6D	
-----	CSDPQ6E/ CSDSQ6E/ CSDCQ6E	
-----	CSDPQ6F/ CSDSQ6F/ CSDCQ6F	
-----	CSDPQ6G/ CSDSQ6G/ CSDCQ6G	
-----	CSDPQ6H/ CSDSQ6H/ CSDCQ6H	
-----	CSDPQ6I/ CSDSQ6I/ CSDCQ6I	
-----	CSDPQ6J/ CSDSQ6J/ CSDCQ6J	
-----	CSDPQ6K/ CSDSQ6K/ CSDCQ6K	
-----	CSDPQ6L/ CSDSQ6L/ CSDCQ6L	
-----	CSDPQ6M/ CSDSQ6M/ CSDCQ6M	
-----	CSDPQ6N/ CSDSQ6N/ CSDCQ6N	
-----	CSDPQ6O/ CSDSQ6O/ CSDCQ6O	
-----	CSDPQ6P/ CSDSQ6P/	

	CSDCQ6P	
-----	CSDPQ6Q/ CSDSQ6Q/ CSDCQ6Q	
-----	CSDPQ6R/ CSDSQ6R/ CSDCQ6R	
-----	CSDPQ6S/ CSDSQ6S/ CSDCQ6S	
CSDPQ6B1/ CSDSQ6B1/ CSDCQ6B1	-----	
CSDPQ6B2/ CSDSQ6B2/ CSDCQ6B2	-----	
CSDPQ7A/ CSDSQ7A/ CSDCQ7A	CSDPb4AA/ CSDSb4AA/ CSDCb4AA/	
CSDPQ7B/ CSDSQ7B/ CSDCQ7B	CSDPb4AD/ CSDSb4AD/ CSDCb4AD/	
CSDPQ7C/ CSDSQ7C/ CSDCQ7C	CSDPb4AK/ CSDSb4AK/ CSDCb4AK/	
CSDPQ7D/ CSDSQ7D/ CSDCQ7D	CSDPb4AG/ CSDSb4AG/ CSDCb4AG/	
CSDPQ7E/ CSDSQ7E/ CSDCQ7E	CSDPb4AJ/ CSDSb4AJ/ CSDCb4AJ/	
CSDPQ7F/ CSDSQ7F/ CSDCQ7F	CSDPb4AB/ CSDSb4AB/ CSDCb4AB/	
CSDPQ7H/ CSDSQ7H/ CSDCQ7H	CSDPb4AC/ CSDSb4AC/ CSDCb4AC/	
CSDPQ7I/ CSDSQ7I/ CSDCQ7I	CSDPb4AH/ CSDSb4AH/ CSDCb4AH/	
CSDPQ7J/ CSDSQ7J/ CSDCQ7J	CSDPb4AF/ CSDSb4AF/ CSDCb4AF/	
CSDPQ7K/ CSDSQ7K/ CSDCQ7K	CSDPb4AL/ CSDSb4AL/ CSDCb4AL/	
CSDPbQ7L/ CSDSbQ7L/ CSDCbQ7L	CSDPb4AI/ CSDSb4AI/ CSDCb4AI/	
-----	CSDPb4AE/ CSDSb4AE/ CSDCb4AE/	
CSDPQ8/ CSDSQ8/	CSDPQ8/ CSDSQ8/	Territories include an additional option ‘15 - Bahai’

CSDCQ8	CSDCQ8	
CMSCbS01	CMSCS01	
CMSCbS02	CMSCS02	
CMSCQ1	CMSCQ01	
CMSCQ2	CMSCQ02	
CMSCQ3	CMSCQ03	
CMSCQ4	CMSCQ04	
CMSCQ5	CMSCQ05	
CMSCQ6	CMSCQ06	
CMSCQ7	CMSCQ07	
CMSCQ8	CMSCQ08	
CMSCQ9	CMSCQ09	
CMSCQ10	CMSCQ10	
CMSCQ11	CMSCQ11	
CMSCQ12	CMSCQ12	
CMSCQ13	CMSCQ13	
CMSCQ14	CMSCQ14	
CMSCQ15	CMSCQ15	
CMSCQ16	CMSCQ16	
CMSCQ17	CMSCQ17	
CMSCQ18	CMSCQ18	
CMSCQ19	CMSCQ19	
CMSCQ20	CMSCQ20	
CMSCQ21	CMSCQ21	
CMSCQ22	CMSCQ22	
CMSCQ23	CMSCQ23	
CMSCQ24	CMSCQ24	
CMSCQ25	CMSCQ25	
CMSCQ26	CMSCQ26	
CMSCQ27	CMSCQ27	

CMSCQ28	CMSCQ28	
CMSCQ29	CMSCQ29	
CMSCQ30	CMSCQ30	
CMSCQ31	CMSCQ31	
CMSCQ32	CMSCQ32	
CMSCQ33	CMSCQ33	
CMSCQ34	CMSCQ34	
CMSCQ35	CMSCQ35	
CMSCQ36	CMSCQ36	
CMSCQ37	CMSCQ37	
CMSCQ38	CMSCQ38	
CMSCQ39	CMSCQ39	
CMSCQ40	CMSCQ40	
CMSCQ41	CMSCQ41	
CMSCQ42	CMSCQ42	
CMSCQ43	CMSCQ43	
CMSCQ44	CMSCQ44	
CMSCQ45	CMSCQ45	
CMSCQ46	CMSCQ46	
CMSCQ47	CMSCQ47	
CMSCQ48	CMSCQ48	
-----	CPACQ01-15	
-----	CPACS01-04	
CWTCW01C	CWTCW01C	
CWTCW01L	CWTCW01L	
CIDHD01	CIDHD01	
CIDCD01	CHILDID	
SHXSECWT	SHXSECWT	
SHLONGWT	SHLONGWT	
SHAREFLAG	SHAREFLAG	

XSECFLG	XSECFLG	
LONGFLG	LONGFLG	

7.0
Weighting of the Sample

The principle behind estimation in a probability sample such as the NLSCY is that each person in the sample "represents", besides himself or herself, several other persons not in the sample. For example, in a 2% simple random sample of the population, each person in the sample represents 50 persons in the population.

The weighting phase is a step that calculates, for each record, what this number is (i.e., the number of individuals in the population represented by this record). As the target population is not the same for the cross-sectional sample and the longitudinal sample, the number of persons each child represents is not the same. Consequently, two series of weights must be calculated: one for the cross-sectional sample, and one for the longitudinal sample. These weights appear on the master files (CWTCW01C – cross-sectional and CWTCW01L - longitudinal), and must be used to derive meaningful estimates of the characteristics measured by the survey. For example, if the number of children living in single-parent families in 1998 is to be estimated, it is done by selecting the records in the cross-sectional sample of Cycle 3 with that characteristic and summing the cross-sectional weights found on those records.

Longitudinal Sample or Cross-sectional Sample?

The choice of which sample to use depends on the type of analysis to be done. The longitudinal sample pertains to the child population at the time this sample was selected (i.e., 1994-95). The sum of the longitudinal weights is equal to the available demographic estimates for July 1994. Only the longitudinal children, i.e., those selected in 1994, are given a longitudinal weight other than 0. For each cycle, the longitudinal weight of the panel is recalculated to take into account the further erosion (non-response) that occurs between the two cycles of the survey, i.e., about two years. It is this one that is usually better suited to longitudinal analysis based on a comparison of the data for more than one year, as it allows for the life courses of the children to be quantified over time.

The cross-sectional sample makes it possible to do estimates based on data from a single cycle. A separate cross-sectional weight is calculated for each cycle. For Cycle 1, the longitudinal sample and the cross-sectional sample have the same target population. As the target populations are identical, only one series of weights was needed for this cycle.

Fluctuations may be calculated using cross-sectional estimates produced for two cycles. However, the fluctuations thus measured are net fluctuations. They are calculated based on a snapshot taken for each reference period. As a result, they mask all transitions that cancel each other out.

The first step involved with assigning the weights for Cycle 3 was to identify which children receive a longitudinal weight and/or cross-sectional weight. The different types of children identified are as follows:

- Longitudinal children from the Cycle 1 sample.
- Longitudinal children from the Cycle 2 top-up.
- Children new to the survey in Cycle 3.

7.1 Longitudinal Weight

An initial longitudinal weight was assigned to longitudinal children based on the cycle 1 weight or the cycle 2 longitudinal weight. This weight is determined from the probability of selection first measured when the sample was drawn in Cycle 1. For children new to the survey in Cycle 2, their probability of selection was unknown and was estimated from an external source for benchmark totals. Their probability of selection was estimated as the ratio of children interviewed over the estimated population of

children from adjusted census counts provided at the time.

Post-Stratification

The longitudinal weights were adjusted to conform to known population totals. For the children who were part of the survey in 1994, post-stratification adjusts the weights so that the sum of the weights is equal to known demographic estimates for July 1994. For the children who were part of the cycle 2 top-up, post-stratification adjusts the weights so that the sum of the weights is equal to know demographic estimates for July 1996. Post-stratification was done by age, sex and territory.

NWT Population Estimates - July 1, 1994			
New Territory	Sex	Age Group	Population
Nunavut	Male	0-4	1,865
		5-9	1,658
		10-11	579
	Female	0-4	1,742
		5-9	1,573
		10-11	532
Western	Male	0-4	2,146
		5-9	2,028
		10-11	750
	Female	0-4	2,084
		5-9	1,904
		10-11	686

Yukon Population Estimates – July 1, 1994		
Sex	Age Group	Population
Male	0-4	1,309
	5-9	1,290
	10-11	530
Female	0-4	1,209
	5-9	1,134
	10-11	489

NWT Population Estimates - July 1, 1996			
New Territory	Sex	Age Group	Population
Nunavut	Male	0-4	1,923
		5-9	1,746
		10-13	1,163
	Female	0-4	1,764
		5-9	1,648
		10-13	1,076
Western	Male	0-4	2,140
		5-9	2,036
		10-13	1,494
	Female	0-4	2,056
		5-9	2,041
		10-13	1,309

Yukon Population Estimates – July 1, 1996		
Sex	Age Group	Population
Male	0-4	1,252
	5-9	1,364
	10-13	1,060
Female	0-4	1,168
	5-9	1,166
	10-13	1,027

7.2 Cross-sectional Weight

In cycle 3 the collection for the territories was the responsibility of the individual territories. As a result we do not know what sample design was used to select the children who were part of the top-up in cycle 3. These households were selected using a non-probabilistic method such as quota or snowball sampling. The absence of this information poses a problem for estimating population totals or averages using the common Horowitz-Thompson estimator since it requires knowledge of inclusion probabilities. By choosing the alternate ratio estimator, the values of the inclusion probabilities are not essential. What is required is auxiliary information known for all population units. The census and demographic update of population estimates are used as this auxiliary information.

For the children who were not part of the top-up in cycle 3 an initial cross-sectional weight was assigned based on the cycle 1 weight or the cycle 2 longitudinal weight. For the children who were new in cycle 3 and between the ages of 2 and 15, their initial weight was set to the average initial weight of the longitudinal children of the same age and sex, who lived in the same territory. This was done to eliminate outlying weight factors that would result from the disparate sampling probabilities assumed for children who were not in the sample initially. For the newborn children in cycle 3 an initial weight of 1 was assigned.

Post-Stratification

The cross-sectional weights were adjusted to conform to known population totals. Post-stratification adjusts the weights so that the sum of the weights is equal to known demographic estimates for July 1998. Post-stratification was done by age, sex and territory.

NWT Population Estimates - July 1, 1998			
New Territory	Sex	Age Group	Population
Nunavut	Male	0-4	1,868
		5-9	1,879
		10-15	1,732
	Female	0-4	1,721
		5-9	1,725
		10-15	1,642
Western	Male	0-4	1,876
		5-9	2,017
		10-15	2,214
	Female	0-4	1,778
		5-9	2,141
		10-15	1,944

Yukon Population Estimates – July 1, 1998		
Sex	Age Group	Population
Male	0-4	1,133
	5-9	1,267
	10-15	1,571
Female	0-4	1,033
	5-9	1,160
	10-15	1,470

8.0
NLSCY Concepts and Definitions

There are many variables and concepts which are critical to the analysis of the NLSCY data. In this section there is a brief discussion regarding the types of analyses that are possible with the NLSCY data. This is followed by a description of key variables which have been derived to explain the living arrangements of the child and the socio-economic conditions under which the child lives.

8.1 Cross-sectional and Longitudinal Estimates

The NLSCY design and sample has been constructed so that it will be possible to produce both **cross-sectional** and **longitudinal** estimates.

The allocation of the Cycle 1, 2 and 3 sample was such that it will be possible to produce estimates at the territorial levels for aggregated age groups. This is true for cross-sectional data as well as longitudinal data. The cross-sectional sample has been at risk for coverage inadequacies since a significant portion of the sample should have been devoted to the larger than usual influx of migrants to the North from one cycle period to the next. In Cycle 3 a small cross-sectional top-up was added to compensate for this. However, because of the non-probabilistic method used to top-up the sample and the limited size of this sample, correction for this coverage problem has been uneven across the territories.

There are two longitudinal cohorts, those who were in the sample beginning with Cycle 1 (aged 0-11 years at cycle 1) and those who were in the sample beginning in Cycle 2 (0-1 year olds). The Cycle 1 longitudinal sample is comprised of all children sampled for Cycle 1 of the survey in responding households (excluding those from the integrated sample (NPHS) and the 3rd and 4th child of each family). The plan is to follow these children over time, and revisit them every two years. Due to the transient nature of the population in the North, which is more noticeable there than anywhere else in Canada, estimations from a sample are particularly vulnerable. In a longitudinal context, this can cause serious reliability problems since children who no longer reside in the North are no longer part of the sample for the North, effectively reducing the sample size significantly over the number of panels in the series. Analyses of these children should allow researchers the opportunity to perform in-depth studies of the long-term impact of risk factors (such as divorce or the onset of a health condition) and protective factors (such as positive interactions with parents or academic success at school) on these children. If a child moves out of the household where he or she was sampled at Cycle 1, that child will be traced to wherever he or she resides during future cycles of the survey, provided that they do not move out of the North. From a longitudinal perspective, the child, not the household, is the statistical unit of analysis.

It should be noted that some children who were participants in Cycle 1 of the NLSCY did not participate in the second cycle or may not participate in subsequent cycles due to a variety of reasons. This is usually referred to as attrition. The numbers of these children is being carefully monitored and we are making every effort to keep these numbers at a minimum. Due to the nature of the population and the data collection in the north the response rates have been lower than expected for the Territories. This negatively affects the reliability of the longitudinal estimates. The future response rates and any further attrition of the sample will determine if the sample will still permit longitudinal research by age cohort at the current levels.

In cycles 2 and 3, the NLSCY added children belonging to age groups no longer covered in the longitudinal sample. For example, in Cycle 2 a panel of children 0 and 1 years of age was added to the Cycle 2 sample. In Cycle 3 a top-up of children of all ages (0 to 15 years old) was also added. This augmented sample allows for ongoing cross-sectional analyses to supplement the primary longitudinal research. As such, at each cycle it is possible to get a snapshot of Canadian children of all ages.

8.2 NLSCY Units of Analyses

The unit of analysis for the NLSCY is intended to be the child and eventually the young adult. For each cycle of the NLSCY, extensive information will be gathered on the child's family, parent(s), and neighbourhood.

It is true that families or households are relatively straightforward units of analysis with cross-sectional data but the situation becomes quite problematic with longitudinal data. Households change composition frequently, due to divorce of parents, or children leaving the parental nest. Attempts have been made in other studies to define "longitudinal households" but the implementation of this concept has never been straightforward. No single definition has been found to be appropriate for most analytic tasks, and many definitions exclude the portion of the population that has undergone the change. Unfortunately, this is often a significant as well as interesting population to study. It has been suggested that a superior alternative is to use the individual as the unit of analysis and present family and household variables as a characteristic of the individual.

Thus the file which has been constructed for this release of NLSCY data consists of child records. In order to understand the family situation, estimates such as of the number of children in single parent families, or the number of children living in low income households, can be produced.

8.3 PMK and Spouse

In each NLSCY household, one child 0 to 15 years of age was selected at random and a question was asked to determine who in the household was the **p**erson **m**ost **k**nowledgeable about this child. This person was labelled as the **PMK**. The intention was that the PMK would provide the information for all selected children in the household and also provide socio-demographic information about herself/himself and his/her spouse/partner. In some rare cases it might have been appropriate to label two different people in a household as PMKs. For example, in the case of a step family, it may have been appropriate to label the mother as the PMK for one child and the father for another. However, in order to simplify the interview procedures, only one PMK was selected per household. It must be noted that, for several reasons, the PMK and his/her spouse could be two different people in the first and third cycles. Therefore, care must be taken when doing longitudinal analyses involving the characteristics of the parents.

The following is the breakdown of the relationship of the PMK to the NLSCY children in the territories for Cycle 3 (using unweighted data) (CDMCD06).

-for 81.96% of responding children, the PMK was the mother (74.23% the biological mother and 7.73% the step, adoptive or foster mother)
-for 15.94% of the children the PMK was the father (biological, step, adoptive or foster father)
-for 2.12% of children the PMK was not a parent.

If the PMK had a partner residing in the household at the time of the interview, this person was labelled as the **spouse**. Spouses included both married, common-law and same-sex partners. Detailed socio-economic information was collected about the spouse/partner in order to describe the family situation of the child.

The following is the breakdown of the relationship of the spouse/partner to the NLSCY children in the territories for Cycle 3 (using unweighted data) (CDMCD06B).

-for 20.34% of the children, the PMK did not have a spouse/partner residing in the household
-for 63.02% of children the spouse/partner was the father (53.42% the biological father and 9.6% the step, adoptive or foster father)
-for 12.05% of children the spouse/partner was the mother (biological, step, adoptive or foster)

-for .97% of children, the spouse/partner was not a parent.
-for 3.60% not stated

8.4 Family Derived Variables

Using NLSCY data, a child's family may be described in several different ways. Many of the family variables that have been used to describe the NLSCY children were derived from what is known as the relationship grid. As part of the household roster some basic demographic information was collected for all members of the child's household. As part of this questionnaire, the relationship of everyone in the household to the PMK was asked. Using this information it was possible to create an extensive set of variables to describe the child's family situation.

The following are some of the family derived variables for the child that exist on this master file for the NLSCY. The names of the derived variable are given in brackets.

Single-parent family

There are two ways of describing the parental situation of children using NLSCY data.

Using the relationship grid, a child's single-parent status was derived. In Cycle 3, there were 78.51% of children living with two parents, 19.36% with one parent, 0.55% without a parent and 1.57% not stated. (CDMCD04) (using weighted data).

A child's parent status can also be defined in terms of the PMK. 79.66% of the NLSCY children were living in a household where the PMK had a spouse/partner; and for 20.34% of children, the PMK did not have a spouse/partner (CDMPD06A) (using weighted data).

The two ways of describing the child's family are very similar. The only reason for the small differences is a result of the few cases where the child lived with a parent, but the parent was not selected to be the PMK.

Step, Blended and Intact Families

Children living with two parents are classified as being members of intact, step and/or blended families based on the relationship of these children to the parents.

Intact family

An intact family consists of a married or common-law couple where **all** children are the natural and/or adopted offspring of both members of the couple.

For the NLSCY children, 70.79% were members of an intact family (CDMCD16) (using weighted data).

Step family

A step family consists of a married or common-law couple residing in the same household, with at least one step child living with them who is the biological or adopted child of one parent but not the other parent. It should be noted that a child who is the biological child of both parents is said to belong to a step family if at least one of these parents has a step child residing in the household.

For the NLSCY children, 5.22% were step children themselves (CDMCD03) and 7.01% lived in a step family (CDMCD15) (using weighted data).

Blended family

Blended families combine children who have different relationships with their parents. A blended family consists of a married or common-law couple living with at least two children, one of whom does not share the same natural and/or adoptive parents as the other child(ren). The following are examples of blended families:

- a couple with biological children of the female partner as well as biological children of the male partner (i.e., hers and his)

- a couple with biological children of the female partner as well as children out of the new union (i.e., hers and theirs).

The blended family is a sub-set of the step family. For the Cycle 3 NLSCY children, 1.95% were members of a blended family (CDMCD14) (using weighted data).

Economic Family

For the NLSCY, an economic family is defined as all family members related by blood, marriage, common-law relationship or adoption; foster children are considered to be part of the economic family. For example, if a woman lives in a household with her spouse and two children as well as her sister and her sister's child then all of these individuals would be part of one economic family. If a boarder also resided in the household with her child then this would constitute a second economic family.

Siblings

For the NLSCY data, siblings include full, half, step, adopted and foster siblings. Only siblings residing in the household have been included in the calculation of the sibling derived variables included on the master file. In the case of common-law relationships, if both members have brought their own children into the relationship then these children are considered as siblings. It should be noted that the classification of siblings was age independent. If an NLSCY child had an adult sibling (for example, 21 years of age) living in the household then this sibling was included in the calculation of the sibling derived variables. The sibling derived variables include total siblings, as well as number of older siblings, younger siblings and siblings of exactly the same date of birth; i.e., twins (CDMCD08, 09, 10 and 11).

8.5 Socio-Economic Derived Variables

There were several derived variables produced from Cycle 3 data to assist analysts in understanding and explaining the socio-economic situation of the child's family.

Socio-economic Status (CINHD09)

Sociologists often use the term "socio-economic status" (SES) to refer to the relative position of a family or individual in an hierarchical social structure, based on their access to, or control over, wealth, prestige and power. In studies of children's academic and social-emotional development, SES is often operationally defined through measures describing the occupational prestige, educational levels, and economic positions of children's parents.

For the first three cycles of the NLSCY a measure of SES was derived for each household in the sample

and the result assigned to each selected child in that household. It was derived from five sources: the level of education of the PMK, the level of education of the spouse/partner, the prestige of the PMK's occupation, the prestige of the occupation of the spouse/partner, and household income. The method of constructing each component of SES, and the construction of the overall SES measure are described below. It should be noted that for the North, family income was collected differently and so the SES calculation is different from what was done in the provinces. The SES was also calculated using the provinces standard scores (for the PINEOLOG, the mean and standard deviation of the provinces were used in the calculation of the score), to allow for comparisons between the territories and the provinces.

Education - Years of School

The education variable used in the construction of SES was years of schooling. Two such variables were derived independently; one for the PMK and one for the spouse/partner (CEDPD04 for the PMK and CEDSD04 for the Spouse/partner). For the PMK the years of schooling variable was derived based on items CEDPQ1 (years of elementary and high school) and CEDPQ4 (highest level of education attained beyond high school). To create a somewhat continuous interval-level education variable, these two items were recoded to form years of schooling in the following :

CEDPD04	Condition	
00	CEDPQ1=1	(no schooling)
03	CEDPQ1=2	(1 to 5 years)
06	CEDPQ1=3	(6 years)
07	CEDPQ1=4	(7 years)
08	CEDPQ1=5	(8 years)
09	CEDPQ1=6	(9 years)
10	CEDPQ1=7	(10 years)
11	CEDPQ1=8	(11 years)
12	CEDPQ1=9	(12 years)
13	CEDPQ1=10	(13 years)
16	CEDPQ4=6	(BA/BSC)
18	CEDPQ4=7	(Masters)
20	CEDPQ4=8 or 9	(MD/PHD)

An extra year was then added to CEDPD04 if the PMK had a diploma from a trade school or community college (i.e., if CEDPQ4= 4 or 5 then CEDPD04 = CEDPD04+1).

The same procedure was used to set up a years of schooling variable for the spouse/partner (CEDSD04).

Occupational Prestige

Occupational status is an important indicator of SES. The occupation variable used in the derivation of SES was a modified version of a scale developed by Pineo, Porter and McRoberts (1977). The classification system groups occupations described in Statistics Canada's 1980 Standard Occupational Classification into 16 somewhat homogeneous categories, ordered from 1 to 16, where code 1 represents the highest level of occupation and code 16 the lowest. The 16-category scale provides a ranking of occupations according to their social standing or prestige. For the NLSCY, for both the PMK and the spouse/partner, a detailed description was taken of the job considered to be his or her main job during the previous 12 months. The information was used to code occupations into the 1980 classification, and in turn into the 16 prestige categories. For the purposes of deriving SES, the order of the Pineo-Porter-McRoberts scale was reversed. The final scale used in the derivation of SES had the following values:

01 Farm labourer
02 Unskilled manual
03 Unskilled clerical/sales/service
04 Semi-skilled manual

05 Semi-skilled clerical/sales
06 Farmer
07 Skilled crafts and trade
08 Skilled clerical/sales/service
09 Foreman/forewoman
10 Supervisor
11 Middle manager
12 Technician
13 Semi-professional
14 High-level management
15 Employed professional
16 Self-employed professional
96 Not-applicable - this was assigned for the spouse/partner for cases where the PMK did not have a spouse/partner or where the PMK and the spouse/partner did not work during the past 12 months.
99 Not stated

This ordinal scale can be used to rank individuals into the various occupation groups but one cannot assume that the intervals between ranks are equal interval. For example, in this scale a middle manager (code 11) is ranked higher than a supervisor (code 10), which in turn ranked higher than a foreman (code 09). However, this does not imply that the difference in occupation between the middle manager and a supervisor is equivalent to the difference between a supervisor and a foreman. By assuming that the underlying latent construct has a particular distribution, one can assign intervals to the various categories. Mosteller and Tukey (1977) propose a logit transformation to re-express ordinal data on an interval scale. To do this, the percentage of individuals in each occupation group is considered a piece of the logistic distribution. The code assigned to each occupation is the centre of its piece in the logistic distribution. This transformation was employed to scale the 16 occupations.

For each occupation group x, the following values were computed in the provinces:

p =the percentage of individuals with an occupation less than occupation x (based on the Pineo-Porter-McRoberts category)

pp =the percentage of individuals with an occupation less than or equal to occupation x (based on the Pineo-Porter-McRoberts category)

phi(p) $=p*\ln(p) + (1-p)*\ln(1-p)$

phi(pp) $=pp*\ln(pp) + (1-pp)*\ln(1-pp)$

The recoded (logit) value for occupation x was assigned to be:

$$PINEOLOG=\frac{phi(pp) - phi(p)}{pp-p}$$

The PINEOLOG calculated in the provinces was associated to the persons in each of the 16 occupation groups (for both the PMK and spouse/partner). This was done to allow comparison between the territories and the provinces. (If a PINEOLOG score had been recalculated in the north, this would have restricted the comparison of the distribution of the SES within the territories only). This PINEOLOG was then used in the derivation of SES.

Household Income

The last variable used in the derivation of SES was household income. In the north, a scale was used to obtain household income. More detail regarding the collection of household income and data quality issues can be found in Section 9.17. To derive SES, the mid-point of the income class was used. The value of the highest mid-income class was set to $100,000.

Final Derivation of SES

Thus the five variables that were used to derive SES were:
- CEDPD04 (years of schooling for the PMK),
- CEDSD04 (years of schooling for the spouse/partner),
-PINEOLOG-PMK (the pineo occupation code for the PMK transformed to the logit distribution),
-PINEOLOG-SP (the pineo occupation code for the spouse/partner transformed to the logit distribution) and
-HHINC (household income in thousands of dollars)

Each of the five variables were standardized to have a mean of zero and a standard deviation of one. The means and the standard deviations obtained from the data in the provinces were used and applied to the results in the territories. This was done to allow for comparisons of the SES in the territories and the SES produced in the provinces. Missing values (i.e., not-stated values) were ignored in the standardization.

In the standardization of the spouse/partner variables (CEDSD04 and PINEOLOG-SP), if the PMK did not have a spouse/partner these records were ignored. The SES composite was then calculated by taking the (unweighted) average of the five standardized variables. If one of the five variables had missing data due to non-response (refusal, don't know, etc.) then the average was taken over the remaining non-missing items. If there was no spouse/partner in the household (i.e., the PMK had no spouse/partner) then the average was taken over the three applicable variables (CEDPD04, PINEOLOG-PMK, and). For two-parent families (i.e., for cases where there was a PMK and a spouse/partner), if two or more out of the five input variables were missing, then SES was set to "not-stated" . For single-parent families (i.e., there was no spouse/partner), if one or more out of the three input variables were missing, then SES was set to "not-stated".

The cycle 3 cross-sectional SES on this release is CINHD09.

Examples of SES

On the Cycle 3 master file, the value for SES ranges from –5.050 to +2.947.

A note of caution should be used when using the SES, especially in the context of the north. SES is a concept implying the average of five components; education of both parents, occupational prestige scale for the occupation and family income. Some of these concepts may not be relevant to the context of the north.

In the provinces, in order to give a flavour for the types of families associated with various SES scores the following examples are given for illustration purposes. It should be noted that the SES scores given in these examples are approximate and do not correspond to actual records on the NLSCY file. Many more examples are possible for each score involving both one and two parent families.

SES SCORE	EXAMPLE
1.5	A family in which: both the PMK and spouse have a university degree (BA/BSC) they are both employed professionals the household income is approximately $77,000

0.5	A family in which: the PMK has a university degree (BA/BSC) and the spouse has grade 13 the PMK is employed as a semi-professional and the spouse is employed in a semi-skilled clerical position household income is approximately $57,000
0.0	A family in which: the PMK has grade 13 and the spouse grade 12 the spouse is employed as a semi-professional position and the PMK is not in the labour force household income is approximately $25,000
-0.5	A family in which: the PMK and spouse have both completed grade12 the PMK is employed in a semi-skilled clerical position and the spouse in a semi-skilled manual position household income is approximately $16,000
-1.0	A family in which: neither the PMK nor the spouse have completed high school the PMK is employed in an unskilled clerical position and the spouse is employed in an unskilled manual position household income is approximately $20,000
-1.5	A family in which: neither the PMK nor the spouse have completed high school neither the PMK nor the spouse are in the labour force household income is approximately $12,000
-2.0	A family in which: there is no spouse the PMK has not completed high school the PMK is not in the labour force the household income is less than $10,000

8.6 Geographical Indicators

The NLSCY sample was allocated so that territorial analyses will be possible for broad age groupings of children. A variable to indicate territory of residence is available on the master file (CGEHD03).

Census metropolitan area (CMA) is available on the NLSCY master file (CGEHD02) as well as an indicator of urban/rural class size (CGEHD01). It should be noted that Yukon respondents were assigned a “not stated” code of 99 for this variable because the method used to derive it did not correctly identify the size of urban area.

9.0
Content and Validation of NLSCY Data

The NLSCY was designed to follow an ecological or holistic approach to measuring child development. The survey captures the diversity and dynamics of the factors affecting children. To ensure that all relevant topic areas affecting child development were adequately addressed by the survey, a multidisciplinary consultation was carried out at the inception of the survey. The selection of specific subject areas, priorities and survey questions was very much a group effort with input and advice from:

- the NLSCY expert advisory group which consists of researchers in the area of child development and the social sciences;
- federal departments;
- representatives from the provinces and territories responsible for child development programs.

It was recommended that the NLSCY cover a broad range of characteristics and factors affecting child growth and development. Extensive information was gathered about the child, as well as information on the child's parent(s), characteristics of the family and the neighbourhood. This section provides an outline of the content for each section of the questionnaire included in this release of NLSCY data.

As part of the NLSCY processing system, there were some basic quality checks performed for each section of the questionnaire. Any items for which there was a high level of non-response or which were frequently involved in edit failures were looked at in detail. Where appropriate, comparisons were made to external data sources and analyses were carried out to investigate possible reasons for differences from these other sources. Any concerns about potential data quality problems for any items in a particular section of the questionnaire are discussed in this section of the documentation.

Before the section by section discussion of content and validation results, the general validation procedures used for the "scale" data are presented.

9.1 Validation of Scale Data

For some of the concepts that were deemed to be important to measure in the 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 measure a certain concept when the answers to the items are put together.

For example, on the child's questionnaire it was determined that it was important to have an assessment of certain parenting behaviours. The Parenting Scale that was employed was one that was proposed by Dr. M. Boyle at Chedoke-McMaster Hospital, based on work by Dr. Ken Dodge (Vanderbilt University) which was an adaptation of Strayhorn and Weidman's Parent Practices Scale. The scale is intended to measure three different constructs or factors related to parenting; positive interaction, hostile/ineffective parenting and consistent parenting.

For each factor measured by a scale, a score is calculated. The score for a particular factor can be used to give an ordering of individuals. For example, for the Parenting Scale, for children with higher scores for the "positive interaction" factor, the PMK reported having more positive encounters with the child (e.g., laughed with them more, praised them more etc.). The score for a particular factor is usually based on a series of items, since one single item usually cannot measure the factor or construct with adequate precision.

During the development of the NLSCY, when consideration was being made of what specific scales should be used to measure a particular concept, as much as possible, scales were selected that had been used in other studies where the psychometric properties of the measures produced by the scale were available with complete references.

However in many instances the wording of certain questions was modified and in some cases new questions were added. Sometimes the scale that was used had not previously been used for children in Canada or had only been used for very small samples. Given these concerns and further concerns regarding interviewing conditions, it was felt that the factorial structures of the scales used in the NLSCY could be different from the ones given in the literature. Therefore the project team felt the need to carry out an extensive evaluation of the scale data to ensure that the psychometric properties found to exist in other studies were also true for the NLSCY experience.

There were three major steps in the analyses of the scale data for the provinces survey. First a new factor analysis was performed on all scales to determine the constructs or factors inherent in each scale. Then scale scores were calculated based on this factor structure. Finally reliability measures were produced. The general procedures that were followed for each of these steps are described in detail on the following pages.

In the Territories, the first of these three steps, that is, the factor analysis, was not performed. Instead, the same factor structures used in the provinces were used to calculate the scale scores in the territories and reliability measures were produced. The idea was to produce scale scores for the territories that would be comparable to the provinces.

The specific details for each scale are discussed later in this section in the appropriate sub-section.

9.1.1 Factor Analyses

The following is a summary of the procedures used in the factor analysis for each scale for the provincial NLSCY.

1 The sample of respondents for each scale (and age group, if the scale used different questions for different groups), was randomly divided into two half-samples. This was done to find out whether different samples would yield the same results.

2 Principal component analysis was carried out separately on each half-sample to find out how many factors should be extracted in the factor analysis performed subsequently. In principle, the same number of factors as was found in the literature was expected. In practice, however, some scales showed a different number of factors because in some cases factors combined while in others new factors emerged.

3 Factor analysis was also done on each half-sample and the factorial structure and loadings of each factor were compared across the half-samples.

Factor analysis requires that the data have the property of interval or ratio data, that is the distance between each answer category of the question should be the same. For example, in scales where the answer choices are: Never, Sometimes, Often, and Always, one must assume that the distance between Never and Sometimes is the same as that between Sometimes and Always in the respondent's perception. It was felt that this was not necessarily the case for the scales used in the NLSCY.

Therefore before performing the factor analysis for each of the NLSCY scales, the data were transformed using optimal scaling. The method used was one proposed by Young and several associates (Young, 1981) which is a variant of Fisher's optimal scaling technique. The method is presented as a means of transforming data which are fundamentally nominal or ordinal in nature to interval or ratio level data so that statistical techniques which are appropriately applied only to interval and ratio data may be utilized.

Initially the factor analysis for each scale to be included in this release of NLSCY data was carried out using unweighted data. At that point in time the final weights had not yet been calculated. Once the weights were available, work started on repeating the factor analyses using the weighted data. With the

weights, the same factor structure was not always observed. At this point in time not all of the scale data included in this release have been re-analysed using the weights. For the PMK Depression Scale, weights are not available since the unit of analyses for these scales are at the PMK and household level.

The factor structure of each scale was determined based on data from the first cycle. The factor structure imposed on the scales already used in the first cycle and repeatedly used in the second cycle of the survey was the result of analyses done based on data from the first cycle.

9.1.2 Calculation of Scores

The results of the factor analyses were used to determine what items "loaded" into each factor (i.e., were a part of each factor). The next step was to calculate a score for each factor. This was done by summing the values for each individual item that made up the factor. The following example illustrates how factor scores were computed.

Example:

One of the constructs that emerged in the factor analysis for the Parenting Scale on the Child's Questionnaire was the positive interaction factor. In the factor analysis four items were found to load into this factor.

> CPRCQ1 How often do you praise your child by saying something like “Good for you!” or “What a nice thing you did!” or “That’s good going!”
> CPRCQ3 How often do you and he/she laugh together?
> CPRCQ6 How often do you do something special with him/her that he/she enjoys?
> CPRCbQ7 How often do you play sports, hobbies, or games with him/her? (Or CPRCbQ7a, “How often do you play games with him/her?”, depending on child’s age).

The answer categories for these items was as follows:

1 - never
2 - about once a week or less
3 - a few times a week
4 - one or two times a day
5 - many times each day

In the calculation of the score for this positive interaction parenting factor, the categories were rescaled to 0 to 4 (i.e., the category "never" was scored as 0, the category "about once a week or less/less than half the time" was scored as 1, and the category "many times each day/all the time" was scored as 4). In order to compute the score these values were summed across the four items involved in the factor resulting in positive/interaction parenting score in the range 0 to 16. A score of 16 represents the absence of a problem and a score of 0 is the highest possible score with respect to problems. For most of the scores calculated for the NLSCY, a score of 0 represents the absence of a problem. However there are exceptions to this (like here) which are noted in the documentation for each particular scale.

The score for the positive interaction parenting factor is labelled as CPRCS01 (for 0 to 23 months) and CPRCS03 (for 2-11 years) on the record layout for the master file. An "S" in the 5th position of the variable name indicates a score.

9.2 Demographic Variables

The demographic variables discussed in this section refer to variables collected on the household roster. As part of the household roster some basic demographic information (e.g., age, gender, marital status) was collected for all members of the child's household. The relationship grid was also completed as part

of this questionnaire i.e., the relationship of everyone in the household to the PMK. Using this information it was possible to create an extensive set of variables to describe the child's family situation. Most of these derived variables are critical to the analyses of NLSCY data and are described in Section 8.

It was necessary to perform an extensive series of edits on the data that were collected.

The following are some examples of the types of editing that were carried out.

- a birth parent should be at least 12 years older (and not more than 55 years older) than a birth child
- the difference in age between a husband and wife should be less than 29 years.

9.3 Medical/Biological

The Medical /Biological Section was completed for children in the 0 to 3 age group. The major objective was to collect information on factors such as gestational age and birth weight. These factors have been shown to have a direct impact on a child's growth and development. For example, in the long term, underweight babies face higher risks of poor health as well as longer-lasting developmental difficulties.

For each child under two, the nature of the delivery, general health of the child at birth and the use of specialized services following the birth were collected in this section. The NLSCY also investigated the biological mother's pregnancy and delivery history, including policy-relevant topics such as the mother's breast-feeding experiences and prenatal lifestyle.

Since birth weight is such an important variable, caution was taken in editing this variable. The records for children with very low birth weights (< 1.5 kilograms) were examined to verify that the response was legitimate. Other variables considered in the edit were the length of the baby at birth, the number of days early of the delivery, the conditions of the delivery (e.g., multiple birth and special medical care) and the health of the child at birth. If there was nothing to collaborate the low birth weight it was set to not-stated.

There were a couple of derived variables created for this section that bear note. Two variables were derived to indicate the gestational age of the child. CMDCD06 gives the gestational age in days and CMDCD07 indicates if the child was born prematurely (gestational age 258 days or less), in the normal range (gestational age 259 to 293 days) or late (gestational age 294 days or later).

A variable was derived (CMDCD08) to indicate if the child was of normal birth weight (2500 grams), moderately low birth weight (1500 to 2499 grams) or very low birth weight (< 1500 grams).

9.4 Education (Child)

The objective of this section was to get some basic information about the child's educational experiences. The amount and type of information collected varied depending upon the age of the child, with more information being collected for the older children who have had greater school experience.

Basic information was collected for all age groups, such as: the child's grade level, type of school and language of instruction, whether the child looks forward to school, absenteeism, number of school changes and residential moves.

For children in grade 1 or higher, additional questions were asked concerning other aspects such as skipping and repeating grades, achievement and special education.

9.5 Behaviour Scale

The objective of the behaviour scale is to assess aspects of the behaviour of children aged 4 to 15 years. The following behaviours of children were measured.

hyperactivity,
emotional disorder,
anxiety,
physical aggression,
inattention,
prosocial behaviour,
conduct disorder

Below is a description of the items that were included on the questionnaire to measure behaviours, the analysis used to construct the scales and the results of these analyses

Questionnaire Items:

Children aged 4 to 15:

- Hyperactivity
  Items include CBECQ6B, Q6I, Q6N, Q6P, Q6S and Q6W from the Ontario Child Health Study (OCHS), and CBECQ6HH from the Montreal Longitudinal Survey.

- Inattention
  For 4 to 7 year olds, items include CBECQ6P from the OCHS and CBECQ6EE and Q6KK and Q6QQ from the Montreal Longitudinal Survey. For 8 to 15 year olds, only item CBECQ6P was used.

- Emotional disorder
  Items include CBECQ6F, Q6K, Q6Q, Q6V, Q6CC, Q6MM and Q6RR. These items were provided by the Ontario Child Health Study (OCHS).

- Anxiety
  Items included some emotional disorder items from the Ontario Child Health Study: CBECQ6F, Q6Q, Q6V and Q6CC.

- Prosocial behaviour
  Items include CBECQ6D, Q6U, Q6BB ,Q6SS and Q6UU were used to measure prosocial behaviour. These items are from the Montreal Longitudinal Survey;.

Analysis of the NLSCY Data

To construct the behaviour scale for the NLSCY data, a factor analysis was conducted to test the theoretical construct. In order to be consistent with the behaviour scale created from the parent questionnaire from the provincial survey, the factor structure which emerged from the behaviour scale was imposed on the territories behaviour scale.

Once the factor structures were analysed and the items included in each factor were determined, the scores were calculated. No imputation for missing values was done. To produce the final scores, 1 was subtracted from each item so that the lowest score would be 0. The score for each factor on the scale

was arrived at by totalling the values of the items that made up the factor. The score was set to “missing” if any of the values of any items included in the factor were unreported. For example, a value may be missing if a parent refused to answer the item. A score of 0 indicates that the child has no problems for all factors in the behaviour scale with the exception of the prosocial factor, where a score of 0 indicates the absence of prosocial behaviour.

Results

As a result of imposed factor analysis, four factors from the provincial survey were identified for the territories for this age group:

hyperactivity-inattention (CBECS06),
prosocial behaviour (CBECS07),
emotional disorder-anxiety (CBECS08), and
property offence (CBECS11) (8 to 15 year olds only).

The items that comprised each factor are listed in the table below.

BEHAVIOUR SCALE FOR 4 TO 15 YEAR-OLDS

FACTOR	SCORE	ITEMS
Hyperactivity - inattention [1]	CBECS06	CBECQ6B, 6I, 6N, 6P, 6S, 6HH and 6W
Prosocial behaviour [2]	CBECS07	CBECQ6D, 6U, 6BB, 6SS and 6UU
Emotional disorder [3]	CBECS08	CBECQ6F, 6K, 6Q, 6V, 6CC, 6MM, 6RR
Property offence (8 to 15 years only)	CBECS11	CBECQ6C, 6E, 6L, 6T, 6DD, 6PP

[1] One additional item is included on the hyperactivity/inattention factor for the provinces that was not asked in the territories.
[2] 10 items were used to construct the prosocial factor for the provinces. Only 5 of those items were asked in the territories and these 5 items were used to construct the factor
[3] One additional item is included on the emotional/anxiety factor for the provinces which was not asked in the territories.

9.6 Relationships

The Relationships Section of the Child's Questionnaire was completed for all children aged 4 to 9 years. The objective was to provide information about the child's relationships with others. Positive relationships with other children and adults may help to counteract other factors that place a child at risk.

The section collects information about how the child gets along with parents, brothers and/or sisters and friends with some variation by age of the child.

The questions on number of days spent doing things with friends, number of friends, and getting along with friends, parents, teachers and siblings (CRLCQ1, Q2, Q9) are based on those in the Ontario Child Health Study.

9.7 Parenting Scale

The objective of the parenting scale is to measure certain parental behaviours. Specifically, one scale was used: to measure positive interaction (0 to 15 year olds), and consistent parenting (2 to 15 year olds).

Questionnaire Items:

Questions CPRCQ1 to CPRCQ16 on positive interaction and consistent parenting were provided by Dr. M. Boyle at Chedoke-McMaster Hospital, based on Dr. Ken Dodge's work (Vanderbilt University) and an adaptation of Strayhorn and Weidman's Parent Practices Scale.

Results

In order to be consistent with the parenting scales created for the provincial survey, the factor structure which emerged from the analysis on the provincial survey was imposed on the territories' parenting scales. All items asked on the provincial survey, and used to construct the scales, were not asked in the territories. Thus, only those questions that were asked in the territories were used in constructing the scale.

For children 0 to 23 months of age:

One factor was constructed for those aged 0 to 23 months: CPRCS01, which measured positive interaction, and used items CPRCQ1, Q3, Q6 and bQ7. In the provincial survey, one additional item (which was not asked in the territories) was used to construct this scale. No imputation took place. To produce this score, 1 was subtracted from each item so that the lowest possible score would be 0. The final score ranges from 0 to 16, with a low score indicating the lack of positive interactions.

PARENTING SCALE FOR CHILDREN 0 to 23 MONTHS OLD

FACTOR	SCORE	ITEMS
Positive interaction	CPRCS01	CPRCQ1, 3, 6, 7

For children 2 to 11 years of age:

Two factors were constructed for those aged 2 to 11 years. The first was CPRCS03, which measured positive interaction and used items CPRCQ1, Q3, Q6 and Q7/7a. In the provincial survey, one additional item (which was not asked in the territories) was used to construct this scale. The second was CPRCS05, which measured consistent parenting and used items CPRCQ10, Q11, Q12 and Q16. In the provincial survey, one additional item (which was not asked in the territories) was used to construct this scale.

No imputation took place in constructing either score. To produce both scores, 1 was subtracted from each item so that the lowest possible score would be 0. The final score for CPRCS03 (positive interaction) ranges from 0 to 16, with a low score indicating a lack of positive interactions. The final score for CPRCS05 (consistent parenting) ranges from 0 to 16, with a low score indicating inconsistent parenting.

PARENTING SCALE FOR 2 to 11 YEAR-OLDS

FACTOR	SCORE	ITEMS
Positive interaction	CPRCS03	CPRCQ1, 3, 6, 7/ 7a
Consistency	CPRCS05	CPRCQ10, 11, 12* and 16*

*These Items were reversed when calculating the score.

9.8 Socio-demographic Characteristics

The objective of the Socio-demographic Section was to gather information on immigration, ethnic background and the language profile of household members. This will allow for analysis of various components of the Canadian population and will permit identification of visible minorities.

As well, there were questions on religious affiliation and frequency of attendance at religious services. Religion, particularly frequency of attendance, is acknowledged as having a positive influence on a child's development.

9.9 Depression Scale

PMK Depression scale:

A Depression scale was administered to the PMK as part of the Parent Questionnaire. Questions for this scale (CDPPQ12A to CDPPQ12L) are a shorter version of the depression rating scale (CES-D), comprising 20 questions, developed by L. S. Radloff of the Epidemiology Study Centre of the National Institute of Mental Health in the United States. This rating scale is used to measure the frequency of symptoms in the public at large. The occurrence and severity of symptoms associated with depression during the previous week are measured. The rating scale was reduced to 12 questions by Dr. M. Boyle of the Chedoke-McMaster Hospital of McMaster University.

The depression rating scale includes twelve questions, each of which contains four response categories. In order that the lowest score value be 0, the value for each question was reduced by 1 in calculating the score. As well, the answer categories were reversed for questions having a negative loading (CDPPQ12F, Q12H, and Q12J). The total score (CDPPS01) may therefore vary between 0 and 36, a high score indicating the presence of depression symptoms.

The sample size for the territories was 1,421 PMKs. Only one factor, using all 12 variables, was created. There was no imputation for the variables in this rating scale.

9.10 Education (Parent)

The Education Section was completed for both the PMK and spouse/partner. The objective was to gather information on the years of school completed, educational attainment, and current attendance at an educational institution.

Research (for example, the Ontario Child Health Study and the National Longitudinal Survey of Youth in the United States) has indicated a link between maternal educational attainment, the home environment and child development. The questions on full-time and part-time school attendance provide an indicator of the main activities of the PMK and the spouse/partner.

9.11 Labour Force

Employment stability impacts the home environment, both in terms of income and stress levels. Research, conducted for the Ontario Child Health Study, indicates that parental unemployment can adversely impact child mental health.

The Labour Force Section was completed for both the PMK and spouse/partner. The main objective of the section was to determine employment stability as an indicator of the continuity of employment income. Questions include, periods of absence from work, reason for the most recent absence, hours worked, and

work arrangements (e.g. shifts) during the previous year. Information was collected on up to three jobs for a one-year period. A series of questions were asked about each of the jobs the PMK and spouse/partner held during the previous year to a maximum of three jobs. For each of these jobs, questions were asked to determine when the job started and ended. As well, in order to address absences within a job the following question was asked as the initial lead-in question to a job:

Did you have that job one year ago, without a break in employment since then?

The intent was that if there had been a break in employment the respondent was supposed to report this situation as two jobs. The start date for the first job would be when the respondent first started working at the job and the end date would be when the break occurred. The second job would have the point at which the respondent returned to the job as the start date.

Respondents were asked to identify what they considered to be their main job over the previous year (if they had more than one job). A complete description was recorded for this main job and industry and occupation coding was carried out (using 1980 Standard Industrial Classification codes and 1980 Standard Occupational Classification codes).

9.11.1 Labour Force Derived Variables

At this point in time, for the Cycle 3 release, 10 labour force derived variables have been created for the PMK and spouse/partner of the PMK. They include:

CLFPD05/CLFSD05: SIC for PMK's /Spouse's main job
CLFPD06/CLFSD06: SOC for PMK's /Spouse's main job
CLFPD07/CLFSD07: Standard industry code for main job – grouped
CLFPD08/CLFSD08: Standard occupation code for main job – grouped
CLFPD09/CLFSD09: Pineo socio-economic classification of main job

9.12 Income

In the Income section of the survey, the sources of income and the income range were collected for each household. The income range was also collected for the PMK income and for the income of the PMK's spouse. This information provides an indicator of the family's economic situation, an essential component of the child's environment.

A cascade question was used to record income ranges. The respondents who did not wish to choose from the ten detailed income classes were able to pick a response from two broad classes or four intermediate groupings. Most respondents provided information by the detailed income ranges.

9.12.1 Income Edit and Imputation

Logical Consistency Edits

Several logical consistency edits were applied to ensure agreement among the income responses. While the use of income ranges reduced the precision of these relationships between income variables, a number of records failed the rules.

The household must have a non-zero income. For records failing this rule, household income was set to "not stated" for imputation.

In a lone PMK household, the PMK income and the household income should be equal unless there is another household member with an income. If the household income was greater than the PMK income

and the household did not have another member aged 15 or over, then the PMK income was imputed to be equal to the household income. In particular, this rule was applied to several cases where the PMK income was reported to be zero.

In a lone PMK household, the PMK income cannot be greater than the household income. When this error condition occurred, the household income was imputed to be equal to the PMK income.

In a two-parent household, the sum of the PMK and spouse incomes cannot be greater than the household income. When this error condition occurred, the household income was imputed using the hot-deck method.

In a two-parent household, the sum of the PMK and spouse income should be equal to the household income unless there is another household member with an income. When the household did not have another member aged 15 or over, then the household income was imputed using the hot-deck method.

If one of the personal incomes is equal to the highest income range, i.e. over $80,000, then the household income must also be over $80,000. When this error condition occurred, the household income was imputed to be over $80,000.

In applying the above rules, it was necessary to use the minimum and maximum values of the income ranges. The rules may be expressed as follows:

If no spouse, then the record fails edit if

min(PMK income)> max(household income)

OR

max(PMK income)< min(household income) and no other members 15 or over.

If there is a spouse, the record fails edit if

min(PMK income)+min(spouse income) > max(household income)

OR

max(PMK income)+max(spouse income) < min(household income) and no other members 15 or over.
The reason for using the hot-deck method to correct most of these errors is that, in many cases, more than one income range could make the record valid. The hot-deck procedure selects imputed values from the distribution of valid values that appear in valid records.

Imputation of Grouped Response Codes

A hot-deck imputation was done for each group code to assign one of the corresponding detailed range codes. For example, a response code=1 (<$20,000) was imputed to one of four detailed codes (3, 4, 5 or 6) and a response code=2 (<$10,000) was imputed to one of two detailed codes (3 or 4).

Each record requiring an imputation was matched to a set of donor records, each of which had a combination of valid responses to the income variables. The procedure was carried out separately for lone PMK and two-parent households and donors were restricted to those whose detailed income code matched the reported group code.

The following variables in order were also matched:
other income variables
other household members aged 15 or over

First Nations status (PMK or spouse or both)
working for pay or profit in last 12 months (PMK and spouse)
main source of income
Urban/rural size (large/small)
Pineo codes (based on occupations of PMK and spouse)

Matches on 1 and 2 were mandatory. Matching on the other variables was relaxed as necessary to find matching donor records. The use of an individual donor record was limited to 4 imputations in order to avoid imputing too many records from a single record.

In cases, where there were no donors, the record was reviewed and a manual imputation done.

Imputation for Missing Values

Again, a hot-deck imputation was performed in a manner similar to that described above. The procedure was carried out separately for lone PMK and two-parent households.

The following variables in order were matched:
-other income variables (if any present)
-other household members aged 15 or over
-working for pay or profit in last 12 months (PMK and spouse)
-First Nations status (PMK or spouse or both)
-main source of income
-Urban/rural size (large/small)
-Pineo code (based on occupations of PMK and spouse)

Matches on 1, 2 and 3 were mandatory.

10.0 Data Quality

The estimates derived from this survey are based on a sample of children. Somewhat different figures might have been obtained if a complete census had been taken using the same questionnaire, interviewers, supervisors, processing methods, etc. as those actually used. The difference between the estimates obtained from the sample and the results from a complete count taken under similar conditions is called the sampling error of the estimate. Sample estimations in the North are particularly vulnerable due to the transient nature of the population which is more noticeable there than anywhere else in Canada. Put in a longitudinal context, this can cause serious coverage issues since children who no longer reside in the North are no longer part of the sample for the North. Moreover, the cross-sectional sample is also at risk for coverage inadequacies since a significant portion of the sample should be devoted to the larger than usual influx of migrants to the North from one cycle period to the next.

Errors that are not related to sampling may occur at almost every phase of a survey operation. Interviewers may misunderstand instructions, respondents may make errors in answering questions, the answers may be incorrectly entered on the questionnaires and errors may be introduced in the processing and tabulation of the data. These are examples of non-sampling errors.

In this section some of the non-sampling errors that occurred in the survey are discussed.

10.1 Overall Response Rates

Due to the fact that the non-probability sampling was carried out in the Territories, the total number of households sampled is not known. Without knowing the total number of households the overall response rates cannot be calculated.

Instead response rates have been calculated for the longitudinal children in Cycle 1 (all those who were part of the survey in Cycle 1), based on the number of responding households from Cycle 1 minus the 170 households that were dropped from the sample in the Yukon. The overall response rate is 71%, which is lower than the rate for the Territories in Cycle 2.

Territory	Cycle 1 Households	Responding Longitudinal Households in Cycle 3	Response Rate
Yukon	487	328	67.4%
Western	399	317	79.4%
Nunavut	331	215	65.0%
Total	1,217	860	70.7%

There are many reasons why some households did not respond to the survey. In some cases the interviewer was unable to make contact with a selected household for the entire collection period. In other cases the household refused to participate in the survey.

It is worth noting that because of the small populations in the Territories, the probability of being selected in any sample survey is much higher than in the provinces. Because of the integration of the NPHS and

the NLSCY in the Territories, a significant number of households were asked to participate in both surveys. This may have caused burden on the respondents and could have adversely affected the participation rate among our longitudinal respondents.

10.2 Child Questionnaire Response Rates

Within responding households, the information on some children was incomplete. Where key questions were not completed for a particular child, the child record was treated as a non-response and not used in compiling the statistics. In Cycle 3 there were 11 such cases in the Northwest Territories and 2 in the Yukon representing 0.8% and 0.2% respectively of the children who responded to the survey.

11.0 Guidelines for Tabulation, Analysis and Release

This section of the documentation outlines the guidelines to be adhered to by users tabulating, analysing, publishing or otherwise releasing any data derived from the survey master file. With the aid of these guidelines, users of the master file should be able to produce the same figures as those produced by Statistics Canada and, at the same time, will be able to develop currently unpublished figures in a manner consistent with these established guidelines.

11.1 Rounding Guidelines

In order that estimates for publication or other release derived from the NLSCY master file correspond to those produced by Statistics Canada, users are urged to adhere to the following guidelines regarding the rounding of such estimates:

a) Estimates in the main body of a statistical table are to be rounded to the nearest hundred units using the normal rounding technique. In normal rounding, if the first or only digit to be dropped is 0 to 4, the last digit to be retained is not changed. If the first or only digit to be dropped is 5 to 9, the last digit to be retained is raised by one. For example, in normal rounding to the nearest 100, if the last two digits are between 00 and 49, they are changed to 00 and the preceding digit (the hundreds digit) is left unchanged. If the last digits are between 50 and 99 they are changed to 00 and the preceding digit is incremented by 1.

b) Marginal sub-totals and totals in statistical tables are to be derived from their corresponding unrounded components and then are to be rounded themselves to the nearest 100 units using normal rounding.

c) Averages, proportions, rates and percentages are to be computed from unrounded components (i.e. numerators and/or denominators) and then are to be rounded themselves to one decimal using normal rounding.

d) Sums and differences of aggregates (or ratios) are to be derived from their corresponding unrounded components and then are to be rounded themselves to the nearest 100 units (or the nearest one decimal) using normal rounding.

e) In instances where, due to technical or other limitations, a rounding technique other than normal rounding is used resulting in estimates to be published or otherwise released which differ from corresponding estimates published by Statistics Canada, users are urged to note the reason for such differences in the publication or release document(s).

f) Under no circumstances are unrounded estimates to be published or otherwise released by users. Unrounded estimates imply greater precision than actually exists.

11.2 Sample Weighting Guidelines for Tabulation

The sample design used for the NLSCY was not self-weighting. When producing simple estimates, including the production of ordinary statistical tables, users must apply the proper sampling weights. (CWTCW01C – cross sectional weight and/or CWTCW01L – longitudinal weight). For the longitudinal children, the longitudinal weight inflates the estimates produced by 1,444 respondents to the total population of children aged 0-11 in the two territories in 1994 (23,508). The cross-sectional weight inflates the estimates produced by 2,355 respondents to the total population of children aged 0-15 in the two territories in 1998 (30,171).

If proper weights are not used, the estimates derived from the master file cannot be considered to be representative of the survey population, and will not correspond to those produced by Statistics Canada. In effect, the weights assigned to each child reflect the number of children represented by a particular respondent. For any analysis dealing with correlation analysis or any other statistics where a significance measure is required, it is recommended that an “analytical” weight be used, which is the original weights rescaled to produce sample counts in lieu of population counts. This weight is obtained by multiplying the population weight (CWTCW01C and/or CWTCW01L) by the number of respondents and dividing this total by the total population that we are estimating for. This produces a mean weight of 1 and a sum of weights equal to the number of respondents.

For example if we were estimating for the two territories for Cycle 3, the number of respondents would be 2,355 and the total population would be 30,171 so the sum of the sample weights would be 2,355. The benefit of this adjusted weight is that an over estimation of the significance (which is very sensitive to sample) is avoided while maintaining the same distributions as those obtained when using the population weight. The disadvantage is that the numerator is not weighted up to the target population and the Approximate Coefficient of Variance Tables described in section 12 and presented in Appendix 1 are no longer useful as a measure of data quality.

We need to point out that this re-scaling of weights for sub-domains resulting from the exclusion of units with partial non-response will not be adjusted for that missing data. This weight re-scaling does not re-distribute the demographic load of the units excluded for having missing data unless you assume that it is truly random.

Users should also note that some software packages may not allow the generation of estimates that exactly match those available from Statistics Canada, because of their treatment of the weight field.

11.2.1 Definitions of Types of Estimates: Categorical vs. Quantitative

It should be pointed out that the NLSCY file has been set up so that the child is the unit of analysis. The weights that can be found on each record (CWTCW01C/CWTCW01L) are "child" weights. Estimates of parents or families cannot be made from the NLSCY master file. A further discussion of units of analyses can be found in Section 8.1 of this document.

Before discussing how the NLSCY data can be tabulated and analysed, it is useful to describe the two main types of point estimates of population characteristics which can be generated from the master file for the NLSCY.

Categorical Estimates

Categorical estimates are estimates of the number, or percentage of the surveyed population possessing certain characteristics or falling into some defined category. The number of children who were born before the due date or the proportion of children who were in excellent health at birth are examples of such estimates. An estimate of the number of persons possessing a certain characteristic may also be referred to as an estimate of an aggregate.

Examples of Categorical Questions:

Q: Was (the child) born before, after or on the due date?

R: Before
After

On due date

Q: Compared to other babies in general, would you say the (the child's) health at birth was:

R: Excellent
Very good
Good
Fair
Poor

Quantitative Estimates

Quantitative estimates are estimates of totals or of means, medians and other measures of central tendency of quantities based upon some or all of the members of the surveyed population. They also specifically involve estimates of the form X mean over Y mean where X mean is an estimate of the surveyed population quantity total and Y mean is an estimate of the number of persons in the surveyed population contributing to that total quantity.

An example of a quantitative estimate is the average number of days of care received by babies who required special medical care following birth. The numerator is an estimate of the total number of days for which babies required special care. The denominator is the number of babies who required special care at birth.

Examples of Quantitative Questions:

Q: For how many days, in total, was this care received?

R: |_|_|_| Days

Q: What was the child's weight at birth in pounds and ounces?

R: |_|_| Pounds |_|_| Ounces

11.2.2 Tabulation of Categorical Estimates

Estimates of the number of children with a certain characteristic can be obtained from the master file by summing the final weights of all records possessing the characteristic(s) of Interest.

Proportions and ratios of the form estimated mean of X over the estimated mean of Y are obtained by:

(a) summing the final weights of records having the characteristic of interest for the numerator X mean (estimated)

(b) summing the final weights of records having the characteristic of interest for the denominator Y mean (estimated)

then

(c) dividing the numerator estimate by the denominator estimate.

11.2.3 Tabulation of Quantitative Estimates

Estimates of quantities can be obtained from the master file by multiplying the value of the variable of interest by the final weight for each record, then summing this quantity over all records of interest.

For example, to obtain an estimate of the total number of days of special care received by infants who were born prematurely

- multiply the number of days for which special care was received by the final weight,

-then sum this value over all records for which the child was born prematurely

To obtain a weighted average of the form estimated X mean over estimated Y mean, the numerator is calculated as for a quantitative estimate and the denominator is calculated as for a categorical estimate. For example, to estimate the average number of days spent in special care by premature babies,

(a) estimate the total number of days as described above,
(b) estimate the number of children in this category by summing the final weights of all records for the babies which were premature, then
(c) divide estimate (a) by estimate (b).

11.3 Guidelines for Statistical Analysis

In Cycle 1, the NLSCY is based upon a sample design with stratification and multiple stages of selection. In Cycles 2 and 3 the survey is based upon a non-probabilistic sample design. Using data from such complex surveys presents problems to analysts because the survey design and the selection probabilities affect the estimation and approximate variance procedures that should be used. In order for survey estimates and analyses to be closer to the population characteristics, the survey weights must be used.

While many analysis procedures found in statistical packages allow weights to be used, the meaning or definition of the weight in these procedures differ from that which is appropriate in a sample survey framework, with the result that while in many cases the estimates produced by the packages are correct, the variance estimates that are calculated are not adequate. Approximate variances for simple estimates such as totals, proportions and ratios (for qualitative variables) are provided in the accompanying approximate Sampling Variability Tables.

For other analysis techniques (for example linear regression, logistic regression and analysis of variance), a method exists which can make the approximate variances calculated by the standard packages more meaningful, by incorporating the unequal probabilities of selection. The method rescales the weights so that there is an average weight of 1.

For example, suppose that analysis of all male children is required. The steps to rescale the weights are as follows:

- select all respondents from the file with SEX=male
- Calculate the AVERAGE weight for these records by summing the original person weights from the master file for these records and then dividing by the number of records with SEX=male
- for each of these records, calculate a RESCALED weight equal to the original person weight divided by the AVERAGE weight
- perform the analysis for these respondents using the RESCALED weight.

However, because the stratification and clustering of the sample's design are still not taken into account, the variances calculated in this way are likely to be under-estimated.

11.4 C.V. Release Guidelines

Before releasing and/or publishing any estimate from the NLSCY, users should first determine the quality level of the estimate. The quality levels are *acceptable*, *marginal* and *unacceptable*. Data quality is affected by both sampling and non-sampling errors as. However for this purpose, the quality level of an estimate will be determined only on the basis of sampling error as reflected by the coefficient of variation as shown in the table below. Nonetheless users should be sure to read Section 10 to be more fully aware of the quality characteristics of these data.

First, the number of children who contribute to the calculation of the estimate should be determined. If this number is less than 30, the weighted estimate should be considered to be of unacceptable quality.

For weighted estimates based on sample sizes of 30 or more, users should determine the coefficient of variation of the estimate and follow the guidelines below. These quality level guidelines should be applied to weighted rounded estimates.

All estimates can be considered releasable. However, those of marginal or unacceptable quality level must be accompanied by a warning to caution subsequent users.

QUALITY LEVEL GUIDELINES

Quality Level of Estimate	Guidelines
1.Acceptable	Estimates have: a sample size of 30 or more, **and** low coefficients of variation in the range 0.0% to 16.5%. No warning is required
2.Marginal	Estimates have: a sample size of 30 or more, **and** high coefficients of variation in the range 16.6% to 33.3%. Estimates should be flagged with the letter M (or some similar identifier).They should be accompanied by a warning to caution subsequent users about the high levels of error, associated with the estimates.
3.Unacceptable	Estimates have: a sample size of less than 30, **or** very high coefficients of variation in excess of 33.3%. Statistics Canada recommends not to release estimates of unacceptable quality. However, if the user chooses to do so then estimates should be flagged with the letter U (or some similar identifier) and the following warning should accompany the estimates: "The user is advised that . . . (specify the data) . . . do not meet Statistics Canada's quality standards for this statistical program. Conclusions based on these data will be unreliable, and

most likely invalid. These data and any consequent findings should not be published. If the user chooses to publish these data or findings, then this disclaimer must be published with the data."

12.0
Approximate Sampling Variability Tables

In order to supply coefficients of variation which would be applicable to a wide variety of categorical estimates produced from this file and which could be readily accessed by the user, a set of Approximate Sampling Variability Tables has been produced. These "look-up" tables, which can be found in Appendix 1, allow the user to obtain an approximate coefficient of variation based on the size of the estimate calculated from the survey data.

The coefficients of variation (C.V) for Cycle 1 are derived using the variance formula for simple random sampling and incorporate a factor which reflects the actual nature of the sample design. This factor, known as the design effect, was determined by first calculating design effects for a wide range of characteristics and then choosing from among these a conservative value to be used in the look-up tables which would then apply to the entire set of characteristics.

For the NLSCY in the Territories, the sample was constructed in order to have a sufficient sample size in each territory, including the new territories of Nunavut and Western, to allow for the production of reliable estimates for all children 0 to 15 years of age. However, the sample was not intended to be large enough to ensure reliable estimates for age groups.

The tables below show the design effects, sample sizes and population counts for Cycle 1, first by territory and then by age groupings which were used to produce the Approximate Sampling Variability Tables.

TERRITORY	DESIGN EFFECT	SAMPLE SIZE	POPULATION *
Yukon	2.8	1,059	5,961
Northwest Territories	2.5	1,345	17,547

TERRITORY	CYCLE 1 AGE GROUP	DESIGN EFFECT	SAMPLE SIZE	POPULATION *
Yukon	0-3 years	1.7	310	2,013
	4-7 years	1.9	362	1,936
	8-11 years	1.9	387	2,012
Northwest Territories	0-3 years	1.7	487	6,587
	4-7 years	1.7	452	6,079
	8-11 years	1.8	406	5,237

* These population counts do not match the population counts found in the Cycle 1 and 2 user's guide as the counts been recently revised.

All coefficients of variation in the Approximate Sampling Variability Tables are approximate and, therefore, unofficial.

Remember: if the number of observations on which an estimate is based is less than 30, the weighted estimate should be classified as “unacceptable” regardless of the value of the coefficient of variation for this estimate. This is because the formulas used for estimating the variance do not hold true for small sample sizes.

12.1 How to use the C.V. tables for Categorical Estimates

The following rules should enable the user to determine the approximate coefficients of variation from the Approximate Sampling Variability Tables for estimates of the number, proportion or percentage of the surveyed population possessing a certain characteristic and for ratios and differences between such estimates. The Approximate Sampling Variability Tables were calculated in Cycle 1, so they only apply to the original longitudinal population. Since the sampling of the top-ups in Cycles 2 and 3 was non-probabilistic, the coefficient of variations could not be calculated.

Rule 1: Estimates of Numbers Possessing a Characteristic (Aggregates)

The coefficient of variation depends only on the size of the estimate itself. On the Approximate Sampling Variability Table for the appropriate geographic area or age group, locate the estimated number in the left-most column of the table (headed "Numerator of Percentage") and follow the asterisks (if any) across to the first figure encountered. This figure is the approximate coefficient of variation.

Rule 2: Estimates of Proportions or Percentages Possessing a Characteristic

The coefficient of variation of an estimated proportion or percentage depends on both the size of the proportion or percentage and the size of the total upon which the proportion or percentage is based. Estimated proportions or percentages are relatively more reliable than the corresponding estimates of the numerator of the proportion or percentage, when the proportion or percentage is based upon a sub-group of the population. For example, the proportion of babies who were of low birth weight (i.e., less than 2500 grams) is more reliable than the estimated number of "babies who were of low birth weight”. Note that in the tables the cv's decline in value reading from left to right.

When the proportion or percentage is based upon the total population of the geographic area or age group covered by the table, the cv of the proportion or percentage is the same as the cv of the numerator of the proportion or percentage. In this case, Rule 1 can be used.

When the proportion or percentage is based upon a subset of the total population (e.g. those in a particular sex or age group within province or territory), reference should be made to the proportion or percentage (across the top of the table) and to the numerator of the proportion or percentage (down the left side of the table). The intersection of the appropriate row and column gives the coefficient of variation.

Rule 3: Estimates of Differences Between Aggregates or Percentages

The standard error of a difference between two estimates is approximately equal to the square root of the sum of squares of each standard error considered separately. That is, the standard error of a difference $(\hat{d} = \hat{X}_1 - \hat{X}_2)$ is

$$\sigma_{\hat{d}} = \sqrt{(\hat{X}_1 \alpha_1)^2 + (\hat{X}_2 \alpha_2)^2}$$

where $\hat{X}_1$ is estimate 1, $\hat{X}_2$ is estimate 2, and alpha 1 and and alpha 2 are the coefficients of variation of $\hat{X}_1$ $\hat{X}_2$ respectively.

The coefficient of variation of $\hat{d}$ is given by $\sigma_{\hat{d}}/\hat{d}$.

. This formula is accurate for the difference between separate and uncorrelated characteristics, but is only approximate otherwise.

Rule 4: Estimates of Ratios

In the case where the numerator is a subset of the denominator, the ratio should be treated as a percentage and Rule 2 applied. This would apply, for example, to the case where the denominator is the number of low birth weight babies and the numerator is the number of low birth weight babies who were born prematurely (gestational age 258 days or less).

In the case where the numerator is not a subset of the denominator, the standard deviation of the ratio of the estimates is approximately equal to the square root of the sum of squares of each coefficient of variation considered separately multiplied by the ratio itself. For example, this would apply to an estimate such as, the ratio of the number of female babies who were of low birth weight as compared to the number of male babies who were of low birth weight. The standard error of such a ratio $(\hat{R} = \hat{X}_1 / \hat{X}_2)$ is:

$$\sigma_{\hat{R}} = \hat{R}\sqrt{\alpha_1^2 + \alpha_2^2}$$

where α_1 and α_2 are the coefficients of variation of $\hat{X}_1$ (the number of low birth weight female babies) and $\hat{X}_2$ (the number of low birth weight male babies) respectively.

The coefficient of variation of $\hat{R}$ is given by $\sigma_{\hat{R}}/\hat{R}$

.The formula will tend to overstate the error, if $\hat{X}_1$ and $\hat{X}_2$ are positively correlated and understate the error if $\hat{X}_1$ and $\hat{X}_2$ are negatively correlated.

Rule 5: Estimates of Differences of Ratios

In this case, Rules 3 and 4 are combined. The cv's for the two ratios are first determined using Rule 4, and then the cv of their difference is found using Rule 3.

12.1.1 Examples of using the C.V. tables for Categorical Estimates

The following are examples using actual NLSCY data (from the provincial survey) to illustrate how to apply the foregoing rules.

Example 1 : Estimates of Numbers Possessing a Characteristic (Aggregates)

Using Cycle 1 NLSCY (provincial) data, 84,085 babies were estimated to be of low birth weight (i.e., less than 2500 grams). How does the user determine the coefficient of variation of this estimate?

(1) Refer to the cv table for children in 0-3 age group. Note that the question on birth weight was applicable only to children in the 0-3 age group and therefore this is the table that should be used to determine the cv for this estimate.

(2) The estimated aggregate (84,085) does not appear in the left-hand column (the 'Numerator of Percentage' column), so it is necessary to use the figure closest to it, namely 85,000.

(3) The coefficient of variation for an estimated aggregate is found by referring to the first non-asterisk entry on that row, namely, 7.3%.

(4) The approximate coefficient of variation of the number of low birth weight babies is estimated to be 7.3%. The finding that there were 84,085 babies that were of low birth weight is “acceptable” and no warning message is required to produce this estimate since the cv for the estimate is in the 0.0%-16.5% range.

Example 2 : Estimates of Proportions or Percentages Possessing a Characteristic

Using Cycle 1 NLSCY provincial data, it is estimated that 70.8% (59,567/84085) of low birth weight babies were born prematurely (gestational age 258 days or less). How does the user determine the coefficient of variation of this estimate?

(1) Refer to the cv table for children in 0-3 age group. Note that the questions on birth weight and delivery time were applicable only to children in the 0-3 age group and therefore this is the table that should be used to determine the cv for this estimate.

(2) Because the estimate is a percentage which is based on a subset of the total population (i.e., low birth weight babies who were born prematurely), it is necessary to use both the percentage (70.8%) and the numerator portion of the percentage (59,567) in determining the coefficient of variation.

(3) The numerator, 59,567, does not appear in the left-hand column (the 'Numerator of Percentage' column) so it is necessary to use the figure closet to it, namely 60,000. Similarly, the percentage estimate does not appear as any of the column headings, so it is necessary to use the figure closest to it, 70.0%.

(4) The figure at the intersection of the row and column used, namely 5.0% is the coefficient of variation to be used.

(5) The approximate coefficient of variation of the percentage of low birth weight babies who were prematurely is estimated to be 5.0%. The finding that 70.8% of low birth weight babies were born prematurely is “acceptable” and no warning message is required to produce this estimate since the cv for the estimate is in the 0.0%-16.5% range.

Example 3 : Estimates of Differences Between Aggregates or Percentages

Using Cycle 1 NLSCY provincial data, it is estimated that 6.1% (45,690/753,203) of female babies were born prematurely, while 4.9% (38,395/791,149) of male babies were born prematurely. How does the user determine the coefficient of variation of the difference between these two estimates?

(1) Using the cv table for the 0-3 age group in the same manner as described in example 2 gives the cv of the estimate for female babies as 10.3%, and the cv of the estimate for male babies as 10.9%.

(2) Using rule 3, the standard error of a difference $(\hat{d} = X_1 - X_2)$ is:

$$\sigma_{\hat{d}} = \sqrt{(\hat{X}_1 \alpha_1)^2 + (\hat{X}_2 \alpha_2)^2}$$

where X_1 is estimate 1 (the percent of low birth weight female babies), X_2 is estimate 2 (the percent of low birth weight male babies) , and α_1 and α_2 are the coefficients of variation of X_1 and X_2 respectively.

That is, the standard error of the difference $\hat{d}$ = (.061-.049) = .012

(3) The coefficient of variation of $\hat{d}$ is given by

$$\sigma_{\hat{d}}/\hat{d} = 0{,}008/0{,}012 = 0{,}667$$

(4) So the approximate coefficient of variation of the difference between the estimates is 66.7%. This estimate is “unacceptable” since the coefficient of variation is over 33.3%. Statistics Canada recommends not to release estimates of unacceptable quality.

Example 4: Estimates of Ratios

Suppose now a user wants to compare the number of low birth weight female babies to the number of low birth weight male babies. The user is interested in comparing these estimates in the form of a ratio. How does the user determine the coefficient of variation of this estimate?

(1) First of all, this estimate is a ratio estimate, where the numerator of the estimate (=X_1) is the number of low birth weight female babies and denominator (=X_2) of the estimate is the number of low birth weight male babies.

(2) Refer to the table for the 0-3 age group. The questions on birth weight were applicable only to children in the 0-3 age group.

(3) The numerator of this ratio estimate is 45,690. The figure closest to it is 45,000. The coefficient of variation for this estimate is found by referring to the first non-asterisk entry on that row, namely, 10.3%.

(4) The denominator of this ratio estimate is 38,395. The figure closest to it is 40,000. The coefficient of variation for this estimate is found by referring to the first non-asterisk entry on that row, namely, 10.9%.

(5) So the approximate coefficient of variation of the ratio estimate is given by rule 4, which is,

$\alpha_{\hat{R}} = \sqrt{\alpha_1^2 + \alpha_2^2}$ where α_1 and α_2 are the coefficients of variation of $\hat{X}_1$ and $\hat{X}_2$ respectively.

That is:

$$\alpha_{\hat{R}} = \sqrt{(0{,}103)^2 + (0{,}109)^2}$$
$$= 0{,}150$$

The obtained ratio of female babies who were of low birth weight versus male babies who were of low birth weight is 45,690/38,395 which is 1.19 : 1. The approximate coefficient of variation of this estimate is 15.0%, which is “acceptable” and no warning message is required to produce this estimate since the cv for the estimate is in the 0.0%-16.5% range.

12.2 How to use the C.V. tables to obtain Confidence Limits

Although coefficients of variation are widely used, a more intuitively meaningful measure of sampling error is the confidence interval of an estimate. A confidence interval constitutes a statement on the level of confidence that the true value for the population lies within a specified range of values. For example a 95% confidence interval can be described as follows:

If sampling of the population is repeated indefinitely, each sample leading to a new confidence interval for an estimate, then in 95% of the samples the interval will cover the true population value.

Using the standard error of an estimate, confidence intervals for estimates may be obtained under the assumption that under repeated sampling of the population, the various estimates obtained for a population characteristic are normally distributed about the true population value. Under this assumption, the chances are about 68 out of 100 that the difference between a sample estimate and the true population value would be less than one standard error, about 95 out of 100 that the difference would be less than two standard errors, and about 99 out 100 that the differences would be less than three standard errors. These different degrees of confidence are referred to as the confidence levels.

Confidence intervals for an estimate, $\hat{X}$, are generally expressed as two numbers, one below the estimate and one above the estimate, as where k is determined depending upon the level of confidence desired and the sampling error of the estimate.

Confidence intervals for an estimate can be calculated directly from the Approximate Sampling Variability Tables by first determining from the appropriate table the coefficient of variation of the estimate $\hat{X}$ and then using the following formula to convert to a confidence interval CI:

$$IC_{\hat{X}} = [\hat{X} - t\hat{X}\alpha_{\hat{X}}),\ \hat{X} + t\hat{X}\alpha_{\hat{X}}]$$

where $\alpha_{\hat{X}}$ is the determined coefficient of variation $\hat{X}$ and

t = 1 if a 68% confidence interval is desired
t = 1.6 if a 90% confidence interval is desired
t = 2 if a 95% confidence interval is desired
t = 3 if a 99% confidence interval is desired.

Note: Release guidelines which apply to the estimate also apply to the confidence interval. For example, if the estimate is “marginal”, then the confidence interval is marginal and should be accompanied by a warning note to caution subsequent users about the high levels of error.

12.2.1 Example of using the C.V. tables to obtain confidence limits

A 95% confidence interval for the estimated proportion of babies who were of low birth weight would be calculated as follows.

estimate of X=5.5%
t=2
alpha estimate of X =7.3% (.073 expressed as a proportion)
is the coefficient of variation of this estimate

With 95% confidence it can be said that between approximately 4.7% and 6.3% of babies who were 0-3 years old at the time of the survey were of low birth weight.

12.3 How to use the C.V. tables to do a t-test

Standard errors may also be used to perform hypothesis testing, a procedure for distinguishing between population parameters using sample estimates. The sample estimates can be numbers, averages, percentages, ratios, etc. Tests may be performed at various levels of significance, where a level of significance is the probability of concluding that the characteristics are different when, in fact, they are identical.

Let X_1 and X_2, be sample estimates for two characteristics of interest.Let the standard error on the difference $X_1 - X_2$ be σ_d.

If $t = \frac{X_1 - X_2}{\sigma_d}$ is between -2 and 2, then no conclusion about the difference between the characteristics is justified at the 5% level of significance.If however, this ratio is smaller than -2 or larger than +2, the observed difference is significant at the 0.05 level.That is to say that the characteristics are significantly different.

12.3.1 Example of using the C.V. tables to do a t-test

Let us suppose we wish to test, at 5% level of significance, the hypothesis that there is no difference between the proportion of female babies who were of low birth weight and the proportion of male babies who were of low birth weight. From example 3 (Section 12.1.1), the standard error of the difference between these two estimates was found to be = .008. Hence ,

$$t = \frac{\hat{X}_1 - \hat{X}_2}{\sigma_d} = \frac{0,061 - 0,049}{0,008} = \frac{0,012}{0,008} = 1,5.$$

Since t = 1.5 is between -2 and 2, no conclusion at the 0.05 level of significance can be made regarding the difference in proportions of male of female babies who were of low birth weight.

12.4 Coefficients of Variation for Quantitative Estimates

For quantitative estimates, special tables would have to be produced to determine their sampling error. Since most of the variables for the NLSCY are categorical in nature, this has not been done.

As a general rule, however, the coefficient of variation of a quantitative total will be larger than the coefficient of variation of the corresponding category estimate (i.e., the estimate of the number of persons contributing to the quantitative estimate). If the corresponding category estimate is not releasable, the quantitative estimate will not be either. For example, the coefficient of variation of the total number of days of special medical care received for low birth weight babies would be greater than the coefficient of variation of the corresponding proportion of babies who were of low birth weight. Hence if the coefficient of variation of the proportion is not releasable, then the coefficient of variation of the corresponding quantitative estimate will also not be releasable.

Coefficients of variation of such estimates can be derived as required for a specific estimate using a technique known as pseudo replication. This involves dividing the records on the master files into subgroups (or replicates) and determining the variation in the estimate from replicate to replicate. Users wishing to derive coefficients of variation for quantitative estimates may contact Statistics Canada for advice on the allocation of records to appropriate replicates and the formulae to be used in these calculations.

12.5 Release cut-off's for the NLSCY

In the tables that follow, cut-off numbers are given for NLSCY estimates in order for them to be of “acceptable”, “marginal” or “unacceptable” quality. Users are encouraged to use these cut-offs when publishing data from the NLSCY. First a table is given to show the cut-offs at the territory level. Then a table is given to show the cut-offs for the various age groups. An interpretation of what is meant by the various cut-off levels can be found in Section 11.4.

For example, an estimate for Yukon of 350 would fall into the “marginal” range. This would mean that the estimate should be flagged and a warning note attached to caution subsequent users about the high level of error associated with the estimate.

GEOGRAPHICAL RELEASE CUT-OFFS

Territory	Acceptable - estimates at or above:	Marginal - estimates between:	Unacceptable - estimates at or below
Yukon	400	200 & 400	100
Northwest Territories	1,100	500 & 1,100	300

AGE GROUP RELEASE CUT-OFFS

Territory	Cycle 1 Age Group	Acceptable - estimates at or above:	Marginal - estimates between:	Unacceptable - estimates at or below
Yukon	0-3 years	300	100 & 300	100
	4-7 years	300	100 & 300	100
	8-11 years	200	100 & 200	100
Northwest Territories	0-3 years	600	300 & 600	200
	4-7 years	700	300 & 700	200
	8-11 years	700	300 & 700	200

13.0 Record Layout

13.1 Technical Information

The file for the NLSCY territories is stored as an ASCII text file. SAS and SPSS record layouts have been included for reading the file. The SAS record layouts are included in Section 13.2 SAS Record Layouts.

For this release of NLSCY territories data, the file consists of complete child records with data for all of the sections of the various questionnaires. The data collected for the PMK and spouse/partner have been replicated for each child in the household. For example if there were three children in the family, the education variables for the PMK and the spouse have been written to all three child records. If a section of a questionnaire (or a complete questionnaire) was not applicable for a child, all of the variables for that section (or questionnaire) have been set to not-applicable. For example if the PMK did not have a spouse, the spouse variables have been set to not-applicable for all children in the family. All variables on the master file which are longer than one byte have been zero-filled. The record layouts included with the release package declare all variables as numeric, except those for gender (character M or F) and for postal code.

The total number of records that appear on the Cycle 3 master file is 2,361. However, six children on the file have a cross-sectional flag (XSECFLG) of 0, which means that they were non-respondents for cycle 3 but kept on the file because they are longitudinal children. These children did not receive either a longitudinal or a cross-sectional weight.

Of the 2,361 records that appear on the master file, 1,981 agreed to share their data and an additional two were included on the file but were out of scope. Thus, the total number of children that appear on the cycle 3 share file is 1,983.

The unit of analysis for all estimates made from the NLSCY file should refer to the child. The sample design used for the NLSCY was not self-weighting. When producing estimates for children, including the production of ordinary statistical tables, users must apply the proper sampling weight. If proper weights are not used, the estimates derived from the master file cannot be considered to be representative of the survey population, and will not correspond to those produced by Statistics Canada. Weights have been included and are labeled CWTCW01C (cross-sectional weight) and CWTCW01L (longitudinal weight).

There are two identification variables on the master file. One identifies children (CIDCD01) and one identifies households (CIDHD01). The child ID (CIDCD01) is unique for each child on the file and can be used to link records between the files. Children on the first master file (Cycle 1) have IDs in the range 100001 to 102404. New records added in Cycle 2 begin with ‘20’, in Cycle 3 with ‘30’ and so on. In subsequent releases of NLSCY data in the years to come, the ID for each child will remain the same. There will be additional children added to the file, but once a child is included on the file, the ID for that child will remain the same.

All children in the same household will have the same household ID (CIDHD01). Over the ensuing years children will not continue to live in the same households and will move out, or families may divide. Therefore the household ID will not remain the same over time. For each cycle there will be a household ID which can be used to determine which children live in similar households for that particular cycle. Again, on the Cycle 2 master file, this ID starts with ‘10’ for records that existed in Cycle 1. New records added in Cycle 2 begin with ‘20’, records in Cycle 3 will start with ‘30’ and so on.

The NLSCY master file documentation system has employed certain standards to label variable names and values. The intent is to make interpretation of the data more straightforward for the user. These standards and examples are provided in Section 6.3 of this guide.

13.2 SAS Record Layouts

13.2.1 Primary File

```
/*******************************************************/
/* NLSCY TERRITORIES MASTER FILE  - CYCLE 3   */
/*******************************************************/

data DATA1;
  infile &DATAIN  LRECL=1112;
  input
@00001   CGEHD01    2.
@00003   CGEHD02    2.
@00005   CGEHD03    2.
@00007   CMMPQ4    3.
@00010   CMMPQ5   $1.
@00011   CMMPQ4A    4.
@00015   CMMPQ4B    2.
@00017   CMMPQ4C    2.
@00019   CMMPQ6    2.
@00021   CMMSQ4    3.
@00024   CMMSQ5   $1.
@00025   CMMSQ4A    4.
@00029   CMMSQ4B    2.
@00031   CMMSQ4C    2.
@00033   CMMSQ6    2.
@00035   CMMCQ4    3.
@00038   CMMCQ5   $1.
@00039   CMMCQ4A    4.
@00043   CMMCQ4B    2.
@00045   CMMCQ4C    2.
@00047   CMMCQ6    2.
@00049   CDMCD01    2.
@00051   CDMHD02    2.
@00053   CDMCD03    2.
@00055   CDMCD04    1.
@00056   CDMCD05    1.
@00057   CDMCD06    2.
@00059   CDMPD06A    1.
@00060   CDMCD06B    2.
@00062   CDMPD06D    2.
@00064   CDMSD06D    2.
@00066   CDMHD06F    2.
@00068   CDMHD07    2.
@00070   CDMCD08    2.
@00072   CDMCD09    2.
@00074   CDMCD10    2.
@00076   CDMCD11    2.
@00078   CDMCD12    2.
@00080   CDMCD13    2.
@00082   CDMCD14    1.
@00083   CDMCD15    1.
@00084   CDMCD16    1.
@00085   CDMCD18    2.
@00087   CDMCD18B    2.
```

```
@00089  CDMCD19   2.
@00091  CDMCD19B  2.
@00093  CDMCD20   2.
@00095  CDMHPC  $5.
@00100  CEDPQ1    2.
@00102  CEDPQ2    1.
@00103  CEDPQ3    1.
@00104  CEDPQ4    2.
@00106  CEDPQ5    1.
@00107  CEDPQ6    1.
@00108  CEDPD01   2.
@00110  CEDPD02   1.
@00111  CEDPD04   2.
@00113  CEDPcQ7A  2.
@00115  CEDPcQ7B  2.
@00117  CEDPcQ7C  2.
@00119  CEDPcQ7D  2.
@00121  CEDHcQ8A  1.
@00122  CEDHcQ8B  1.
@00123  CEDHcQ8C  1.
@00124  CEDHcQ8D  1.
@00125  CEDHcQ8E  1.
@00126  CEDSQ1   2.
@00128  CEDSQ2   1.
@00129  CEDSQ3   1.
@00130  CEDSQ4   2.
@00132  CEDSQ5   1.
@00133  CEDSQ6   1.
@00134  CEDSD01  2.
@00136  CEDSD02  1.
@00137  CEDSD04  2.
@00139  CEDScQ7A  2.
@00141  CEDScQ7B  2.
@00143  CEDScQ7C  2.
@00145  CEDScQ7D  2.
@00147  CLFPQ2   1.
@00148  CLFPQ4A   1.
@00149  CLFPQ6A   1.
@00150  CLFPQ11A  1.
@00151  CLFPQ4B   1.
@00152  CLFPQ6B   1.
@00153  CLFPQ11B  1.
@00154  CLFPQ4C   1.
@00155  CLFPQ6C   1.
@00156  CLFPQ11C  1.
@00157  CLFPQ12   1.
@00158  CLFPQ12A  2.
@00160  CLFPQ12B  2.
@00162  CLFPQ12C  1.
@00163  CLFPQ16   1.
@00164  CLFPc18A  1.
@00165  CLFPc18B  2.
@00167  CLFPc18C  2.
@00169  CLFPb8_4  1.
@00170  CLFPb84A  1.
@00171  CLFPb84B  2.
```

```
@00173  CLFPb17B   2.
@00175  CLFPD05    4.
@00179  CLFPD05B   4.
@00183  CLFPD06    4.
@00187  CLFPD06B   $4.
@00191  CLFPD07    2.
@00193  CLFPD08    2.
@00195  CLFPD09    2.
@00197  CLFSQ2     1.
@00198  CLFSQ4A    1.
@00199  CLFSQ6A    1.
@00200  CLFSQ11A   1.
@00201  CLFSQ4B    1.
@00202  CLFSQ6B    1.
@00203  CLFSQ11B   1.
@00204  CLFSQ4C    1.
@00205  CLFSQ6C    1.
@00206  CLFSQ11C   1.
@00207  CLFSQ12    1.
@00208  CLFSQ12A   2.
@00210  CLFSQ12B   2.
@00212  CLFSQ12C   1.
@00213  CLFSQ16    1.
@00214  CLFSc18A   1.
@00215  CLFSc18B   2.
@00217  CLFSc18C   2.
@00219  CLFSb8_4   1.
@00220  CLFSb84A   1.
@00221  CLFSb84B   2.
@00223  CLFSb17B   2.
@00225  CLFSD05    4.
@00229  CLFSD05B   4.
@00233  CLFSD06    4.
@00237  CLFSD06B   $4.
@00241  CLFSD07    2.
@00243  CLFSD08    2.
@00245  CLFSD09    2.
@00247  CINHD01A   2.
@00249  CINHD01B   2.
@00251  CINPD02    2.
@00253  CINSD02    2.
@00255  CINHD07    2.
@00257  CINHD09    6.3
@00263  CDPPQ12A   1.
@00264  CDPPQ12B   1.
@00265  CDPPQ12C   1.
@00266  CDPPQ12D   1.
@00267  CDPPQ12E   1.
@00268  CDPPQ12F   1.
@00269  CDPPQ12G   1.
@00270  CDPPQ12H   1.
@00271  CDPPQ12I   1.
@00272  CDPPQ12J   1.
@00273  CDPPQ12K   1.
@00274  CDPPQ12L   1.
@00275  CDPPS01    2.
```

@00277 CSDPD01 3.
@00280 CSDPD02 3.
@00283 CSDPD02B 1.
@00284 CSDPD03 2.
@00286 CSDPD04 2.
@00288 CSDPD05 2.
@00290 CSDPD05B 1.
@00291 CSDPD06 2.
@00293 CSDPD06B 1.
@00294 CSDSD01 3.
@00297 CSDSD02 3.
@00300 CSDSD02B 1.
@00301 CSDSD03 2.
@00303 CSDSD04 2.
@00305 CSDSD05 2.
@00307 CSDSD05B 1.
@00308 CSDSD06 2.
@00310 CSDSD06B 1.
@00311 CSDCD01 3.
@00314 CSDCD02 3.
@00317 CSDCD02B 1.
@00318 CSDCD03 2.
@00320 CSDCD04 2.
@00322 CSDCD05 2.
@00324 CSDCD05B 1.
@00325 CSDCD06 2.
@00327 CSDCD06B 1.
@00328 CMDCQ1A 1.
@00329 CMDCQ1B 1.
@00330 CMDCQ1C 1.
@00331 CMDCQ3 1.
@00332 CMDCQ4 2.
@00334 CMDCQ5A 1.
@00335 CMDCQ5B 1.
@00336 CMDCQ5C 1.
@00337 CMDCQ6 2.
@00339 CMDCQ7 1.
@00340 CMDCQ8A 1.
@00341 CMDCQ8B 1.
@00342 CMDCQ8C 1.
@00343 CMDCQ9A 1.
@00344 CMDCQ9B1 1.
@00345 CMDCQ9B2 1.
@00346 CMDCQ9B3 1.
@00347 CMDCQ10A 1.
@00348 CMDC10B1 1.
@00349 CMDC10B2 1.
@00350 CMDC10B3 1.
@00351 CMDCQ12A 1.
@00352 CMDCQ12B 3.
@00355 CMDCQ13A 5.3
@00360 CMDCQ14A 2.
@00362 CMDCQ16 1.
@00363 CMDCQ17 1.
@00364 CMDCc18B 1.
@00365 CMDCc18C 1.

@00366 CMDCc18D 1.
@00367 CMDCc18E 1.
@00368 CMDCQ21A 1.
@00369 CMDC21B1 1.
@00370 CMDC21B2 1.
@00371 CMDC21B3 1.
@00372 CMDC21B4 1.
@00373 CMDCQ21C 3.
@00376 CMDCQ22 2.
@00378 CMDCQ24A 1.
@00379 CMDCQ25 1.
@00380 CMDCQ26 1.
@00381 CMDCQ27 2.
@00383 CMDCQ28A 1.
@00384 CMDCQ28B 1.
@00385 CMDCQ28C 1.
@00386 CMDCQ28D 1.
@00387 CMDCQ28E 1.
@00388 CMDCQ28F 1.
@00389 CMDCQ28G 1.
@00390 CMDCQ28H 1.
@00391 CMDCQ28I 1.
@00392 CMDCQ28J 1.
@00393 CMDCQ28K 1.
@00394 CMDCQ28L 1.
@00395 CMDCQ28M 1.
@00396 CMDCbQ29 1.
@00397 CMDCc29a 3.
@00400 CMDCb29B 2.
@00402 CWBCc29C 1.
@00403 CWBCc29D 1.
@00404 CMDCD01 1.
@00405 CMDCD02 2.
@00407 CMDCD03 2.
@00409 CMDCD04 2.
@00411 CMDCD05 2.
@00413 CMDCD06 3.
@00416 CMDCD07 1.
@00417 CMDCD08 1.
@00418 CLTCcQ1 1.
@00419 CLTCcQ2 2.
@00421 CLTCcQ3 2.
@00423 CLTCcQ4 2.
@00425 CLTCcQ5 2.
@00427 CLTCcQ6A 1.
@00428 CLTCcQ6B 2.
@00430 CLTCcQ7 2.
@00432 CLTCcQ8 2.
@00434 CLTCcQ12 2.
@00436 CLTCcQ13 2.
@00438 CLTCcQ14 2.
@00440 CEDCD01 2.
@00442 CEDCQ6 1.
@00443 CEDCD03 2.
@00445 CEDCQ12A 2.
@00447 CEDCcQ13 2.

@00449 CEDCQ14A 2.
@00451 CEDCb14A 2.
@00453 CEDCQ14B 2.
@00455 CEDCQ14C 2.
@00457 CEDCb14C 2.
@00459 CEDCQ14D 2.
@00461 CEDCQ15 2.
@00463 CBECQ6B 1.
@00464 CBECQ6C 1.
@00465 CBECQ6D 1.
@00466 CBECQ6E 1.
@00467 CBECQ6F 1.
@00468 CBECQ6G 1.
@00469 CBECQ6I 1.
@00470 CBECQ6K 1.
@00471 CBECQ6L 1.
@00472 CBECQ6N 1.
@00473 CBECQ6O 1.
@00474 CBECQ6P 1.
@00475 CBECQ6Q 1.
@00476 CBECQ6S 1.
@00477 CBECQ6T 1.
@00478 CBECQ6U 1.
@00479 CBECQ6V 1.
@00480 CBECQ6W 1.
@00481 CBECQ6X 1.
@00482 CBECQ6AA 1.
@00483 CBECQ6BB 1.
@00484 CBECQ6CC 1.
@00485 CBECQ6DD 1.
@00486 CBECQ6EE 1.
@00487 CBECQ6FF 1.
@00488 CBECQ6HH 1.
@00489 CBECQ6JJ 1.
@00490 CBECQ6MM 1.
@00491 CBECQ6PP 1.
@00492 CBECQ6QQ 1.
@00493 CBECQ6RR 1.
@00494 CBECQ6SS 1.
@00495 CBECQ6UU 1.
@00496 CBECQ7A 1.
@00497 CBECQ7B 1.
@00498 CBECQ7C 1.
@00499 CBECQ7D 1.
@00500 CBECQ7E 1.
@00501 CBECQ7F 1.
@00502 CBECS06 2.
@00504 CBECS07 2.
@00506 CBECS08 2.
@00508 CBECS11 2.
@00510 CRLCQ1 2.
@00512 CRLCQ2 2.
@00514 CRLCQ4 1.
@00515 CRLCQ9 2.
@00517 CPRCQ1 2.
@00519 CPRCQ3 2.

@00521 CPRCQ6 2.
@00523 CPRCbQ7A 2.
@00525 CPRCbQ7 2.
@00527 CPRCQ10 2.
@00529 CPRCQ11 2.
@00531 CPRCQ12 2.
@00533 CPRCQ16 2.
@00535 CPRCS01 2.
@00537 CPRCS03 2.
@00539 CPRCS05 2.
@00541 CRSPD01 1.
@00542 CRSSD01 1.
@00543 CRSPQ1A 1.
@00544 CRSPQ1B 1.
@00545 CRSPQ1C 1.
@00546 CRSPQ1D 1.
@00547 CRSPQ1E 1.
@00548 CRSSQ1A 1.
@00549 CRSSQ1B 1.
@00550 CRSSQ1C 1.
@00551 CRSSQ1D 1.
@00552 CRSSQ1E 1.
@00553 CCHPD01 1.
@00554 CCHSD01 1.
@00555 CCHPQ1A 1.
@00556 CCHPQ1B 1.
@00557 CCHPQ1C 1.
@00558 CCHPQ1D 1.
@00559 CCHPQ1E 1.
@00560 CCHPQ1F 1.
@00561 CCHPQ1G 1.
@00562 CCHPQ1H 1.
@00563 CCHPQ1I 1.
@00564 CCHPQ1J 1.
@00565 CCHPQ1K 1.
@00566 CCHPQ1L 1.
@00567 CCHPQ1M 1.
@00568 CCHPQ1N 1.
@00569 CCHPQ1O 1.
@00570 CCHPQ1P 1.
@00571 CCHPbQ1V 1.
@00572 CCHPQ1R 1.
@00573 CCHPQ1S 1.
@00574 CCHPQ1T 1.
@00575 CCHPbQ1W 1.
@00576 CCHPbQ1U 1.
@00577 CCHSQ1A 1.
@00578 CCHSQ1B 1.
@00579 CCHSQ1C 1.
@00580 CCHSQ1D 1.
@00581 CCHSQ1E 1.
@00582 CCHSQ1F 1.
@00583 CCHSQ1G 1.
@00584 CCHSQ1H 1.
@00585 CCHSQ1I 1.
@00586 CCHSQ1J 1.

@00587 CCHSQ1K 1.
@00588 CCHSQ1L 1.
@00589 CCHSQ1M 1.
@00590 CCHSQ1N 1.
@00591 CCHSQ1O 1.
@00592 CCHSQ1P 1.
@00593 CCHSbQ1V 1.
@00594 CCHSQ1R 1.
@00595 CCHSQ1S 1.
@00596 CCHSQ1T 1.
@00597 CCHSbQ1W 1.
@00598 CCHSbQ1U 1.
@00599 CHLPQ1 2.
@00601 CHLPQ2 1.
@00602 CHLPQ3 2.
@00604 CHLPcQ5 2.
@00606 CHLPcQ5A 2.
@00608 CHLPbQ6 2.
@00610 CHLPQ5B 1.
@00611 CHLPQ5C 1.
@00612 CHLPQ5D 2.
@00614 CHLPcQ7 1.
@00615 CHLSQ1 2.
@00617 CHLSQ2 1.
@00618 CHLSQ3 2.
@00620 CHLScQ5 2.
@00622 CHLScQ5A 2.
@00624 CHLSbQ6 2.
@00626 CHLSQ5B 1.
@00627 CHLSQ5C 1.
@00628 CHLSQ5D 2.
@00630 CHLScQ7 1.
@00631 CHLMQ8 2.
@00633 CHLMQ9 2.
@00635 CHLMQ11 2.
@00637 CHHHQ1 1.
@00638 CHHHQ3 2.
@00640 CHHHQ6 2.
@00642 CHLCQ1 2.
@00644 CHLCQ2 2.
@00646 CHLCQ3B 4.2
@00650 CHLCQ4A 7.3
@00657 CHLCb4C1 5.3
@00662 CHLCQ5 2.
@00664 CHLCbQ6 1.
@00665 CHLCbQ7 1.
@00666 CHLCQ8 1.
@00667 CHLCQ9 1.
@00668 CHLCQ10 1.
@00669 CHLCQ11 1.
@00670 CHLCQ12 1.
@00671 CHLCQ13 1.
@00672 CHLCQ14 1.
@00673 CHLCQ15 1.
@00674 CHLCQ16 1.
@00675 CHLCQ17 1.

@00676 CHLCQ18 1.
@00677 CHLCQ19 1.
@00678 CHLCbQ20 1.
@00679 CHLCQ21 1.
@00680 CHLCbQ22 1.
@00681 CHLCQ23 1.
@00682 CHLCQ24 1.
@00683 CHLCQ25 1.
@00684 CHLCQ26 1.
@00685 CHLCQ27 1.
@00686 CHLCQ28 1.
@00687 CHLCQ29 1.
@00688 CHLCQ30 1.
@00689 CHLCQ31 2.
@00691 CHLCQ32 1.
@00692 CHLCQ33 2.
@00694 CHLCQ34 1.
@00695 CHLCQ35 1.
@00696 CHLCQ36 1.
@00697 CHLCQ37 1.
@00698 CHLCQ38 2.
@00700 CHLCQ39 2.
@00702 CHLCQ40 2.
@00704 CHLCQ41 2.
@00706 CHLCQ42 2.
@00708 CHLCb43A 1.
@00709 CHLCQ43B 1.
@00710 CHLCQ43C 1.
@00711 CHLCQ44 1.
@00712 CHLC45AA 1.
@00713 CHLC45AB 1.
@00714 CHLC45AC 1.
@00715 CHLC45AD 1.
@00716 CHLC45AE 1.
@00717 CHLC45AF 1.
@00718 CHLC45AG 1.
@00719 CHLC45AH 1.
@00720 CHLC45AI 1.
@00721 CHLC45AJ 1.
@00722 CHLC45AK 1.
@00723 CHLCQ45B 1.
@00724 CACCQ3A 2.
@00726 CACCb3AA 2.
@00728 CACCbQ3B 2.
@00730 CACCQ3C 2.
@00732 CACCQ3D 2.
@00734 CACCcQ4A 2.
@00736 CACCcQ4B 4.1
@00740 CACCbQ8 2.
@00742 CACCQ9 1.
@00743 CCSCQ1 2.
@00745 CCSCQ2A 2.
@00747 CCSCQ2B 4.
@00751 CCSCQ3A 2.
@00753 CCSCQ3B 4.
@00757 CCSCQ4A 2.

@00759 CCSCQ4B 4.
@00763 CCSCQ5 2.
@00765 CCSCQ6A 1.
@00766 CCSCQ6B 1.
@00767 CCSCbQ6C 1.
@00768 CCSCQ7 1.
@00769 CCSCQ8 2.
@00771 CCSCQ9A 1.
@00772 CCSCQ9B 1.
@00773 CCSCQ10A 2.
@00775 CCSCQ10B 2.
@00777 CCSCc10C 1.
@00778 CCSCQ11 1.
@00779 CCSCQ12 1.
@00780 CCSCQ13 1.
@00781 CCSCQ14 2.
@00783 CINHQ1A 1.
@00784 CINHQ1B 1.
@00785 CINHQ1C 1.
@00786 CINHQ1D 1.
@00787 CINHQ1E 1.
@00788 CINHQ1F 1.
@00789 CINHQ1G 1.
@00790 CINHQ1H 1.
@00791 CINHQ1I 1.
@00792 CINHQ1J 1.
@00793 CINHQ1K 1.
@00794 CINHQ1L 1.
@00795 CINHQ1M 1.
@00796 CINHQ2 2.
@00798 CINHD02B 1.
@00799 CINHQ3 2.
@00801 CINPQ4B 2.
@00803 CINSQ4B 2.
@00805 CSFHcQ1 2.
@00807 CSFHcQ2 2.
@00809 CSFHcQ3 1.
@00810 CSFHcQ4 2.
@00812 CSFHcQ5A 1.
@00813 CSFHcQ5B 1.
@00814 CSFHcQ5C 1.
@00815 CSFHcQ6A 1.
@00816 CSFHcQ6B 1.
@00817 CSFHcQ6C 1.
@00818 CSFHcQ6D 1.
@00819 CSFHcQ6E 1.
@00820 CSFHcQ7A 1.
@00821 CSFHcQ7B 1.
@00822 CSFHcQ7C 1.
@00823 CSFHcQ7D 1.
@00824 CSFHcQ7E 1.
@00825 CSFHcQ7F 1.
@00826 CSDPQ1 2.
@00828 CSDPQ2A1 1.
@00829 CSDPQ2A2 1.
@00830 CSDPQ2A3 1.

@00831 CSDPQ2A4 1.
@00832 CSDPQ2B 1.
@00833 CSDPQ3 4.
@00837 CSDPQ4A 1.
@00838 CSDPQ4B 1.
@00839 CSDPQ4C 1.
@00840 CSDPQ4D 1.
@00841 CSDPQ4E 1.
@00842 CSDPQ4F 1.
@00843 CSDPQ4G 1.
@00844 CSDPQ4H 1.
@00845 CSDPQ4I 1.
@00846 CSDPQ4J 1.
@00847 CSDPQ4K 1.
@00848 CSDPQ4L 1.
@00849 CSDPbQ4S 1.
@00850 CSDPQ4M 1.
@00851 CSDPQ4N 1.
@00852 CSDPQ4O 1.
@00853 CSDPQ4P 1.
@00854 CSDPQ4Q 1.
@00855 CSDPQ4R 1.
@00856 CSDPQ5A 1.
@00857 CSDPQ5B 1.
@00858 CSDPQ5C 1.
@00859 CSDPQ6A1 1.
@00860 CSDPQ6A2 1.
@00861 CSDPQ6A3 1.
@00862 CSDPQ6B1 1.
@00863 CSDPQ6B2 1.
@00864 CSDPQ7A 1.
@00865 CSDPQ7F 1.
@00866 CSDPQ7H 1.
@00867 CSDPQ7B 1.
@00868 CSDPQ7J 1.
@00869 CSDPQ7D 1.
@00870 CSDPQ7I 1.
@00871 CSDPbQ7L 1.
@00872 CSDPQ7E 1.
@00873 CSDPQ7C 1.
@00874 CSDPQ7K 1.
@00875 CSDPQ8 2.
@00877 CSDSQ1 2.
@00879 CSDSQ2A1 1.
@00880 CSDSQ2A2 1.
@00881 CSDSQ2A3 1.
@00882 CSDSQ2A4 1.
@00883 CSDSQ2B 1.
@00884 CSDSQ3 4.
@00888 CSDSQ4A 1.
@00889 CSDSQ4B 1.
@00890 CSDSQ4C 1.
@00891 CSDSQ4D 1.
@00892 CSDSQ4E 1.
@00893 CSDSQ4F 1.
@00894 CSDSQ4G 1.

@00895 CSDSQ4H 1.
@00896 CSDSQ4I 1.
@00897 CSDSQ4J 1.
@00898 CSDSQ4K 1.
@00899 CSDSQ4L 1.
@00900 CSDSbQ4S 1.
@00901 CSDSQ4M 1.
@00902 CSDSQ4N 1.
@00903 CSDSQ4O 1.
@00904 CSDSQ4P 1.
@00905 CSDSQ4Q 1.
@00906 CSDSQ4R 1.
@00907 CSDSQ5A 1.
@00908 CSDSQ5B 1.
@00909 CSDSQ5C 1.
@00910 CSDSQ6A1 1.
@00911 CSDSQ6A2 1.
@00912 CSDSQ6A3 1.
@00913 CSDSQ6B1 1.
@00914 CSDSQ6B2 1.
@00915 CSDSQ7A 1.
@00916 CSDSQ7F 1.
@00917 CSDSQ7H 1.
@00918 CSDSQ7B 1.
@00919 CSDSQ7J 1.
@00920 CSDSQ7D 1.
@00921 CSDSQ7I 1.
@00922 CSDSbQ7L 1.
@00923 CSDSQ7E 1.
@00924 CSDSQ7C 1.
@00925 CSDSQ7K 1.
@00926 CSDSQ8 2.
@00928 CSDCQ1 2.
@00930 CSDCQ2A1 1.
@00931 CSDCQ2A2 1.
@00932 CSDCQ2A3 1.
@00933 CSDCQ2A4 1.
@00934 CSDCQ2B 1.
@00935 CSDCQ3 4.
@00939 CSDCQ4A 1.
@00940 CSDCQ4B 1.
@00941 CSDCQ4C 1.
@00942 CSDCQ4D 1.
@00943 CSDCQ4E 1.
@00944 CSDCQ4F 1.
@00945 CSDCQ4G 1.
@00946 CSDCQ4H 1.
@00947 CSDCQ4I 1.
@00948 CSDCQ4J 1.
@00949 CSDCQ4K 1.
@00950 CSDCQ4L 1.
@00951 CSDCbQ4S 1.
@00952 CSDCQ4M 1.
@00953 CSDCQ4N 1.
@00954 CSDCQ4O 1.
@00955 CSDCQ4P 1.

@00956 CSDCQ4Q 1.
@00957 CSDCQ4R 1.
@00958 CSDCQ5A 1.
@00959 CSDCQ5B 1.
@00960 CSDCQ5C 1.
@00961 CSDCQ6A1 1.
@00962 CSDCQ6A2 1.
@00963 CSDCQ6A3 1.
@00964 CSDCQ6B1 1.
@00965 CSDCQ6B2 1.
@00966 CSDCQ7A 1.
@00967 CSDCQ7F 1.
@00968 CSDCQ7H 1.
@00969 CSDCQ7B 1.
@00970 CSDCQ7J 1.
@00971 CSDCQ7D 1.
@00972 CSDCQ7I 1.
@00973 CSDCbQ7L 1.
@00974 CSDCQ7E 1.
@00975 CSDCQ7C 1.
@00976 CSDCQ7K 1.
@00977 CSDCQ8 2.
@00979 CMSCbS01 2.
@00981 CMSCbS02 3.
@00984 CMSCQ1 1.
@00985 CMSCQ2 1.
@00986 CMSCQ3 1.
@00987 CMSCQ4 1.
@00988 CMSCQ5 1.
@00989 CMSCQ6 1.
@00990 CMSCQ7 1.
@00991 CMSCQ8 1.
@00992 CMSCQ9 1.
@00993 CMSCQ10 1.
@00994 CMSCQ11 1.
@00995 CMSCQ12 1.
@00996 CMSCQ13 1.
@00997 CMSCQ14 1.
@00998 CMSCQ15 1.
@00999 CMSCQ16 1.
@01000 CMSCQ17 1.
@01001 CMSCQ18 1.
@01002 CMSCQ19 1.
@01003 CMSCQ20 1.
@01004 CMSCQ21 1.
@01005 CMSCQ22 1.
@01006 CMSCQ23 1.
@01007 CMSCQ24 1.
@01008 CMSCQ25 1.
@01009 CMSCQ26 1.
@01010 CMSCQ27 1.
@01011 CMSCQ28 1.
@01012 CMSCQ29 1.
@01013 CMSCQ30 1.
@01014 CMSCQ31 1.
@01015 CMSCQ32 1.

```
@01016   CMSCQ33    1.
@01017   CMSCQ34    1.
@01018   CMSCQ35    1.
@01019   CMSCQ36    1.
@01020   CMSCQ37    1.
@01021   CMSCQ38    1.
@01022   CMSCQ39    1.
@01023   CMSCQ40    1.
@01024   CMSCQ41    1.
@01025   CMSCQ42    1.
@01026   CMSCQ43    1.
@01027   CMSCQ44    1.
@01028   CMSCQ45    1.
@01029   CMSCQ46    1.
@01030   CMSCQ47    1.
@01031   CMSCQ48    1.
@01032   SHAREFLG    1.
@01033   XSECFLG    1.
@01034   LONGFLG    1.
@01035   SC_ID94   $10.
@01045   SC_ID96   $10.
@01055   SC_ID98   $10.
@01065   CIDHD01    6.
@01071   CHILDID    6.
@01077   CWTCW01C    9.4
@01086   CWTCW01L    9.4
@01095   SHXSECWT    9.4
@01104   SHLONGWT    9.4;
```