

Youth in Transition Survey (YITS)
Cohort B – 26 to-28 year-olds
Cycle 5

USER GUIDE

June 2009

Statistics Canada Statistique Canada

Canada

TABLE OF CONTENTS

1.0 Introduction 4

2.0 Background 5
2.1 YITS Component 5
2.2 Reference Documentation 6
2.3 Objectives 6

3.0 Concepts and Definitions 7
3.1 YITS Concepts 7
3.1.1 Move to the U.S./Return to Canada (Module A) 8
3.1.2 Education and School Activities – high school, junior high or elementary (Module B) 8
3.1.3 Postsecondary Education (Modules H, K, M and N) 12
3.1.4 Postsecondary Engagement (Module K) 15
3.1.5 Financing Postsecondary Education (Modules L and M): 15
3.1.6 Education and Work Aspirations (Module M) 16
3.1.7 Loans, Tax Incentives and Debt (Modules L and M) 16
3.1.8 Health – Activity Limitations (Module M) 16
3.1.9 Support from Others (Module N) 16
3.1.10 Employment (Modules P1, P2 and P5) 16
3.1.11 Courses or Training Programs Related to Job or Career (Module P6) 21
3.1.12 GAPS - Months not in school full-time and not working (Module PS) 21
3.1.13 Volunteering (Module Q) 22
3.1.14 Skills (Module R) 22
3.1.15 Personal Characteristics and Family Background (Modules U and UNK) 22
3.1.16 Income (Modules V, VI et VIT) 24

4.0 Data Collection and Processing 26
4.1 The Questionnaire 26
4.2 Training 26
4.3 Supervision and Control 26
4.4 The Interview 26
4.5 Follow-up of Non-response 26
4.6 Data Capture 27
4.7 Minimum Completion Requirements 27
4.8 Computer Assisted Interview (CAI) Editing 27
4.9 Head Office Processing 28

5.0 Derived Variables and Codebooks 31
5.1 Cycle 5 Codebooks 31
5.1.1 Person Level Main File Codebook 32
5.1.2 Education Above High School – Institution Roster Codebook 42
5.1.3 Education Above High School – Program Roster Codebook 43
5.1.4 Postsecondary Engagement Roster Codebook 45
5.1.5 Confirmation of Open Jobs from Cycle 4 Roster Codebook 46
5.1.6 Job Roster Codebook 47
5.1.7 Job Details Roster Codebook 48
5.1.8 Dependent Children Codebook 49

6.0 YITS Scales 50
6.1 Defining Scales and Their Uses 50
6.1.1 What is a Scale? 50
6.1.2 Why Use a Scale? 50
6.1.3 What Type of Scales are Used in YITS? 50
6.1.4 Response Bias 51
6.1.5 Negative and Positive Questions 51
6.2 Scale Development 52
6.2.1 Investigation of Model Validity 52
6.2.2 Estimation of the Scores 54
6.2.3 Scale Reliability 55
6.2.4 Testing Scale Validity 55
6.3 Social Support Scale 56
6.3.1 Description of Social Support Scale 56
6.3.2 Model Validity 56
6.3.3 Estimating Scores 57
6.3.4 Scale Reliability 57
6.3.5 Testing Scale Validity 58
6.4 Scale References 59

7.0 Survey Methodology 61
7.1 LFS Sample Design 61
7.2 LFS Target Population 61
7.3 LFS Stratification 61
7.4 LFS Cluster Delineation and Selection 62
7.5 LFS Dwelling Selection 63
7.6 LFS Sample Rotation 63
7.7 Household Members Eligible for the LFS 63
7.8 YITS Target Population and Sample Design 63
7.9 YITS Sample Size By Province 64

8.0 Data Quality ..... …65
8.1 The Frame 65
8.2 Measuring Sampling Error 68
8.3 Non-sampling Error 69
8.4 Response Rates 70

9.0 Imputation of Missing Data for Income and Earnings Variables 72

10.0 Guidelines for Tabulation, Analysis and Release 74
10.1 Rounding Guidelines 74
10.2 Sample Weighting Guidelines for Tabulation 75
10.2.1 Definitions of Types of Estimates: Categorical vs. Quantitative 75
10.2.2 Tabulation of Categorical Estimates 76
10.2.3 Tabulation of Quantitative Estimates 76
10.3 Guidelines for Statistical Analysis 76
10.4 CV Release Guidelines 77

11.0 Weighting……..... 79

12.0 Variance Estimation 80

13.0 Working with YITS Files 81
13.1 Roster and Flat Files 81
13.2 Youth In Transition Survey: Data Extraction Tool 81
13.2.1 About Youth in Transition Survey (YITS)..... 81
13.2.2 Statistical Activity..... 81
13.2.3 Purpose of the Application..... 82
13.2.4 Saving and Loading Queries..... 89

APPENDICIES

APPENDIX A – Cycles 1 to 5 – New "Other – Specify" Categories..... 90

APPENDIX B - Cycle 1 - Module H variables..... 92

LINKS TO REFERENCE DOCUMENTS 94

OTHER DOCUMENTATION AVAILABLE ON REQUEST..... 95

1.0 Introduction

The Youth in Transition Survey (YITS) is a longitudinal survey designed to provide policy-relevant information about school-work transitions and factors influencing pathways among education, training and work. Cycle 5 of the survey - for the cohorts aged 23 and 26 to 28, was conducted by Statistics Canada between January and June 2008 with the co-operation and support of Human Resources and Skills Development Canada (HRSDC).

This User Guide for cycle 5 of the YITS 18-20 year-olds is developed for the fifth release of the microdata file. Throughout this document, this cohort may also be referred to as Cohort B or Cohort 2, 18 to 20 year-olds or 26 to 28 year-olds (their age as of December 2007).

A separate User Guide has been written for the Cycle 5, Cohort A - 15 year-olds who were 23 years of age as of December 2007.

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

At Statistics Canada:

Client Services
Centre for Education Statistics, Statistics Canada
2000 Main Building
150 Tunney's Pasture Driveway
Ottawa, Ontario
K1A 0T6
Telephone: (613) 951-7608
Toll free: 1-800-307-3382
Fax: (613) 951-9040
E-mail: educationstats@statcan.gc.ca

At Human Resources and Skills Development Canada:

National Learning Policy Research Group
Learning Policy Directorate
Strategic Policy and Planning
140 promenade du Portage
Phase IV, 3rd Floor
Gatineau, Quebec
K1A 0J9
Tel. : (819) 994-4566
Fax: (819) 997-5433
E-mail: NC-SP-PS-LPDR-DGPAR-GD

2.0 Background

Starting in 1996, Human Resources, Skills Development Canada (HRSDC), and Statistics Canada began developing the Youth in Transition Survey (YITS). Consultations took place with representatives from federal government departments with an interest in youth policy, provincial ministries and departments of education and labour, practitioners working directly with youth (teachers, counsellors, school board personnel and social workers), employers, business and education associations, academic researchers, youth and parents[1]. The result of these consultations was the development of the YITS as a longitudinal survey to collect policy-relevant information on the school-work transitions of young people, and the factors that influence such transitions.

The Youth in Transition Survey is composed of a family of surveys. These surveys provide a set of information from which complex data analysis between the various files and cycles can be undertaken.

The 15 year-old respondents of the Reading Cohort (2000) participated in both PISA and YITS. They are followed-up longitudinally every two years by YITS, which includes the cohort of 18 to 20 year-olds. Note that in each cycle the cohorts will have aged two years.

2.1 YITS Component

YITS is designed to examine the patterns of, and influences on, major transitions in young people's lives, particularly with respect to education, training and work. Human Resources and Skills Development Canada and Statistics Canada have been developing the YITS in consultation with provincial and territorial ministries and departments of labour and education. Content includes all formal educational experiences and most labour market experiences, achievement, aspirations and expectations.

The results from the Youth in Transition Survey have many uses. Human Resources and Skills Development Canada can use them to aid policy and program development. Other users of the results include educators, social and policy analysts, and advocacy groups. The information will show how young adults are making their critical transitions into their adult years.

These researchers and analysts will have access to important information that can be used in developing programs to deal with both short-term and long-term problems or barriers that young adults may face in their pursuit of higher education or in gaining work experience. Information from the survey will help to evaluate the effectiveness of existing programs and practices, to determine the most appropriate age at which to introduce programs, and to better target programs to those most in need.

Young adults themselves will be able to see the impact of decisions relating to education or work experiences. They will be able to see how their own experiences compare to those of other young adults.

[1]For more information about the consultation process and other aspects of YITS, see *Youth in Transition Survey Project Overview – T-00-5E (September 2000)* (Ottawa: Human Resources and Skills Development Canada, 2000, Cat. No. MP32-30/00-5E/F)

2.2 Statistics Canada Reference Documentation

Reference documentation for both cohorts for Cycles 1 through 5, can be found on the Statistics Canada website www.statcan.gc.ca. The Meta database provides information for both cohorts under 'Definitions, Data Sources and Methods':

- The 15 year-olds Reading Cohort (2000) PISA (Survey 5060) and YITS (Survey 5058);
- 18-20 year-olds cohort YITS (Survey 4435)

2.3 Objectives

The broad objectives of the Youth in Transition Survey are:

- to examine key transitions in the lives of youth, such as the transition from high school to postsecondary schooling and the initial transition from schooling to the labour market;
- to better understand educational and labour market pathways and the factors influencing these pathways;
- to identify educational and occupational pathways that provide a smoother transition to the labour market;
- to examine the incidence, characteristics, factors and effects of leaving school;
- to understand the impact of school effects on educational and occupational outcomes;
- to examine the contribution of work experience programs, part-time jobs, and volunteer activities to skill development and transition to the labour market;
- to study the attitudes, behaviours, and skills of young people entering the labour market;
- to gain a better understanding of the determinants of postsecondary entry and postsecondary retention, including education financing;
- to better understand the role of educational and labour market aspirations and expectations in investment in further education and career choice; and,
- to explore the educational and occupational pathways of various sub-groups, particularly youth "at risk".

Objectives for Cycle 5 and future cycles are:

- to confirm data at attendance of educational institutions collected in previous cycle;
- to follow youth as they move to accommodate the attendance at educational institutions and acceptance of employment;
- to identify use of government sponsored student loans
- to determine the awareness and use of tax incentives

3.0 Concepts and Definitions

Major data elements

This section outlines the major concepts and definitions of interest to the users of the YITS microdata file. The reference period for the data collected in the YITS is up to December 2007. The only exceptions are the personal characteristics and family background data, because respondents were asked for their current situation, in other words, as of the date of the interview.

Section 3.1 describes the concepts included in the YITS and *Section 5.0* describes derived variables included on the data file and in codebooks.

3.1 YITS Concepts

There are many concepts that are pertinent to understanding the YITS questionnaire. They have been categorised under different questionnaire modules pertaining to education, health, training, employment, volunteer, skills, income, and personal characteristics and family background.

Cycle 5 of YITS brings forward historical data from Cycle 4 and asks respondents to confirm the information or to update to previous information (names or dates). Comparison can then be made between the cycles' different reference periods (Cycle 1 – as of December 1999; Cycle 2 – January 2000 to December 2001; Cycle 3 – January 2002 to December 2003; Cycle 4 – January 2004 to December 2005; and Cycle 5 – January 2006 to December 2007).

Sub-Section	Description	Modules
3.1.1	Move to the United States/Return to Canada	A
3.1.2	High school, junior high or elementary education Status	B
3.1.3	Postsecondary education	H, K, M and N
3.1.4	Postsecondary Engagement	K and KE
3.1.5	Financing Postsecondary Education	L and M
3.1.6	Loans and education and work aspirations	M
3.1.7	Loans, Tax Incentives and Debt	L and M
3.1.8	Health – activity limitations	M
3.1.9	Support from others	N
3.1.10	Employment: Labour market and job roster Employment details Reservation Wage	 P1 P2 P5
3.1.11	Courses or Training Programs Related to Job or Career	P6
3.1.12	GAPS - months not in school full-time and not working	PS
3.1.13	Volunteering	Q
3.1.14	Skills	R
3.1.15	Personal Characteristics and Family Background	U and UNK
3.1.16	Income	V, V1 and V1T

3.1.1 Move to the U.S./Return to Canada (Module A)

Respondents who live in the United States, or lived there anytime between January 2006 and December 2007 were interviewed. They may have moved to attend school, to work or to accompany a parent or spouse/partner. Visits and temporary vacation periods were not included. The dates of moves are essential to assist in clarifying factors such as education funding, level of income, educational levels based on institutions and programs enrolled in and support systems. Respondents may have had more than one move to and from the U.S., for example if they attended school and then returned to work in Canada during breaks. As well, they may have had periods of work in both the United States and Canada. Information on resident status is also collected.

Temporary Resident Status: the respondent may be living in the U.S. for a specified period to attend school on a student visa, to be in training with a business, or to work for a company (also referred to as "non-immigrant" status).

Permanent Resident: the respondent is entitled to remain in the U.S. for an undetermined period and to work and/or attend school (also referred to as a 'green card holder').

3.1.2 Education and School Activities – high school, junior high or elementary (Module B)

The high school education section of the YITS contains questions on the respondent's educational attainment and experiences.

Respondents are first asked to confirm information collected for the Cycle 4 reference period 2004 to 2005, then to provide information on their enrolment status from January 2006 to December 2007). Youth who report that they were not enrolled provide their reason for leaving school and the date at which this separation occurred. The respondents also state the highest grade level they have taken, the highest grade level they have completed, and whether they met the requirements for high school graduation. In addition, the survey asks all youth if they ever dropped-out of high school. Those who dropped-out of high school state the number of times they have done so.

Adult high school: education taken by adults to obtain their high school diploma or equivalencies, and can be taken in an alternative format or non-traditional setting (internet, at malls, television registration, correspondence or distance learning).

Alternate programs: high school programs flexible enough to accommodate students who, otherwise might dropout, students returning to school, pregnant teens, and adults. Students study at their own pace and sometimes study at non-traditional school sites, such as shopping malls or office buildings.

Alternative school: a school offering a provincially approved curriculum that uses different teaching methods or places, with the emphasis on teaching cultural identity. Parents frequently work with the teachers in the classroom and in planning programs.

Continuers – high school: respondents who were continuing their studies at high school as of December 2007. Respondents continuing with postsecondary are not classified as high school continuers.

Correspondence courses (Distance Education): an educational or training activity that does not require students to physically attend a school, college or university. Mail, radio and television or other media communications such as the Internet are methods used to deliver the instruction. These courses are recognised by the province and exams are conducted in accordance with provincial standards.

Education: Elementary / Junior high / High school: the responsibility for education in Canada rests with provincial and territorial governments. Each province and territory has developed its own system

for education, and the structure can differ from jurisdiction to jurisdiction. The following table illustrates the similarities and differences for most schools, as of 2006.

Province/ Territory	Pre-Grades	Primary/ Elementary	Junior High/ Intermediate/ Middle	Senior High/ Secondary/High School
Newfoundland and Labrador	5 year-old Kindergarten	1-6	7-9	Levels I-III (10-12)
Prince-Edward-Island		1-6	7-9	10-12
Nova Scotia	Pre-grade 1	1-6	7-9	10-12
New-Brunswick (English Sector)	5 year-old Kindergarten	1-5	6-8	9-12
New-Brunswick (French Sector)	5 year-old Kindergarten	1-8		9-12
Quebec General	4 and 5 year-old Kindergarten	1-6		Secondary 1-5
Quebec Vocational				Secondary 3-5
Ontario	4 and 5 year-old Kindergarten	1-8		9-12/ OAC*
Manitoba	4 and 5 year-old Kindergarten	1-8		Senior 1-4 (9-12)
Saskatchewan	5 year-old Kindergarten	1-5	6-9	10-12
Alberta	5 year-old Kindergarten	1-6	7-9	10-12
British Columbia	5 year-old Kindergarten	1-7		8-12
Yukon	5 year-old Kindergarten	1-6	7-9	10-12
Northwest-Territories	5 year-old Kindergarten	1-6	7-9	10-12
Nunavut	5 year-old Kindergarten	1-6	7-9	10-12

* Grade 13/OAC was phased out in 2002 – 2003

Elementary school: the educational structure varies across the provinces and territories (see definition for ***Education: Elementary / Junior high / High school***). The elementary school level is the first level of instruction of children in the current school system. In general, at the elementary grade level, education is general and basic, and as a minimum includes kindergarten through grade six.

Ever dropped out: question is asked of: high school graduates who at some point dropped out of school, but returned to continue their education until graduation; high school continuers who at some point dropped-out of school, but returned to continue their education; and school leavers are those who have never graduated.

Full-time/part-time school status: full-time/part-time status is determined by the educational institution. All schools classify their students as being full-time or part-time depending on the number of courses in which they are enrolled.

Full-time schooling: full-time schooling is schooling or courses taken as a full-time student (see ***Full-time/part-time studies***).

Graduates – high school: respondents who have completed the minimum requirements for a high school graduation certificate, diploma or equivalent are considered to have graduated (see definition for ***High school graduation***). Some people might still take courses even after they have obtained their graduation certificate, for reasons such as upgrading marks, or taking courses not taken previously.

High school: the educational structure varies across the provinces and territories (see definition for ***Education: Elementary / Junior high / High school***). In general, at the high school level there is usually a choice of at least two programs: academic or vocational. Some secondary schools may specialise in vocational training (technical and commercial) but most high schools offer both academic courses (preparatory to university) and vocational courses, which prepare students either for an occupation or for further postsecondary non-university education.

High school graduation - diplomas and equivalencies: the following table lists the graduation diploma, certificate or equivalency awarded by province or territory:

Province/Territory	Graduation Diploma, Certificate or Equivalent
Newfoundland and Labrador	Senior High Graduation Diploma Provincial High School Graduation Certificate High School Equivalency Diploma General Educational Development (GED)
Prince Edward Island	High School Graduation Certificate General Educational Development (GED)
Nova Scotia	High School Graduation Diploma N.S. Secondary School Equivalency Certificate General Educational Development (GED)
New Brunswick (English and French Sectors)	High School Diploma or Certificate Adult High School Diploma General Educational Development (GED)
Quebec	Diplôme d'études secondaires (DES)/Secondary School Diploma (SSD, general education) Diplôme d'études professionelles (DEP)/ Diploma of Vocational Studies (DVS) (up to 1998 known as Secondary School Vocational Diploma (SSVD)) Attestation de formation professionelle (AFP)/Secondary School Vocational Certificate (SSVC) Attestation de spécialisation professionnelle (ASP)/Attestation of Vocational Studies (AVS) Attestations d'équivalence du niveau de scolarité (AENS/AESS)
Ontario	Ontario Secondary School Diploma (OSSD) General Educational Development (GED)
Manitoba	Provincial Diploma Adult Basic Education Certificate General Educational Development (GED)
Saskatchewan	Record of Secondary Level standing High School Equivalency Diploma General Educational Development (GED)

Province/Territory	Graduation Diploma, Certificate or Equivalent
Alberta	Alberta High School Diploma Certificate of Achievement High School Equivalency Diploma for Mature Students General Educational Development (GED)
British Columbia	Certificate of Graduation or "Dogwood Diploma" School Completion Certificate General Educational Development (GED)
Northwest Territories	Northwest Territories Secondary School Graduation Diploma General Secondary School Graduation Certificate (phased out as of August 31, 1998) Advanced Secondary School Graduation Certificate (phased out as of August 31, 1998)
Nunavut	At this time Nunavut was still using the Alberta graduation standards
Yukon	Yukon Graduation Certificate School Completion Certificate

Junior high / Intermediate / Middle school: the educational structure varies across the provinces and territories (see definition for ***Education: Elementary / Junior high / High school***). A school forming a link between elementary and secondary education usually consists of grades 7 to 9, which is not common to all provinces/territories.

Leavers / Non-completers - high school: respondents who had not completed the high school graduation requirements, and were not attending elementary, junior high or high school as of December 31, 2007.

Private Elementary/ Secondary (high) school: in contrast to public schools, private schools are not publicly supported but receive funding from private individuals or groups (e.g., student tuition fees, religious groups). These schools, whether church-affiliated or non-sectarian, are operated and administered by private individuals or groups.

Secondary School Vocational Diploma (SSVD/DVS): This was a category for the highest level of education completed in Module B for respondents taking schooling in Québec. As of 1998, the Secondary School Vocational Diploma was replaced by Diploma of vocational studies (DVS). The same information applies, enabling a respondent to go on to practise a skilled or semiskilled trade.

It is considered to be at the high school level. In cycle 1, only a few respondents reported this in Module B, whereas the information was reported and captured in Module H as to the type of program a respondent was working towards, or in which they had received a degree, diploma or certificate. In subsequent cycles SSVD is collected in Module B only.

Also referred to as Diplôme d'études professionnelles (DEP) in Quebec, SSVD now DVS can take from 6 months to two years to complete. Requirements to begin this diploma are completion of Secondary III, although some exceptions may be granted. This education prepares people for employment in specific industries, such as Buildings and Public works, Motorized Equipment Maintenance, Forestry and Pulp and paper, Health Services and Beauty Care.

Work experience programs: programs or courses, combined with high school that provide students with hands-on experience while spending time with an employer, outside the classroom environment. They are part of the student's curriculum. The student receives credit for participation, and may or may not be paid for the work they do. These programs do not include field trips.

Work experience/preparation programs have various names by province or territory such as school-to-work program, Co-op education, entrepreneurship education, youth apprenticeship, bridges transition-to-work programs, practical and applied arts program, work study component, trade program or information technology. Province/territory-specific examples are provided in the questions.

3.1.3 Postsecondary Education (Modules H, K, Mand N)

In Module H, respondents are first asked to confirm information on postsecondary education (PSE) from Cycle 4, to determine eligible institutions and programs for the reference period 2006 to 2007 which they may:

- still be enrolled in;
- have graduated from;
- have left; or
- have changed (program or main field of study).

Eligible institutions and programs from Cycle 4 begin the roster of postsecondary education in Module H and are referred to as "open".

The Postsecondary Education sections contain questions on the respondent's participation in any education, above the high-school level and more than 3 months in duration, that could be counted towards a degree, certificate or diploma from an educational institution. An **eligible Postsecondary Program** is one that:

- is above the high school level;
- is towards a diploma, certificate or degree;
- would take someone three months or more to complete; and
- the respondent should have started the program before January 2006.

All respondents, who in December 2007, were no longer in high school, or who had completed the high school graduation requirements were asked if they had taken any postsecondary education prior to January 2008.

The order of institutions and programs, within an institution, were provided by the respondent. Respondents were asked to identify institutions and programs beginning with the most recent. Researchers may want to select one institution in particular on which to do analyses, and may want to look at a particular derived variable such as HLATTD5 that indicates the status of the respondent at that institution as being a "continuer" or "non-continuer" (see *section 5.0, Derived Variables*).

Trades programs offered through apprenticeship, vocational schools or private trade schools do not always require high school graduation. Such education is considered as postsecondary.

Module H collects information on the type of postsecondary education:

1) Trades certificate or diploma from a vocational or apprenticeship training;
2) Non-university certificate or diploma from a community college, CEGEP, school of nursing, etc.;
3) University certificate below bachelor degree;
4) Bachelor degree; and
5) University degree or certificate above bachelor degree.

Youth, who report being enrolled in a postsecondary program, are asked for the number of institutions attended, and the number of programs taken altogether. Information collected includes the type of degree sought and the duration of the program. For each program, the survey gathers data on primary and secondary fields of study. Survey staff used this information to derive the Major Field of Study using Classification of institutions and programs (CIP)[2] coding. CIP codes are available on the data file. The youth also report on attitudes and behaviours while taking postsecondary education.

[2] See section 4.9

Apprenticeship Programs: lead to journeyman status in several designated trades. Skills and knowledge are provided through on-the-job experience (components) with short periods of formal instruction. Some examples of apprenticeship trades are auto mechanic, hairdresser, boilermaker, steamfitter, millwright, electrician, plumber, machinist, chef.

On-the-job training and the formal schooling are all counted as part of the formal education in order to obtain a license in a trade.

College Post-Diploma or Graduate Program: is a relatively new type of program offered by some colleges. Students usually require a previous college diploma or university degree for admission. Examples of these types of programs are "a certificate in telecommunication management" or "a certificate in international business administration".

Commercial school: private schools that receive no public funding and are licensed by a province or territory. They engage in providing professional and vocational training for profit.

Community Colleges: includes community colleges, colleges of applied arts and technology (CAATS in Ontario), "collèges classiques" or CEGEPS in Quebec, technical institutes, hospital and regional schools of nursing, or teachers' college and establishments providing technological training in specialised fields. Community colleges offer career programs of one to four years. Some also provide one- or two-year academic programs which prepare a student to proceed to university.

Continuers – Postsecondary: respondents who were continuing their studies towards the completion of a postsecondary program in December 2007.

Degrees - First Professional: a first professional degree may be taken part way through or after a university bachelor's degree. Examples of this type of degree are medicine, dentistry, veterinary medicine, law, optometry and divinity. Engineering is not considered a first professional degree, rather it is a professional licence.

Degrees versus Diplomas: are different types of PROGRAMS, but the word diploma is sometimes used (incorrectly in English) to refer to either a degree or diploma. Most degrees (but not all) are for a program of study at a university. If the official name of the qualification contains the word "degree", "Bachelor's, Master's or Doctor of"; it is a degree. Diplomas are less common from a university, but more common from other institutions such as colleges.

Distance Education or Correspondence program: an educational or training activity that does not require students to physically attend a school, college or university. Mail, radio and television or other media communications such as the Internet are methods used to deliver the instruction. These courses are recognised by the province or territory and exams are conducted in accordance with provincial/territorial standards.

Eligible program: to be deemed eligible, a postsecondary program must meet the following criteria: the program is above the high school level; the program is towards a diploma, certificate or degree; the program would take someone three months or more to complete; and the respondent started taking the program before January 2008. If at least one program within a given institution has been deemed eligible, then the institution itself is deemed eligible.

- Eligible programs include: diplomas, degrees, certificates or licenses obtained through professional associations such as in accounting, banking, real estate or insurance.
- Ineligible programs include: non-professional health certificates such as St. John's First Aid, Red Cross; continuing education or personal interest courses.

Ever dropped out: question is asked of:

- Postsecondary graduates who at some point dropped out of their program, but returned to continue their education until graduation;
- Postsecondary continuers who at some point dropped-out of their program but returned to continue their education; and
- school leavers - those who never graduated.

Fellowship: A Fellowship is a position in a university held by a graduate student having teaching duties as part of his or her educational program.

Full-time/part-time school status: full-time/part-time status is determined by the educational institution. All schools classify their students as being full-time or part-time depending on the number of courses in which they are enrolled. Hence, whether a person was marked full-time or part-time depends on how he/she was classified by the institution attended.

Full-time schooling: full-time schooling is schooling or courses taken as a full-time student (see ***Full-time/part-time studies***).

Graduates – Postsecondary: respondents who have completed the graduation requirements towards a diploma, certificate or degree.

Leavers / Non-completers – Postsecondary: respondents who had not completed the graduation requirements for their program, and were no longer taking courses toward the completion of the program in December 2007.

Licence (Quebec): Licence, Licentiate, Testamur are credentials awarded mainly by religious programs in Quebec. The term 'Licence' does not include professional association licences, and are to be specified under "other". Interviewers were asked to identify whether the licence was at a graduate or post-graduate level.

On-the-job experience program (Module H): programs or courses, combined with postsecondary study, which provide students with hands-on experience while spending time with an employer, outside the classroom environment. They are part of the student's curriculum. The student receives credit for participation, and may or may not be paid for the work they do. For respondents who have participated in such a program, the type of program is collected (e.g., Co-op program, Apprenticeship, Trade/vocational, or another type).

Private training institution (Module H): privately owned schools that are profit oriented and are engaged in providing professional and vocational training, and are licensed by the province or territory.

Programs – Postsecondary (Module H): includes programs lasting three months or longer and are above the high school level.

- University programs leading to bachelors, master's or doctoral degrees, or specialised certificates or diplomas.
- Programs offered at CEGEPs, community colleges, technical schools, hospital schools of nursing and similar institutions (towards a diploma, certificate or degree) normally requiring secondary school completion or its equivalency for admission.
- Police Academies; RCMP colleges and training camps; Firefighters' training.

Trade /vocational certificate or diploma: this term is used to classify skill courses that prepare trainees for occupations NOT at the professional or semi-professional levels. A trade-vocational program prepares people for employment in a specific occupation such as a heavy equipment operator, automotive mechanic and upholstering. Many community colleges or technical institutes offer certificates or diplomas at the trade level.

University: an independent institution granting degrees in at least arts and sciences.

University College: A University College is a college that is an integral part of a university, governed by the university Administration. Respondents would attend university-level courses at the college.

Vocational or Trade School: Technical and trades training varies between and within provinces or territories. It is offered in both public and private institutions such as community colleges, institutes of

technology, trade schools and business colleges. It may also take place on the job, in apprenticeship programs or in industry training programs.

3.1.4 Postsecondary Engagement (Modules K and KE)

For Cycle 2, Module K was revised to accommodate the possible collection of zero, one or two sets of postsecondary engagement questions. The goal, originally, was to acquire information with respect to the respondent's first postsecondary experience. For many of the respondents, this would simply be the first institution above the high school level that they have attended. However, based on information already obtained during the YITS Cycle 1 data collection, it was found that students attending CEGEP institutions and students attending NON-CEGEP institutions (e.g., university, community college outside of Québec, etc) have two distinct postsecondary experiences (in terms of cost and distance away from home). This motivated the decision to attempt to collect two sets of postsecondary engagement questions for students studying in Québec (1 CEGEP and 1 first non-CEGEP).

3.1.5 Financing Postsecondary Education (Modules L and M):

All respondents who had taken a postsecondary program were asked about their sources of income and the amounts of income used to finance their postsecondary education.

Bursary: refers toa monetary award to assist a student in the pursuit of his/her studies based on financial need and satisfactory achievement.

Grants: a gift (usually a sum of money) made by a government or corporation (as an educational or charitable foundation) to a beneficiary on the condition that certain terms are accepted or certain engagements fulfilled.

Scholarships, award or prizes: refers to monetary award to assist a student in the pursuit of his/her studies, based usually on outstanding academic achievement rather than on financial need.

Funding also includes government-sponsored student loans, money from family or relatives that does not have to be paid back; money from trust funds, RESPs or RRSPs; and/or money from jobs or from personal savings.

Respondents are asked, if they have government student loans if they have started to pay them back, frequency of payments, if they are aware of and have used government programs available to assist in repayment of their debt..

Assistance from government on repayment of student loans may include Interest Relief, Loan Remissions, Debt Reduction and Revision of Terms. Cycle 5 is the first cycle of YITS to collect this information which was focus group tested in April 2007.

Information on these assistance programs is available on the HRSDC website at www.hrsdc.gc.ca under Policies and Programs.

3.1.6 Education and Work Aspirations (Module M)

The survey asks youth the level of education they would like to get, and the level they think they will get. Respondents are then asked if there is anything standing in their way of going as far in school as they would like. The list includes barriers such as financial situation, not being accepted into a program, wanting to stay close to home, caring for children, etc.

3.1.7 Loans, Tax Incentives and Debt (Modules L and M)

Respondents are asked about government sponsored student loans to see if there have been barriers to further education based on loans not being approved or not receiving a sufficient amount to cover school costs. As well, beginning in Cycle 5, respondents are asked if they are aware of and have used certain fiscal incentives like tax credits for post-secondary education. Debt load is collected on cost of living such as payment for residence, personal loans, or credit card or line of credit debt.

Tax credits for PSE include Education Tax Credits, Tuition Fee Tax Credits, Textbook Tax Credits and Student Loan Interest Tax Credits. This information was obtained from the Canada Revenue Agency website at www.cra-arc.gc.ca under Registered Education Savings Plans and Tuition, Education and Textbook Amounts (Line 323).

3.1.8 Health – Activity Limitations (Module M)

All respondents provide information about any long term physical condition(s), mental condition or health problem(s) that limit the kind or amount of activity they can do at school or at work**. Long term** condition(s) were defined as those that have lasted or are expected to **last six months or more**. The purpose of the disability questions is not to determine the nature of the condition so much as the impact on activities, particularly at school and at work.

3.1.9 Support from Others (Module N)

Support from others has an impact on success in education and in the labour market. This section asks respondents to consider types of support they may receive. (*See Section 6.0 – YITS Scales*)

3.1.10 Employment (Modules P1, P2 and P5)

Determining the type of labour market data to be collected by the YITS presented a challenge. For example, to measure school-work transitions, it was necessary to collect the first job at which the respondent worked after leaving full-time schooling. Measuring all jobs since this event was not feasible. However, as most youth of the younger cohort are still in full-time schooling or had left not too long ago, measuring work activities during the current cycle's reference period represented a good alternative. For those who had left full-time schooling prior to 2008, additional questions on the first job after leaving full-time schooling were asked in order to capture the "transition" job. In addition, for those who had not worked in the current cycle reference period, and had not held a job after leaving full-time schooling, questions were asked to determine if they had ever worked and if so, certain details on that job were collected.

Three different categories of labour market activity were collected: employee jobs, self-employment jobs, and jobs at the farm or business of a family member living in the same household.

Youth's labour market experiences were captured, in addition to the first job after leaving full-time schooling, if this event occurred prior to 2007. "Open" jobs from the previous cycle were confirmed for eligibility. Start and stop dates were collected for a maximum of six jobs the respondent worked at between January 2006 and December 2007. (Module P1). This job count included previous cycle "open" and eligible jobs and the current cycle eligible jobs.

Open jobs were those at which a respondent worked during the previous cycle. The respondent may have still been employed but not actually working at the job prior to January 2006 (temporary layoff, business slowdown, etc.). Respondents are asked if they are still working with that employer in Cycle 5 and/or whether they had returned to work for the employer. (See *Eligibility, job*).

Further job details were collected on a maximum of four jobs the respondent worked at during the Cycle 5 reference period (Module P2). These restrictions were imposed to limit the time of interview and minimise respondent burden.

Apprenticeship: The employer undertakes, by contract, to employ and train an apprentice under the supervision of a qualified journeyman. To become an apprentice, there is usually a formal registration process with a provincial or territorial Ministry or Department or Trade organization.

Bonuses (Module P2): in some situations, wages are paid in the form of both regular pay cheques and periodic bonuses based on work performance. In these cases, the bonuses should be averaged over the period for which it applies and included with the wages or salary reported. (See ***Wages or Salary***).

Business, for self-employed persons (Module P2): for self-employed persons, a business exists when one or more of the following conditions are met:

- an office, store, farm or other place of business is maintained and is used exclusively for conducting the business;
- or the enterprise is incorporated (see definition of *Incorporated Business*);
- or the self-employed person usually has paid help;
- or land, buildings, machinery or equipment in which the person has invested money is used by respondents or their employees solely in conducting the business.

Examples of self-employed persons **WITH a business** would be:

- a person with their own beauty salon(s);
- a person with a medical practice;
- someone who sub-contracts from someone else.

Examples of self-employed persons **WITHOUT a business** would be:

- a cleaning person working for a number of people in their homes;
- a freelance writer, a tutor, general handyman or a babysitter who regularly works for a number of people.

Breaks from jobs: Although these questions were removed in Cycle 4, they were reinstated in Cycle 5. If breaks are longer than 4 weeks, respondents are asked for a maximum of 3 breaks per job between January 2006 and December 2007.

Class of Worker: There are three main categories of worker that are defined further on in this section – Paid Worker (an employee); Self-employed Worker; Unpaid worker in the family farm or business.

Dates of jobs:

Start date of job (Module P1): if the respondent...

- works for the same employer on a “seasonal” basis, the date first started work is the date of the most recent period of uninterrupted work, not the date when he/she first began to work for this employer.
- is a paid worker, who works strictly on-call, the date first started work is the date in which the most recent period of work began. Note: Paid on-call workers are only considered to have a job in those months in which some work was done. Any period of one month or more in which no work was done is considered a break in employment for on-call workers, and hence, counted as separate jobs.
- is a paid worker who seeks and obtains employment only at certain times of the year (e.g., students who only work in the summer months), the date first started work is the date when

he/she last began to work for this employer, even is he/she has worked for the same employer previously (e.g., last three summers).
- is self-employed with a business, the date first started work is the date when he/she created or acquired the business. For self-employed persons with a business, periods of inactivity are not considered as breaks in employment.
- is self-employed without a business, the date first started work is the date in which the most recent period of continuous work began. Self-employed persons without a business are only considered to have a job in those months in which some work was done. Any period of a month or more in which no work was done is considered a break in employment, and hence, the next work period becomes a separate job.

End date (Module P1): if the respondent…
- is a paid worker with a definite schedule of work, the date last worked is the month and year the respondent last worked at his/her job prior to January 2008. Respondents who have had paid leaves from their employer, such as vacation, training or sick leave are included as having worked.
- is a paid worker without a definite schedule of work, the date last worked is the month and year the respondent last worked at this job.
- is self-employed with a business, farm or professional practice, the date last worked is the month and year they ceased the operations of their business, or the business closed down, or December 2007 if they still operated the business at that time.
- is self-employed without a business, the date last worked is the month and the year in which they last did any work.

Eligible Job (Module P1): to be deemed eligible, a job collected for the 2006 to 2007 reference period must meet the following criteria. If the respondent was still working at that job from Cycle 4 ("open" job) in the Cycle 5 reference period or if the job began in 2006 or 2007, the respondent had to be able to provide the job's start and end dates. If, at the time of interview, the respondent was still working at the job, the end date for that job was set to December 31, 2007. Any eligible jobs would remain "open" for the next cycle.

Employee, Paid (Modules P1 and P2): a person who works for others (i.e. works for an employer) and receives a wage or salary. The employer usually deducts and remits from the wage or salary income tax, Canada/Quebec Pension Plan premiums, etc. There are cases where persons receive a wage or salary but no deductions are made for tax or EI/CPP because the wages earned are too low. (See ***Self-employed*** and ***Unpaid family worker***.)

Employers (Modules P1 and P2): Are those persons or companies/businesses for whom the respondent did any paid jobs whether part-time or full-time.

First Job *(Module P1):* First job after leaving full-time schooling identifies the job a respondent held at the time of leaving full-time schooling or the job s/he first started after leaving full-time schooling. During survey collection, respondents who were no longer full-time students in December 31, 2007 were asked to report the first job they worked at after leaving full-time schooling. A procedure was then created to validate the job reported by respondents as their first job after leaving full-time schooling and/or when not reported, to identify one of the other jobs reported as being "first job"

Full-time employment (Modules P1 and P2): consists of persons who usually work 30 hours or more per week at their job.

Hours of work (Module P2) – Usual number of hours worked -
- Number of paid hours usually worked is asked of employees.
- Number of hours usually worked (paid not part of the question) is asked of self-employed workers and unpaid workers in the family farm or business.

For people who do not work a fixed number of hours, usual hours of work mean the average number of hours during a four-week period. In the survey, usual hours of work are collected for two reference

periods. The first reference period is when the respondent last worked at his/her job, and the second is when the respondent first worked at his/her job.

For self-employed workers, number of hours worked include time spent on work-related activities in addition to time actually spent on producing goods or providing services. These related activities include: time spent actively looking for potential clients; preparing estimates, quotes or tenders; time spent on operating a business; professional practice or farm even if no sales were made; no professional services were rendered or nothing was actually produced; time spent on activities related to establishing a new business, farm or professional practice; and/or time spent by a person who owns and manages his/her business or farm even though he/she is physically unable to do the actual work.

Incorporated business (Module P2): refers to the legal status of a business, farm, or in some cases, professional practice. It is directed at persons who were self-employed. An incorporated business is a business or farm, which has been formed into a legal corporation, having a legal entity under federal or provincial/territorial laws. An unincorporated business or farm has no separate legal entity, but may be a partnership, family business or owner-operated business.

Industry (Module P2): the general nature of the business carried out by the employer for whom the respondent worked (when an employee, or an unpaid worker in the farm or business of the family), or for their own business (when self-employed).

Job leavers (Module P2): persons who were not working at their job as of December 31, 2007 and left that job voluntarily. That is, the employer did not initiate the termination. Detailed reasons collected are: own health reasons, pregnancy or caring for own children, personal or family responsibilities, going to school, changed residence, dissatisfied with job, found a new job, to concentrate on another job, or another reason.

Job losers (Module P2): persons who were not working at their job as of December 31, 2007 and left the job involuntarily, that is the employer initiated the termination. Detailed reasons collected are: company moved; company went out of business; seasonal or non-seasonal layoff; strike, fired, end of contract; or another reason.

Job/Work (Modules P1, P2, P5, P6 and PS) Any activity carried out by the respondent during the reference period for pay or profit, includes 'payment in kind' (payment in goods or services rather than money) whether actual payment was received during the reference period. Work includes time spent:

- actively looking for work or clients, preparing estimates, quotes or tenders, establishing a new business;
- operating a business, professional practice or farm even if no sales were made, no professional services were rendered or nothing was actually produced;
- as the owner or manager of a business even though the person is physically unable to do the actual work;
- on-the-job training; unpaid work for a family business or farm; odd jobs.

Method found job (Module P2): identifies the method used through which the respondent found the job. Methods include: through placement or posting at school, public employment agency, private employment agency, contacted employers directly or sent out resumes, through friends or relatives, placed an ad, answered a job add, or through another method.

Net income - ***for self-employed workers*** (Module P2): total earnings for all of 2007 are collected for the self-employed. Net business income is income after all business expenses have been deducted. (See ***Wages or Salary*** for employee.)

Occupation (Module P2): refers to the kind of work the person was doing at his/her job, as determined by the kind of work reported and the description of the most important duties.

Odd jobs (Module P1): odd jobs may be any type of work for pay and are defined as jobs done on the side to make money, or extra money. These jobs are mostly intermittent such as babysitting, tutoring, yard work, housecleaning, newspaper delivery, etc. **Note**: When a person baby-sits for more than one family, this is considered as one job only. Another example, if a person does many different "odd

jobs" to earn extra money, for example, baby-sits and mows neighbours lawns; this is also considered as one job only.

Paid worker (Modules P1 and P2): someone who works for others (i.e., works for an employer) and receives a wage or salary. The employer usually deducts and remits from the wage or salary income tax, Canada/Québec Pension Plan premiums, etc. There are cases where persons receive a wage or salary but no deductions are made for tax or EI/CPP because the wages earned are too low.

Part-time employment (Modules P1 and P2): consists of persons who usually work less than 30 hours per week at their job.

Permanent employees (Module P2): permanent employees work at a job for which there is/was no indication that the job would end at some definite point in time, e.g., hired permanently with no specified term. (See ***Temporary employee***.)

Reasons for leaving job (Module P2): asked for all jobs that ended prior to December 2007. (See ***Job losers*** and ***Job leavers***.)

Reservation Wage (Module P5): The questions asked in this section are required to understand if a respondent is willing to work just for money, or wants to have work that will pay them what they feel they are worth (from other job experience or from acquired educational skills). They may feel that although there are better job opportunities in another locale, they would not be willing to move to improve their job or career prospects.

Self-employed (Modules P1 and P2): includes:

- persons for whom the job consisted of operating a business or professional practice, alone or in a partnership. This includes operating a farm whether the land is rented or owned, working on a freelance or contract basis to do a job (e.g., architects, private duty nurses). It also includes operating a direct distributorship selling and delivering products such as cosmetics, newspapers, brushes and soap products, and fishing with own equipment or with equipment in which the person has a share.
- persons who do not have a business but who are paid directly by a client such as a child care giver; house cleaner; dog walker - with one or more clients, who provides these services on a formal or informal contractual basis.
- persons who work at "odd jobs" such as occasional babysitting, tutoring, shovelling neighbours entrances.

Temporary employee (Module P2): is an employee for whom there was a definite indication that the job would terminate at some specified point in time. For example, hired for a six month term or a student hired by the same employer during his/her summer holidays or school breaks. Often referred to as term or contract job by respondents. (See ***Permanent employee***)

Temporary help agency (Module P2): arranges for the job and the respondent is paid by the agency. For example: Bob does clerical work for Briggs Inc. He obtained this position through Bradshaw Associates, a temporary placement agency. He receives his pay from Bradshaw Associates, not Briggs Inc.

Tips and commissions (Module P2): paid workers may receive tips, commissions or bonuses in addition to their wage. However, it is likely that the tips, commissions or bonuses are paid on a less frequent basis than the regular wages or salary (e.g. weekly, monthly, etc.). In this case, the value of tips, commissions or bonuses earned are averaged over the period for which the respondent reported their wages or salary. For example, an hourly amount is determined by adding up the total amount of tips, commissions or bonuses received and dividing this by the number of hours worked in that period. This amount is included as part of the hourly rate of pay. (See ***Wages or Salary***)

Unpaid family worker (Modules P1 and P2): someone who works without pay on a farm or in a business owned and operated by another family member living in the same household. The work done must contribute directly to the operation of a family farm or family business. Excluded are respondents

who perform regular household chores around the house or yard (e.g. cutting the lawn, painting the house, cleaning the home). (See ***Employee* and *Self-employed***)

Unpaid leave from work (Module P2): the term unpaid leave from a job denotes a period of not-working during which the respondent did not receive any pay from the employer. The period was defined, for the purpose of the YITS as four consecutive weeks or longer. The respondent would normally receive a wage or salary from the employer had he/she worked, and may, during the unpaid break receive compensation from some other source such as Workers' Compensation. They are still considered an employee during that time. Unpaid leave periods from jobs were collected from "paid workers", i.e. they were not asked for self-employed jobs and unpaid work in family business or farm jobs.

Wages or Salary (Module P2): for employees, wages or salary are before taxes and deductions (i.e. employment insurance (EI), government pension plans (CPP/QPP), union dues, etc). The respondent chooses the pay period that makes it easier for him/her to give accurate data. For those respondents who choose to report on a yearly basis, the earnings must correspond to an **entire** year, even if the respondent has not worked for the full year (e.g., a respondent started a job a few months ago). The amount entered should reflect what the respondent would normally earn, had he/she worked for a full year.

The category "other" under method of reporting wages and salaries includes persons earning straight commission from their work.

Income from tips, bonuses and commissions are included and averaged over the period for which they apply and included with the reported wages or salary.

"Usual" wage or salary: "Usual" refers to a typical pay period. Where situations are unclear, "usual" pertains to a four week period. If the four week period was not representative of a usual month because the person was on holiday or sick, the respondent is asked for the average earnings under normal circumstances. (For income of the self-employed, see ***Net Income***).

Workfare (Ontario): is a provincial program in which participants exchange their labour services for social assistance payments. Such a worker is classified as an "employee".

3.1.11 Courses or Training Programs Related to Job or Career

(Module P6)

Regardless of whether a respondent worked in the last two years, they are asked for information on any courses or training programs related to a job or career. These programs might be sponsored by an employer or may have been taken to have better job opportunities in a current job, or in the labour market. If a program was made ineligible in Module H, the respondent is asked in Module P6 if the program might be applicable. A maximum of four courses or training programs are flagged for collection (2 for employer organized/sponsored and 2 for any other training related to a job or career).

Training, outside of formal educational programs and training courses taken to acquire skills for a job or career, might include reading books, manuals or other written materials, using materials available electronically; or watching others work, receiving advice or assistance from others.

3.1.12 GAPS - Months not in school full-time and not working

(Module PS)

For the reference period of January 2006 to December 2007, it was possible to determine during the interview the months in which the respondent had not been working or was not in school full-time – referred to as "gap" months. These "gap" months were derived from the start and end dates of jobs held during the year and from the months during 2006 and 2007, when the respondent was not in school full-time (high school or postsecondary).

For the last gap month, the respondent was asked whether he/she had done anything to look for work, and if so, to name the type(s) of activities.

3.1.13 Volunteering (Module Q)

Volunteer worker: someone who gives his/her unpaid time to a group or an organisation such as charities, schools, religious organisations or community associations. This includes unpaid community service that was done voluntarily, or as a school program, or in order to obtain assistance, or as part of a court sentence. Informal voluntary activities such as painting a neighbour's house or looking after someone's children or pets as a favour is excluded.

3.1.14 Skills (Module R)

Given the changing nature of the workplace and the emphasis on human resource development, the importance of providing skill assessment measures on the YITS was recognised. Due to survey length, such assessment had to be short in duration. Respondents were asked to self-assess (self-evaluate) six skills often used in the workplace, and those generally sought by employers. The skills assessed are ability to use a computer, ability to solve new problems, mathematical abilities, and writing, reading, and oral communication skills.

3.1.15 Personal Characteristics and Family Background (Modules U and UNK)

***Citizenship* :** refers to legal citizenship status of the respondent. Persons who are citizens of more than one country were asked to report this information. The concept of citizenship stems from the *Citizenship Act.* Persons may be Canadian by birth and yet hold the citizenship of another country. Persons may also be Canadian by naturalisation and hold citizenship of their country of birth or some other country.

Canada by birth
Persons born in Canada, and those born outside of Canada, if at the time of their birth, one or both parents were Canadian citizens **and** this person has retained Canadian citizenship.

Canada by naturalisation (citizenship process)
Persons who were landed immigrants and have been issued a Canadian Citizenship Certificate are considered Canadian citizens

Other country
Persons who hold citizenship of another country.

Cultural or racial background: refers to the ethnic or cultural group(s) to which the respondent's ancestors belong.

Dependent children: children for whom the respondent has sole or joint custody. Included are children for whom the respondent is financially responsible on a regular basis, even if they have infrequent or no contact. This includes birth, adopted, step and foster children from the previous and current YITS cycle.

Household: Household refers to a person or group of persons who occupy the same dwelling and do not have a usual place of residence elsewhere in Canada. It may consist of a family group with or without other unrelated persons, of two or more families sharing a dwelling, of a group of unrelated persons, or of one person living alone. Each person is a member of one and only one household.

Landed immigrant: persons who have been granted the right to live in Canada permanently by immigration authorities, but have not obtained Canadian citizenship. These persons are referred to as "permanent residents" under the Immigration Act.

Language: spoken well enough to conduct a conversation - languages in which the respondent can carry on a conversation at some length on various topics.

Although respondents may have declared that they learned two or more languages simultaneously, interviewers attempted to have these respondents choose one language over the other. However, in the few circumstances where respondents could not choose between English and French as their first language, the cases have been included in the derived variable in the category "Other: English and French."

Marital status: marital status (conjugal status) of respondent at time of interview. Marital status from cycle 3 is confirmed if respondent indicated either married or living common-law. For cycle 5 the respondent is asked, if there is a change in his marital status, how many relationships he has entered into since January 2004 and the date when his marital status changed. The categories are as follows:

Single (that is never married)
Persons who have never married and persons whose marriage has been annulled and who have not remarried.

Married
Persons who are legally married, and whose spouse is living.

Living Common-law or with a partner (girlfriend, boyfriend)
Refers to persons who live together as a couple but who are not legally married to each other. These persons can be of the opposite sex or of the same sex.

Widowed
Persons who have lost their spouse through death and who have not remarried.

Separated (still legally married)
Persons currently married, but who are no longer living with their spouse (for any reason other than illness or work), and have not obtained a divorce.

Divorced
Persons who have obtained a legal divorce and who have not remarried.

Permanently moved out: Respondents who have moved away from their family home and have a permanent address different from that residence are considered to have permanently moved out. They may, in time move back, or move to different addresses.

Province/Territory: the data file contains the province or territory that the respondent considers to be their address (PROVD5) and the province or territory where they attended a postsecondary institution (Module H – PSPROVD5).

Spouse or partner background: If a respondent is legally married or living common-law, they are asked about their spouse's/partner's highest level of education completed and for their spouse's/partner's current main activity. Details of kind of business are also required if the respondent's spouse or partner is employed.

Usual Place of Residence:

School Residence: rooms or apartments that are registered with the institution the respondent is attending.

House, apartment or other private dwellings: single detached dwellings, doubles or duplexes, row or terrace homes, low or high-rise apartments;

Somewhere else: institutions such as penitentiaries, group homes, nursing homes for the aged, hospitals, homes of religious orders, convents; and/or boarding houses, mobile homes, camps, colonies, houseboats, motor homes, hostels, hotels/motels, tourist homes.

Visible Minority: the concept of visible minority applies to persons who are identified according to the *Employment Equity Act* as being non-Caucasian in race or non-white in colour. Under the *Act*, Aboriginal persons are not considered to be members of visible minority groups.

3.1.16 Income (Modules V, VI, VIT)

Information collected in the income module is income **received** from all sources **during the year** by the respondent. It is not limited to monies that are taxable. The information refers to income or monies received in 2007 only (January 1, 2007 to December 31, 2007).

Canada Child Tax Benefit or Provincial/Territorial Child Tax Benefits or Credits: Reported only for the parent who received the cheque (the person with the lowest income).

Employment Insurance: refers to total Employment Insurance benefits received during the year, before tax deductions. It includes benefits for unemployment, sickness, maternity, paternity, adoption, job creation, work sharing, retraining and benefits to self-employed fishermen. As well, include retraining and retirement benefits received under the Human Resources and Skills Development Canada (employment insurance program).

Family Business – Farm or non-Farm self-employment NET income: This is receipts minus operating expenses, depreciation and capital costs allowances. If it is a partnership, the respondent only reports their own share. If the farm is incorporated, the income is reported in Wages and Salaries and/or Dividends. Net rent from farms/property leased to others is reported in Other Investment Income.

Goods and Services Tax Credit (GST) or Harmonized Tax Credit (HST): A person applies for this credit on their income tax return. A person may apply for the credit if they are 19 years of age or older, had a spouse, or were a parent. The credit is based on their net income, added to the net income of their spouse, if applicable.

Income from other government sources: refers to total income from transfer payments from federal, provincial, territoriall or municipal governments not reported individually, and received during the year. Included are benefits received under the Canada or Quebec Pension Plan (retirement pensions, survivor's benefits, disability pensions – lump sum death benefits are excluded); and Worker's Compensation.

Income from other non-government sources: included are sums received from investments, child support, and other non-government sources not reported in separate categories.

> **Income from investments**: includes interest from bank accounts and other deposits, net dividends and other investment income.
>
> **Income from child support**: includes payments made by a spouse during separation or by an ex-spouse following a divorce; all regular payments and occasional contributions towards the child's maintenance.
>
> **Other income**: includes income from royalties on books; rental income from other properties; income from roomers and boarders; non-refundable scholarships and bursaries; alimony; and strike pay.
>
> **Excludes:** tax-free RRSP withdrawals used for purchasing a home, proceeds from the sale of property, businesses, financial assets or personal belongings, loans received and repaid to you as a lender, and refunds of contributions to work-related pension plans.

Money from parents, guardians or other people: included are sums of money received from a parent or guardian or other people that the respondent does not have to repay. Monies received as loans (regardless of when they are to be repaid) are excluded from income.

Scholarships, grants or bursaries: income from scholarships, grants or bursaries may be provided either to the respondent or to the institution to pay for tuition. Money received from fellowships (a position in a university held by a graduate student having teaching duties as part of his or her educational program) is included.

> **Scholarship**: A monetary award to assist a student in the pursuit of studies, based usually on outstanding academic achievement rather than on financial need
>
> **Grant**: a gift (usually a sum of money) made by a government or corporation (as an educational or charitable foundation) to a beneficiary on the condition that certain terms are accepted or certain engagements fulfilled.
>
> **Bursary**: a monetary award to assist a student in the pursuit of studies based on financial need and satisfactory achievement.

Self-employment income: refers to net income (gross income minus expenses of operation such as wages, rents and depreciation) received during the year from the respondent's business or professional practice. In the case of partnerships, only the respondent's share was reported. Also included is net income from persons babysitting, operators of direct distributorships such as those selling and delivering cosmetics, as well as from freelance activities of artists, writers, music teachers, hairdressers, etc.

Social Assistance (welfare) or Provincial/Territorial Income Supplements: include payments from provincial, territorial or municipal programs for persons in need such as social assistance or welfare for:

- mothers with dependent children
- persons temporarily or permanently unable to work
- the blind and disabled
- benefits covering basic needs (food, fuel, shelter, clothing)
- benefits for special needs (education, respite care, companion services)
- payments from work incentive programs.

Wages and salaries: refers to gross wages and salaries (including income from commissions, tips and bonuses) before deductions for such items as income tax, pensions and Employment Insurance. Earnings received from all paid worker jobs held during the year are to be reported.

4.0 Data Collection and Processing

Data collection for Cycle 5 of the YITS took place between January and June 2008.

4.1 The Questionnaire

The YITS questionnaire was developed for cycle 1, using CASES software for a computer assisted telephone interview (CATI). In cycle 2 and for subsequent cycles the software was changed from CASES to BLAISE Some advantages of CATI are: question flows are controlled by the computer which allows for a more complex interview with little added burden to the interviewer; data capture occurs during the interview with on-line edits verifying that the data have been captured properly; and the system is able to take care of scheduling appointments, accounting for time zones, etc.

4.2 Training

Given the large survey sample size, it was decided to conduct the YITS from four Statistics Canada regional offices: Winnipeg, Sherbrooke, Sturgeon Falls and Halifax. A large number of interviewers and senior interviewers were required to work on the survey so the team implemented a two-phase classroom training plan. Prior to the classroom training, the participants received their self-study materials, which included interviewer and content manuals.

The first phase of classroom training took place in Head Office in Ottawa, where up to two senior interviewers and one project manager from each regional office were invited to a two day training session followed by a 1-day self-tutorial on the survey definitions and collection procedures. The second phase of the training took place in the regional offices. Senior interviewers and project managers, who were trained during the first phase, provided the same training to interviewers in the regional offices.

4.3 Supervision and Control

All interviewers are under the supervision of a staff of senior interviewers who are responsible for ensuring that interviewers are familiar with survey concepts and procedures, periodically monitoring their interviewers and reviewing their work. Senior interviewers ensured that prompt follow-up action was taken for refusals and other non-response cases. The senior interviewers are, in turn, under the supervision of the Regional Office project managers.

4.4 The Interview

In Cycle 1 a single respondent was identified as a potential respondent for the YITS. In Cycle 2, attempts were made to contact respondents who had completed the interview in Cycle 1. The sample of respondents for subsequent cycles was determined the same way. In order to facilitate contacting the selected respondent, the sample file included the respondent's name, address and telephone number, as well as additional addresses and phone numbers where possible. This provided the interviewer additional "leads" to find the respondent, should attempts with the original telephone number prove unsuccessful.

If the selected respondent could not be located, then the interview could not take place. Proxy reporting was not permitted.

By the end of data collection, the total average time spent per case/unit (including contact, tracing, call-backs and interview) was overall 66 minutes - Cohort A: 70 minutes and Cohort B: 61 minutes)

4.5 Follow-up on Non-response

Interviewers were instructed to make all reasonable attempts to obtain the YITS interviews with the selected respondent. For cases in which the timing of the interviewer's call was inconvenient, an appointment was arranged to call back at a more convenient time. For cases in which there was no response, there was follow-up.

Non-response can be viewed from a number of perspectives. In the context of the YITS, there are two types of non-response:

Refusals

If a respondent refused to participate, then the case was coded as a "refusal". An interviewer specifically responsible for refusal conversions would then access the case and call back the respondent to attempt to persuade him/her to participate.

Partial interviews

If a respondent failed to complete the interview, the case was assigned a partial code. Call backs to convert the case to a complete were only made if the survey response rate was not met or a particular segment of the target population experienced low response rates (as determined by Methodology). For example, Cohort A, high school-leavers.

4.6 Data Capture

As mentioned previously, responses to the questionnaire were captured directly by the interviewers at the time of collection, using computer-assisted telephone interviewing. A partially edited file was transmitted to Ottawa for further post-capture processing.

4.7 Minimum Completion Requirements

For all surveys it is essential that a minimum number of key fields are completed. In the case of the YITS, records were retained so long as high school status could be derived (see Derived Variables and Codebooks - ***Section 5.0***).

4.8 Computer Assisted Interview (CAI) Editing

CAI editing occurred in the Regional Offices during data collection. The data were collected in a telephone interview using a CAI application. As such, it was possible to build various edits and checks into the questionnaire in order to ensure high quality of the information collected. Below are specific examples of the types of edits used in the YITS computer-assisted interviewing application:

Review Screens (Confirmation Screens)

Review screens were created for important and complex information. For example in Module H both an Institution Confirmation Screen and Program Confirmation Screen were used to identify eligible institutions and programs. In Module P1, the Job Confirmation Screens identified eligible jobs. All review screens provide essential information to assist interviewers and respondents in recall and verification.

Range Edits

Range edits were built into the CAI system for questions asking for numeric values. If values entered were outside the range, the system generated a pop-up window that stated the error and instructed the interviewer to make corrections to the appropriate question. Range edits were provided for years, number of months, weeks, days or hours of work, and monetary values through out the questionnaire. For example, please see B_Q20Y. This question asked the respondent what year they were last in high school and the acceptable range was 1983 – 2007. If the respondent answered outside of this range the interviewer was to prompt for a more accurate date.

Flow Pattern Edits

All flow patterns were automatically built into the CAI system. For example, in Module A, at A_Q01, respondents were asked if they had lived in the United States at any time in the last two years. If not, they flowed immediately to Module B for questions on high school status. Whereas the group that lived in the U.S. were asked a number of questions related to their stay in the U.S. before going to Module B.

Consistency Edits

Consistency edits (indicated within a variable name as “E” or “D”) were used to identify inconsistencies in responses. These edits also identified missing information from previous variables. They were included as part of the CAI system to allow interviewers to return to previous questions to correct for inconsistencies. Interviewer instructions were displayed for handling or correcting problems such as incomplete or incorrect data. Variable B_E46d asks the respondent if they have completed the SSVD graduation requirements. This question is asked to correct SSVD graduation status.

4.9 Head Office Processing

The main outputs of the Youth in Transition Survey (YITS), Cycle 5 are "clean" data files. This section presents a brief summary of some of the processing steps involved in producing these files.

The processing of the YITS Cycle 5 data was done using the Generalized Processing System. This is a generic system that follows a series of steps to “clean” a file from beginning to end. The main steps were:

- Clean-up
- Pre-edit
- Flow edits
- Coding
- Derived variables
- Computer Generated Edits

To facilitate processing, the file was split into smaller files, which corresponded to the survey modules. Rostered files were also created in order to process the data more efficiently.

The YITS prepared both flat files and rostered files for Cycle 5. The rostered data files contain as many records for a given longitudinal respondent as the number of events, such as for the number of employers and the employment history, the number of institutions attended and programs taken at those institutions. (The employer questions would be programmed in a roster, which would be, repeated the appropriate number of times.)

Clean Up

The purpose of this step is to drop full-duplicate records and split-off records with duplicate identification numbers for examination. The data is then split between response and non-response based on pre-determined criteria. For more information, see *Section 7.0*).

Pre-edits

For all records where values were missing (blank) from the collection, the value of “9, 99, 999…” was inserted during Head Office processing to indicate that no information was collected. The “Don’t know” values returned by the CAI application as code “9” are changed to “7” in the pre-edits. As well, the “Mark all that apply” questions were de-strung and values converted to Yes (1) or No (2) responses. Finally, all text answers were removed from the processing file and set aside to be handled separately.

Recoding Other Specify

A number of questions in each module of the YITS have a category of “other specify”. Interviewers manually type in a “long answer” response that was not easily categorized during the interview or could not be coded into one of the already pre-assigned categories for that question. During the pre-edit, the “other specify” responses were reviewed and when possible coded back up to already existing categories, or new categories were created, when they met specific criteria. For example, in B_Q47 and B_Q52 (for which the derived variable DRED2 (cycle 2) was created) there was a high frequency of “moved” as a response under the other category.

Questions containing a response category of "other specify" may not be consistent across cycles. While the original response categories as appeared on the questionnaire are always included, it is possible that additional categories may be generated based on the frequency of responses found in the “other specify” category.

General rules were applied:
1. If a new category existed in a previous cycle, it will retain the same category value in the current cycle.

2. A new category in a cycle will always retain its value regardless of whether it is used in a future cycle. The "other-specify" (general) would take on the last value. Users should be cautious when comparing "other specify" counts across cycles as they may not always contain similar items.

For example:
For variable X we have values 1-5 and "other" is 6. Beginning with cycle 3, after Other Specify Long Answer (OSLA) coding, two new categories were added and were given the values 6 and 7 and "other" moved down to 8. Next cycle, if we only use new category 6, and 7 is left off, the "other" stays as 8. In this context the value "8" is dissimilar across cycles.

Flow Edits

Standards have been developed for the coding structure of data in order to explain certain situations in a consistent fashion. These standard codes are applied at the flows editing stage of processing of the YITS. The following describes these various situations and the codes used to describe the situation.

Valid Skip

In some cases a question, a series of questions, or an entire module was not applicable to the survey respondent. For example, a respondent reporting he/she has no children will have all questions related to dependent children set to a valid skip during processing. A code "6", "96" "996" was used on the data file to indicate that a question is a valid skip. In cases where an entire module of the questionnaire was not applicable to the respondent, all the variables of the module were set to a "Valid Skip".

An item which was coded as "Valid Skip" is indicated by a code "6". For a variable that is one digit long the code is "6", for a two-digit variable "96", for a three-digit variable "996", etc.

Not stated

The not stated code indicates that the answer to the question is unknown. Not stated codes were assigned for the following reasons.

- As part of the CAI interview, the interviewer was permitted to enter a "Refusal" or "Don't know" code. When this happened, the CAI system was often programmed to skip out of this particular section of the questionnaire. As part of the Generalized Processing System, all of the subsequent questions of this section are assigned a "Not stated" code. A not stated code means that the question was not asked of the respondent. In some cases it is not even known if the question was applicable to the respondent.
- In some cases a questionnaire 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 a questionnaire was only partially completed but enough information had been collected to consider to meet the preset criteria to stay in the sample, then all remaining unanswered questions on the questionnaire were set to "Not stated". The one exception was that if it was known that a certain question or section was not applicable, then these questions were set to "Valid Skip".

As with the Valid Skip, an item which was coded as "Not stated" is indicated by a code "9". For a variable that is one digit long the code is "9", for a two-digit variable "99", for a three-digit variable "999", etc.

Coding of Open-ended Questions

A few data items on the YITS questionnaire were recorded by interviewers in an open-ended format. For example, in the Employment Section (Module P1), respondents who had worked in the reference period were asked a series of open-ended questions regarding their employment:

- What kind of business, service or industry is/was this?
- What kind of work are/were you doing?
- At this work, what are/were your most important duties or activities?

The answer provided by the respondent was entered as text by the interviewer. At Head Office, the open-ended questions were coded using various standard classifications. Occupation questions were coded using the 1991 Standard Occupational Classification codes (SOC) and the industry questions were coded using the 1997 North American Industry Classification System (NAICS). Other open-ended questions for education (type of institution and program) were also collected and codes were assigned for Education (CIP).

Imputation

For a few variables on the YITS file, rather than using a special non-response code, imputation was carried out. Imputation is the process whereby missing or inconsistent items are "filled in" with plausible values. For the YITS, imputation was carried out for sections of the survey pertaining to employment income and household income. See *Section 9.0* for more details on imputation.

Derived Variables

Once the data are “clean”, derived variables (see *Section 5.0*) are programmed following specifications written by subject matter staff of (CTCES) and programmed by the processing team of Special Surveys Division. The standard codes of valid skip and not stated are also applied in the Derived Variable stage. In a few instances, a category has been set to 5, 95, 995 or 9995. This reserve code refers to Not Applicable. Furthermore, 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”.

Final Processing Files

The final processing files created for the YITS, Cycle 5 included both flat and rostered files. Users should see *Section 13.0* for working with the file.

5.0 Derived Variables and Codebooks

Answers from certain questions from the survey were combined to form specific concepts. Among them are: high school student status; postsecondary school status; and labour force status in December 2007. Such concepts are called *derived* variables (DVs). A derived variable is the result of combining the answers from a number of questions that pertain to a specific concept into a single variable.

The following tables identify the derived variables and the respective codebooks. The derived variables are listed under specific headings and in the order in which they appear in the codebooks. There are modules for which derived variables were not required. The universe for each derived variable indicates who responded to the questions contained in the derived variables, and may be for both cohort A (15 year-olds – referred to as 23 year-olds in Cycle 5) and cohort B (18-20 year-olds – referred to as 26-28 year-olds in Cycle 5), or either cohort. In the cases where a DV is for cohort A only, it will not be listed in this User Guide but can be found in the Cycle 5, 23 year-olds User Guide.

5.1 Cycle 5 Codebooks

The codebooks for Cycle 5 were developed depending on whether a module contained historical data (brought forward from Cycle 4) or were Cycle 5 reference period only. For those modules containing Cycle 4 and Cycle 5 data, rostered codebooks were produced. The following is a table of codebooks all of which can be found on the Statistics Canada website at ***www.statcan.gc.ca*** under the Youth in Transition Survey (project 4435), Cycle 5.

Input from researchers using YITS data from all cycles required that changes be made to codebooks in Cycle 1 (Cohort B only), and for both cohorts in Cycles 2 to 4. Changes have been highlighted on the first page of each revised codebook. As well, new derived variables have been produced and additional codebooks developed. Each cycle's documentation has been updated on the website at www.statcan.gc.ca

Cycle 5 changes included:

- adding questions in Module L to ask respondents who took PSE at any time between 2007 and 2008: (10 new DVS have been developed in Cycle 5 for these questions)
  - for amounts of funding received from different sources for postsecondary education taken between 2007 and 2008 (L_Q02A to L_Q04C). This included more information on use of RESPs (L_Q04A to L_Q04C)
  - about loan remission, interest relief, debt reduction and revision of terms available for government student loans (L_Q07B3A to L_Q07B3C).
  - about payments on government student loans (L_Q07B4 to L_Q07B5D).
  - if respondent received assistance from the governments or lenders in making repayment on loans (L_Q07B5E to L_Q07B5G)
  - if the respondent defaulted on loans (L_Q07B5H)
- adding questions in Module L to ask respondents who were not students during the reference period but had a government sponsored student loan at any time:
  - if they were aware of government aid programs such as loan remission, interest relief, debt reduction or revision of terms for government student loans (L_Q07D1A to L_Q07D1C)
  - if they made payments on their government student loan (L_Q07D1E to L_Q07D1H)
  - if they received assistance from governments or lenders in making repayment on these student loans (L_Q07D1I to L_Q07D1K)
  - if they defaulted on loans (L_Q07D1L)
- adding questions in Module M to ask respondents:
  - if they had to pay for housing, had a car loan or lease; a line of credit; credit cards or other personal loans; average monthly payments and outstanding balance (M_Q10A to M_Q10F1)
  - if they were aware of tax incentives (education, tuition, textbook or student loan interest tax credits and if they took advantage of these credits (M_Q11a to M_Q14)

- eliminating questions in Module P1 on activity during months when last in school full-time and when started first job
- reinserting questions in Module P2 on unpaid leave (questions were cut in cycle 4); recreating DV on Working Job Status WSTPmmD5 (mm = 01 to 24) and WKFPTPD5 for WSTP24D5 (full-time/part-time status for respondent at all jobs in December 2007)
- Module R asked of both Cohorts

Section	File Type	Descriptive File Name	Modules	Cycle
5.1.1	Flat file	Person Level Main File	A, B, H, L, M, N, Q, R, U, V	Cycle 5
5.1.2	Institution Roster	Education above High School (Institutions)	H	Cycles 4 and 5
5.1.3	Program Roster	Education above High School (Programs)	H	Cycles 4 and 5
5.1.4	Roster	Post Secondary Engagement Roster	K and KE	Cycle 5
5.1.5	Roster	Confirmation of Open Jobs from Cycle 4 Roster	P1	Cycle 4
5.1.6	Roster	Job Roster	P1 and P2	Cycle 5
5.1.7	Roster	Job Details Roster	P2	Cycle 5
5.1.8	Roster	Dependent children	U	Cycle 5

5.1.1 Person Level Main File Codebook

Module Name and Identifier	Derived Variable	Description	Universe
Entry	RecordID	Respondent identification number	All respondents
	AGED5	Age of respondent as of Dec. 2007	(as above)
	BYEARD5	Respondent's year of birth (1979 – 1984)	(as above)
	BMONTHD5	Respondent's month of birth	(as above)
	GENDERD5	Respondent's gender	(as above)
	PROVD5	Province /territory code of residence for the household as of date of interview (2001 Census geography)	(as above)
	URMIZD5	Indicator of rural or urban geography, based on the Statistical Area Classification (SATYPE), based on the 2001 Census geography.	(as above)
Moved to the United States – Module A	No derived variables		
High School Status – Module B	ACMD5	Respondent's age at date completed high school diploma requirements or equivalent	Respondents who have completed high school diploma or equivalent.
	AFTD5	Respondent's age when last in elementary or secondary school full-time prior to January 2008	All respondents
	AHSD5	Respondent's age when last in elementary or secondary school prior to January 2008	(as above)
	DNOD5	Number of times respondent dropped out of elementary or secondary school	(as above)
	DRED5	Main reason for having ever dropped out of elementary or secondary school prior to January 2008	Respondents who have ever dropped out of school regardless of whether they were a high school continuer, graduate or leaver as of December 2007. For respondents who have dropped out more than once, this variable is the main reason for the last time they have dropped out
	DRPD5	Variable indicating if respondent has ever	All respondents

Module Name and Identifier	Derived Variable	Description	Universe
		dropped out of elementary or secondary school	
	FPTLESD5	Full-time/part-time status when respondent was last in elementary or secondary school	(as above)
	HG9D5	Variable to indicate if respondent has gone past grade 9 in elementary or secondary school as of December 2007	(as above)
	HGCD5	Highest grade respondent has completed in elementary or secondary school as of December 2007	(as above)
	HSDIPMD5	Date (month) respondent completed high school diploma requirements or equivalent	Respondent who completed their high school diploma, Secondary School Vocational diploma (SSVD) requirements or equivalent
	HSDIPYD5	Date (year) respondent completed high school diploma requirements or equivalent	(as above)
	HSSTATD5	High school status as of December 2007	(as above)
	LESMTD5	Date (month) respondent was last in elementary or secondary school prior to January 2008	(as above)
	LESYRD5	Date (year) respondent was last in elementary or secondary school prior to January 2008	(as above)
	LFTESMD5	Date (month) respondent was last in elementary or secondary school full-time prior to January 2008	(as above)
	LFTESYD5	Date (year) respondent was last in elementary or secondary school full-time prior to January 2008	(as above)
	LGED5	Grade enrolled in elementary or secondary school as of December 2007 or the date last in school (elementary or secondary)	Respondents who have not completed the minimum requirements for a high school graduation certificate, diploma or its equivalent as of December 2007 or the date last in school. Excludes those last enrolled in an SSVD program in Quebec
	NUMHSD5	Number of different high schools attended between January 2006 and December 2007	Respondents who have taken any high school, junior high or elementary school sometime between January 2006 and December 2007
	SCIPD5	Respondent's main field of study or specialization - secondary school vocational diploma program (CIP Codes)	Respondents who took some education in an SSVD program between January 2006 and December 2007
	SCIPRD5	Respondent's field of study or specialization - secondary school vocational diploma program (primary grouping)	(as above)
Education and Training Above High School– Module H	AGEPSD5	Respondent's age at start of first postsecondary program	All respondents except high school continuers who had not graduated
	DLFPSMD5	Date (month) respondent was last taking	(as above)

Module Name and Identifier	Derived Variable	Description	Universe
		postsecondary education on a full-time basis ever, prior to January 2008	
	DLFPSYD5	Date (year) respondent was last taking postsecondary education on a full-time basis ever, prior to January 2008	(as above)
	DLPSMD5	Date (month) respondent was last taking postsecondary education ever, prior to January 2008	Respondents who took some postsecondary education between January 2006 and December 2007
	DLPSYD5	Date (year) respondent was last taking postsecondary education ever, prior to January 2008	(as above)
	EDTPSMD5	Date (month) respondent first started postsecondary education prior to January 2008	Respondents who took some postsecondary education prior to January 2008
	EDTPSYD5	Date (year) respondent first started postsecondary education prior to January 2008	(as above)
	FPSPD5	Variable identifying respondent's first postsecondary institution and program attended	All respondents except high school continuers who had not graduated
	HEDATD5	Highest certificate, diploma or degree attained (or graduated from) as of December 2007	(as above)
	HEDLD5	Highest education level taken as of December 2007	(as above)
	HGDAD5	Highest graduation diploma attained as of December 2007	All respondents
	HLPSD5	Highest level of postsecondary education taken across all programs and institutions as of December 2007	All respondents except high school continuers who had not graduated
	LPSATD5	Overall postsecondary status as of December 2007	(as above)
	MHSPSD5	Duration of time, in months, from the date last in elementary/secondary school to the time started first postsecondary program. MHSPSD5 is duration in absolute value. MHSPSFD5 indicates whether duration is negative or positive	Respondents who took some postsecondary education
	MHSPSFD5	Duration of time, in months, from the date last in elementary/secondary school to the time started first postsecondary program. MHSPSFD5 indicates whether duration is negative or positive. MHSPSD5 is duration in absolute value	(as above)
	NEPRCD5	Number of eligible postsecondary programs taken between January 2006 and December 2007	Respondents who attended at least one postsecondary program between January 2006 and December 2007
	NINDID5	Number of postsecondary institutions attended between January 2006 and December 2007	(as above)
	ULPSATD5	University postsecondary status as of December 2007	Respondents who ever attended a university by December 2007
	CLPSATD5	College postsecondary status as of December 2007	Respondents who ever attended a college by December 2007
	OLPSATD5	Other institution postsecondary status as	Respondents who ever attended a

Module Name and Identifier	Derived Variable	Description	Universe
		of December 2007	postsecondary institution other than a university or college by December 2007
Funding – Module L	AINRED5	Respondents who received assistance in the form of interest relief from federal or provincial or territorial governments	Respondents who received money from a government sponsored student loan in the current cycle and those who owed on government sponsored student loans at the end of the previous cycle, who have started to repay their loan and received assistance from governments or lenders in making repayments
	AMBWQD5	Amount of the monthly, bi-weekly, weekly or quarterly payment for respondents who have started to repay their government student loan between January 2006 and December 2007	Respondents who received money from a government sponsored student loan in the current cycle and those who owed on government sponsored student loans at the end of the previous cycle, who have started to repay their loan according to a monthly, bi-weekly, weekly or quarterly schedule
	AMOAIDD5	Amount of the loan remission, interest relief, debt reduction or revision of terms received between January 2006 and December 2007	Respondents who received money from a government sponsored student loan in the current cycle and those who owed on government sponsored student loans at the end of the previous cycle, who are aware and received a loan remission, interest relief, debt reduction or revision of terms
	AOTHERD5	Amount of the occasional, annual or other schedule, for respondents who have started to repay their government student loan between January 2006 and December 2007	Respondents who received money from a government sponsored student loan in the current cycle and those who owed on government sponsored student loans at the end of the previous cycle, who have started to repay their loan according to an occasional or annual schedule (includes those who refused or did not know the method of repayment of their loan)
	AWARGAD5	Respondent's awareness of government aid programs such as loan remission, interest relief, debt reduction or revision of terms available for government student loans	Respondents who received money from a government sponsored student loan in the current cycle and those who owed on government sponsored student loans at the end of the previous cycle
	DECLOAD5	Respondents who received assistance in the form of decreased loan payments that were negotiated with the lender	Respondents who received money from a government sponsored student loan in the current cycle and those who owed on government sponsored student loans at the end of the previous cycle, who have started to repay their loan and received assistance from governments or lenders in making repayments
	DEFAULD5	Respondents who have ever defaulted on any of their student loans.	Respondents who received money from a government sponsored student loan in the current cycle and those who owed on government sponsored student loans at the end of the previous cycle, who have started to repay their loan

Module Name and Identifier	Derived Variable	Description	Universe
	GOVASSD5	Respondents who received any assistance from governments or lenders in making repayment of government sponsored student loans	Respondents who received money from a government sponsored student loan in the current cycle and those who owed on government sponsored student loans at the end of the previous cycle, who have started to repay their loan
	PAYGSLD5	Respondents who made payments to repay their government student loan between January 2006 and December 2007	Respondents who received money from a government sponsored student loan in the current cycle and those who owed on government sponsored student loans at the end of the previous cycle
	PAYOFFD5	Year respondent expects to have the government-sponsored student loans paid off	Respondents who received money from a government sponsored student loan in the current cycle and those who owed on government sponsored student loans at the end of the previous cycle, who have started to repay their loan
	OWEGOVD5	Total amount of money owed to sponsored student loans as of December 2007, to fund post-secondary education	Respondents who received a loan from the government as of December 2007 in order to fund their post-secondary education
	OWEOTHD5	Total amount of money owed to bank loans, lines of credit, parents or family as of December 2007, to fund post-secondary education	Respondents who received a loan from a bank, a line of credit or from parents or family as of December 2007 in order to fund their post-secondary education
	RECAIDD5	Respondent's who received a loan remission, interest relief, debt reduction or revision of terms on their government student loan between January 2006 and December 2007	Respondents who received money from a government sponsored student loan in the current cycle and those who owed on government sponsored student loans at the end of the previous cycle, who are aware of government aid programs such as loan remission, interest relief, debt reduction or revision of terms programs
	SCHPAYD5	Payment schedule of respondents who have started to repay their government student loan between January 2006 and December 2007	Respondents who received money from a government sponsored student loan in the current cycle and those who owed on government sponsored student loans at the end of the previous cycle, who have started to repay their loan
	TMOD5	Total amount of money owed to student loans, bank loans, lines of credit, parents or family as of December 2007, to fund postsecondary education	Respondents who received a loan from the government, a bank, a line of credit or from parents or family as of December 2007 in order to fund their postsecondary education
	TMRD5	Total amount of money received from scholarships, awards, prizes, grants or bursaries between January 2006 and December 2007, to fund postsecondary education	Respondents who received a scholarship, award, prize, grant or bursary between January 2006 and December 2007 in order to fund their postsecondary education
Loans and Aspirations – Module M	M1BMD5	Date (month) respondent was last in school full-time	All respondents
	M1BYD5	Date (year) respondent was last in school full-time	(as above)

Module Name and Identifier	Derived Variable	Description	Universe
	JOA30RD5	2 digit occupational code (SOC 1991) for job or occupation respondent would be interested in having around the age of 30	All respondents
	JOA30D5	4 digit occupational code (SOC 1991) for job or occupation respondent would be interested in having around the age of 30	All respondents
Social Support Scales and Standard Error – Module N	YSHSUPS5	This variable measures how much social support a respondent receives from friends, family and other sources	All respondents
	YSHSUPE5	Standard error for the score on how much social support respondent receives from friends, family and other sources	(as above)
Work related Questions – Module P1	ELJBVD5	Indicates the number of eligible jobs (during this cycle) up to a maximum of 7 jobs (based on first 6 jobs + first job after leaving full-time schooling)	All respondents
	ELJBNVD5	Contains the remaining number of jobs worked at (during this cycle) for which no verification was possible	(as above)
	FJ_AGED5	Respondent's age when started the first job s/he worked at after leaving full-time schooling	Respondents who were in full-time school between January 2006 and November 2007; or who were last in school full-time prior to January 2006; who had not yet had a first job after leaving full-time school
	FSTJOBD5	Identifies the job the respondent held at time of leaving full-time schooling, or the job first started after leaving full-time schooling (jobs 1 to 7 collected), for those respondents who were no longer full-time students in December 2007	(as above)
	FTSFJDD5	Duration of time, in months, from the date left full-time schooling to the time started first job after full-time schooling. FTSFJDD5 is duration in absolute value. (Use in combination with FTSJFLD5 to determine if duration is positive or negative)	(as above)
	FTSJFLD5	Determine if the value of FTSFJDD5 is positive or negative	(as above)
	JBST01D5 to JBST24D5	Employment status - Flag for each month (January through December) in 2006 and 2007 indicating if respondent was employed at at least one job	Respondents who had at least one job between January 2006 and December 2007
Reservation Wage – Module P5	RSWGD5	Reservation wage - the lowest wage or salary a respondent would accept to begin a new job immediately, full-time - expressed in dollars and cents per hour	All respondents
Training – Module P6	ERTD5	Number of employer related training courses or programs taken between January 2006 and December 2007	Respondents who worked during the reference period (January 2006 to December 2007)
	CRTD5	Number of career or job related training courses or programs taken between January 2006 and December 2007	All respondents

Module Name and Identifier	Derived Variable	Description	Universe
	TTH_1D5 and TTH_2D5	Total number of training hours - Employer organized training - courses 1 and 2	Respondents who reported at least one training course or program taken between January 2006 and December 2007
	TTH_3D5 and TTH_4D5	Total number of training hours- Job/career related training - courses 1 and 2	(as above)
	TTHERD5	Total number of training hours- Employer organized training	(as above)
	TTHJCD5	Total number of training hours- Job/career related training	(as above)
	TTHD5	Total number of training hours	(as above)
GAPS – Module PS	FEDS01D5 to FEDS24D5	Full-time student status - Flag for each month in 2006-2007 indicating if respondent was a full-time student	All respondents
	FTES01D5 to FTES24D5	Whether the respondent was in elementary, secondary or postsecondary full-time schooling for each month during 2006-2007	Respondents who took some education (elementary, secondary school or postsecondary education) between January 2006 and December 2007
	LGMD5	Last Gap Month. Last month in the reference period when the respondent was not employed at a job and was not a student full-time	Respondents who had at least one month in which they were not working or at school between January 2006 and December 2007
	LWLGMD5	Looking for Work Last Gap Month. Variable to indicate if respondent was looking for a job during their last gap month of the reference period	Respondents who had at least one month in which they were not working or at school between January 2006 and December 2007
	LFW01D5 to LFW24D5	Variable to indicate if the respondent had done anything in looking for work in the indicated month	Respondents who had at least one month in which they were not working or at school between January 2006 and December 2007
	MLTJBYD5	Variable to indicate if respondent was employed at 2 or more jobs during one month or more at any time between January 2006 and December 2007	Respondents who had at least one job between January 2006 and December 2007
	MTJOB01D5 to MTJOB24D5	Variable identifies, for each month from January 2006 to December 2007, the number of jobs the respondent was employed at during the month	Respondents who had at least one job between January 2006 and December 2007
	JBFPTPD5	Full-time/part-time status among all jobs that the respondent HAD in December 2007	Respondents who had at least one job between January 2006 and December 2007
	WSTP01D5 to WSTP24D5	Working at job status - Flag for each month in 2006-2007 indicating if respondent was employed and working at at least one job	Respondents who had at least one job between January 2006 and December 2007
	WKFPTPD5	Full-time/part-time status for respondent at all jobs in December 2007	(as above)
Volunteer Activities – Module Q	OVRD5	Respondent's volunteer activities in 2007	All respondents
Skills – Module R	No derived variables		
Demographics – Module U	AGMPD5	Respondent's age, in years, at time of moving out permanently from the home of parents or guardians	Respondents who reported that they have moved out permanently from the home of their parents or guardians as of December 2007

Module Name and Identifier	Derived Variable	Description	Universe
	CITZEND5	This variable updates the information on citizenship collected in the previous cycle. It is possible that respondents may have obtained Canadian citizenship or acquired or relinquished citizenship from other countries	All respondents
	DEPCHD5	Number of dependent children	All respondents
	LANIMMD5	Variable indicates whether or not respondents who are not Canadian by birth have ever been a landed immigrant	Respondents who reported that they were not Canadian by birth
	LCCCD5	Language(s) spoken well enough to conduct a conversation	All respondents
	MARSTAD5	Marital status	All respondents
	SPINDD5	4 digit industry code (NAICS 1997) for spouse's job	Respondents who reported a partner/spouse
	SPINDRD5	2 digit industry code (NAICS 1997) for spouse's job	(as above)
	SPOCCD5	4 digit occupation code (SOC 1991) for spouse's job	(as above)
	SPOCCRD5	2 digit occupation code (SOC 1991) for spouse's job	(as above)
Income - Module V	INCEID5	Income received in 2007 from Employment Insurance benefits	(as above)
	INCEISD5	Income respondent's spouse or partner received in 2007 from Employment Insurance benefits	Respondents who had a partner or spouse in 2007
	INCGSD5	Income received in 2007 from other government sources such as Worker's Compensation or Canada Pension Plan or Quebec Pension Plan	All respondents
	INCGSSD5	Income respondent's spouse or partner received in 2007 from other government sources such as Worker's Compensation or Canada Pension Plan or Quebec Pension Plan	Respondents who had a partner or spouse in 2007
	INCNRD5	Income received in 2007 from parents or other people that did not have to be repaid (excludes loans).	All respondents
	INCNRSD5	Income respondent's spouse or partner received in 2007 from parents or other people that did not have to be repaid (excludes loans)	Respondents who had a partner or spouse in 2007
	INCOND5	Income received in 2007 from other non-government sources including income from dividends, interest and other investment income, employer pensions, annuities or rental income	All respondents
	INCONSD5	Income respondent's spouse or partner received in 2007 from other non-government sources including income from dividends and other investment income, employer pensions, annuities or rental income	Respondents who had a partner or spouse in 2007
	INCSCD5	Income received in 2007 from spousal support or child support	All respondents

Module Name and Identifier	Derived Variable	Description	Universe
	INCSCSD5	Income respondent's spouse or partner received in 2007 from spousal support or child support	Respondents who had a partner or spouse in 2007
	INCSED5	Income received in 2007 from self-employment	All respondents
	INCSESD5	Income respondent's spouse or partner received in 2007 from self-employment	Respondents who had a partner or spouse in 2007
	INCSGD5	Income received in 2007 from scholarships, grants or bursaries	All respondents
	INCSGSD5	Income respondent's spouse or partner received in 2007 from scholarships, grants or bursaries	Respondents who had a partner or spouse in 2007
	INCSPD5	Income received in 2007 from Social Assistance or Provincial /Territorial Income Supplements	All respondents
	INCSPSD5	Income respondent's spouse or partner received in 2007 from Social Assistance or Provincial/Territorial Income Supplements	Respondents who had a partner or spouse in 2007
	INCSTD5	Income received in 2007 from the Goods and Services Tax Credit (GST) or Harmonized Sales Tax Credit (HST) or Quebec Sales Tax Credit (QST)	All respondents
	INCSTSD5	Income respondent's spouse or partner received in 2007 from the Goods and Services Tax Credit (GST) or Harmonized Sales Tax Credit (HST) or Quebec Sales Tax Credit (QST)	Respondents who had a partner or spouse in 2007
	INCTBD5	Income received in 2007 from Canada Child Tax Benefit or provincial/territorial child tax benefits or credits	All respondents
	INCTBSD5	Income respondent's spouse or partner received in 2007 from Canada Child Tax Benefit or provincial/territorial child tax benefits or credits	Respondents who had a partner or spouse in 2007
	INCWSD5	Income received in 2007 from wages and/or salaries	All respondents
	INCWSSD5	Income that respondent's spouse or partner received in 2007 from wages and/or salaries	Respondents who had a partner or spouse in 2007
	TINCD5	Total 2007 income from all sources before taxes and deductions	All respondents
	TINCSD5	Total 2007 income for respondent's spouse or partner from all sources before taxes and deductions	Respondents who had a partner or spouse in 2007
	DECEASE5	Respondents deceased on or before December 31, 2007	
	Weight	Weight : decimal in 4th byte of the field	

The following geography derived variables were not released in the codebook

CMA27_D5- Area consisting of one or more adjacent municipalities situated around a major urban core. To form a census metropolitan area, the urban core must have a population of at least 100,000. To form a census agglomeration, the urban core must have a population of at least 10,000.
CMACA_D5 - Identifies the type of census metropolitan area (CMA) or census agglomeration (CA) in which the enumeration area is located. This field will be left blank where the enumeration area is not part of a CMA or CA.
CSDC_D5 - Census subdivision (CSD) is the general term for municipalities (as determined by provincial /territorial legislation) or areas treated as municipal equivalents for statistical purposes (for example, Indian reserves, Indian settlements and unorganized territories).
DA_D5 - Dissemination Area (formerly Enumeration Area (EA)). The postal codes are linked to the Postal Code Conversion File (PCCF) using the latest version corresponding to the 2001 Census Geography
ECR_D5 – Economic Region (ER) of residence for the household at time of interview
EIER_D5 – Employment Insurance Economic Region (2001 Census Geography)
SATYP_D5 - Category assigned to a municipality not included in either a census metropolitan area (CMA) or a census agglomeration (CA). (A CMA or CA is an area consisting of one or more adjacent municipalities situated around a major urban core. To form a CMA, the urban core must have a population of at least 100,000. To form a CA, the urban core must have a population of at least 10,000.) A municipality is assigned to one of four categories depending on the percentage of its residents who commute to work in the urban core of any census metropolitan area or census agglomeration. Strong MIZ: more than 30% of the municipality's residents commute to work in any CMA or CA. Moderate MIZ: from 5% to 30% of the municipality's resident's commute to work in any CMA or CA. Weak MIZ: from 0% to 5% of the municipality's residents commute to work in any CMA or CA. No MIZ: fewer than 40 or none of the municipality's residents commute to work in any CMA or CA
REG_D5 – Region of residence for the household as of date of interview

5.1.2 Education Above High School – Institution Roster Codebook

Module Name and Identifier	Derived Variable	Description	Universe
Entry	RecordID	Respondent identification number	All respondents
	INST_ID	This number given to the institution corresponds to the order in which the respondent reported it	Respondents with at least one postsecondary institution
Education and Training – Module H	DSAINMD5	Date (month) respondent started postsecondary education at this institution, prior to January 2008	Respondents who took some postsecondary education between January 2006 and December 2007
	DSAINYD5	Date (year) respondent started postsecondary education at this institution, prior to January 2008	(as above)
	DLINMD5	Date (month) respondent was last at this institution between January 2006 and December 2007	(as above)
	DLINYD5	Date (year) respondent was last at this institution between January 2006 and December 2007	(as above)
	DLFINMD5	Date (month) respondent was last taking postsecondary education at this institution on a full-time basis prior to January 2008	(as above)
	DLFINYD5	Date (year) respondent was last taking postsecondary education at this institution on a full-time basis prior to January 2008	(as above)
	FPLIND5	Full-time or part-time student when last at this institution between January 2006 and December 2007	(as above)
	HLATTD5	Postsecondary status at this institution as of December 2007	(as above)
	NEPRPID5	Number of eligible postsecondary programs taken at this institution between January 2006 and December 2007	(as above)
	INSCDD5	Institution Code	(as above)
	PSCMD5	Campus Code	(as above)
	PSPROVD5	Province/territory of postsecondary institution	(as above)
	PSIPOSID	Postsecondary institution position identifier which identifies the cycle and position where the data in this cycle for this program was collected	Respondents with at least one postsecondary institution
	PSILNGID	Postsecondary institution longitudinal identifier which permits following an institution across cycles	Respondents with at least one postsecondary program at this institution
	TYPEID5	Type of postsecondary institution	(as above)
	ICYID	Postsecondary institution cycle identifier, which identifies the cycle in which data were first collected for this institution	(as above)

The postsecondary institution longitudinal identifier, PSILNGID, allows the user to follow an institution across the cycles. The identifier has 2 digits which can take the following values:

First digit: represents the cycle in which the institution information was first collected; can take values of 1 to 4

Second digit: represents the institution attended by respondent ; can take values of 1 to 4

This identifier does not exist on the cycle 1 data file. If you need to retrieve a variable from the cycle 1 file you will need to refer to the table on Appendix B which indicates which variables belong to which institution. For example if you need question H8, and PSILNGID=12 then variable H8b, is associated with institution 2.

5.1.3 Education Above High School – Program Roster Codebook

Module Name and Identifier	Derived Variable	Description	Universe
Entry	RecordID	Respondent identification number	All respondents
	INST_ID	This number given to the institution corresponds to the order in which the respondent reported it	Respondents with at least one postsecondary institution
	PROG_ID	This number given to the program within the institution corresponds to the order in which the respondent reported it	(as above)
Education and Training – Module H	INELIGD5	An ineligibility flag indicating the reason why an open program and/or institution from cycle 4 was deemed ineligible in cycle 5	Respondents with an open program/institution from cycle 4
	INELGHD5	Flag indicating whether or not an open program and/or institution from cycle 4 was deemed ineligible in cycle 5	Respondents with an open program/institution from cycle 4
	LVPRD5	Level of postsecondary program	Respondents who took some postsecondary education between January 2006 and December 2007
	CLGPRD5	Postsecondary status in this program as of December 2007	(as above)
	DSPRMD5	Date (month) respondent started this postsecondary program, prior to January 2008	(as above)
	DSPRYD5	Date (year) respondent started this postsecondary program, prior to January 2008	(as above)
	DLPRMD5	Date (month) respondent was last taking this postsecondary program between January 2006 and December 2007	(as above)
	DLPRYD5	Date (year) respondent was last taking this postsecondary program between January 2006 and December 2007	(as above)
	FPLPRD5	Full-time or part-time student when last in this program, between January 2006 and December 2007	Respondents who were taking a postsecondary program between January 2006 and December 2007
	DLFPRMD5	Date (month) respondent was last taking this postsecondary program, on a full-time basis prior to January 2008	Respondents who took some postsecondary education between January 2006 and December 2007
	DLFPRYD5	Date (year) respondent was last taking this postsecondary program, on a full-time basis prior to January 2008	(as above)
	SIPRD5	For postsecondary programs which are ongoing or completed as of December 2007, whether respondent has stopped or interrupted their education between January 2006 and December 2007	(as above)
	AGEPSD5	Respondent's age at start of postsecondary program	Respondents who were taking a postsecondary program between January 2006 and December 2007
	NUMDURD5	Time spent taking a postsecondary program, as of December 2007 (months)	Respondents who participated in a postsecondary program between January 2006 and December 2007

Module Name and Identifier	Derived Variable	Description	Universe
	RSIPRD5	For programs in which respondents participated between January 2006 and December 2007, reason for stopping or interrupting program if the respondent stopped or interrupted their program	Respondents who took some postsecondary education in programs which are ongoing or completed and who have ever stopped or interrupted their studies
	CIP1D5	Respondent's first main field of study or specialization	Respondents who had a valid postsecondary program
	CIP2D5	Respondent's second main field of study or specialization	(as above)
	CIP1RD5	Respondent's first main field of study or specialization (primary grouping)	(as above)
	CIP2RD5	Respondent's second main field of study or specialization (primary grouping)	(as above)
	THEPSD5	Total time spent with an employer in a co-op, apprenticeship, trade/vocational training or another program (e.g. practicum, internship or clinical) for this program	Respondents who attended an eligible postsecondary program between January 2006 and December 2007 and participated in a program which included on the job experience and/or time spent in a workplace
	OPSPD5	Chronological order of postsecondary programs attended by respondent during 2006 and 2007	Respondents who took some postsecondary education between January 2006 and December 2007
	PSPPOSID	Postsecondary program position identifier which identifies the cycle and position where the data in this cycle for this program was collected	Respondents with at least one postsecondary program
	PSPLNGID	Postsecondary program longitudinal identifier which permits following a program across cycles	(as above)
	ICYID	Postsecondary institution cycle identifier, which identifies the cycle where data were first collected for this institution	(as above)

The postsecondary program longitudinal identifier, PSPLNGID, allows the user to follow a program across the cycles. The identifier has 4 digits which can take the following values:

First digit : represents the cycle in which the program started; can take values 1 to 4
Second digit : represents the institution in which the program was taken ; can take values from 1 to 4
Third digit is always zero
Fourth digit : represents the program taken; can take values from 1 to 3.

This identifier does not exist on the cycle 1 data file. If you need to retrieve a variable from the cycle 1 file you will need to refer to the table on Appendix B which indicates which variables belong to which institution and which program. For example you need to find question H_Q21, if the PSPLNGID is 1301 then variable H21c1, is associated with institution 3 program 1.

5.1.4 Postsecondary Engagement Roster Codebook

Module Name and Identifier	Derived Variable	Description	Universe
Entry	RecordID	Respondent identification number	All respondents
Postsecondary Engagement – Module K	HPDPSD5	Average number of hours of paid work per week during the first year of postsecondary school	Respondents whose 1st and/or 2nd post-secondary experience occurred in cycle 5.
	HUWPSD5	Average number of hours of unpaid work in family's business or farm per week during the first year of postsecondary school	(as above)
	KEXPIDD5	Indicates if it is the respondent's 1st or 2nd post-secondary experience, in Quebec or non Quebec, in a CEGEP or non-CEGEP institution.	(as above)
	KINSTD5	Post-secondary institution identifier which identifies which institution this experience is associated with.	(as above)

The two derived variables KEXPIDD5 and KINSTD5, provides information on the type of postsecondary experience: first or second, in Quebec or not in Quebec, a CEGEP experience or a non-CEGEP experience. It also provides a link between the postsecondary experience and the institution to which that experience is associated.

Starting with cycle 2, KINSTD (postsecondary experience institution identifier) matches PSILNGID, the postsecondary institution longitudinal identifier.

If the first postsecondary experience was in cycle 1, the link cannot be established with PSILNGID. This identifier does not exist on the cycle 1 data file. If you need to retrieve a variable from the cycle 1 file you will need to refer to the table on Appendix B which indicates which variables belong to which institution. For example if you need question H8, and KINSTID=12 then variable H8b, is associated with institution 2.

In cycle 1, information on postsecondary experiences was not rostered. Appendix B is a table including all variables names (including derived variables) at the institution level, the program level and the person level for cycle 1.

5.1.5 Confirmation of Open Jobs from Cycle 4 Roster Codebook

Module Name and Identifier	Derived Variable	Description	Universe
Entry	RecordID	Respondent identification number	All respondents
	P1UNID	Longitudinal job identifier which permits following a job across cycles	Respondents who had a job in December 2005 (Cycle 4)
Work Related Questions – Module P1	INELJBD5	Respondents were asked details about jobs they reported in Cycle 4 that they either worked at in December 2005 or jobs they had in December 2005 but had not worked at during that period. Some of these jobs became ineligible during Cycle 5 collection because of respondent recall, respondents reporting that they did not return to work at the job in 2006-2007, or the job became not eligible during Cycle 5 collection because the respondent was not able to provide key information about the Cycle 4 job. INELJBD5 notes the reason why this job became ineligible	Respondents who worked at a job in December 2005or had a job but did not work at it in December 2005 (Cycle 4)

5.1.6 Job Roster Codebook

Module Name and Identifier	Derived Variable	Description	Universe
Entry	RecordID	Respondent identification number	All respondents
	P1JOBID	Unique job identifier, indicates the position where data in this cycle for this job were collected	Respondents who worked at an eligible job between January 2006 and December 2007
	P1UNID	Longitudinal job identifier which permits following a job across cycles	Respondents who had a job between January 2006 and December 2007.
Work Related Questions – Module P1	OJOBD5	Chronological order of jobs	Respondents who had at least one job between January 2006 and December 2007
	TENURED5	Total number of months respondent employed at job (regardless of unpaid leaves)	(as above)
	TNURD5	Total number of months in 2006-2007 respondent employed at job (regardless of unpaid leaves)	(as above)
	JOBOCCD5	4 digit occupation code (SOC 1991) for eligible jobs	(as above)
	JOBOCRD5	2 digit occupation code (SOC 1991) for eligible jobs	(as above)
	JOBINDD5	4 digit industry code (NAICS 1997) for eligible jobs	(as above)
	JOBINRD5	2 digit industry code (NAICS 1997) for eligible jobs	(as above)
	JSTDATD5	Start date of the job (year/month)	(as above)
	JBFTPTD5	Full-time/part-time status for respondent who HAD the job in December 2007	(as above)
Not applicable in Cycle 4	WKST01D5 to WKST24D5	Working at job status - Flag for each month in 2006-2007 indicating if respondent was employed and working at the job - for jobs 1 to 7	(as above)
Not applicable in Cycle 4	WKFTPTD5	Full-time/part-time status for respondent who was working at a job in December 2007	(as above)

5.1.7 Job Details Roster Codebook

Module Name and Identifier	Derived Variable	Description	Universe
Entry	RecordID	Respondent identification number	All respondents
	P1JOBID	Unique job identifier, indicates the position where data in this cycle for this job were collected	Respondents who worked at an eligible job between January 2006 and December 2007
Employment – Module P2	EPHS15	Earnings per hour when first started job	Respondents who had a job at any time between January 2006 and December 2007 and who were paid employees or self-employed when they first started this job
	EPWSI5	Earnings per week when first started job	(as above)
	EPMSI5	Earnings per month when first started job	(as above)
	EPHEI5	Earnings per hour when last worked at job	(as above)
	EPWEI5	Earnings per week when last worked at job	(as above)
	EPMEI5	Earnings per month when last worked at job	(as above)
	NHWPMSI5	Number of hours usually worked per month when first started working at job	Respondents who were employed at a job between January 2006 and December 2007
	NHWPMEI5	Number of hours usually worked per month when last worked at job	(as above)
	NWWPMSI5	Number of weeks usually worked per month when first started at job	(as above)
	NWWPMEI5	Number of weeks usually worked per month when last worked at job	(as above)
	NMW03D5	Number of months in the cycle where respondent did some work at job (i.e., total months employed at job less number of months respondent had unpaid leaves, if there were any) (for each listed job)	(as above)
	HWSD5	Indicates whether the respondent usually worked 30 or more hours per week when first started working at this job	Respondents who had at least one job between January 2006 and December 2007
	HWED5	Indicates whether the respondent usually worked 30 or more hours per week when last worked at this job	(as above)

5.1.8 Dependent Children Codebook

Module Name and identifier	Derived variable	Description	Universe
Entry	RecordID	Respondent identification number	All respondents
	CBDYMD5	Date (month) of birth of all dependent children.	All respondents who reported dependent children.
	CBDYYD5	Date (year) of birth of all dependent children.	All respondents who reported dependent children.
	HPMCHCD5	Reason dependent children live with respondent most or part of the time.	Respondents with dependent children who live in the same house with the child most or part of the time.
	LVECHD5	Status of living arrangement of dependent children in the household.	Respondents with dependent children.
	RELCHCD5	Relationship of dependent children to respondent.	Respondents with dependent children.

6.0 YITS Scales

In YITS cycle 5, the concept social support was measured through the use of a scale. Social support was assessed for all members of the 15-year-old Reading Cohort and the 18-20 year-old Cohort. For more theoretical details about any of the mathematical/statistical concepts discussed in this chapter, please see the Statistics Canada technical document *Analysis of Scales for YITS Cycle 1 Main Survey*.

This chapter is divided into four sections. The section 6.1 provides a justification for scaling and describes the type of scaling applied within YITS. The section 6.2 discusses the theoretical procedure that the YITS team used to form the scales. The results and analysis of an adapted version of the social support scale are provided in section 6.3. Finally, a list of the references cited or consulted for this chapter is provided in section 6.4.

6.1 Defining Scales and Their Uses

6.1.1 What is a Scale?

For the purposes of social science research, a scale is a type of composite measure consisting of several items (questions) that share an empirical or logical structure. A scale can be regarded as a set of possible scores for combining answers to a group of questions. The term scale is also used within the context of this chapter to refer to the theoretical concept upon which the scales are derived.

6.1.2 Why Use a Scale?

The use of scales in data analysis allows researchers to estimate a measure of a particular underlying (latent) concept when the items measuring the concept are put together. A scale is created by assigning scores to patterns of responses that enable the analyst to assess the relative weakness or strength of each indicator. The use of scales is advantageous in that scales can demonstrate the intensity of the relationships that may exist among concepts.

For each factor (concept) measured by a scale, a latent score value is estimated for each individual surveyed within the sample of eligible respondents. This estimated score is based upon appropriate combinations of a number of responses to a group of survey questions (items). The score for a particular factor may be used to order individuals with reference to the factor or to illustrate differences between individuals or groups with respect to that factor.

A scale has a higher level of reliability (see Section 6.2.3) than do individual items. Indeed, a scale increases in reliability as the number of items contained within it increases. An item can contain information about the construct being measured (signal) and confounding variance due to measurement error and information uniquely associated with that item (noise). Using a scale helps to reduce the effects of noise and increases the amount of information available for analysis. Therefore, a multiple-item scale provides more information to analysts than does a single item. Scales are useful in social science research because they facilitate the efficient reduction of large amounts of data into manageable and meaningful sources of information for the analyst.

6.1.3 What Type of Scales are Used in YITS?

All of the scales used in cycle 1 to cycle 5 of YITS are modeled after the Likert Scale (Likert, 1932). This type of scale is valued for the ordinality of its multiple response categories. This allows researchers to compare the relative strength of agreement of survey respondents to any particular item. For example, a particular question with four categories may require respondents to express their views on an issue from four ordinal scale values such as 1 (strongly disagree), 2 (disagree), 3 (agree) and 4 (strongly agree).

A Likert-type scale is a highly reliable tool for rank-ordering people when attempting to measure attitudes or opinions that they hold on a topic. It is one of the most commonly used and most easily recognizable formats in the area of questionnaire design. Likert scales can have any number of categories. Indeed, the more categories provided to a respondent, the more precise the distinction between the categories and the more information available to distinguish between respondents on the measured construct. However, in practice, respondents may not be able to respond meaningfully when there are too many categories given to them.

Moreover, there is additional burden to the respondent in terms of the time required to make such fine distinctions. The application of scale-type questions is an inexact science; their use is somewhat subjective. The number of item categories should be chosen based upon situation-specific judgment including knowledge of the item content, the underlying scale construct and the respondents. Scales need not have the same number of item categories for all of the items in the scale.

In many of the items that make up the various YITS scales, an item category representing a neutral response such as "*neither agree nor disagree"* or "*sometimes"* has been excluded from the categories available to the respondent. Many of the questions asked in YITS scales deal with topics that are not often considered by respondents; thus, respondents could have tended toward a neutral response, which would have reduced the variability in responses to each item and therefore would have reduced the utility of the scale.

Removing the neutral category, in this case, makes it more likely that the scale will detect tendencies of respondents, with respect to an item, even if these tendencies are slight. Some experts in questionnaire design feel that if a respondent does not know whether they are leaning to the positive or the negative end of a statement that he or she should indicate a response of "*don't know*". Other experts, however, think that the neutral response category is necessary for those respondents who truly do not have an opinion on the statement presented to them. The theory and design of survey scale items is discussed in more detail in *Survey Research Methods* (Fowler, 1995).

If one chooses to use a previously constructed scale in a survey, then it is important to consider whether this scale addresses the issues that the survey is attempting to measure. This is especially important in research domains where there are debates within the social science literature as to what the concept should measure. One such discussion occurs when there are many different scales that have varying degrees of specificity and/or focus upon particular aspects of a concept. Although differences between global measures towards a concept and measures of specific attitudes or facets of a concept are often overlooked in research, their measures and their behaviours are not necessarily equivalent (Rosenberg et al., 1995).

6.1.4 Response Bias

The systematic tendency for participants to respond to rating items independent of their content (what the item is designed to measure) is referred to as response bias. This tendency is also referred to as a response set or as a response style depending upon the context. A response set is a temporary reaction to a situational demand. These demands can include time pressure or expected public disclosure. Bias could also result from context issues such as the format of the item or the nature of previous items in the questionnaire. If an individual displays bias consistently over time and situations, then this bias is regarded as his or her response style (Paulhus, 1991, p. 17).

6.1.5 Negative and Positive Questions

It is recommended that both negatively and positively worded questions be included in widely-used rating scales within the discipline of psychology in order to reduce a variety of response bias including acquiescence[3] (Marsh, 1996). This is done under the working assumption that positively and negatively worded items represent the same concept.

Sometimes, however, factor analytic techniques (Section 6.2.1) indicate different separate factors resulting from the positive and negative worded questions. The crucial question that must be answered in these cases is whether this distinction between the positively and negatively worded item factors is in fact substantively meaningful. Alternatively, it is possible that this distinction is merely an artefact of a person's response style (Paulhus, 1991, p. 48). These two explanations have quite different implications; however, distinguishing between them can be difficult (Marsh, 1996).

3 Acquiescence is the tendency to agree rather than disagree with item statements (Paulhus, 1991, p. 46). Some individuals referred to as "*yea-sayer*", tend to agree with statements whereas other individuals referred to as "*naysayers*" tend to disagree with statements.

6.2 Scale Development

There were three main steps in the estimation of the scale scores. The first step was to use linear factor analysis to investigate if the underlying theoretical structure of items and subscales was supported by the data (Section 6.2.1). Assuming that the theoretical structure was valid, the second step in development was to estimate the scale scores using an item response theory (IRT) model (Section 6.2.2). Finally, reliability and validity checks were performed on the estimated scores (Section 6.2.3 and Section 6.2.4). The general procedures that were adopted for each of these steps are described below.

6.2.1 Investigation of Model Validity

Factor Analysis: Strategy

Factor analysis was used to determine whether the theoretical construct of the scale was supported by the data. Factor analysis is also one way to help achieve the goals of an item analysis. Item analysis is the verification that items are related to their proposed constructs and that the strength of these relationships are adequate for measurement purposes (Gorsuch, 1997).

According to Comrey and Lee, the goal of a factor analysis is to isolate constructs that have a greater intrinsic value for the purpose of describing the functional relationships between the variables in the field. However, all sets of variables are not equally good in representing this relationship. Moreover, there is not necessarily only one "correct" concept or "real" factor for a given domain of interest (Comrey and Lee, 1992, p. 245).

The relationship of each variable to each of the factors, referred to as the loading on a factor, provides a way for the analyst to quantitatively assess how an item interacts with other items. The strength of these loadings on different factors indicates to the analyst whether an item is related to none of the proposed factors, to only one of the factors or to more than one factor (multiple loadings). The greater the level of loading of a variable on a factor, the greater the amount of overlapping (common) variance[4] between a data variable and a factor and the more an item is a pure measure of this factor.

Items most clearly associated with only one factor can become part of a scale for the construct underlying that factor. Items not strongly associated with any of the factors can be dropped from further analysis. Items can be strongly associated with more than one factor. These items may indicate a relationship between the factors on which the loadings occurred, in which case many items should load on more than one factor and the multiple loadings can be explained mathematically. Alternately, they may indicate the presence of an unknown or confounding factor that is related to the proposed factors. When this is the case the item or items in question should be dropped from the scale.

A question that often arises when examining factor loadings is how high the correlation between an item and a factor must be before the item is considered "significant". Although no formal test to determine significance has been developed, Comrey and Lee provide a benchmark to use in interpreting variable-factor correlations[5] (Comrey and Lee, 1992, p. 243). They base their benchmark upon an examination of the percentage of variance of the item common to the factor (see Table 6.1).

4 The variance in responses to any particular item can be described by two main components: the specific variance and the common variance. Specific variance represents the differences between people related to the unique characteristics of the item. Common variance, or communality, refers to differences that can be explained by the common factor related to all items in a scale.
5 Orthogonal factor loadings or structure coefficients

Table 6.1 Scale of Variable-Factor Correlations

Loading	Percentage of Variance Explained	Quality of Loading
Above 0.71	Above 50	Excellent
Above 0.63	Above 40	Very Good
Above 0.55	Above 30	Good
Above 0.45	Above 20	Fair
Above 0.32	Above 10	Poor
Below 0.32	Inconsequential	Trivial (not to be interpreted).

Examining factor loadings is not in and of itself sufficient for factor interpretation. A full explanation of a factor requires a theoretical understanding of why the items fit together and how the identified factor or factors relates to other previously identified factor structures within the domain of interest.

Factor Analysis: Procedures

For the YITS data, consideration was given to the effect of the language of interview on responses. Any changes in the presentation of items may introduce substantial bias in responses (Fowler, 1995, p.74). In this case, the bias may affect the responses due to imprecise translations. This problem could create potentially different interpretations of the question. This, in turn, may lead to different response patterns on a question between those respondents who were administered the questionnaire in English and those who were administered it in French.

To examine the possibility of translation bias, the data was first divided up into two samples according to the language of questionnaire administration: English and French. Comparison of the results from each linguistic sample was undertaken. Greater dissimilarities between the results would indicate a greater translation bias.

The data from both the English and the French samples were further randomly split into two half-samples. The first half-sample for both the French and the English sample was considered as a test sample and the second half-sample was regarded as a verification sample. The verification sample was used to independently confirm the structure identified in the first half-sample.

A principal component based factor analysis was carried out separately on each linguistic half-sample to determine how many factors should be extracted from the data. Factor loadings of each variable were compared between the half-samples and across language. Loadings were examined under a rotation of the factor loadings when the concept had two or more factors associated with it. A rotation involves a shift in the coordinate axis of the loadings matrix. When it is not easy to interpret a loading, a more readily interpreted solution may be obtained from a rotation. Under a rotation, one would hope to obtain a pattern of factor loading such that a variable loads highly on one factor and has only low to moderate loadings on the other factor or factors.

For every factor analysis presented within this document, the value of the items for each respondent was multiplied by the respondent's normalized weight in order to obtain a design-consistent estimate of the variance-covariance matrix. A normalized weight was calculated by dividing a respondent's cycle 5 survey weight (see ***Section 11.0*** for a discussion on weighting in YITS) by the average weight of all eligible respondents in the sample. Thus, in theory, the sum of the normalized weights is equal to the sample size of the eligible respondents.

Within a questionnaire, some questions are positively oriented such as "*There are people I can count on in times of trouble".* Other questions are negatively oriented such as "*There is no one I feel close to"*. In order for the effects on a scale of negatively orientated items not to cancel out the effects of positively oriented items, the negative items were reoriented to make them positive. Letting **k** be the number of categories for a particular item, a positively oriented item is obtained from a negatively oriented one by subtracting the value of the item from **k** + 1((**k**+1) - item value). On a four point scale, if a respondent indicated a value of 2 on a negatively oriented question, then the positively oriented response value would be 3 (3 = 5-2). This reorientation has to be done before the score is calculated in order to properly estimate the scale item internal consistency (Cronbach's alpha) and to estimate scale scores (see **Section 6.2.3** for a discussion of Cronbach's alpha).

In addition to the estimated scores, the items for each scale are included on the micro data file. This will allow researchers to consider alternate factor structures. The original values, before any reversal of values took place, have been retained for all of these items.

6.2.2 Estimation of the Scores

The results from the factor analysis were used to determine what items loaded onto each factor (**Section 6.2.1**). Once the factor structures were analysed and the items to be included in each factor were verified, the next step was to estimate the respondent's latent construct score for each factor. Two approaches were investigated, factor scores, a linear composition based upon the linear factor loadings developed under the factor analysis framework (the standard classical method) and non-linear maximum likelihood estimation based upon item response theory (IRT). Both methods were examined using normalized survey weights. Weights were incorporated into the analysis processes in order to derive design-consistent point estimates of the scores.

Factor analysis requires that the scale test data have the property of interval or ratio data. However, some people argue that the test data only have the properties of ordinal data. Whether psychological test scores should be considered ordinal or interval data is a subject of some debate within the social science community (Crocker and Algina, 1986, pp. 61-63). Generally, it is agreed that if it can be demonstrated that the scores produce more useful information when they are treated as interval data, then they should be treated as such. On the other hand, if treating the data as interval level measurement does not improve, or actually decreases, their usefulness, then only their rank-order information should be used (Crocker and Algina, 1986, p. 61).

IRT is able to control better for the ordinal nature of the data than is factor analysis. The software program PARSCALE[6] (based upon the theory developed by Eiji Muraki and Darrell Bock, 1997) was applied to calculate the IRT scores and the estimates of the score's measurement errors. Scores released for YITS scales were based upon a parametric IRT approach.

The IRT scores and their respective standard errors were estimated using weighted maximum likelihood (see Warm, 1989) and applying a generalized partial credit model. The generalized partial credit model is an extension of the two parameter logistic distribution to polytomous (categorical) data (Muraki, 1992). With this particular extension, one obtains an overall slope parameter for each item and a difficulty parameter for each category within an item. The YITS team adopted a particular form of this model where the difficulty parameter is spilt into a category parameter (a common parameter to all items within a specific block) and an item-specific location parameter.

For estimating IRT scores, the population distribution of the scores was specified to have a mean of zero and a standard deviation of one. Once standardized, the respondent's estimated score, in this case, can be interpreted as the number of standard deviations of the population of interest above (if positive) or below (if negative) the mean.

A response pattern of mostly extreme positives (i.e., strongly agree to all positively-worded questions and strongly disagree to all negatively worded questions) is most likely to have been produced by an individual with a highly positive standard score. Conversely, a response pattern of mostly extreme negative values is most likely to have been produced by an individual with a highly negative score. A typical mix of responses (few extreme responses) is likely produced by an individual who has a score on the scale of around zero. A response pattern that results from choosing the option that is the most extreme, in most cases, may be due to an internal bias by the respondent towards extreme responses or it may indicate a strong opinion by the respondent on the subject of inquiry (Paulhus, 1991, p. 49).

[6] For more information about PARSCALE, please see its user guide. (Du Toit, 2003).

6.2.3 Scale Reliability

Reliability, when discussing scales, refers to the accuracy, dependability, consistency or repeatability of score results. More technically, it refers to the degree to which the scores are free from measurement errors. It can be interpreted as a holistic measure of the accuracy of a scale, in that it describes the proportion of the population variance in scores that can be explained by the population variance in the underlying construct. Two measures of reliability are commonly used in examining scales: Cronbach's Alpha and the Index of Reliability.

Cronbach's Alpha and its Interpretation

Cronbach's alpha is a measure of the internal consistency of the items within a factor. It is based upon the average covariance of items within the factor and assumes that the items within a factor are positively correlated with each other.

Cronbach's alpha has several interpretations. It is theoretically related to the correlation between the scale being used and all other possible scales containing the same number of items that could be constructed from a hypothetical universe of items that measure the same characteristic of interest. With this measure, the analyst is able to obtain the expected relationship between the scale that was used and all other possible scales that measure the same concept. Since Cronbach's alpha can be interpreted as a coefficient of determination, its values range from 0 to 1. Cronbach's alpha can be regarded as a lower bound on reliability; the true reliability of the scale is at least as high as the value of reliability calculated using this measure.

One common misconception about Cronbach's alpha is that a relatively high alpha value for a factor indicates that the administered test items are unidimensional (the items are associated with only one common underlying factor). Indeed, as "[Cronbach's] alpha is a function of item covariances, and high covariance between items can be the result of more than one common factor, [Cronbach's] alpha should not be interpreted as a measure of the test's unidimensionality" (Crocker and Algina,1986, p. 142).

Index of Reliability

While Cronbach's alpha estimates the reliability as the internal consistency of a scale's items, a more accurate estimate of reliability is the index of reliability, η_j which incorporates the IRT characteristics of each item.

Similar to Cronbach's alpha, values of this index closer to 1 indicate a greater accuracy and denote better measurement properties of the scale (Crocker and Algina, 1986, p. 352).

$$\eta_j = \sqrt{1 - \frac{\sigma^2_{E_j}}{\sigma^2_j}} \quad (6.1).$$

For a given scale j, $\sigma^2_{E_j}$ is the weighted average measurement variance across the sample, and σ^2_j is the estimated variance of all scores in scale j. Although the value of η_j obtained will be similar in magnitude to that of Cronbach's alpha, it is a more accurate measure of the reliability of the final scores that have been produced.

6.2.4 Testing Scale Validity

In order to assess whether the estimated scale scores behave according to the theoretically proposed conception of the model, validity tests were performed on the scales. The validation process checks to see if the construct appears to be the same as it is commonly defined. This is despite modifications that may be made to the number or wording of items in the scale for operational constraints. These tests involve evaluating the proposed scales or subscales by comparing their estimated scores on the scales to scores on other relevant scales or to the values of other relevant variables (criterion validity). They can also involve the

comparison of different identifiable groups of respondents on the scale of interest (known-group validity). It can also be important to show that a scale does not have high correlation with attitudes that it is not designed to measure (discriminant validity). Testing for scale validity involves knowledge of the subject matter involved in the analysis and in particular, which variables or scales are expected to be related or not related to the scale of interest and the form that this relationship is expected to take.

6.3 Social Support Scale

6.3.1 Description of the Social Support Scale

Most conceptualizations of support include the following ideas: emotional sustenance, self-esteem building, provision of information and feedback and tangible assistance (Russell and Cutrona, 1987). A number of different social support measures have been developed. These measures differ widely and on multiple dimensions on how they model social support. These instruments differ on whether they assess 1) structure or the function of support, 2) subjective or objective support, 3) availability or adequacy of support, 4) individual structures or functions or global indices, 5) several individual structures, 6) the role of people available to provide support or simply whether support is available, 7) the number of people available to provide support or merely the availability of support (irrespective of the number of people) (Cohen and Syme, 1985, p. 15).

While social support does not have a unique concept or an empirical concept, it is still widely used by researchers; “The term [social support] connotates enough that it has proved fruitful even in the absence of denotation” (House and Kahn, 1985, p. 84). It has been suggested that the reason for this usage is that even without a single concept, it captures a common theme in many seemingly diverse phenomena.

Social support is a concept that may help in the interpretation of the differences observed in people’s responses to common problems. Conditions that create distress in some people do not seem to affect others. Researchers theorize that factors exist that can mediate between difficult life conditions and outcomes. Social support is one of these coping mechanisms (Pearlin and Schooler, 1978); (Pearlin, 1985, p. 57).

The central goal of the social support model proposed by Carolyn Cutrona and Daniel Russell is to understand the processes through which interpersonal relationships enhance or retain both psychological and physiological well-being. The objective of the measure for YITS was to determine the availability of social supports, via friends, family and other sources for the youth. The social support scale used in YITS is a modified version of the Social Provisions Scale developed by Cutrona and Russell. It was based upon similar modifications to the scale adapted for the Canadian NLSCY (Microdata User Guide (2003)) [7].

The aspects of social support measured in YITS include three aspects of the original model and are classified under the broad category of assistance-related provisions. They are reliable alliance (the assurance that others can be counted upon for practical help), attachment (emotional closeness) and guidance (advice or information). These aspects are most directly related to problem-solving within the context of stress. Two items were proposed to measure each of these aspects for a total of six items. All respondents in cycle 5 were eligible to receive the social support questions.

6.3.2 Model Validity

No strong differences were found between the factor loadings on the English and the French samples and all of the proposed items were kept. The items that make up the factor, their description and their factor loadings are provided below.
No strong differences were found between the factor loadings on the English and the French samples for either the 15-year-old Reading Cohort or the 18-20 year-old Cohort and all of the proposed items were kept. The items that make up the factor, their description and their factor loadings for the 18-20 year-old Cohort are provided below.

7 Statistics Canada Microdata User Guide (2003) – *National Longitudinal Survey of Children and Youth – Cycle 4*. Statistics Canada

Table 6.2 Social Support Items Description and Loadings

Item Code Codebook	**Item Description**	**18-20 year-old Cohort Factor Loadings**
N5Q11	If something went wrong, no one would help me	0.69
N5Q12	I have family and friends who help me feel safe, secure and happy	0.80
N5Q13	There is someone I trust whom I would turn to for advice if I were having problems	0.80
N45Q14	There is no one I feel comfortable talking about problems with	0.73
N45Q15	There is no one I feel close to	0.80
N45Q16	There are people I can count on in times of trouble	0.82

The loadings for the Social Support scale, according to the Comrey and Lee benchmark for rating scale loadings (Section 6.2.1), with values from 0.6970 to 0.823, range from very good to excellent.

6.3.3 Estimating Scores

These scale scores have the codeYSHSUPS5 on the dataset and their standard error has the code YSHSUPE5. Due to a few respondents not answering any of the questions upon which the scale was based**,** the score could not be estimated for 266respondents for the 18-20 year-old Cohort. In all of the cases of missing scores, for this scale, the scores and the standard error of the scores were assigned a value of 99.99999.

The item parameters for the social support scale are provided below.

Table 6.3a. Social Support Item Specific Parameters 18-20 year-old Cohort

Item Code Codebook	Slope Parameter	Location Parameter
N5Q11	1.23616	-1.56108
N5Q12	1.93302	-1.59451
N5Q13	1.92132	-1.62485
N5Q14	1.25144	-1.35053
N5Q15	1.84990	-1.43637
N5Q16	2.23513	-1.51454

Table 6.3b. Social Support Category Parameters 18-20 year-old Cohort

Category 01	Category 02	Category 03	Category 04
0.00000	0.34050	0.67724	-1.01775

6.3.4 Scale Reliability

Two common measures of reliability, Cronbach's alpha and the index of reliability were estimated. The value of Cronbach's alpha for the items in the Social Support scale is 0.87 for the 18-20 year-old Cohort and the value of the index of reliability is 0.9089. Researchers should use these reliability estimates and the standard errors of the scores provided with the micro data file to determine whether or not this scale is reliable enough for their purposes.

6.3.5 Testing Scale Validity

Subsection 6.3.5 refers to scale validity checks that were performed on previous cycles. The scale was validated by comparison to module F in cycles 1 and 2 and for the 15 year -old reading cohort in cycle 3. The particular validity check based on question responses to module F was not possible for cycle 5 as this module was dropped from the questionnaire.

6.4 Scale References

Bowlby, J.W. and McMullen, K. (2002). *At a Crossroads: First Results for 18 to 20-Year-old Cohort of the Youth in Transition Survey*. Catalogue No. RH64-12/2002E. Statistics Canada.

Cohen, S. and Syme, S. L. (1985). "Issues in the Study and Application of Social Support", (pp. 1-22) in *Social Support and Health*. Cohen, S. and Syme, S. L. (Eds.). San Diego, California: Academic Press.

Comrey, A.L. and Lee, H.B. (1992). *A First Course in Factor Analysis.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Crocker, L. and Algina, J. (1986). *Introduction to Classical and Modern Test Theory.* Belmont, California: Wadsworth Group.

Cutrona, C.E. and Russell, D.W. (1987). "The Provisions of Social Relationships and Adaptation to Stress", *Advances in Personal Relationships*, Vol. 1, 37-67.

Du Toit, M., (ed.) 2004, Muraki, E. *IRT from SSI*, Chicago, Illinois: Scientific Software International.

Documentation of the Scales used in the National Longitudinal Survey of Children and Youth, Cycles 1-3, Internal Document. Statistics Canada.

Fowler, F.J. (1995). *Survey Research Methods: second edition.* London, England: Sage Publications.

Goodenow, C. (1993). "The Psychological Sense of School Membership among Adolescents: Scale Development and Educational Correlates", *Psychology in the Schools*. Vol. 30, 79-90.

Gorsuch, R.L. (1997). "Exploratory Factor Analysis: Its Role in Item Analysis", *Journal of Personality Assessment.* Vol. 68, 532-560.

Glass G. and Hopkins, K. (1996). *Statistical Methods in Education and Psychology 3rd ed.* Boston, Massachusetts: Allyn and Bacon.

House, J. L. and Kahn, R. L. (1985). "Measures and Concepts of Social Support" (pp. 83-108). In *Social Support and Health*. Cohen, S. and Syme, S. L. (Eds.). San Diego, California: Academic Press.

Likert, R. (1932). "A Technique for the Measurement of Attitudes", *Archives of Psychology*. No. 140, 1-55.

Marsh, H.W. (1996). "Positive and Negative Global Self-Esteem: a Substantively Meaningful Distinction or Artifactors?", *Journal of Personality and Social Psychology* . Vol. 70, No. 4, 810-819.

Microdata User Guide. (2003). - *National Longitudinal Survey of Children and Youth - Cycle 4*. Statistics Canada.

Muraki, E. (1992). "A Generalized Partial Credit Model: Application of an EM Algorithm." (Research Reports Educational Testing Services RR-92-06) Princeton, New Jersey: Educational Testing Services.

Paulhus, D. L. (1991). *"*Measurement and Control of Response Bias*",* (pp. 291-372). In *Measures of Personality and Social Psychological Attitudes: Volume 1 of Measures of Social Psychological Attitudes.* Robinson, J., Shaver, P., Wrightsman, L. (Eds.), San Diego, California: Academic Press.

Pearlin, L.I. and Schooler, C. (1978). "The Structure of Coping", *Journal of Health and Social Behaviour,* Vol. 19, 2-21.

Pearlin, L.I. (1985). *"Social* Structure and Processes of Social Support", (pp. 43-60). in *Social Support and Health*. Cohen, S. and Syme, S. L. (Eds.), San Diego, California: Academic Press.

Rosenberg, M. et al (1995). "Global Self-Esteem and Specific Self-Esteem: Different Concepts, Different Outcomes", *American Sociological Review.* Vol. 60, 141-156.

Statistics Canada. (2000). *T-00-5E (September 2000)Youth in Transition Survey – Project Overview .* Catalogue no. MP32-30/00 – 5E. Statistics Canada

Warm, T. (1989). “Weighted Likelihood Estimation of Ability in Item Response Theory”, *Psychometrika.* Vol. 54, 427-450.

7.0 Survey Methodology

Definition of the YITS population

The YITS target population for the 26 to 28 year-olds cohort comprises individuals who were born in the years 1979 to 1981 and who resided in one of the ten provinces when cycle 1 was designed, A large portion of the questionnaire for Cycle 5, which was administered from January to June 2008, is devoted to profiling these individuals' education and labour market activities during the reference years of 2006 to 2007, when the respondents were 26 to 28 years old.

The sample design of the YITS 26-28 cohort was determined to a large degree by the sample design of the Labour Force Survey (LFS). This applies to the stratification, allocation and selection of the sample as well as the estimation methods.[1]

Note that the YITS is strictly a longitudinal survey. The initial sample of 18-20 year-olds selected at cycle 1 will continue to be surveyed every two years for five cycles. No attempts were made to top-up the sample at Cycle 5 to ensure a cross-section representation of the population. Therefore, what follows is a summary of the methodology that was used at cycle 1 to design the survey. Note also that the YITS uses a funnel approach meaning that non-respondents at a specific sample are not followed-up for subsequent cycles of the survey.

7.1 LFS Sample Design

The LFS is a monthly survey that collects labour market data from a national sample of about 60,000 dwellings. The LFS design used for the YITS was implemented at the end of 1994 following a redesign program that included a reassessment of the survey's principal role as a provider of current labour market information as well as a central vehicle for conducting household surveys within Statistics Canada.

7.2 LFS Target Population

The LFS sample is representative of the civilian, non-institutionalised population 15 years of age or older in Canada's ten provinces. Specifically excluded from the survey's coverage are residents of the Yukon, Nunavut and Northwest Territories, persons living on Indian Reserves, full-time members of the Canadian Armed Forces and inmates of institutions. These groups together represent an exclusion of approximately 2% of the population aged 15 or over.[2]

7.3 LFS Stratification

The LFS sample is based upon a stratified, multistage design employing probability sampling at all stages of the design. The design principles of the LFS are the same for each province.

Primary strata

Provinces are divided into economic regions and employment insurance economic regions. Economic regions (ERs) are geographic areas of more or less homogeneous economic structure formed on the basis of federal provincial agreements. They are relatively stable over time. Employment insurance economic regions (EIER) are also geographic areas, and are roughly the same size and number as ERs, but they do not share the same definitions. Labour force estimates are produced for the EIER regions for the use of Human Resources and Skills Development Canada.

[1] For a detailed account of the LFS methodology, see Gambino, J.G., Singh, M.P., Dufour, J., Kennedy, B. and Lindeyer, J. (1998), *Methodology of the Canadian Labour Force Survey*, Statistics Canada, Catalogue no. 71-526.

[2] Since 1992, the LFS has been administered in the Yukon, using an alternative methodology that accommodates some of the operational difficulties inherent to remote locales. To improve reliability due to small sample size, estimates are available on a three-month average basis only. In 2000, the survey was extended to the Northwest Territories and Nunavut. No estimates for any of the territories are included in national totals.

The intersections of the two types of regions form primary strata for the LFS. Note that a third set of regions, Census Metropolitan Areas (CMAs), is also respected by stratification in the current LFS design, since each CMA is also an EIER.

Types of areas

The primary strata (ER/EIER intersections) are classified into three types of areas: rural, urban, and remote areas. Urban and rural areas are loosely based on the Census definitions of urban and rural, with some exceptions. Urban areas include the largest CMAs down to the smallest villages categorised by the 1991 Census as urban (1,000 people or more), while rural areas are made up of areas not designated as urban or remote.

All urban areas are further classified into two types: those using an apartment list frame and an area frame, as well as those using only an area frame.

Approximately 1% of the LFS population is found in remote areas of provinces that are less accessible to LFS interviewers than other areas. For administrative purposes, this portion of the population is sampled separately through the remote area frame. Places with fewer than 10 households or 25 persons and Census Enumeration Areas (EA) with fewer than 25 households are omitted from the design.

Secondary strata

In urban areas with sufficiently large numbers of apartment buildings, the strata are grouped according to those based on apartment frames and those based on area frames. The apartment list frame is a register that is based upon information supplied by Central Mortgage and Housing Corporation (CMHC) and is maintained in the 18 largest cities across Canada. The purpose of this frame is to ensure better representation of apartment dwellers in the sample as well as to minimise the effect of cluster growth due to construction of new apartment buildings. In the major cities, the apartment strata are further stratified into low-income strata and regular strata.

Where it is possible and/or necessary, the urban area frame is further stratified into regular strata, high-income strata, and low population density strata. Most urban areas fall into the regular urban strata, which, in fact, cover the majority of Canada's population. High-income strata are found in major urban areas, while low-density urban strata consist of small towns that are geographically scattered.

In rural areas, the population density can vary greatly from relatively high population density areas to low population density areas, resulting in the formation of strata that reflect these variations. The different stratification strategies for rural areas were based not only on concentration of population, but also on cost-efficiency and interviewer constraints. Also, within each of the secondary strata in rural areas further stratification is carried out where necessary to reflect differences among a number of socio-economic characteristics within each stratum.

The remote area frame is stratified only by province.

7.4 LFS Cluster Delineation and Selection

Households in final strata are not selected directly. Instead, each stratum is divided into clusters, and then a sample of clusters is selected within the stratum. Dwellings are then sampled from selected clusters. Different methods are used to define the clusters, depending on the type of stratum.

Within each urban stratum in the urban area frame, a number of geographically contiguous groups of dwellings, or clusters, are formed based upon 1991 Census counts. These clusters are generally a set of one or more city blocks or block faces. The selection of a sample of clusters (always six or a multiple of six clusters) from each of these secondary strata represents the first stage of sampling in most urban areas. In some other urban areas, Census EAs are used as clusters. In the low-density urban strata, a three-stage design is followed. Under this design, two towns within a stratum are sampled and then a multiple of six clusters within each town are sampled.

For urban apartment strata, instead of defining clusters, the apartment building is the primary sampling unit. Apartment buildings are sampled from the list frame with probability proportional to the number of units in each building.

Other procedures are applied in rural and remote areas. Within each rural stratum, six EAs or two or three groups of EAs are sampled as clusters, whereas remote settlements within each province are sampled proportional to the number of dwellings in the settlement.

7.5 LFS Dwelling Selection

In all three types of areas (urban, rural and remote areas) enumerators in the field first visit selected clusters and a listing of all private dwellings in the cluster is prepared. From the listing, a sample of dwellings is then selected. The sample yield depends on the type of stratum. For example, in the urban area frame, sample yields in regular strata within major urban areas are either six or eight dwellings, depending on the size of the city. In the urban apartment frame, each cluster yields five dwellings, while in the rural areas and EA parts of cities, each cluster yields ten dwellings. In all clusters, dwellings are sampled systematically. This represents the final stage of sampling.

7.6 LFS Sample Rotation

The LFS employs a panel design whereby the entire monthly sample of dwellings consists of six panels or rotation groups, of approximately equal size. Each of these panels can be considered to be representative of the entire LFS population. Dwellings are in the LFS for six consecutive months. Each month a new panel of dwellings selected from the same or similar clusters replaces the sample dwellings in one of the rotation groups.

This rotation pattern has the statistical advantage of providing a common sample base for month-to-month comparisons of LFS characteristics. It also ensures that the sample of dwellings constantly reflects changes in the current housing stock and helps to minimise the respondent burden and non-response that could result if households were to remain in the sample longer than six months. Surveys that use the LFS frame or sample can take advantage of the rotation group feature to use larger or smaller sample sizes than that of the LFS.

7.7 Household Members Eligible for the LFS

The first month a dwelling is in the LFS a roster containing information on the household composition is completed. Demographic information including the name, sex, date of birth and education level is obtained for all persons for whom the selected dwelling is the usual place of residence. LFS information is obtained for all civilian household members 15 years of age or older.

When the dwelling is contacted in subsequent months the roster is updated to reflect changes in household membership from the previous month. Response burden is minimised for the elderly (70 years of age or older) by carrying forward their responses for the initial interview to the subsequent five months in the survey.

7.8 YITS Target Population and Sample Design

As noted at the beginning of *Section 7.0*, the target population for the 26-28 year-old cohort of the YITS comprises residents of the ten provinces of Canada who were born in 1979 to 1981. These individuals turned 26, 27 or 28 during 2007, the reference year for Cycle 5. The LFS exclusion of full-time members of the armed forces and persons living on Indian reserves or in northern and remote areas also applies to the YITS.

The design implemented for the 26-28 year-olds cohort of the YITS is based on certain groups of households that were in the LFS between January 1997 and December 1999. A sample of 29,950 individuals was selected for Cycle 1 from a combination of households in 30 LFS groups that rotated out

from January 1997 to November 1999 and households from all six rotation groups in the December 1999 LFS.

To select the YITS cohort B sample, the set of in-scope households within each of the 36 groups was identified. In-scope households were those with at least one household member with a high probability of being in the YITS target population. The majority of individuals were identified by year of birth - for household members without a year of birth on the LFS roster, the age and the month of the LFS were used to calculate an estimated age as of December 31, 1999. Thus households were included in the initial YITS sample if at least one household member had

- a specified year of birth from 1979 to 1981, or
- an unspecified year of birth and an estimated age of 17 to 21 years[1]

From the initial sample of 29,950 households, 786 were treated as a priori non-response and were excluded from the final sample selection stage for the YITS. Two-thirds of these households were eliminated from the YITS sample because they already had a household member in the longitudinal sample for the National Longitudinal Survey of Children and Youth and/or the National Population Health Survey. Most of the other excluded households were eliminated because they had refused forthcoming LFS interviews or their LFS interviews were conducted in person rather than by telephone.

The final stage of sampling for the YITS was the selection of one household member in the YITS target population from each of the remaining 29,164 in-scope households. In the minority of households that did have more than one person in the YITS target population, one was selected with equal-probability systematic sampling.

7.9 YITS Sample Size by Province

The following table 7.1 shows the sample sizes by province and by cycle. Note that some respondents may have moved to another province or outside of Canada between cycles. The table below shows the sample distribution based on the cycle 1 province or residence. Note also that since our target population is comprised of individuals living in Canada at cycle 1, out-of-country respondents remain in-scope for future cycles. Although they may be more difficult to contact and trace, they were still sent out for Cycle 5 collection. There were few such cases.

Table 7.1 Sample Sizes by Province and Cycle

Province	Sample Size at cycle 1	Sample Size at cycle 2	Sample Size at cycle 3	Sample Size at cycle 4	Sample Size at cycle 5
Newfoundland and Labrador	1,411	1,198	994	778	636
Prince Edward Island	780	632	521	417	362
Nova Scotia	1,826	1,471	1,204	958	831
New Brunswick	1,715	1,309	989	783	667
Quebec	5,881	4,411	3,785	3,043	2,614
Ontario	8,520	6,423	5,549	4,206	3,412
Manitoba	1,952	1,595	1,378	1,077	968
Saskatchewan	2,105	1,727	1,432	1,237	1,033
Alberta	2,380	1,755	1,407	1,172	953
British Columbia	2,594	1,857	1,482	1,082	884
Total	**29,164**	**22,378**	**18,741**	**14,753**	**12,360**

[1] An age interval of 17 to 20 was applied for the rotate-out group from January 1997 to avoid flooding the sample file with a large number of cases that had a high probability of being out of scope. This was the last rotate-out group based on the old LFS questionnaire that collected age rather than date of birth. The actual rotate-out month combined with the year-of-birth definition of the YITS target population resulted in a very low probability of a person with estimated age 21 actually being born in 1979 to 1981 so these persons were classified as not eligible for the YITS sample selection.

8.0 Data Quality

8.1 The Frame

The frame for this survey was constructed during the design stage prior to cycle 1 collection. Since the YITS is strictly a longitudinal survey, they were no frame issues specific to Cycle 5. What follows is a description of the frame as it was constructed in cycle 1. Although what follows is also included in the User Guides for all previous cycles, it is repeated here so that the reader does not need to refer to the other User Guides.

Given the important differences among the educational systems in Canada, the YITS 18-20 survey sought to generate estimates at the province level. For example, from the inception of the YITS, there was a clear desire for province-level estimates of high-school dropout rates among 20-year-old men and women and more generally, for estimates of proportions of at-risk groups by province. The study was also designed to support analysis of characteristics of at-risk groups at a regional level over possibly five cycles of a longitudinal panel. Interest in such characteristics influenced the data quality targets specified at the outset of work on the sample design. In concrete terms the objectives were:

- a CV less than or equal to 16.5% for Cycle 1 province-level leaver rates among 20-year-olds by sex
- in cycle 3, for a characteristic found in a proportion of at least 20% of leavers, a CV less than or equal to 16.5% for that proportion at a regional level.

With some initial work on a sample design it became apparent these CV levels would be extremely difficult to attain for many of the province-level leaver rates. First, there was the question of a suitable frame to survey this relatively rare and highly mobile target population. Although this population comprises about 1.2 million individuals, they are found in only about one in ten households and have a one-year mobility rate of 22%.[9] With the proposed data quality levels it was estimated that if the sample were selected from an up-to-date household frame, the initial sample size would be about 20,000 persons. A sample selected from an older frame would have to be much larger to take into consideration the mobility rates of 18-20 year-olds.

Several options were examined in the search for a sample design based on a single frame. Unfortunately for the YITS, the frame used for the 1991 School Leavers Survey was no longer available because the Family Allowance program, a universally available benefit, had been replaced by the means-based Child Tax Credit. It should be noted that building a new frame was not an option for the YITS 18-20 cohort, due to the large amount of time and resources this would have required.

Single-frame designs based on the Labour Force Survey (LFS), the 1996 Census, random-digit dialling, electoral lists and tax files were considered, each with its own advantages and drawbacks. For the pilot survey a dual-frame dwelling-based design was proposed to combine the strengths of the LFS and Census frames. Central to the dual-frame sample design was the requirement to correctly classify respondents as movers or non-movers, according to whether their usual place of residence at the time of the YITS was different from that at the time of the 1996 Census. Analysis of the data from the pilot survey as well as mobility and school-leaving estimates from the 1996 Census data led us to reject the dual-frame sample design for the main survey. The primary factor in this decision was a projected large downward bias in important survey estimates such as leaver rates, resulting from a high incidence of mover misclassification.[10]

The final choice of frame for the 18-20 year-olds Cohort B main survey was the LFS. It was decided that a sample would be selected from persons in households currently in the LFS sample as well as those that had been in the LFS within the past three years. This restriction limited the YITS sample size, but it was important to recognise the potential risk of trying to trace the highly mobile target population in more dated LFS samples. The design required effective tracing and high response rates to control sampling error and potential bias in estimates for characteristics correlated with mobility. As for the data quality targets, various scenarios based on alternative assumptions for leaver rates, tracing success rates and response

[9] Based on data from the 1996 Census of Population and the Labour Force Survey, 1997 to 2001.

[10] For more details on the sample designs considered see Moloney, Joanne C. and Hidiroglou, Mike C. (2000), Youth in Transition Survey: A Case Study in Design and Development, *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, pages 132-141.

rates were examined. For example, leaver rates from the 1991 School Leavers Survey and the 1996 Census of Population were both included in these analyses. Although the data quality from the proposed design appeared unlikely to meet the initially proposed standards (as did all of the alternatives), the expected levels of sampling error and potential bias were judged to be acceptable.

8.2 Measuring Sampling Error

The estimates derived from this survey are based on a sample of households. 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.

Since it is an unavoidable fact that estimates from a sample survey are subject to sampling error, sound statistical practice calls for researchers to provide users with some indication of the magnitude of this sampling error. This section of the documentation outlines the measures of sampling error which Statistics Canada commonly uses and which it urges users producing estimates from this microdata file to use also.

The basis for measuring the potential size of sampling errors is the standard error of the estimates derived from survey results.

However, because of the large variety of estimates that can be produced from a survey, the standard error of an estimate is usually expressed relative to the estimate to which it pertains. This resulting measure, known as the coefficient of variation (CV) of an estimate, is obtained by expressing the standard error of the estimate as a percentage of the estimate.

Much of the work on the sample design for the 18-20 year-olds cohort of the YITS focused on the desire to obtain reliable estimates of provincial leaver rates for 20-year-old men and women, as well as adequate numbers of respondents to support regional analysis of young adults in at-risk groups. The following tables relate to these survey objectives.

In general, the estimated CVs are consistent with the range predicted by scenarios examined under the final sample design. However, a quick look at analysis and interpretation of the YITS data should include statistical significance tests, and these must be based on estimates of sampling error that take into account the survey design. Caution is required particularly for analysis of rare characteristics in a population subgroup and for any characteristics within small subgroups.

The following tables give an example of sampling error for a handful of variables collected through the YITS. In order to see how estimates and sampling errors are affected by the reduction of sample size for Cycle 5 (due to non-response), the tables show estimates for cycles 1 to 5.

Table 8.1 High-school leaver rates among Cohort B respondents

High school province		**Cycle 1 % (CV)**	**Cycle 2 % (CV)**	**Cycle 3 % (CV)**	**Cycle 4 % (CV)**	**Cycle 5 % (CV)**
10 (NFLD)	Men	13.44 (17.74)	14.85 (19.25)	16.53 (25.90)	15.73 (31.66)	9.65 (31.20)
	Women	5.24 (19.93)	5.08 (23.03)	4.66 (25.52)	4.22 (32.49)	3.53 (39.21)
11 (P.E.I)	Men	14. (29.75)	15.19 (36.22)	20.92 (45.11)	21.24 (53.28)	26.12 (58.17)
	Women	7.7 (24.23)	10.01 (24.91)	9.43 (28.4)	10.03 (29.50)	11.60 (31.74)
12 (N.S,)	Men	13.12 (12.24)	11.47 (14.67)	11.44 (17.89)	11.21 (19.71)	9.81 (24.89)
	Women	5.25 (19.34)	4.95 (23.24)	5.11 (27.81)	4.51 (29.89)	5.10 (31.18)
13 (N.B.)	Men	10.08 (15.51)	8.04 (18.45)	7.2 (25.26)	7.29 (29.46)	7.35 (36.36)
	Women	5.82 (18.48)	6.21 (21.23)	4.98 (27.39)	3.45 (38.20)	3.04 (53.94)
24 (PQ)	Men	20.91 (5.75)	20.26 (6.38)	19.86 (7.38)	19.18 (8.08)	19.58 (9.25)
	Women	10.57 (9.89)	10.19 (11.20)	10.72 (12.85)	9.7 (14.86)	9.83 (16.75)
35 (ON)	Men	9.68 (7.26)	9.08 (8.47)	8.9 (10.76)	8.38 (12.94)	7.26 (15.86)
	Women	7.92 (8.75)	7.69 (9.99)	8.4 (13.46)	8.71 (15.17)	9.76 (17.75)
46 (MAN)	Men	15.73 (11.32)	14.56 (13.00)	14.56 (15.83)	14.92 (16.67)	13.76 (20.72)
	Women	13.39 (13.94)	12.79 (13.23)	11.44 (16.80)	11.01 (17.69)	11.63 (19.38)
47 (SASK)	Men	9.39 (13.17)	8.48 (13.40)	8.19 (14.40)	7.63 (15.54)	7.16 (18.97)
	Women	5.08 (16.45)	4.93 (19.27)	4.79 (20.93)	4.16 (25.27)	4.54 (27.59)
48 (ALTA)	Men	16.12 (13.91)	16.2 (15.80)	15.59 (17.59)	14.99 (15.84)	15.82 (17.56)
	Women	12.15 (11.02)	12.53 (12.30)	12.74 (13.67)	13.27 (14.80)	13.28 (17.45)
59 (B.C)	Men	12.88 (13.46)	14.04 (15.23)	14.6 (17.00)	13.26 (17.82)	11.69 (20.42)
	Women	8.02 (13.58)	7.95 (17.11)	8.26 (18.09)	8.07 (20.71)	9.33 (22.88)
All provinces	Men	13.87 (3.74)	13.52 (4.26)	13.4 (5.03)	12.79 (5.50)	12.19 (6.28)
	Women	8.88 (4.69)	8.72 (5.35)	9.08 (6.55)	8.88 (7.37)	9.50 (8.68)

Table 8.2 Self-assessment of high-school leavers' communication skills

Region		Cycle 1 % (CV)	Cycle 2 % (CV)	Cycle 3 % (CV)	Cycle 4 % (CV)	Cycle 5 % (CV)
Atlantic	Poor/Fair	16.30 (10.63)	15.40 (12.61)	13.76 (15.96)	13.21 (18.13)	13.35 (23.31)
	Good	8.80 (9.60)	7.73 (10.84)	7.77 (12.87)	7.41 (15.81)	7.39 (19.16)
	Very good/ Excellent	6.06 (12.17)	5.76 (13.85)	5.38 (17.67)	4.66 (21.25)	3.64 (22.52)
Quebec	Poor/Fair	26.63 (9.48)	24.25 (11.30)	24.48 (12.49)	22.71 (14.34)	22.80 (15.26)
	Good	16.31 (8.01)	16.27 (9.01)	16.67 (10.69)	16.30 (11.65)	16.55 (13.13)
	Very good/ Excellent	12.31 (7.87)	12.07 (8.72)	11.84 (9.76)	10.76 (10.72)	11.21 (11.91)
Ontario	Poor/Fair	12.40 (13.49)	11.99 (16.16)	14.17 (20.40)	14.86 (23.60)	15.79 (29.98)
	Good	9.76 (8.64)	9.22 (10.17)	8.86 (14.33)	8.86 (17.29)	8.38 (17.35)
	Very good/ Excellent	7.28 (9.48)	7.04 (10.58)	7.28 (12.09)	7.01 (13.96)	6.54 (17.48)
Prairies	Poor/Fair	18.33 (15.99)	18.68 (19.44)	18.72 (20.75)	17.25 (15.78)	18.52 (19.44)
	Good	12.77 (9.05)	12.58 (10.34)	13.24 (11.83)	14.07 (13.27)	14.06 (13.58)
	Very good/ Excellent	10.74 (8.81)	10.77 (9.76)	10.26 (11.63)	10.06 (13.29)	9.39 (16.83)
British Columbia	Poor/Fair	10.61 (24.77)	10.41 (30.72)	15.62 (33.85)	9.99 (33.44)	13.41 (35.26)
	Good	10.65 (11.87)	12.51 (12.70)	12.47 (16.48)	12.25 (18.93)	13.26 (22.53)
	Very good/ Excellent	9.16 (15.18)	8.8 (18.64)	6.94 (21.73)	6.51 (25.59)	5.98 (32.18)
All regions	Poor/Fair	17.31 (6.43)	16.45 (7.82)	17.81 (9.00)	16.58 (9.58)	17.58 (11.48)
	Good	11.93 (4.40)	11.89 (5.10)	11.99 (6.16)	12.03 (7.03)	12.05 (7.74)
	Very good/ Excellent	9.21 (4.71)	8.98 (5.24)	8.7 (5.94)	8.19 (6.52)	7.84 (7.96)

Table 8.3 Self-assessment of high-school leavers' problem-solving skills

Region		Cycle 1 % (CV)	Cycle 2 % (CV)	Cycle 3 % (CV)	Cycle 4 % (CV)	Cycle 5 % (CV)
Atlantic	Poor/Fair	15.14 (10.42)	14.45 (12.38)	13.89 (15.65)	14.4 (16.99)	13.42 (22.63)
	Good	8.97 (8.98)	7.94 (11.14)	7.67 (13.53)	7.06 (16.26)	6.60 (17.05)
	Very good/ Excellent	5.86 (12.47)	5.63 (14.85)	5.08 (18.72)	4.17 (21.07)	3.77 (23.97)
Quebec	Poor/Fair	29.75 (10.16)	29.33 (12.24)	26.97 (15.47)	25.38 (18.01)	24.31 (19.15)
	Good	17.21 (7.12)	16.37 (7.96)	16.87 (9.33)	15.81 (11.09)	17.15 (11.97)
	Very good/ Excellent	11.47 (8.35)	11.52 (9.36)	11.96 (9.98)	11.37 (10.91)	10.96 (12.87)
Ontario	Poor/Fair	12.11 (14.20)	11.01 (18.37)	11.89 (24.83)	12.48 (30.72)	8.36 (28.22)
	Good	10.21 (7.36)	10.08 (8.37)	10.16 (11.17)	10.46 (13.28)	11.30 (16.02)
	Very good/ Excellent	6.53 (10.35)	6.09 (11.87)	6.48 (14.15)	5.99 (16.89)	5.28 (21.36)
Prairies	Poor/Fair	21.57 (16.99)	22.9 (20.28)	24.35 (20.27)	21.95 (17.27)	24.24 (19.69)
	Good	11.7 (7.68)	11.79 (8.56)	11.53 (10.15)	11.94 (11.62)	13.33 (14.04)
	Very good/ Excellent	11.41 (9.67)	10.98 (11.31)	10.97 (13.48)	11.08 (15.05)	8.52 (15.50)
British Columbia	Poor/Fair	11.72 (22.61)	12.37 (26.55)	9.82 (34.09)	12.89 (34.40)	17.45 (35.80)
	Good	10.16 (12.56)	11.03 (14.49)	12.62 (18.33)	10.56 (18.57)	11.12 (21.70)
	Very good/ Excellent	9.06 (17.27)	9.1 (19.52)	7.93 (22.65)	6.57 (27.38)	5.76 (35.82)
All regions	Poor/Fair	18.17 (7.11)	17.73 (8.63)	17.33 (10.13)	17.47 (10.71)	17.09 (10.83)
	Good	11.95 (3.83)	11.76 (4.39)	12.07 (5.50)	11.65 (6.41)	12.55 (7.57)
	Very good/ Excellent	8.92 (4.99)	8.69 (5.60)	8.79 (6.41)	8.27 (7.29)	7.35 (8.90)

8.3 Non-sampling error

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 and errors may be introduced in the processing and tabulation of the data. These are all examples of non-sampling errors. Considerable time and effort were used to reduce non-sampling errors in the survey. Quality assurance measures were implemented at each step of the data collection and processing cycle to monitor the quality of the data. These measures included the use of highly skilled interviewers, extensive training of interviewers with respect to the survey procedures and questionnaire, observation of interviewers to detect problems of questionnaire design or misunderstanding of instructions, procedures to ensure that data capture errors were minimised and coding and edit quality checks to verify the processing logic (see ***Section 4.0***).

A major source of non-sampling errors in surveys is the effect of non-response on the survey results. The extent of non-response varies from partial non-response (failure to answer just one or some questions) to total non-response. Total non-response occurred because the interviewer was either unable to contact the respondent, or the respondent refused to participate in the survey. Total non-response was handled by adjusting the weight of households who responded to the survey to compensate for those who did not respond.

In most cases, partial non-response to the survey occurred when the respondent did not understand or misinterpreted a question, refused to answer a question, or could not recall the requested information.

Partial non-response was generally low for the YITS Cycle 5. Table 8.4 summarizes partial non-response for the YITS Cycle 5 questionnaire. In general, item non-response was not a significant problem, although

there are a few questions for which the rate was very high. Note that these rates do not include income variables and derived variables that are usually dependent on more than one question and could therefore have slightly higher non-response rates. In addition, table 8.4 only presents item non-response rates for variables with less than 50% of values coded as valid skip, and for modules of the questionnaire where there are at least 5 variables that meet these criteria. Codebooks should be used to get non-response rates for specific variables.

Table 8.4 Item Non-Response Rates (NR)

Modules	Number of variables	Min item NR rate	Max item NR rate	Mean item NR rate	Median item NR rate
B	15	0.86	6.28	2.33	1.31
H	48	0.00	29.72	4.69	1.34
K	23	0.00	9.14	1.21	0.63
L	43	1.08	15.52	2.19	1.30
M	8	1.26	1.34	1.28	1.28
P1	69	0.00	1.27	0.39	0.02
P2	16	0.00	0.47	0.08	0.00
P6	11	1.56	1.99	1.90	1.94
PS	75	0.90	1.55	1.08	0.90
Q	9	2.06	2.10	2.09	2.09
R	9	2.17	2.98	2.31	2.19
U	25	0.05	4.54	2.04	2.29
V	24	0.85	1.65	1.25	1.25

8.4 Response rates

Tables 8.5 to 8.7 present response rates for different subgroups of the sample. These rates are not weighted and use, as a base, the initial YITS sample count excluding a priori non-response households. Since YITS is longitudinal using a funnel approach, this means that only responding units from Cycle 4 were followed in Cycle 5. As a result, the initial sample size for Cycle 5 was 12,360 for Cohort B. The respondent count includes persons who were interviewed, persons contacted but confirmed to be outside the YITS target population by year of birth and persons whom a household contact confirmed as deceased. (The latter two groups are included in the respondent count because they provided all the relevant information, given their special status). A derived variable DECEASE5 was created to provide for these cases. – see *Section 11.0*. Table 8.5 shows provincial response rates for cycles 1 through 5.

Note that the response rates reported for cycles 3 and 4 in the User Guides for these cycles did not include respondents confirmed as deceased (as indicated above). This accounts for any discrepancies between previously published response rates and those found in Table 8.5.

The province-level response rates in Table 8.5 show considerable variation. Between the two demographic characteristics age and sex, age seems to be a more influential factor, although the age-sex response rates (Table 8.6) are more uniform than the province-level rates. All of these factors, among others, were examined for a possible role in the non-response adjustment of the weights (see *Section 11*.0 on weighting). Potential bias due to differing response rates by age and sex were also taken into account by the post-stratification step in the weighting process. Table 8.7 presents the longitudinal response rate over the four five cycles of the survey, by age and gender. Although there are slight differences between the age groups, what is important to note is that we are left with roughly 36% of our original cycle 1 sample at the conclusion of Cycle 5.

Table 8.5 Provincial Response Rates – Cycles 1 through 5

Province	Response rates for each cycle (%)					Longitudinal Response Rate (%)
	1	2	3	4	5	
Newfoundland & Lab	84.9	83.0	78.4	81.9	74.3	33.6
Prince-Edward Island	81.0	82.6	81.0	87.0	79.9	37.7
Nova-Scotia	80.6	81.9	79.8	87.1	78.3	35.9
New-Brunswick	76.3	75.7	79.3	85.2	75.2	29.4
Quebec	75.0	86.0	80.4	86.1	87.2	39.0
Ontario	75.4	86.5	76.0	81.2	75.9	30.5
Manitoba	81.7	86.6	78.3	89.9	82.4	41.1
Saskatchewan	82.0	83.1	86.5	83.7	80.9	40.0
Alberta	73.7	80.6	83.4	81.5	81.0	32.7
British-Columbia	71.6	80.0	73.2	81.8	78.9	27.0
All provinces	76.7	83.9	78.9	83.9	80.0	34.1

Table 8.6 Response rates for Cycle 5 by sex and age as of December 2007

Age (at Cycle 5)	Response rate (%)		
	Men	Women	Both sexes
26	79.4	80.1	79.8
27	78.9	80.7	79.8
28	78.9	81.9	80.5
All ages	79.1	80.8	80.0

Table 8.7 Longitudinal response rate by age and gender

Age (at Cycle 5)	Response rate (%)		
	Men	Women	Both sexes
26	35.8	38.5	37.1
27	35.0	37.8	36.4
28	31.9	36.9	34.4
All ages	34.3	37.7	36.0

9.0 Imputation of Missing Data for Income and Earnings Variables

For quantitative variables such as wages and total earnings, imputation was carried out rather than using special non-response codes. Imputation is the process by which missing or inconsistent items are “replaced” with plausible values. When carried out properly, imputation can improve data quality by reducing non-response bias. It also has the advantage of producing a complete data set for those variables being imputed.

The first step in the imputation process was a within-record imputation where missing information was replaced with values derived from the respondent’s answer to other questions in the questionnaire using deterministic edit rules. In a few cases, “capping” was used, meaning that a respondent’s answer was changed to a preset maximum or minimum allowable value for that variable. The remaining missing data were imputed using nearest-neighbour donor imputation. This is a widely used technique for treating item non-response. It aims at replacing missing information for a respondent with values provided from another respondent which is “similar” to him. Rules for identifying the respondent most similar to the non-respondent can vary depending on the variable being imputed. Donor imputation methods have good properties and generally will not alter the distribution of the data which is a drawback of many other imputation techniques. Once the nearest neighbour imputation was done, within-record editing was performed again to ensure consistency of the data.

For the YITS, a list of the variables for which imputation was carried out can be found in Table 9.1 that follows. Essentially, imputation was done for all of the earnings and income variables, for each of jobs 1 through 4 and job 7. The table shows an overall imputation rate for all jobs combined for each of the variables. Note that although imputation generally improves overall data quality, the artificial data created are used in estimation and can lead to underestimation of the sampling errors. This would only be a concern for variables with high imputation rates.

Table 9.1: Imputation rates for income and earnings

Note: Earning variables are derived variables and are considered as being imputed if at least one of the components in deriving the earning was imputed.

Derived Variable	Cohort B		
	Number of imputed values	# records where questions apply	Imputation rate (%)
From module P2:			
Ephei5	624	16 512	4.3
Epwei5	624	16 512	4.3
Epmei5	624	16 512	4.3
Ephsi5	828	16 512	10.7
Epwsi5	828	16 512	10.7
Epmsi5	828	16 512	10.7
Nwwpmei5	139	16 512	1.2
Nhwpmei5	231	16 512	2.0
Nwwpmsi5	338	16 512	8.2
Nhwpmsi5	379	16 512	8.8
From module V:			
Incwsd5	0	9 861	0.0
Incsed5	0	9 861	0.0
Incsgd5	182	9 861	1.8
Inceid5	277	9 861	2.8
Incstd5	436	9 861	4.4
Incspd5	177	9 861	1.8
Incscd5	170	9 861	1.7
Incnrd5	271	9 861	2.7
Inctbd5	225	9 861	2.3
Incgsd5	180	9 861	1.8
Incond5	228	9 861	2.3
Tincd5	747	9 861	7.6
Incwssd5	503	5 068	9.9
Incsesd5	132	5 068	2.6
Incsgsd5	75	5 068	1.5
Inceisd5	182	5 068	3.6
Incstsd5	305	5 068	6.0
Incspsd5	43	5 068	0.8
Incscsd5	43	5 068	0.8
Incnrsd5	107	5 068	2.1
Inctbsd5	137	5 068	2.7
Incgssd5	54	5 068	1.1
Inconsd5	109	5 068	2.2
Tincsd5	882	5 068	17.4

10.0 Guidelines for Tabulation, Analysis and Release

This section of the documentation outlines the guidelines to be applied by users tabulating, analysing, publishing or otherwise releasing any data derived from the survey microdata files. With the aid of these guidelines, users of microdata 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.

10.1 Rounding Guidelines

In order that estimates for publication or other releases derived from these microdata files 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 are then 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. In normal rounding to a single digit, if the final 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 increased by 1.

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.

10.2 Sample Weighting Guidelines for Tabulation

The sample design used for the YITS was not self-weighting. When producing simple estimates 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 microdata files cannot be considered to be representative of the survey population, and will not correspond to those produced by Statistics Canada.

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

10.2.1 Definitions of Types of Estimates: Categorical vs. Quantitative

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

Categorical Estimates

Categorical estimates are estimates of the number, or percentage of the surveyed population possessing certain characteristics or falling into some defined category. Whether the respondent has had any education in Canada or has worked more or less than 30 hours per week at a job 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: Did you attend an elementary, junior high or high school in Canada
R: Yes / No

Q: Did you usually work 30 hours or more per week or less than 30 hours per week at this job?
R: 30 hours or more / Less than 30 hours

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 $\hat{X} / \hat{Y}$ where $\hat{X}$ is an estimate of the surveyed population quantity total and $\hat{Y}$ is an estimate of the number of persons in the surveyed population contributing to that total quantity.

An example of a quantitative estimate is calculating the average number of hours worked per day by the surveyed population when they first started their job. The numerator could be an estimate of the total number of hours worked per week when they first started, and the denominator could be an estimate of the total number of days worked per week when they first started.

Examples of Quantitative Questions:

Q: When you first started this job, how many hours did you usually work per week?
R: |_|_| hours

Q: When you first started this job, how many days per week did you usually work?
R: |_|_| days

10.2.2 Tabulation of Categorical Estimates

Estimates of the number of people with a certain characteristic can be obtained from the microdata file by summing the final weights of all records possessing the characteristic(s) of interest. Proportions and ratios of the form X/Y are obtained by:

(a) summing the final weights of records having the characteristic of interest for the numerator (X),
(b) summing the final weights of records having the characteristic of interest for the denominator (Y), then
(c) dividing the numerator estimate by the denominator estimate.

10.2.3 Tabulation of Quantitative Estimates

Estimates of quantities can be obtained from the microdata 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 hours worked per week for those people who work part-time, multiply the value of total number of hours worked per week by the final weight for the record, and then sum this value over all records who reported working part-time.

To obtain a weighted average of the form X/Y, the numerator (X) is calculated as for a quantitative estimate and the denominator (Y) is calculated as for a categorical estimate. (Note: This applies when Y represents a subgroup of the survey population, but the characteristic Y could also be a quantitative estimate, as in the example above (in Section 10.2.1) for average number of hours worked in a day.) For example, to estimate the number of hours worked per week for those people who work part-time,
(a) estimate the total number of hours per week as described above,
(b) estimate the number of people in this category by summing the final weights of all records who reported working part-time
(c) divide estimate (a) by estimate (b).

10.3 Guidelines for Statistical Analysis

The YITS is based upon a complex sample design, with stratification, multiple stages of selection, and unequal probabilities of selection of respondents. Using data from such complex surveys presents problems to analysts because the survey design and the selection probabilities affect the estimation and variance calculation procedures that should be used. In order for survey estimates and analyses to be free from bias, 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 variances that are calculated are poor.

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

For example, suppose that analysis of all male respondents is required. The steps to re-scale the weights are as follows:

- select all respondents from the file who reported SEX=male
- calculate the AVERAGE weight for these records by summing the original person weights from the microdata file for these records and then dividing by the number of respondents who reported SEX=male
- for each of these respondents, 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 design are still not taken into account, the variance estimates calculated in this way are likely to be under-estimates.

For more information on calculating variance estimates for the YITS, see ***Section 12.0***.

10.4 CV Release Guidelines

Before releasing and/or publishing any estimate from the YITS, 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 discussed in ***Section 8.0***. However, for this purpose, the quality level of an estimate will be determined only based on sampling error as reflected by the coefficient of variation as shown in the table below. Nonetheless, users should be sure to read ***Sections 8.0 through 8.4*** to be more fully aware of the quality characteristics of these data.

First, the number of respondents 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. (The figure “30” is for use with LFS based surveys and other surveys with generally small sampling fractions. From time to time, a lower figure may be appropriate for surveys with higher sampling fraction.)

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 provided in Table 10.1. 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.

Table 10.1 Quality Level Guidelines

Quality Level of Estimate	**Guidelines**
1. Acceptable	Estimates have: a sample size of 30 or more, and coefficients of variation in the range 0.0% - 16.5% No warning is required.
2. Marginal	Estimates have: a sample size of 30 or more, and coefficients of variation in the range 16.6% - 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 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 a warning should accompany the estimates.

Note: Applies only to weighted rounded estimates

11.0 Weighting

The YITS 18-20 cohort weighting strategy takes the final weight from the previous cycle and applies a non-response adjustment. Therefore, the base weight for all cycles is the cycle 1 weights. As the sample design of the YITS 18-20 cohort was determined to a large degree by the sample design of the LFS, so was the cycle 1 weighting. A complete description of the cycle 1 weighting procedure can be found in *Section 9.0* of the Cycle 1 User Guide for this cohort (Project 4435). The starting weight for the Cycle 5 records was the final Cycle 4 weight before post-stratification was applied. This weight took into account the LFS sample design as well as the YITS sample design. The purpose of the Cycle 4 weight adjustment is strictly to account for non-response that occurred during Cycle 5. To account for people who participated in Cycle 4 and did not participate in Cycle 5, the final Cycle 4 weights (before post-stratification) of those who participated in both cycles were proportionally increased so that the sum of their adjusted weights would equal the sum of the Cycle 4 final weights. Although the response rate for Cycle 5 was high, analysis of non-response patterns showed that the non-response adjustments should take into consideration certain variables. The adjustments were made separately within response classes defined by those variables. Variables used included province, income, gender, language of interview and family structure.

Note also that respondents who were “deceased” in Cycle 5 were treated as respondents since they represented others in the target population that have died since cycle 1but were not in the sample. For these deceased respondents, their Cycle 5 data was set to missing and a flag called DECEASE5 was created in order to easily identify these cases. The deceased do have a longitudinal weight and the reason they are kept on the file is so that the sum of the weights is consistent with the sum of the weights from Cycle 4 These records should be removed from most analyses since they do not contain any information from the Cycle 5 questionnaire.

The final weighting adjustment for the YITS 18-20 year-olds cohort is post stratification. The adjustment was applied within post-stratification classes defined by year of birth, sex and province. Official post-censal demographic projections for the LFS target population are created for single-year age groups, but not by year of birth. The control totals used for the YITS are averages of population counts derived from the LFS age control totals for the 12 months of 2007. Since YITS is strictly longitudinal and we wish to represent the population of 18-20 year-olds as they were defined at cycle 1, the control totals we calibrated to be the same ones as we calibrated to in cycle 1.

The weights resulting form the previous steps were adjusted proportionally so that the sum of their final weights would be equal to the control total in their respective post-stratification classes.

12.0 Variance Estimation

Due to the complexity of the YITS sample design a re-sampling technique was chosen to calculate estimates of the variance. For the 18-20 cohort, the bootstrap re-sampling method is used. This technique is popular among surveys with a large number of strata and multiple primary sampling units (PSU) per stratum. Unlike the Jackknife method the bootstrap does not suffer from inconsistent estimates for population estimates such as percentiles.

The idea behind the bootstrap method is to select random sub-samples from the full sample in such a way that each of the sub-samples (or replicates) follows the same design as the full sample. The final weights for units in each replicate are recalculated, following the same weighting steps used for the full sample (see *Section 11.0*)**.** These bootstrap weights are used to calculate a population estimate for each replicate. The variance among the replicate estimates for a given characteristic is an estimate of the sampling variance of the full-sample population estimate.

For the YITS 18-20 cohort a total of 1,000 replicates were created at Cycle 4 and those same replicate weights (again before post-stratification) are the starting point for deriving the Cycle 5 replicate weights. Each replicate was initially formed by sampling independently within each stratum. If there were *n* PSUs in a stratum, (n-1) were selected by simple random sampling with replacement. While sampling with replacement to create the bootstrap samples is a departure from the full-sample design for the YITS, this is a common practice in large surveys with small first-stage sampling fractions because it greatly simplifies the variance estimation process at the expense of overestimating the true variance slightly.

13.0 Working With YITS Files

13.1 Roster and Flat Files

File or Roster	Cohort B – 18-20 year-olds (26-28 year-olds in 2007)
Main Flat File	Cycle 5 - Cohort B - Person Level Main File
K roster	Cycle 5 - Cohort B - Post Secondary Engagement Roster
Hinst roster	Cycle 5 - Cohort B - Education above High School (Institution Roster)
Hprog roster	Cycle 5 - Cohort B - Education above High School (Program Roster)
P1cycle4 roster	Cycle 5 - Cohort B - Confirmation of Open Jobs from Cycle 4 Roster
P1cycle5 roster	Cycle 5 - Cohort B - Job Roster
P2 roster	Cycle 5 - Cohort B - Job Details Roster
Kids Roster	Cycle 5 - Cohort B - Dependent children

13.2 Youth In Transition Survey: Data Extraction Tool

13.2.1 About Youth in Transition Survey (YITS)

The Youth is Transition Survey (YITS) is designed to examine the patterns of, and influences on, major transitions in young people's lives, particularly with respect to education, training and work. Human Resources and Skills Development Canada and Statistics Canada have been developing the YITS in consultation with provincial and territorial ministries and departments of labour and education. Content includes measurement of major transitions in young people's lives including virtually all formal educational experiences and most about-market experiences, achievement, aspirations and expectations, and employment experiences. The implementation plan encompasses a longitudinal survey of each of two cohorts, ages 15 and 18-20, to be surveyed every two years.

The results from the Youth in Transition Survey will have many uses. Human Resources and Skills Development Canada will use them to aid policy and program development. Other users of the results include educators, social and policy analysts, and advocacy groups. The information will show how young adults are making their critical transitions into their adult years.

These researchers and analysts will have access to important information that can be used in developing programs to deal with both short-term and long-term problems or barriers that young adults may face in their pursuit of higher education or in gaining work experience. Information from the survey will help to evaluate the effectiveness of existing programs and practices, to determine the most appropriate age at which to introduce programs, and to better target programs to those most in need.

Young adults themselves will be able to see the impact of decisions relating to education or work experiences. They will be able to see how their own experiences compare to those of other young adults.

13.2.2 Statistical Activity

PISA/YITS is one project, which consists of two parallel survey programs: the Programme for International Student Assessment (PISA) and the Youth in Transition Survey (YITS).

PISA is an international assessment of the skills and knowledge of 15 year-olds which aims to assess whether students approaching the end of compulsory education have acquired the knowledge and skills that are essential for full participation in society.

YITS is designed to examine the patterns of, and influences on, major transitions in young people's lives, particularly with respect to education, training and work. Human Resources and Skills Development Canada and Statistics Canada have been developing the YITS in consultation with provincial and territorial ministries and departments of labour and education. Content includes measurement of major transitions in young people's lives including virtually all formal educational experiences and most about-market experiences, achievement, aspirations and expectations, and employment experiences. The implementation plan encompasses a longitudinal survey of each of two groups, ages 15 and 18-20, to be surveyed every two years.

The 15 year-old respondents to the Reading Cohort (conducted in 2000) participated in both PISA (record number 5060) and YITS (record number 5058). Starting in 2002, they will be followed up longitudinally by YITS (record number 4435).

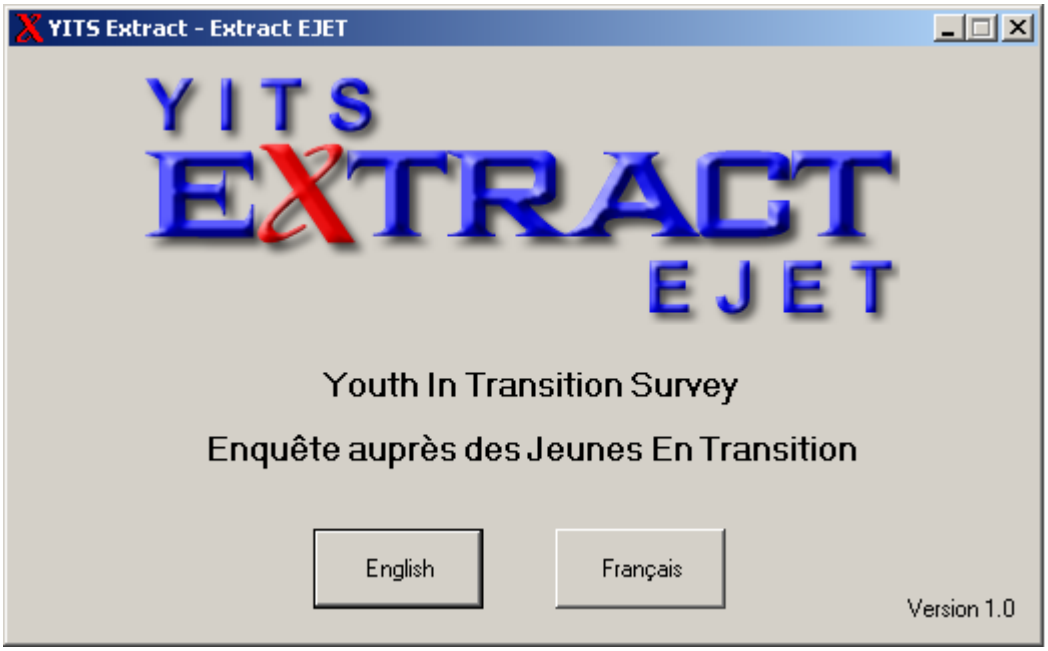
13.2.3 Purpose of the Application

The YITS data sets are many, large and are stored in two different formats (normalized data for the main file with rostered or un-normalized data for information collected that have many iterations). As the survey continues and more cycles of information are available the number of files and complexity in figuring out how to use the files will grow to a point where it will take more time programming and running a merge sequence than actually researching.

The YITS: Data Extraction Tool will facilitate the process users go through to create their files used in research and analysis. The application provides an intuitive and direct interface for users to select the specific variables they need in order to produce their findings. Behind the interface the program will normalize the information that is in roster format, merge each of the individual data files from which variables were selected and assign the weight file based on the cycle and cohort choices of the user. Due to the complexity of assigning weights (i.e. depending on what variables and cohort are selected there is a lot of room for error) the program will also automatically assign the appropriate bootstrap weights to the file. The design of the application also takes into account the longitudinal aspect of the survey when merging the data files, the number of observations in the final data file will be based on the most recent cohort selected by the user.

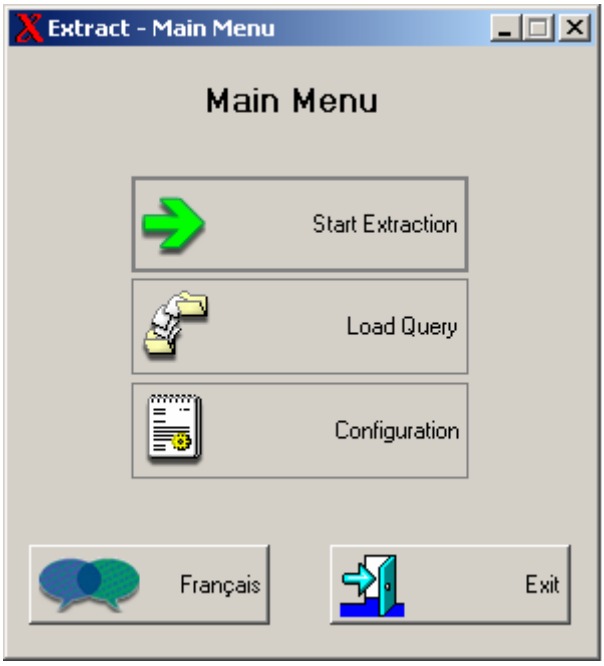
This application will do much to facilitate the initial use of YITS data, unfortunately it isn't a tool to explain how to use the survey in analysis or research. The data files are merged according to the design of the survey; to understand why the tool was necessary along with the overall design of PISA & YITS the researcher must consult other materials. Using the data extraction tool in conjunction with the codebooks, questionnaires and user guides will provide a researcher with a solid foundation for their work. Other reference materials that may be used are available on the Statistics Canada website under the link for definitions, data sources and methods section (http://www.statcan.gc.ca/english/sdds/0020t.htm).

Splash Page



Welcome to the Extraction Tool: The language selection buttons serve two purposes: selecting the language of the application and determining the language of the output file and formats

Main Menu

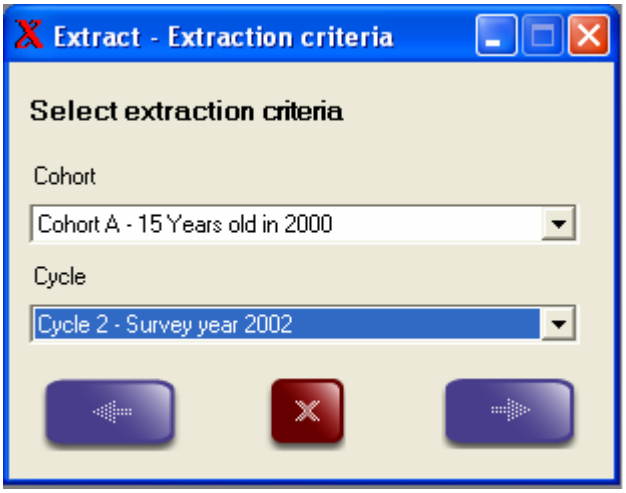


Main Menu: This is the central menu for the application. From this point you can start a new extraction, load a previously saved extraction routine, change the language of the extraction, close down the application and go into the configuration settings. Please note that the users query can be saved in the "Select variables menu". Also, a query that had previously been saved in a particular

language will work even if the language of the application has changed. When the user cancels their extraction in subsequent menus the application automatically brings them back to the main menu. The configuration settings are for the local administrator to use.

Users must be linked to the YITS Data Extraction Tool using the installation procedures found in the YITS Administration Documentation.

Extraction Criteria



Primary PISA/YITS Extraction Criteria: The first of two Extraction Criteria Menus where users can specify their population of interest.

Cohorts:

- Cohort A – 15 year-olds in year 2000
- Cohort B – 18-20 year-olds in year 2000

Cycle:

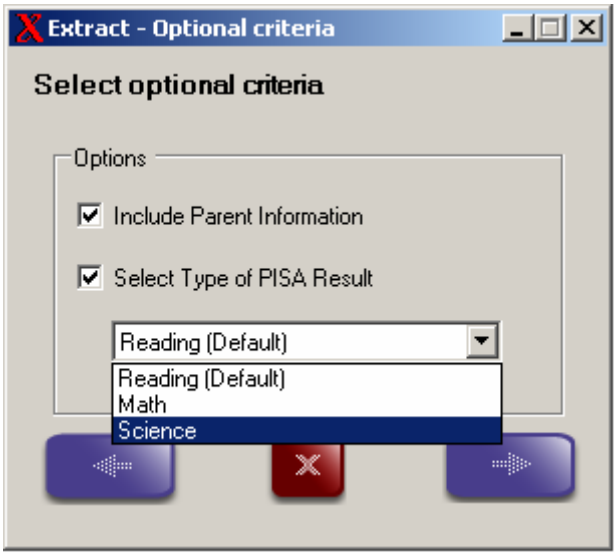
- Cycle 1 – Survey year 2000
- Cycle 2 – Survey year 2002
- Cycle 3 – Survey year 2004
- Cycle 4 – Survey year 2006
- Cycle 5 – Survey year 2008

Navigation Buttons:

- Back arrow – Main menu
- X – Main menu
- Forward arrow:
    - If cohort A selected – Optional Criteria menu
    - If cohort B selected – Select Data menu

For more information on the cohorts and cycles please refer to this User Guide or under Definitions, Data Sources and Methods on the Statistics Canada website.

Optional Extraction Criteria



Optional Extraction Criteria Menu: If Cohort A (15 year-olds) is selected from the Primary Extraction Criteria menu the Optional Criteria menu opens.

Users can:

- Select the parent information
- Choose which PISA test results they wish to use

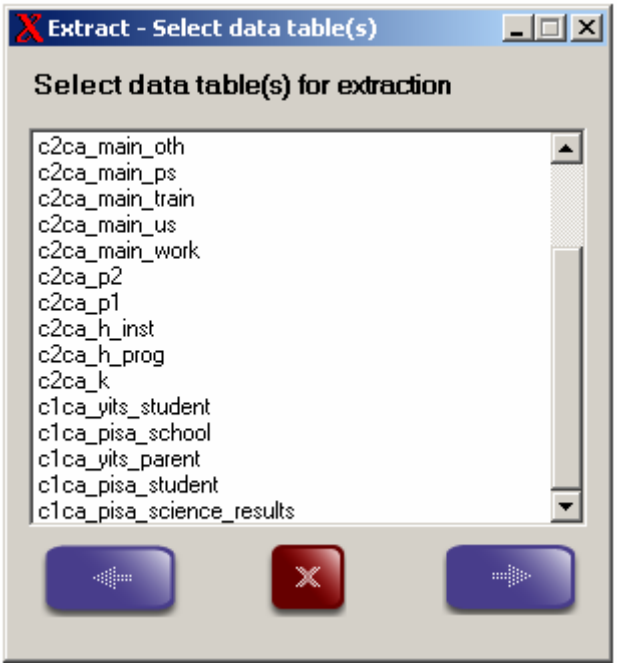
Navigation Buttons:

- Back arrow – Extraction Criteria menu
- X – Main menu
- Forward arrow - Select Variables menu

For each cycle of PISA/YITS there are seven sample weights to choose from, each with 1000 bootstrap weights for calculating the variance estimates. The Extraction Criteria menus allow the user to select any of the possible paths a respondent could have followed and then assigns the appropriate weight to the output data file based on what the user has chosen to look at.

It is important to note that if a user does not wish to have parent information or test results included in their analysis file they do not have to select anything. The reading weights are assigned to the file by default (according to the design of the survey). Selecting parent information or either the Mathematics or Science results will have a significant impact on the sample size of the output data file. Each of the options presented in the Optional Extraction Criteria Menu are a sub sample of the PISA/YITS population. For more information on how each of the selections may affect the sample extracted please refer to Section 8.0 Data Quality in this User Guide or under Definitions, Data Sources and Methods on the Statistics Canada website (www.statcan.gc.ca) under Record Numbers 4435, 5058 or 5059.

Select Data Tables



Selecting Data Files: Once the user has completed the Extraction Criteria menus the application presents the user with a list of data files to choose from. If the user chose to use information for a later cycle of the survey, all data files from previous cycles of PISA/YTIS will be made available for browsing.

For the Record Number 4435 (YITS), the Main Person Level file has been broken down into multiple data sets of about 100 variables each for ease of use with this application. This has been done for Cohort B (18 to 20 year-olds) in Cycle 1 and for both Cohorts in subsequent cycles. Each of the data files have been loosely grouped around the modules of the YITS questionnaire. For example, the data file "c2ca_main_us" contains variables related to the questions about moving to the US – Module A.

For a complete list of the data files and the variables contained within refer to the Statistics Canada website under Definitions, Data Sources and Methods for Record Numbers 4435, 5058 or 5059.

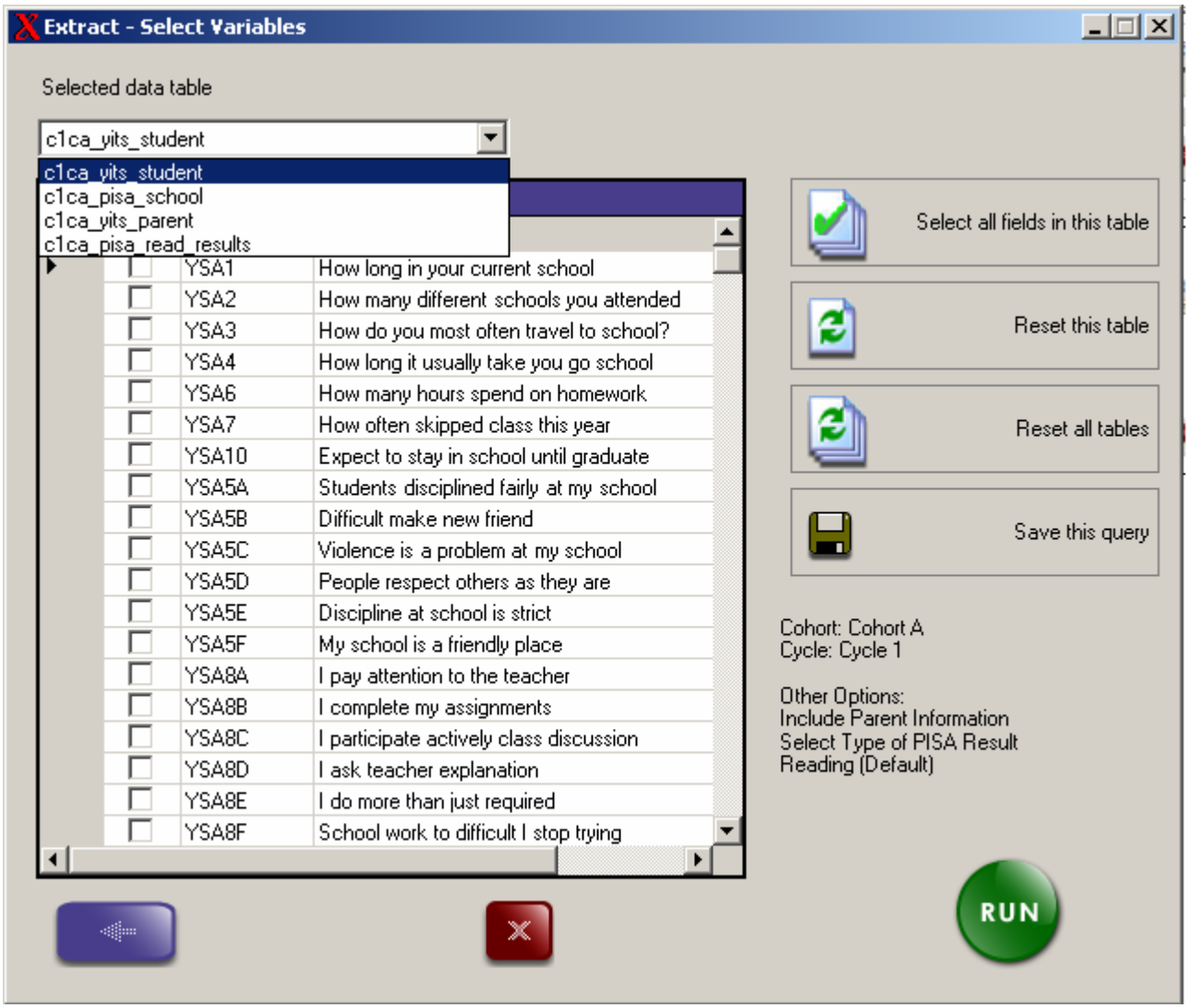
Naming Convention For the Data Files:

- **c5cb** - **C**ycle 5 **C**ohort **B;**
- **c4cb** - **C**ycle **4** **C**ohort **B;**
- **c3cb** - **C**ycle **3** **C**ohort **B;**
- **c2cb** - **C**ycle **2** **C**ohort **B;**
- **c1ca** – **C**ycle **1** **C**ohort **A**
- _Main - indicates that data file is part of the main respondent information (Note if _Main isn't included in the cycle 2 file name it is a rostered file)
- _dem, _fund, _work etc. – indicates which subject or module is represented in the particular data file
- _yits or _pisa – are for cohort A indicating whether the information is from the PISA or YITS component of the survey.

Navigation Buttons:

- Back arrow
  - If cohort A selected – Optional Criteria menu
  - If cohort B selected – Extraction Criteria menu
- X – Main menu
- Forward arrow - Select Variables menu

Select Variables



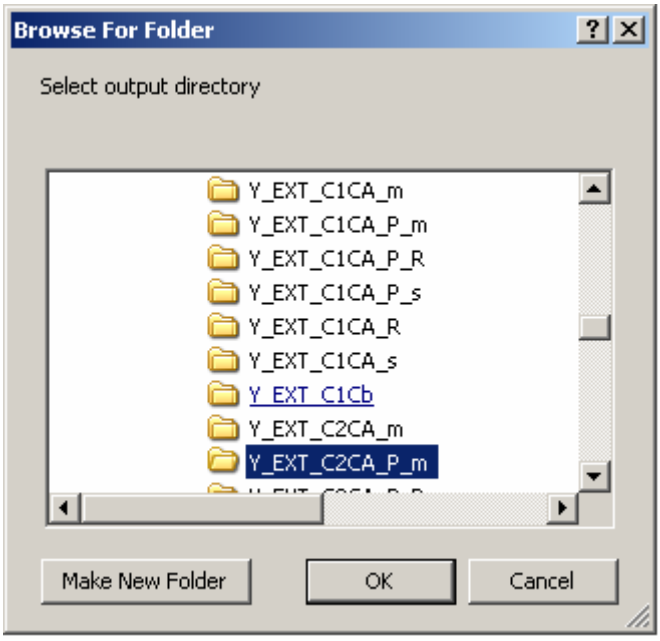
Selecting Variables for Output:

- Drop down menus allows user to select the file from which they can choose their variables.
    - There is a counter next to the data file name to show how many variables have been selected.
- Some buttons have been provided to aid the user in:
    - Selecting all variables in the selected file
    - Resetting their choices for the selected file
    - Resetting the entire variable selection process.
- Save Query button allows the user to save their current extraction set-up so that they can update their extraction when further information is necessary.
- Notes are provided for the user to document from which cycle and cohort data are being extracted.

Navigation Buttons:

- Back arrow – Select Data menu
- X – Main menu
- Run – Starts the extraction process, opens the Browse for output folder. Please Note: Pressing the RUN button without selecting any variable from any table creates a file containing only the weights and the bootstrap weights.

Output Folder



Selecting the Output Directory:

- Allows users to output the information to a specific directory
- If there are data files or other information already in the output directory there will be a warning that the program may overwrite a data file.
- The output is in the form of
    - SAS & SPSS cards, set to run in the output folder selected.
    - An ASCII file (.dat) that contains the variables selected and along with the appropriate weights (the sample weight and 1000 bootstrap weights).

13.2.4 Saving and Loading Queries

Using the Output Files

The extraction tool produces four output files for each run:

- One data file in ascii format containing the variables of interest and the weight variables. The name of the data file is generated automatically by the program
- Two SAS programs are generated:
  - the program with the prefix “Create” will generate the SAS data file with all of the available formats & labels applied to it.
  - the SAS program with the prefix “Lrecl” contains the record layout for the ascii file and is referenced by the first SAS program.
- The SPSS program also generated by the extraction tool will create an SPSS file from the ascii data file.

Example of output
Ascii file containing information from: Cycle 2 Cohort A (C2CA), selected parent variables (_PAR), selected PISA variables (_PISA), English labels and formats (_E), with Math results (_F)

- C2CA_PAR_PISA_E_M.DAT

The Sas files :

- Create_C2CA_PAR_PISA_E_M.sas
- Lrecl_C2CA_PAR_PISA_E_M.sas

The SPSS Cards

- Create_C2CA_PAR_PISA_E_M.sps

APPENDIX A – Cycles 1 to 5 – New "Other – Specify" Categories

One or more new categories, which were not present at the time of interview (not on the questionnaire used during data collection), were generated from frequency of responses to "Other - Specify". New categories were added to the codebooks on release and then added to the next cycle questionnaire, only up to cycle 4. No new categories were created for Cycle 5.

			Cycle 2		Cycle 3		Cycle 4	
Modules	**# of Responses**	**Variable**	**Categories**	**Text**	**Cat.**	**Text**	**Cat.**	**Text**
Module B	Mark one	**B_Q47**	11	Other - Moved	11	Other - Moved	11	Other - Moved
			12	Other Specify	12	Other specify	12	Other specify
	Mark one	**B_Q52**	11	Other - Moved	11	Other - Moved	11	Other - Moved
			12	Other Specify	12	Other specify	12	Other specify
Module D	Mark one	**D_Q05**	11	Other - To gain experience/knowledge	11	Other - To gain experience/ knowledge		Module Dropped in Cycle 4
			12	Other - Money		Category not added this cycle		
			13	Other Specify	13	Other Specify		
	Mark one	**D_Q06**	9	Other - academic workload heavy/more important	9	Other - academic workload heavy/more important		Module Dropped in Cycle 4
			10	Other - already participated	10	Other - already participated		
			11	Other - conflict with courses/schedule	11	Other - conflict with courses/ schedule		
			12	Other - no time	12	Other - no time		
			13	Other - transportation		Category not added this cycle		
			14	Other Specify	14	Other Specify		
Module F	Mark all	**F_Q63F**	F_Q63FI - Did not need to work	Other - Did not need to work		Category not added this cycle		Module Dropped in Cycle 4
			F_Q63FJ - Could not get to work/ location	Other - Could not get to work/location		Category not added this cycle		
			F_Q63FK - Foreign student (no VISA, no work permit)	Other - Foreign student (no VISA, no work permit)		Category not added this cycle		
			F_Q63FL - Other	Other	F_Q63FL	Other		
Module H	Mark one	**H_Q420**	4	Another program (pract/inter/clini/work place)	4	Another program (Practicum, internship, clinical)	4	Another program (Practicum, internship, clinical)
					5	Another program with a work placement	5	Another program with a work placement
	Mark one	**H_Q430**	11	Other - To gain experience	11	Other - To gain experience/ knowledge	~~11~~	Category not added this cycle
			12	Other Specify	12	Other Specify	11	Other Specify

Modules	# of Responses	Variable	Cycle 2		Cycle 3		Cycle 4	
			Categories	**Text**	**Cat.**	**Text**	**Cat.**	**Text**
	Mark one	**H_Q441**	09	Other - Offered only later/higher grades	09	Other - Offered only later/higher grades	~~9~~	Category not added this cycle
			10	Other Specify	10	Other Specify	9	Other Specify
Module L	Mark one	**L_Q03A**		Personal Savings				No action needed
				Note: category added to codebook in C2,				
				question added in cycle 3 (no need to add cat.)				
Module M	Mark one	**M_Q02**	10	Other - will apply		Category not added this cycle	~~10~~	Category not added this cycle
			11	Other - future undecided		Category not added this cycle	~~11~~	Category not added this cycle
			12	Other Specify	12	Other Specify	10	Other Specify
Module P2	Mark one	**P2_Q45**	08	Other Specify	08	Other - Multiple job holder	08	Other - Multiple job holder
					09	Other - Characteristics/nature of the job	09	Other - Characteristics/nature of the job
					10	Other Specify	10	Other Specify
	Mark one	**P2_Q77**	10	Other Specify	10	Other - Worked there previously	10	Other - Worked there previously
					11	Other Specify	11	Other Specify
Module P5	Mark one	**P5_Q06**	12	Other - Satisfied/Likes where currently lives	~~12~~	Other - Satisfied/Likes where currently lives	~~12~~	Category not added this cycle
			13	Other - House / lease	~~13~~	Category not added this cycle	~~13~~	Category not added this cycle
			14	Other - Don't want / need to	~~14~~	Category not added this cycle	~~14~~	Category not added this cycle
			15	Other - Age	~~15~~	Category not added this cycle	~~15~~	Category not added this cycle
			16	Other Specify	12	Other - Specify	12	Other Specify
Module P6	Maximum 3	**P6_Q25K**	11	Other - Age	11	Other - Age	11	Category not added this cycle
		P6_Q25L	12	Other Specify	12	Other - Specify	12	Other Specify
Module PS	Mark one	**PS_Q01**		Variable not released (in a DV ReasmmD2)		This question does not exist in Cycle 3.		No action needed
Module U	Mark all	**U_Q01**		no new categories		no new categories	3	Other Canada-unknown if by birth or naturalization (used in DV)
							4	Other

APPENDIX B - Cycle 1 - Module H Variables

The following tables refer to Cycle 1 variables and derived variables which have been renamed in subsequent cycles.

Cycle 1, Module H Variables

Person Level Variables

HGDA	HGDAA	HLPS	HEDAT	HEDL	DLPSM
DLPSY	DLPSFM	DLPSFY	NINDI	NPRDI	LPSAT
MHSPS	MHSPSFLG	AGSPS	FPSP	EDTPSM	EDTPSY
ULPSAT (Cohort B only in Cycle 1) (created at Cycle 4 and written for all cycles)		CLPSAT (Cohort B only in Cycle 1) (created at Cycle 4 and written for all cycles)		OLPSAT (Cohort B only in Cycle 1) (created at Cycle 4 and written for all cycles)	

Institution Level Variables

Questionnaire Variables

Institution 1	Institution 2	Institution 3	Institution 4
H8a	H8b	H8c	H8d
H9a	H9b	H9c	H9d
H10Aa_1	H10Ab_1	H10Ac_1	H10Ad_1
H10Aa_2	H10Ab_2	H10Ac_2	H10Ad_2
H12a	H12b	H12c	H12d

NPRPI1	NPRPI2	NPRPI3	NPRPI4
DSAINM_1	DSAINM_2	DSAINM_3	DSAINM_4
DSAINY_1	DSAINY_2	DSAINY_3	DSAINY_4
DLINM_1	DLINM_2	DLINM_3	DLINM_4
DLINY_1	DLINY_2	DLINY_3	DLINY_4
FPLIN_1	FPLIN_2	FPLIN_3	FPLIN_4
DLFINM_1	DLFINM_2	DLFINM_3	DLFINM_4
DLFINY_1	DLFINY_2	DLFINY_3	DLFINY_4
HLATT_1	HLATT_2	HLATT_3	HLATT_4

Program Level Variables

Questionnaire Variables

Institution 1			Institution 2			Institution 3		Institution 4
program 1	program 2	program 3	program 1	program 2	program 3	program 1	program 2	program 1
H18a1_B	H18a2_B	H18a3_B	H18b1_B	H18b2_B	H18b3_B	H18c1_B	H18c2_B	H18d1_B
H18a1_C	H18a2_C	H18a3_C	H18b1_C	H18b2_C	H18b3_C	H18c1_C	H18c2_C	H18d1_C
H21a1	H21a2	H21a3	H21b1	H21b2	H21b3	H21c1	H21c2	H21d1
H22a1	H22a2	H22a3	H22b1	H22b2	H22b3	H22c1	H22c2	H22d1
H23a1	H23a2	H23a3	H23b1	H23b2	H23b3	H23c1	H23c2	H23d1
H26Aa1	H26Aa2	H26Aa3	H26Ab1	H26Ab2	H26Ab3	H26Ac1	H26Ac2	H26Ad1
H26Ba1	H26Ba2	H26Ba3	H26Bb1	H26Bb2	H26Bb3	H26Bc1	H26Bc2	H26Bd1
H29a1	H29a2	H29a3	H29b1	H29b2	H29b3	H29c1	H29c2	H29d1
H30_1Mth	H30a2_Mt	H30a3_Mt	H30b1_Mt	H30b2_Mt	H30b3_Mt	H30c1_Mt	H30c2_Mt	H30d1_Mt
H30_1Yr	H30a2_Yr	H30a3_Yr	H30b1_Yr	H30b2_Yr	H30b3_Yr	H30c1_Yr	H30c2_Yr	H30d1_Yr
H39a1	H39a2	H39a3	H39b1	H39b2	H39b3	H39c1	H39c2	H39d1
H42a1	H42a2	H42a3	H42b1	H42b2	H42b3	H42c1	H42c2	H42d1
H43Aa1	H43Aa2	H43Aa3	H43Ab1	H43Ab2	H43Ab3	H43Ac1	H43Ac2	H43Ad1
H43Ba1	H43Ba2	H43Ba3	H43Bb1	H43Bb2	H43Bb3	H43Bc1	H43Bc2	H43Bd1
H44a1	H44a2	H44a3	H44b1	H44b2	H44b3	H44c1	H44c2	H44d1
H45a1	H45a2	H45a3	H45b1	H45b2	H45b3	H45c1	H45c2	H45d1
H48Aa1	H48Aa2	H48Aa3	H48Ab1	H48Ab2	H48Ab3	H48Ac1	H48Ac2	H48Ad1
H48Ba1_1	H48Ba2_1	H48Ba3_1	H48Bb1_1	H48Bb2_1	H48Bb3_1	H48Bc1_1	H48Bc2_1	H48Bd1_1
H48Ba1_2	H48Ba2_2	H48Ba3_2	H48Bb1_2	H48Bb2_2	H48Bb3_2	H48Bc1_2	H48Bc2_2	H48Bd1_2
H48Ba1_3	H48Ba2_3	H48Ba3_3	H48Bb1_3	H48Bb2_3	H48Bb3_3	H48Bc1_3	H48Bc2_3	H48Bd1_3
H48Ba1_4	H48Ba2_4	H48Ba3_4	H48Bb1_4	H48Bb2_4	H48Bb3_4	H48Bc1_4	H48Bc2_4	H48Bd1_4
H48Ba1_5	H48Ba2_5	H48Ba3_5	H48Bb1_5	H48Bb2_5	H48Bb3_5	H48Bc1_5	H48Bc2_5	H48Bd1_5
H48Ba1_6	H48Ba2_6	H48Ba3_6	H48Bb1_6	H48Bb2_6	H48Bb3_6	H48Bc1_6	H48Bc2_6	H48Bd1_6
H49Aa1	H49Aa2	H49Aa3	H49Ab1	H49Ab2	H49Ab3	H49Ac1	H49Ac2	H49Ad1
H49Ba1	H49Ba2	H49Ba3	H49Bb1	H49Bb2	H49Bb3	H49Bc1	H49Bc2	H49Bd1

Derived Variables

LVPR_11	LVPR_12	LVPR_13	LVPR_21	LVPR_22	LVPR_23	LVPR_31	LVPR_32	LVPR_41
CLGPR_11	CLGPR_12	CLGPR_13	CLGPR_21	CLGPR_22	CLGPR_23	CLGPR_31	CLGPR_32	CLGPR_41
DLPRM_11	DLPRM_12	DLPRM_13	DLPRM_21	DLPRM_22	DLPRM_23	DLPRM_31	DLPRM_32	DLPRM_41
DLPRY_11	DLPRY_12	DLPRY_13	DLPRY_21	DLPRY_22	DLPRY_23	DLPRY_31	DLPRY_32	DLPRY_41
FPLPR_11	FPLPR_12	FPLPR_13	FPLPR_21	FPLPR_22	FPLPR_23	FPLPR_31	FPLPR_32	FPLPR_41
DLFPRM11	DLFPRM12	DLFPRM13	DLFPRM21	DLFPRM22	DLFPRM23	DLFPRM31	DLFPRM32	DLFPRM41
DLFPRY11	DLFPRY12	DLFPRY13	DLFPRY21	DLFPRY22	DLFPRY23	DLFPRY31	DLFPRY32	DLFPRY41
SIPR_11	SIPR_12	SIPR_13	SIPR_21	SIPR_22	SIPR_23	SIPR_31	SIPR_32	SIPR_41
RSIPR_11	RSIPR_12	RSIPR_13	RSIPR_21	RSIPR_22	RSIPR_23	RSIPR_31	RSIPR_32	RSIPR_41
NMDUR_11	NMDUR_12	NMDUR_13	NMDUR_21	NMDUR_22	NMDUR_23	NMDUR_31	NMDUR_32	NMDUR_41
DSPRM_11	DSPRM_12	DSPRM_13	DSPRM_21	DSPRM_22	DSPRM_23	DSPRM_31	DSPRM_32	DSPRM_41
DSPRY_11	DSPRY_12	DSPRY_13	DSPRY_21	DSPRY_22	DSPRY_23	DSPRY_31	DSPRY_32	DSPRY_41
AGEPS_11	AGEPS_12	AGEPS_13	AGEPS_21	AGEPS_22	AGEPS_23	AGEPS_31	AGEPS_32	AGEPS_41
OPSP_1	OPSP_2	OPSP_3	OPSP_4	OPSP_5	OPSP_6	OPSP_7	OPSP_8	OPSP_9
I1p1MFS1	I1p2MFS1	I1p3MFS1	I2p1MFS1	I2p2MFS1	I2p3MFS1	I3p1MFS1	I3p2MFS1	I4p1MFS1
I1p1MFS2	I1p2MFS2	I1p3MFS2	I2p1MFS2	I2p2MFS2		I3p1MFS2		

LINKS TO REFERENCE DOCUMENTS

The Statistics Canada website is:
http://www.statcan.gc.ca/english/concepts/index.htm

Instructions to access survey documentation:
Access the Survey "*Definitions, Data Sources and Methods*" under "*Survey Information*", either alphabetically or by subject:

Alphabetically
- Youth in Transition Survey (YITS) – Project Codes 4435, 5058 and 5059 for Cycles 1 to 5

Subject
- Education (click on "surveys" and the list is alphabetical)

The project codes for YITS are in brackets.

Documentation Available on the website www.statcan.gc.ca:

Classification codes for:
- Industry (SIC and NAICS)
- Occupation (SIC and NAICS)
- Education Programs (CIP)

Questionnaires:

YITS – 18-20 year-olds (Cycle 1), 17 and 20-22 year-olds (Cycle 2), 19 and 22-24 year-olds (Cycle 3), 21 and 24-26 year-olds (Cycle 4), and 23 and 26-28 year-olds (Cycle 5) (4435)

YITS – 15 year-olds Reading Cohort (5058)
- 15 year-olds Reading Cohort Questionnaire (Canadian Longitudinal Youth in Transition Survey)
- Parent Questionnaire (Canadian Longitudinal Youth in Transition Survey)

YITS – 15 year-olds Mathematics Cohort (5059)
- 15 year-olds Mathematics Cohort Questionnaire (Canadian Longitudinal Youth in Transition Survey)
- Parent Questionnaire (Canadian Longitudinal Youth in Transition Survey)

Codebooks:

YITS (4435) – Cycles 1 to 5

YITS (5058) – Cycle 1
- Student Codebook
- Parent Codebook

YITS (5059) – Cycle 2

YITS Data Extraction Tools:

YITS Data Extraction Excel spreadsheets – 4435, 5058, 5059 (provide all file and roster names and variable names for all cycles)

OTHER DOCUMENTATION AVAILABLE ON REQUEST :

YITS Project Overview (5058 and 4435) – Cycles 1 to 5

The Survey/Project Overview is presented as a mapping document with subject matter themes and also provides the comparison of questions/variables between each cohort of YITS. This document is updated for each cycle of YITS.

Coding Lists for Institutions

Codes for postsecondary institutions for all provinces and territories.

YITS Data Extraction Tool

Administrative Documentation (to be used for installation of the YITS Data Extraction Tool)