LONGITUDINAL SURVEYS OF AUSTRALIAN YOUTH

TECHNICAL REPORT 74

2009 cohort user guide

Longitudinal Surveys of Australian Youth

Longitudinal Surveys of Australian Youth (LSAY)

2009 cohort user guide

National Centre for Vocational Education Research

LONGITUDINAL SURVEYS OF
AUSTRALIAN YOUTH

TECHNICAL PAPER 74

Date created: October 2012
Version: 1.0

Publisher's note

Additional information relating to this publication is available from NCVER's website <http://www.lsay.edu.au/publications/2547.html>.

This document should be attributed as NCVER 2012, *Longitudinal Surveys of Australian Youth (LSAY) 2009 cohort user guide*, NCVER, Adelaide.

This work has been produced by NCVER through the Longitudinal Surveys of Australian Youth (LSAY) Program, on behalf of the Australian Government and state and territory governments, with funding provided through the Australian Department of Education, Employment and Workplace Relations.

The views and opinions expressed in this document are those of author and do not necessarily reflect the views of the Australian Government or state and territory governments.

COVER IMAGE: GETTY IMAGES/THINKSTOCK

TD/TNC 109.28

Published by NCVER
ABN 87 007 967 311

Level 11, 33 King William Street, Adelaide SA 5000
PO Box 8288 Station Arcade, Adelaide SA 5000, Australia

P +61 8 8230 8400 **F** +61 8 8212 3436 **E** lsay@ncver.edu.au **W** <http://www.lsay.edu.au>

User guide updates

Date	Version	Update
October 2012	1.0	Original version of user guide.

Contents

Tables and figures 7

Background 8

Using this guide 9

- Registration 9

The Y09 cohort 10

- Questionnaires and frequency tables 10
- Cohort reports 11
- Other technical papers 12

Accessing the data 13

- Specific data requests 13
- LSAY data releases 13
- Data restrictions 14

Overview of the questionnaires 16

- Programme for International Student Assessment 16
  - National options 17
  - The PISA data 17
- Plausible values 17
- How do I use plausible values? 17
- The LSAY questionnaires 18

The LSAY data 19

- Topic areas 19
- Topic maps 21
- Data elements 22
- Variable listing/metadata workbook 23
- Variable selection 23

Variable naming conventions 25

- PISA variables 25
  - Plausible values and replicate weights 25
  - Simple and scale indices 26
- LSAY standard variables 27
- LSAY non-standard variables 27

Derived variables 29

Sample and survey design 30
Response rates 30
Sources of error 30
Non-sampling error 31
Non-response 31
Weights 31
Sampling error 33
Standard errors 33
Confidence intervals 33
Relative standard errors 33
Examples 34
Classifications and code frames 35
Education 35
Occupation 36
Industry 36
Institution 36
Topic maps 37
Topic map 1: Demographics – Student 39
Topic map 2: Demographics – Parent 40
Topic map 3: Education – School 42
Topic map 4: Education – School transition 49
Topic map 5: Education – Post-school 51
Topic map 6: Employment – Current 58
Topic map 7: Employment – Job history and training 61
Topic map 8: Employment – Seeking employment 63
Topic map 9: Employment – Not in the labour force 65
Topic map 10: Social – Health, living arrangements and finance 66
Topic map 11: Social – General attitudes 68
Appendix A: Updates to the Y09 data file 69

Tables and figures

Tables

1 Technical documents: questionnaires and frequency tables 11

1 PISA technical documents 16

3 Topic maps 22

4 User guide data element documents 22

5 Summary of PISA variable naming conventions 26

6 Summary of LSAY non-standard variable naming conventions 27

7 Derived variables 29

8 Sample sizes and response rates 30

9 Weight variables 32

10 Estimates, standard errors, RSEs and confidence limits for highest school level completed, Y09 cohort in 2010 for a large sample (all respondents) 34

11 Estimates, standard errors, RSEs and confidence limits for highest school level completed, Y09 cohort in 2010 for a small sample (remote respondents) 34

12 Summary of classifications and code frames used in the LSAY Y06 dataset 35

13 Summary of changes made to the Y09 data file 69

Figures

1 Cohort reports 12

2 LSAY hierarchical levels 19

3 Major topic area 1 — Demographics 20

4 Major topic area 2 — Education 20

5 Major topic area 3 — Employment 21

6 Major topic area 4 — Social 21

7 Identifying related topic areas 24

8 PISA variable naming convention 25

9 LSAY standard variable naming convention 27

Background

The Longitudinal Surveys of Australian Youth (LSAY) is a research program that tracks young people as they move from school into further study, work and other destinations. It uses large, nationally representative samples of young people to collect information about education and training, work and social development.

It includes surveys conducted from the mid-1970s through to the mid-1990s: the Youth in Transition (YIT) program; the Australian Longitudinal Survey (ALS); the Australian Youth Survey (AYS); and the current LSAY collection, which began in 1995.

Survey participants in the current LSAY collection (collectively known as a 'cohort') enter the study at age 15 years or, as was the case in earlier studies, when they were in Year 9. Individuals are contacted once a year for up to 12 years, but respondents can miss one survey wave and still remain in the survey. Studies began in 1995 (Y95 cohort), 1998 (Y98 cohort), 2003 (Y03 cohort), 2006 (Y06 cohort) and more recently in 2009 (Y09 cohort). Over 10 000 students start out in each cohort.

Since 2003, the initial survey wave has been integrated with the Organisation for Economic Co-operation and Development (OECD) Programme for International Student Assessment (PISA).

The LSAY research program provides a rich source of information to enable a better understanding of young people and their transitions from school to post-school destinations; it also explores their social outcomes, such as wellbeing. Information collected as part of the LSAY program covers a wide range of school and post-school topics, including: student achievement, student aspirations, school retention, social background, attitudes to school, work experiences and what students do when they leave school.

LSAY is managed and funded by the Australian Government Department of Education, Employment and Workplace Relations, with support from state and territory governments. On 1 July 2007, the National Centre for Vocational Education Research (NCVER) was contracted to provide LSAY analytical and reporting services. NCVER is undertaking this service for the department in collaboration with the Australian National University's Social Policy Evaluation, Analysis and Research Centre.

Between 1995 and 2007 the LSAY analytical and reporting services were provided by the Australian Council for Educational Research (ACER) jointly with the Department of Education, Science and Training.[1]

More information can be obtained from the LSAY website, or by contacting the LSAY team at NCVER:

Toll free: 1800 825 233 Email: <lsay@ncver.edu.au>

Telephone: +61 8 8230 8400 Website: <http://www.lsay.edu.au>

Facsimile: +61 8 8212 3436

[1] Replaced in December 2007 by the Department of Education, Employment and Workplace Relations.

Using this guide

This *User guide* has been developed for users of the LSAY data. The guide endeavours to consolidate existing technical documentation and other relevant information into a single document, thereby improving data accessibility and promoting wider use of the LSAY data.

To promote effective use of the data, the guide aims to address all aspects of LSAY data, including information about: how to access the data, data restrictions, variable naming conventions, the structure of the data (using topic areas, topic maps and data elements), classifications and code frames used, weights and derived variables.

A series of additional documents (*Data elements A* to *D*) complement this *User guide*. Data elements represent variables that are common within and between waves. These documents contain information about the data elements, including the variables they cover, the valid values (or response options) for each variable and additional notes (where applicable). Information about the data elements documentation is contained in the section, 'The LSAY data', sub-section, 'Data elements'.

Users may also find the metadata workbook useful. The workbook provides a listing of all variables in the Y06 dataset, as well as basic information about each variable. Data can be filtered and inspected by wave/year, questionnaire section, topic area(s) and/or data element. See the section, 'The LSAY data', sub-section 'Variable listing/metadata workbook', for further information. The metadata workbook can be accessed at: <www.lsay.edu.au/publications/2547.html> under the 'supporting documents' tab.

If you have any feedback or issues finding the information you need in this guide, please do not hesitate to contact the LSAY branch at NCVER.

Toll free: 1800 825 233 Email: <lsay@ncver.edu.au>

Telephone: +61 8 8230 8400 Website: <http://www.lsay.edu.au>

Facsimile: +61 8 8212 3436

Registration

You need to register for the LSAY website to access LSAY resources and materials. Registration is free and gives you:

- web access to LSAY cohort reports, technical documents and questionnaires
- web access to the full text of LSAY research reports and briefing papers
- email alerts to keep you informed about the latest research and data releases from LSAY.

The following link can be used to register for the LSAY website:
<http://www.lsay.edu.au/subscribe.html>.

Further information about registering for the LSAY website can be found at:
<http://www.lsay.edu.au/newsevents/subscribe.html>.

The Y09 cohort

In 2009, a nationally representative sample of 14 251 students aged 15 years was selected to participate in the Programme for International Student Assessment, conducted by the Organisation for Economic Co-operation and Development. This sample became the fifth cohort of the LSAY program. This is referred to as the LSAY Y09 cohort.

The PISA sample was constructed by randomly selecting students aged 15 years from a sample of schools designed to represent all states and sectors. In Australia, 353 schools and 14 251 students participated in the Programme for International Student Assessment. Assessments in mathematical literacy, reading literacy and scientific literacy were administered in schools to provide information on student achievement. Students also completed a background questionnaire about their families, reading activities, English lessons, libraries, strategies used in reading and understanding texts, educational career, life at school, educational and vocational plans, attitudes to school and learning, work experience, workplace learning, and part-time work.

In 2010, members of the Y09 cohort were contacted for their annual LSAY telephone interview (conducted by the Wallis Consulting Group) and have been contacted annually since then. The questionnaire for their 2010 interview included questions on school, transitions from school, post-school education and training, work, job history, job search history, non-labour force activities, health, living arrangements and finance, and general attitudes. Subsequent surveys asked similar questions, but with the emphasis changing from school to post-school education, training and work, depending on the person's circumstances.

Due to both population shifts over time and survey attrition, care needs to be taken when comparing individual waves of the cohort with other samples drawn from different populations. For example, it can be misleading to compare the LSAY Y09 wave 3 (2011) information with information about 17-year-olds from other surveys in the same year.

Prior to the development of this *User guide*, technical papers (including questionnaires, frequency tables and code books) contained information about the LSAY cohorts. Information from the technical papers has been consolidated in the series of user guides, providing a single source for technical information. These technical documents are discussed below.

Questionnaires and frequency tables

The following six questionnaire instruments were used in PISA 2009:

- student questionnaire
- school questionnaire
- parent questionnaire
- information communication technology questionnaire
- education career questionnaire
- reading for school questionnaire.

Parent, information communication technology and education career questionnaires were offered as national options, with Australia participating in all of these with the exception of the parent questionnaire.

The 2009 PISA questionnaires and code books are available from the OECD PISA website: <http://pisa2009.acer.edu.au/downloads.php>.

The LSAY questionnaires and frequency tables and can be accessed at: <www.lsay.edu.au/data/31281.html>. Table 2 provides a summary of the LSAY Y09 questionnaires and frequency tables.

Table 2 Technical documents: questionnaires and frequency tables

Wave/year	Technical report/paper
Wave 1/2009	Technical report no. 70
Wave 2/2010	Technical report no. 71
Wave 3/2011	Technical report no. 72

Cohort reports

The Y09 cohort reports provide a longitudinal snapshot of the activities of the Y09 cohort from 2009 to the latest survey wave. They are updated annually as new waves of data become available.

The content of the cohort reports focuses on the areas of educational attainment, employment, measures of engagement in study and work, and social outcomes. The cohort reports present a series of tables for each of the indicators. Each series of tables can be filtered by a range of demographic variables and be downloaded into Excel.

The Y09 cohort reports can be accessed at <http://www.lsay.edu.au/cohort/2009/101.html>, and are particularly useful for cross-validation for data users. See figure 1 for an illustration of the cohort reports.

Figure 1 Cohort reports

Longitudinal Surveys of Australian Youth, Y09 cohort to 2011, released August 2012

Table 2: Education Indicators for Y09 LSAY cohort, 2009 - 2011.

Year	2009	2010	2011
Wave	1	2	3
Average age of respondents at 30 June (years)	15.7	16.7	17.7
Number of respondents	14251	8759	7626
Current school level (%)			
Year 12	0.1*	20.0	58.9
Year 11	18.6	61.0	9.1
Year 10	70.8	10.1	0.2*
Year 9 or below	10.5	0.1*	0.0
At school - year level unknown	0.0	0.0	0.0
Not at school	0.0	8.9	31.8
Current qualification level (%)			
Certificate I	0.0	0.4	0.9
Certificate II	0.0	0.6	1.6
Certificate III	0.0	1.5	4.3

Other technical papers

Other technical papers that may be useful include sampling and weighting methodology and the PISA technical reports, data analysis manuals and country reports.

Technical paper number 61, *Weighting the LSAY Programme for International Student Assessment cohorts*, can be accessed at: <http://www.lsay.edu.au/publications/2429.html>.

The PISA 2009 technical report, data analysis manuals and country report provide all the information required to understand the PISA 2009 data (contained in the first wave of the Y09 cohort) and to perform analyses in accordance with the complex methodologies used to collect and process the data. Because the same methods were applied to the PISA 2009 data as for previous cycles, a PISA 2009 data analysis manual was not produced and the PISA 2006 data analysis manuals should be referenced instead.

- The *PISA 2009 technical report* is available from: <http://www.oecd.org/pisa/pisaproducts/pisa2009/50036771.pdf>.
- The *PISA 2006 data analysis manual* (for both SAS and SPSS users) is available from: <http://www.oecd.org/pisa/pisaproducts/pisa2006/pisadataanalysismanualspssandsassecondedition.htm>.
- The PISA 2009 country report (Australia): *Challenges for Australian education: results from PISA 2009* is available from: <http://www.acer.edu.au/documents/PISA-Report-2009.pdf>.

Accessing the data

LSAY data files are deposited annually with the Australian Data Archive (ADA) at the Australian National University in Canberra. Permission to use the data and access requirements are managed by the Australian Data Archive. Data access requires authorisation from the Data Archive Manager.

The data can be accessed by:

- completing the 'Application to access LSAY' restricted data and the LSAY 'User undertaking' forms, available from the Australian Data Archive LSAY information page: <http://www.ada.edu.au/longitudinal/lsay>
- returning the completed forms via email to the Australian Data Archive at <ada@anu.edu.au>.

Part of NCVER's role is to promote and encourage the use of the LSAY data. If you have any feedback or queries about the data and how to access it, please contact:

NCVER

Email: <lsayrequests@ncver.edu.au>

LSAY hotline: 1800 825 233

Australian Data Archive

Email: <ada@anu.edu.au>

Telephone: 02 6125 2200

Fax: 02 6125 0627

Specific data requests

A specific data request allows you to request specific tables and/or data analysis to be undertaken by NCVER without having to obtain full sets of the data. A specific data request can be made to <lsayrequests@ncver.edu.au>.

There are fees and charges applicable for all data requests that require more than one hour to prepare. Please refer to NCVER's policy on charging: <http://www.ncver.edu.au/statistic/21075.html#protocols>.

LSAY data releases

Information about the latest LSAY data releases is available from the LSAY website: <www.lsay.edu.au/data/latest.html>.

You may also request to be notified of recent LSAY releases, which include publications and data releases, by subscribing to NCVER's LSAY alert page at: <http://www.lsay.edu.au/subscribe.html>. For further information, see the section on 'Using this guide: registration'.

Data restrictions

Data use is restricted to research and are not to be used for commercial or financial gain.

Further conditions of use are outlined in the LSAY 'User undertaking' form, which is available from the Australian Data Archive LSAY information page: <http://www.ada.edu.au/longitudinal/lsay>. The conditions of use are as follows:

1 Use of the material is restricted to statistical purposes. This means the user can only use the material to produce information of a statistical nature. Examples of such uses are:

   a the manipulation of data to produce means, correlations or other descriptive summary measures

   b the estimation of population characteristics from sample data

   c the use of data as input to mathematical models and for other types of analyses (for example, factor analysis)

   d the provision of graphical and pictorial representation of characteristics of the population or sub-sets of the population.

2 The material is not to be used for any non-statistical purposes, or for commercial or financial gain without the express written permission of the Australian Data Archive National Manager. Examples of non-statistical purposes include (but are not limited to):

   a transmitting or allowing access to the data in part or whole to any other person, department, or organisation not a party to this undertaking

   b attempting to match unit record data in whole or in part with any other information for the purposes of attempting to identify individuals.

3 Statistical tables, graphs etc. obtained from analysis of these data may be further disseminated, provided that the user:

   a identifies the primary investigators, data series and version number, and data distributors by including the bibliographic reference for the data file

   b acknowledges another archive where the data file is made available through the Australian Data Archive by another archive

   c declares that those who carried out the original analysis and collection of the data bear no responsibility for the further analysis or interpretation of them.

4 Use of the material is solely at the user's risk and the user must indemnify the Australian Data Archive and the ADA consortium members (the Australian National University, the University of Melbourne, the University of Queensland, the University of Technology, Sydney and the University of Western Australia), the National Centre for Vocational Education Research and the Commonwealth of Australia against any liability, loss or expense incurred by the ANU, ADA, NCVER or Commonwealth arising from any action taken against them resulting from unauthorised use or duplication of material, or any other breach of conditions set out in this undertaking.

5 The Australian National University, the Australian Data Archive, the National Centre for Vocational Education Research and the Commonwealth of Australia shall not be held responsible for the accuracy and completeness of the material supplied.

6 Where applicable:

   a the user must draw the terms and conditions of the undertaking to the attention of persons within the department/organisation who shall make use of the material

   b the Australian Data Archive and the ADA consortium members (the Australian National University, the University of Melbourne, the University of Queensland, the University of Technology, Sydney and the University of Western Australia), the National Centre for Vocational Education Research and the Commonwealth of Australia shall not be held liable for any breach of this undertaking.

7 LSAY student achievement information cannot be reported at the school sector aggregate for the LSAY 2003 and 2006 cohorts.

8 Where research findings based on LSAY are published, or otherwise placed in the public arena, the user must agree to provide the Australian Data Archive and the National Centre for Vocational Education Research with the bibliographic details and, where available, online links to any published work (including journal articles, books or book chapters, conference presentations, theses or any other publications or outputs) based wholly or in part on the material.

Overview of the questionnaires

Programme for International Student Assessment

The first wave of the LSAY Y09 cohort was incorporated into the OECD's Programme for International Student Assessment, as was the case with the LSAY Y03 and Y06 cohorts. It is therefore important to understand the PISA 2009 dataset when using the LSAY Y09 cohort data. The following section briefly describes some of the nuances of the PISA dataset, but users are also encouraged to read the PISA technical documents as outlined in table 2.

Table 3 PISA technical documents

Technical report/paper	Web address
PISA 2006 data analysis manual	<http://www.oecd.org/pisa/pisaproducts/pisa2006/pisadataanalysismanualspssandsassecondedition.htm>
PISA 2009 technical report	<http://www.oecd.org/pisa/pisaproducts/pisa2009/50036771.pdf>
PISA 2009 Australian country report: *Challenges for Australian education – results from PISA 2009*	<http://www.acer.edu.au/documents/PISA-Report-2009.pdf>
The role of plausible values in large-scale surveys	<https://mypisa.acer.edu.au/images/mypisadoc/plausiblevaluesinsee.pdf>

As part of PISA 2009, students were assessed in mathematical literacy, reading literacy and scientific literacy to provide information on school achievement. In addition, a short questionnaire, 'Reading for School', was included at the end of the cognitive booklets to collect information about reading curriculum and pedagogy. Students also completed a background questionnaire about their families, reading activities, time spent learning, their school, language (English) lessons, libraries, strategies used in reading and understanding texts, information communication technology, and their educational career.

PISA 2009 covered three domains: reading literacy, mathematical literacy and scientific literacy. For each PISA data collection, one of these domains is chosen as a major domain, while the others are considered minor domains. A major domain is tested more thoroughly in the year of collection. The major domain for PISA 2009 was reading literacy.

The PISA 2009 assessments consisted of a self-completion written test. Examples of items from the PISA 2009 assessment are available in *PISA 2009 assessment framework: key competencies in reading, mathematics and science* available at: http://www.oecd.org/dataoecd/11/40/44455820.pdf.

The *PISA 2009 assessment framework* presents the guiding principles of the PISA 2009 assessment, which are described in terms of the skills students need to acquire, the processes that need to be performed and the contexts in which knowledge and skills are applied. It also illustrates the assessment domains with a range of sample tasks.

National options

Countries participating in PISA are able to introduce country-specific questions into PISA questionnaires, referred to as 'national options' questions. PISA 2009 national options data items administered in Australia include: time spent learning, out-of-school activities, life at school, post-school study plans, views on science, work, work experience, courses at school (for example, the International Baccalaureate and vocational education and training). For this reason, in addition to the publicly available PISA international dataset, a separate national dataset is created for Australia that includes these national options questions. Some variables available from the international dataset are omitted from the national dataset (for example, country). In addition, some minor differences may exist between the two versions of the dataset, for example, the way missing or not applicable values have been assigned to observations, or whether the variables are in numeric or character format.

The PISA data

The PISA international student and school datasets are available from the OECD PISA database: <http://pisa2009.acer.edu.au/downloads.php>. LSAY data can be matched to the PISA international datasets by filtering for Australian records using the country identifiers (CNT, COUNTRY), and using student and school identifiers (STIDSTD and SCHOOLID).

It is recommended that data users wishing to make international comparisons using PISA data download the international dataset available from the OECD 2009 PISA international database located at: <http://pisa2009.acer.edu.au/downloads.php>.

Data users are encouraged to read the documents outlined in table 2 to better understand the PISA variables and data.

Plausible values

In PISA, student assessment is undertaken using 13 different test booklets, and students are randomly assigned one of the booklets. In order to counteract any biases resulting from the use of different text booklets, the OECD calculates plausible values. Plausible values allow for the fact that there is measurement error at the individual level (through differing questionnaires), and the determination of these plausible values takes this error into account.

For each student, five plausible values have been calculated for each of the three domains (reading, mathematics and science), and for each of the five reading sub-domains (access and retrieve, integrate and interpret, reflect and evaluate, continuous text, non-continuous text).

Data users are encouraged to read the documents outlined in table 2 to better understand the construction and use of plausible values in LSAY.

How do I use plausible values?

Unbiased estimates of achievement will only be obtained if plausible values are incorporated appropriately. The following are some key points:

- Averaging plausible values over individuals will lead to biased estimates and incorrect standard errors.

- Analysis should be repeated for each plausible value (five times), and any subsequent estimate (for example, coefficients and/or standard errors) combined in an appropriate way to obtain population estimates.
- Plausible values are correlated within a domain and, as such, an analysis may be undertaken using only a single plausible value, but being aware that standard errors may be incorrect.

Users are reminded that plausible values are not equivalent to the achievement scores in the LSAY Y95 and Y98 cohorts, nor are they equivalent to an individual's raw test scores.

The LSAY questionnaires

From 2010 (wave 2), students have been contacted annually by telephone and asked a range of questions across the following sections:

- Section A: School
- Section B: Transition from school
- Section C: Post-school study
- Section D: Work
- Section E: Job history
- Section F: Job search activity
- Section G: Not in the labour force
- Section H: Living arrangements, finance and health
- Section J: General attitudes.

The Y09 questionnaires can be accessed at: <www.lsay.edu.au/data/31281.html>. Table 1 provides a summary of the technical papers available.

The LSAY data

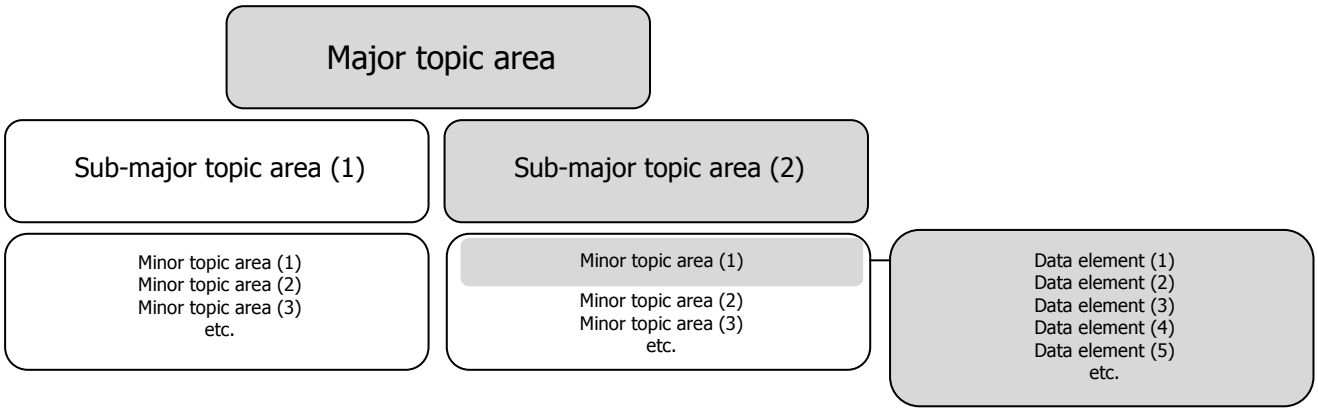
The LSAY data files are large and particularly complex. About 700 variables are collected (on average) across each wave, culminating in close to 7000 variables across the entire data file. To improve accessibility of the LSAY data, variables have been grouped into common themes called 'topic areas'.

Topic areas

The topic areas comprise four hierarchical levels:

- *Major topic areas* are the broadest topic area. There are four major topic areas.
- *Sub-major topic areas* are subdivisions of the major topic areas. There are 11 sub-major topic areas.
- *Minor topic areas* are subdivisions of the sub-major topic areas. There are close to 80 sub-major topic areas.
- *Data elements* are subdivisions of the minor topic areas. There are about 800 data elements.

Figure 2 LSAY hierarchical levels



The four major topic areas are *Demographics*, *Education*, *Employment* and *Social*. The divisions of these major topic areas into sub-major topic areas and minor topic areas are illustrated in figures 3 to 6.

Figure 3 Major topic area 1 – Demographics

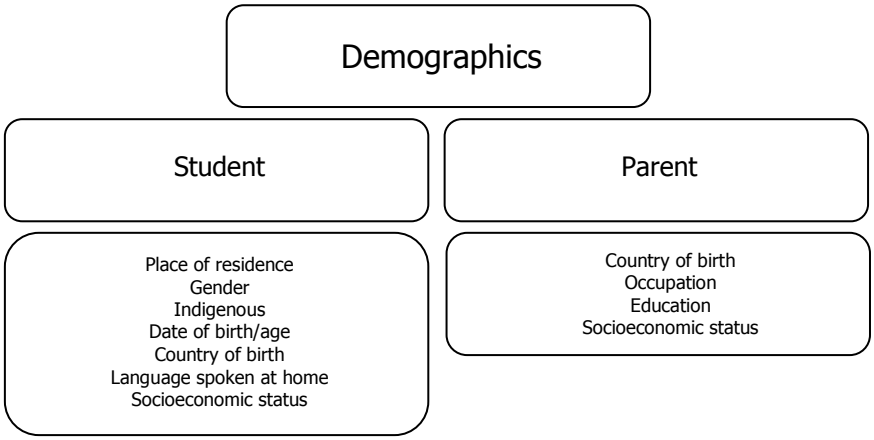


Figure 4 Major topic area 2 – Education

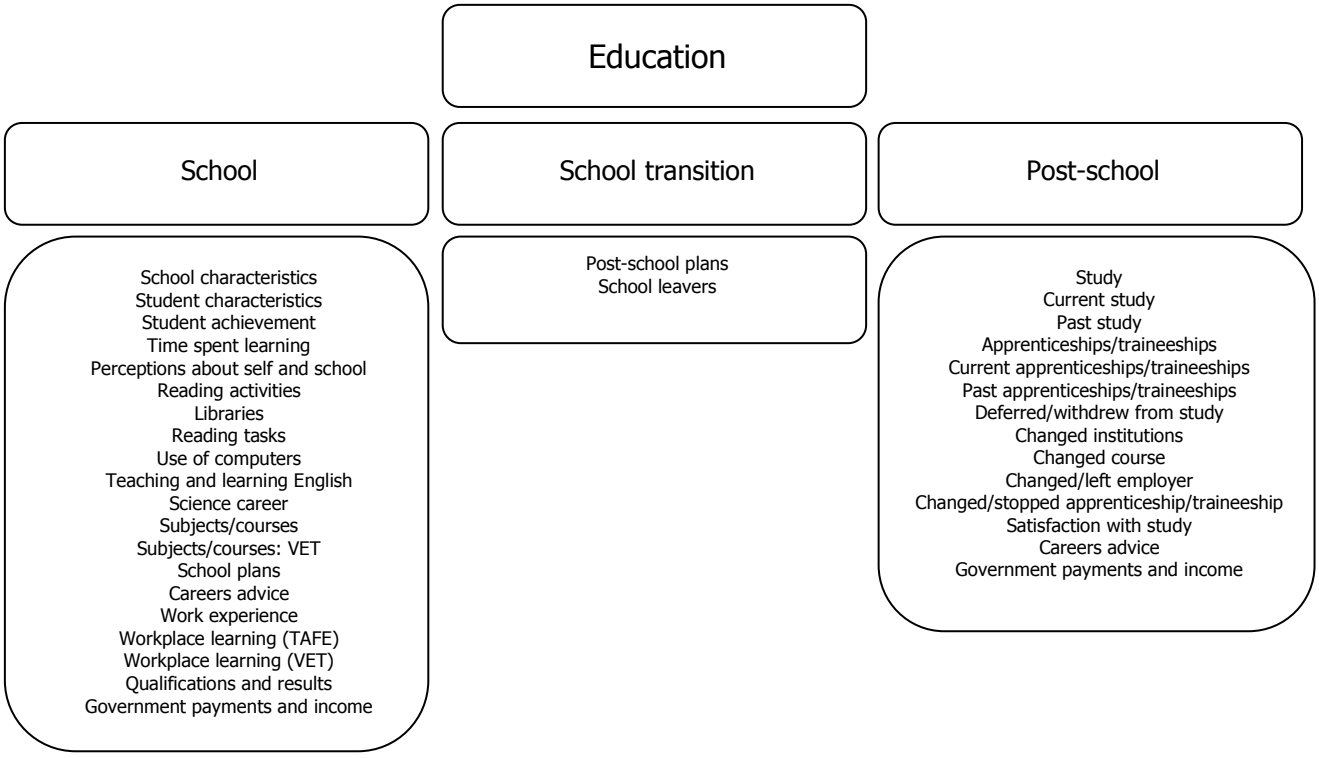


Figure 5 Major topic area 3 – Employment

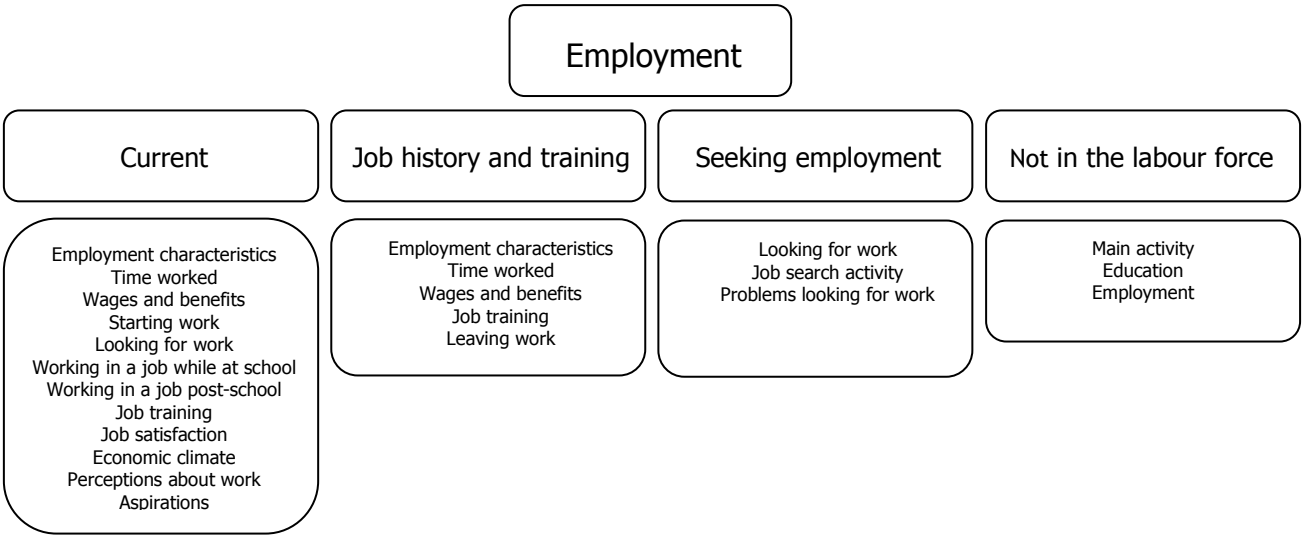
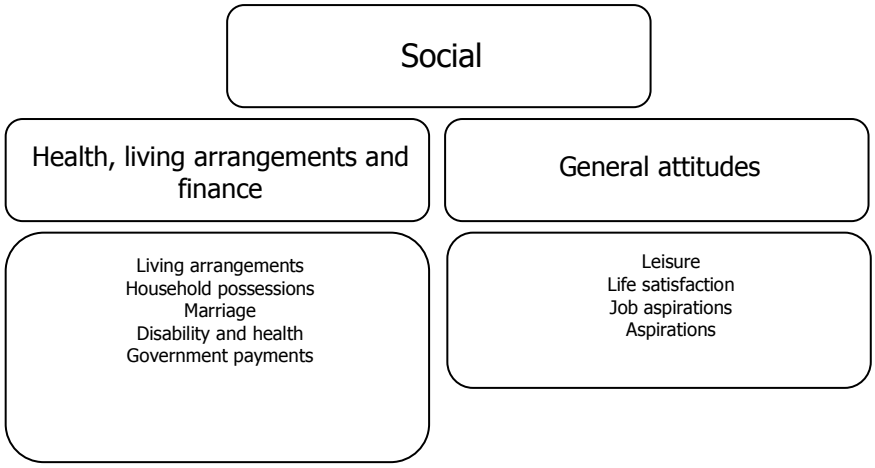


Figure 6 Major topic area 4 – Social



Topic maps

Topic maps have been developed for each of the 11 sub-major topic areas. The topic maps aim to improve accessibility of the LSAY data by linking common questions (or variables) within and between waves. These common variables are identified as *data elements*.

Topic maps by sub-major topic area can be found in the 'Topic maps' section of this *User guide*. A summary of the topic maps appears in table 3.

Table 4 Topic maps

Major topic area	Topic map	Sub-major topic area
Demographics	1	Student
	2	Parent
Education	3	School
	4	School transition
	5	Post-school
Employment	6	Current
	7	Job history and training
	8	Seeking employment
	9	Not in the labour force
Social	10	Health, living arrangements and finance
	11	General attitudes

Data elements

Data elements represent variables that are common within and between waves. In some instances, a data element may represent a single variable (when not collected across multiple waves). Information about each data element is contained in the supplementary sections (*Data elements A* to *D*) of this *User guide*. They can be accessed at: <www.lsay.edu.au/publications/2547.html> under the 'supporting documents' tab.

This series of data element documents are identified by their major and sub-major topic area. An overview of these data element documents is given in table 4.

Table 5 User guide data element documents

User guide	Major topic area	Sub-major topic area(s)
Data element A	Demographics	Student
		Parent
Data element B1	Education	School
		School transition
Data element B2	Education	Post-school
Data element C	Employment	Current
		Job history and training
		Seeking employment
		Not in the labour force
Data element D	Social	Health, living arrangements and finance
		General attitudes

For each data element, the following information is provided (where applicable):

- *Data element* — the data element name
- *Purpose* — the information provided by the data element
- *Variables* — the variable name(s) which correspond to this data element
- *Variable type* — whether the variable(s) is/are in numeric or character format
- *Variable label* — includes the question number (where applicable) and a short description of the variable(s)
- *Question* — the question wording for the variable(s)
- *Values* — the possible values the variable(s) can take and corresponding formats
- *Base population* — a description of and the syntax for the number of respondents required to answer the question
- *Notes* — other information.

Variable listing/metadata workbook

To further assist in the use of the LSAY data, an Excel metadata workbook has been developed. It provides a complete listing of all the variables in the Y09 dataset, as well as information about each variable. The information contained in this workbook is similar to that contained in the topic maps and data elements documents, but can be manipulated using filters to search for and to group variables. Data can be filtered and inspected by wave/year, questionnaire section, topic area(s) and/or data element.

The metadata workbook can be accessed at: <www.lsay.edu.au/publications/2547.html> under the 'supporting documents' tab.

There are three worksheets included in the metadata workbook: *Variables, Values* and *Purpose*. The first worksheet, *Variables*, includes the variable type, variable label, question (wording) and base population. The second worksheet, *Values*, lists each variable and the values that variable can take (where applicable). The third worksheet, *Purpose*, lists the data elements in the data file, and provides a purpose statement for each, along with some additional notes.

The *Variables* and *Values* worksheets list each variable in the order it appears in the data file, and the *Purpose* worksheet lists the data elements in the order they appear in the data elements documents. Major, sub-major and minor topic areas as well as data elements are provided for each variable. The wave/year, questionnaire section and variable label are also included (where applicable).

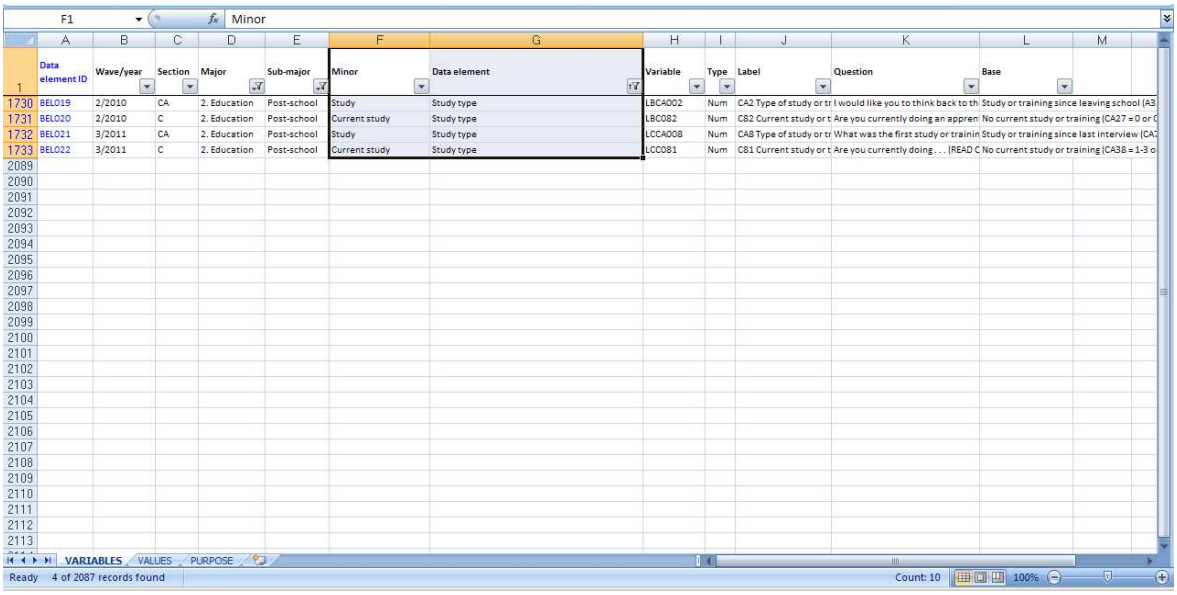
Variable selection

Not all variables assigned to a data element are directly comparable. Additional attributes such as question wording, values, classifications used and base populations must be considered when selecting variables and analysing the data.

Data elements have been created to assist in grouping, thereby simplifying variable selection. They are unique within a minor topic area but may not be unique across broader topic areas.

For example, the data element, *Study type*, exists under the major and sub-major topic area *Education: Post-school*. This data element appears under two different minor topic areas: *Study* and *Current study*. The *Study* minor topic area may include both past and current study (depending on the questionnaire sequencing). When identifying a data element and/or variable for use, it is important to consider other related data elements that may be located in a different topic area. This is illustrated in figure 7 using an excerpt from the metadata workbook.

Figure 7 Identifying related topic areas



To identify variables for analysis and to promote accurate variable selection, refer to the topic maps contained in the 'Topic maps' section. Relevant data elements can be identified by:

- navigating to a major topic area of interest (for example, *Education*)
- identifying a sub-major topic area of interest (for example, *Post-school [education]*)
- identifying a minor topic area of interest (for example, *Current study*)
- inspecting the data elements available within that minor topic area (for example, *Month started study*).

The number of times that data element appears within a wave is shown in the column corresponding to the particular wave.

Before using and/or analysing the variables/data elements selected, it is important to consider:

- variable attributes such as question wording, variable values, classifications used and base populations
- data elements which appear more than once in a wave
- data elements which appear more than once across waves (for longitudinal analysis)
- data elements of the same name across other topic areas (if applicable)
- other data elements that may be closely linked in a topic area or across other topic areas.

Variable naming conventions

PISA variables

PISA variables only exist in wave 1 of the Y09 cohort and have a separate variable naming convention. Naming conventions for different types of PISA variables are summarised in tTable 6.

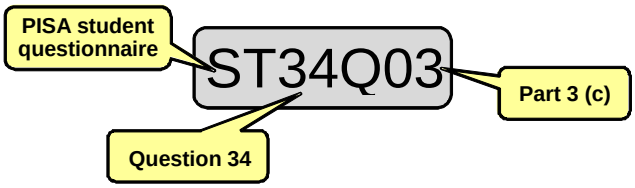
The student questionnaire instruments for PISA 2009 are comprised of the following components:

- the student questionnaire (ST)
- the information communication technology questionnaire (IC)
- the education career questionnaire (EC)
- the reading for school questionnaire (RFS).

Most PISA variables are named using the following convention: questionnaire component, question number, and question part (where applicable). For example, the variable:

- ST16Q01 is question number 16 from the student questionnaire
- ST34Q03 is question number 34 (part c) from the student questionnaire
- IC01Q01 is question number 1 from the information communication technology questionnaire.

Figure 8 PISA variable naming convention



Countries are also able to introduce country-specific questions in the PISA questionnaires, referred to as 'national options' questions. These are denoted by the character 'N' (for example, ST60N01), rather than the character 'Q'.

Plausible values and replicate weights

Plausible values are used to report student achievement in PISA. There are five plausible values for each of the domains and sub-domains[2] and the PISA student achievement variables take this information into account in the variable name. For example, the variable:

- PV1MATH points to the first plausible value in the maths domain
- PV4SCIE points to the fourth plausible value in the science domain
- PV3READ1 points to the third plausible value in the first reading sub-domain: access and retrieve

[2] The Australian PISA 2009 major assessment domains are reading, mathematics and science. The PISA 2009 reading sub-domains are access and retrieve, integrate and interpret, reflect and evaluate, continuous text, and non-continuous text.

- PV4READ5 points to the fourth plausible value in the fifth reading sub-domain: non-continuous text.

Replicate weights have been used to estimate sampling variances for population estimates derived from a complex sample design. The weights are simply named chronologically from W_FSTR1 to W_FSTR80. The variable W_FSTUWT is the final student weight.

Detailed information about plausible values and replicate weights is available from the OECD PISA 2006 data analysis manuals located at: <http://www.oecd.org/pisa/pisaproducts/pisa2006/pisadataanalysismanualspssandsassecondedition.htm>.

Simple and scale indices

Two types of indices are provided in the PISA dataset: simple indices and scale indices. Simple indices are constructed by arithmetically transforming or recoding one or more items, for example, age. Scale indices combine several answers provided by students or principals to build a broader, not directly observable, concept. For example, CULTPOSS is a student-level scale index derived from cultural possessions such as classic literature, books of poetry and works of art.

Simple and scale indices appear towards the end of the wave 1 data and tend to be descriptive rather than carrying a variable naming convention.

Table 6 Summary of PISA variable naming conventions

PISA variable	Examples of PISA variable names	Description
Standard variables	ST16Q01 IC05Q01 ST34Q03	The first two characters indicate the questionnaire instrument. The PISA questionnaire instruments are the student questionnaire (ST), and the information communication technology questionnaire (IC). The following two digits indicate the question number (e.g. ST16 is question 16 from the student questionnaire). The final three characters are the question part or sub-section. So ST34Q03 is part 3 of question 34 from the student questionnaire.
National options	ST60N01	The fifth character 'N' (rather than 'Q') indicates that the question is a national options question.
Student achievement/ plausible values	PV1SCIE PV4READ PV4READ5	The first two characters 'PV' indicate the variable is a plausible value. The next character indicates whether it is the first plausible value up to the fifth plausible value. The next four characters indicate the domain or sub-domain. • PV1SCIE indicates that the variable is the first plausible value from the science domain • PV4READ indicates that the variable is the fourth plausible value from the reading domain • PV4READ5 points to the fourth plausible value in the fifth reading sub-domain. *For further information on plausible values, see section,* 'Overview of the questionnaires: Plausible values'.
PISA weights	W_FSTR1 W_FSTR80 W_FSTUWT CNTFAC	Replicate weights are identified using the characters 'W_FSTR' followed by a chronological number. W_FSTUWT is the final student weight. CNTFAC are country weight factors for equal weights. *For further information on PISA weights, see the* PISA 2006 data analysis manual.
Indices	AGE HISCED CULTPOSS	Student and school-level simple and scaled indices tend to be descriptive rather than adopting a naming convention.

LSAY standard variables

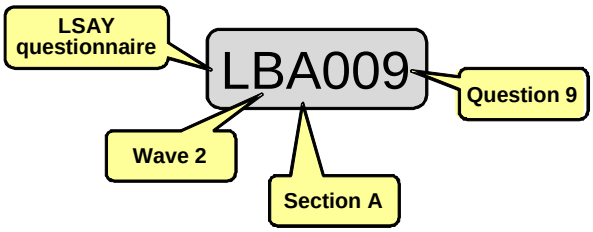
Most variable names are constructed using four pieces of information: the questionnaire instrument, the survey wave, the questionnaire section and the question number.

The character 'L' is used to identify the survey instrument, where L represents the LSAY survey instrument (as opposed to the PISA survey instrument). A wave identifier is used to identify the survey wave from wave 2 (when the LSAY survey instrument is first used). The second survey wave is allocated a B, the third survey wave is allocated a C, etc. The section identifier is used to identify the section of the questionnaire. The question identifier is used to identify the question number.

For example, the variable LBA009 refers to:

- the LSAY survey instrument, denoted by the first character 'L'
- wave 2, denoted by the second character 'B'
- section A, denoted by the third character 'A
- question 9, denoted by the last three characters '009'.

Figure 9 LSAY standard variable naming convention



LSAY non-standard variables

There are a series of other variables that do not take the standard variable naming convention mentioned above. These variables are summarised in the following table.

Table 7 Summary of LSAY non-standard variable naming conventions

Non-standard variable	Examples of non-standard variable names	Description
Demographics	INDIG	Some demographic variables, such as Indigenous status, tend to be descriptive rather than carrying a naming convention.
School characteristics	STATE SECTOR	School characteristics, such as state of the school and school sector, tend to be descriptive rather than carrying a naming convention.
Derived variables	XLFS2009 XCEL2010	Derived variables have been constructed across all waves to summarise key information such as labour force status and current education level. *For further information about derived variables see the section,* 'Derived variables'.
IN flag	IN2009 IN2011	IN flags have been created for each survey year to indicate whether a respondent participated in the survey in that year. If the value of the IN flag is equal to 1, this indicates that the respondent participated in the survey for that year. IN flag variables are denoted by the two characters 'IN' followed by four digits for the survey year.

Non-standard variable	Examples of non-standard variable names	Description
Interview dates	LBWID LBWIM LBWIY INTDAT09 INTSAS09	Day of interview, month of interview, and year of interview are collected each survey year and consolidated into an interview date variable. Interview date variables use the same variable naming convention for the first two characters, followed by the two characters 'WI', and then 'D' for day of interview, 'M' for month of interview, or 'Y' for year of interview. The INTDAT and INTSAS variables are the consolidated interview date variables (in both character and SAS® date format respectively), followed by two digits for the survey year.
Postcode	PC2008 PC2009	Respondents' home postcodes are indicated by the first two characters 'PC' followed by the year of interview.
Sample and derived items	LBWSAM01 LCWSAM08 LDWDV01	Sample and derived items look at information from surveys of previous years. They have been created to enable more efficient and effective direction of questions. For example, the variable LCWSAM08 looks at whether the respondent had a job at the previous interview. Questions about whether respondents have the same job as reported at their last interview would only be asked of those who were recorded as being employed at the previous interview. Sample items are denoted by: ▪ the first character 'L' (to indicate the LSAY survey instrument was used) ▪ followed by the wave identifier (A to F) ▪ followed by the character 'W' ▪ followed by the characters 'SAM', or 'DV' for items derived by the field contractor ▪ followed by two digits denoting the sample/derived item.
Weights	WT10GEN ACH10WT WT2010 WT10GEN_P ACH10WT_P WT2010_P	Weight variables are denoted by the two characters 'WT', either at the beginning or end of the variable name. Two sets of weight variables are produced: the first reproduces the sample sizes in each wave, and the second (denoted by '_P') reproduces the population size at each wave. *For further information about weights see section,* 'Weights' *in the chapter* 'Sample and survey design'.

Derived variables

A series of derived variables has been developed to simplify use of the LSAY data and provide useful measures/indicators for analysis. The derived variables focus on the areas of educational attainment, employment, measures of engagement in study and work, and social indicators. Table 7 summarises the series of derived variables available on the Y09 dataset.

Derived variables are denoted by the character X, followed by three characters uniquely identifying the derived variable, followed by four digits for the survey year.

Detailed technical documentation outlining how the variables are derived as well as their properties is available. This can be accessed at: <http://www.lsay.edu.au/publications/2551.html>.

Table 8 Derived variables

Indicators	Derived variable	Variable name
Education	Current school level	XCSL*YYYY*
	Current qualification level	XCEL*YYYY*
	Highest school level completed	XHSL*YYYY*
	Highest qualification level completed	XHEL*YYYY*
	Study status in VET	XVET*YYYY*
	Study status in bachelor degree or higher	XBAC*YYYY*
	Full-time or part-time study status	XFTS*YYYY*
	Completed Year 12 or certificate II or higher	X122*YYYY*
	Completed Year 12 or certificate III or higher	X123*YYYY*
Employment	Labour force status	XLFS*YYYY*
	Full-time or part-time employment status	XFTP*YYYY*
	Permanent or casual employment	XEMP*YYYY*
	Status in apprenticeship/traineeship	XATR*YYYY*
	Job mobility during last year	XMOB*YYYY*
	Occupation (1 digit ANZSCO first edition)	XOCC*YYYY*
	Average weekly pay	XWKPY*YYYY*
	Average hourly pay	XHRP*YYYY*
	Average weekly hours worked	XHRS*YYYY*
	Any spell of unemployment during the year	XUNE*YYYY*
	In full-time employment or full-time education	XFTE*YYYY*
Social	Marital status	XMAR*YYYY*
	Living with parent(s)	XATH*YYYY*
	Living in own home	XOWN*YYYY*
	Number of dependent children	XCHI*YYYY*

Sample and survey design

In 2009, a nationally representative sample of 15-year-old students was selected to participate in the Programme for International Student Assessment conducted by the OECD; 14 251 students were selected. The initial LSAY survey wave (wave 1) for 2009 was integrated with PISA, and this group of young people became the fifth LSAY cohort.

The 2009 PISA sample comprised 353 schools from all states and territories. This sample was designed to be representative of students across Australia, using state/territory, school sector, geographic location, gender and socioeconomic background as strata. Within each school, 48 non-Indigenous students were selected at random, and all age-eligible Indigenous students were sampled. In schools with fewer than 48 students, all 15-year-olds were selected. Smaller jurisdictions and Indigenous students were oversampled to ensure that reliable results could be produced by state and Indigenous status.

These students were contacted in 2010 to undertake follow-up telephone interviews as part of the LSAY program. This interview collected further information on the respondent's school experience, school and post-school intentions, school leavers and their transitions from school, post-school study, employment, living arrangements, finance, health and general attitudes. Since 2010, respondents have been contacted annually using computer-assisted telephone interviews.

Further information about the survey design for PISA 2009 can be found from the:

- *PISA 2009 technical report* which can be accessed at: <http://www.oecd.org/pisa/pisaproducts/pisa2009/50036771.pdf>
- PISA 2009 Australian country report: *Challenges for Australian education: results from PISA 2009* which can be accessed at: <http://www.acer.edu.au/documents/PISA-Report-2009.pdf>.

Response rates

Table 9 shows the sample sizes and response rates for each wave of the LSAY Y09 cohort from 2009.

Table 9 Sample sizes and response rates

	Wave/year		
	1/2009	2/2010	3/2011
Age at 30 June (years)	15.7	16.7	17.7
Sample size (n)	14 251	8 759	7 626
% of wave 1	100	61.5	53.5
% of previous wave	na	61.5	87.1

Sources of error

Estimates based on sample surveys have two major sources of error: non-sampling and sampling error. A brief description of the two types and an outline of what can be done to overcome the effects of these errors are given below.

Non-sampling error

Non-sampling error arises from inaccuracies in collecting, recording and processing the data. Some common examples of non-sampling error include: non-response, incorrect responses, missing responses, and interviewer and processing error. Non-sampling error can be accounted for, in part, by using weighted estimates to adjust for non-response. However, there are no statistical measures to accurately adjust for other types of non-sampling error. Nevertheless, other types of non-sampling error can be minimised through good questionnaire design, training and monitoring of interviewers, the use of computer-assisted interviews and effective data-checking and processing procedures.

Non-response

All surveys suffer from error related to non-response. Non-response is a form of non-sampling error that can be taken into account in the analysis of survey data. There are typically two forms of survey non-response:

- *Item non-response* occurs when a respondent does not answer all the questions in the survey.
- *Unit non-response* occurs when not all respondents answer the survey due to, for example, refusal to participate, or inaccurate contact details.

Item non-response can be minimised with the use of computer-assisted telephone interviews, which can forward-feed information from previous interviews. Item non-response is generally treated using imputations. There are currently no imputed data for missing values in LSAY. However, data users can apply a number of techniques to help make the data more complete. The use of statistical modelling techniques, such as multiple imputation, allows data users to estimate item non-response, along with their respective standard errors.

Unit non-response (also called attrition) can lead to biased population estimates and incorrect standard errors, particularly if certain groups of the sample drop out at differing rates. Survey attrition is counteracted by attempting to maximise the year-on-year response rate, appropriate statistical modelling techniques, and/or the application of appropriate survey weights.

Weights

In order for the LSAY sample to more accurately represent the population of Australian 15-year-olds in 2009, the collected sample must be weighted to account for differences in the sampling distributions from the original population distribution that may have arisen during the sampling process.

In 2010, NCVER reviewed the weighting methodology used for the LSAY Y03 cohort. As a result of this review, a logistic regression approach to weighting has been adopted. This methodology is consistent with the approach taken to calculate the Y06 and Y09 weights. These weights are provided in the data files deposited with the Australian Data Archive.

Further detailed information regarding the current weighting methodology used is available from technical paper number 61, *Weighting the LSAY PISA cohorts* which can be accessed at: <http://www.lsay.edu.au/publications/2429.html>.

There are two weighting procedures applied to the LSAY data:

1. *Sample weights* reflect the original sample design and ensure that the sample matches the population distribution from which the original sample was drawn. In the Y06 cohort, two sampling weights have been created. The first weights sum to the sample size for that given wave. For example, the sample weights add to 14 251 in wave 1, 8759 in wave 2, etc. In the second set of weights, the sum of the weights equals the original population from which the sample was drawn (240 851). Students from states and territories with smaller numbers of 15-year-olds are over-sampled and students from jurisdictions with larger numbers of 15-year-olds are under-sampled. In order for the sample to more accurately represent the population of Australian 15-year-olds, the sample is weighted so that sample sizes within strata are proportional to the original population sizes of the states and territories (that is, strata).
2. *Attrition weights* are used to address unit non-response by ensuring that the distribution of the sample matches the distribution of the sample population. Attrition weights used in LSAY account for wave-on-wave attrition from the first wave.

In calculating attrition weights, a non-response analysis was undertaken to determine the factors that contributed to attrition. The use of attrition weights ensures that distributions in each wave match those obtained in PISA (for the factors identified as contributing to attrition). Logistic regressions have been used to calculate attrition weights. The response variable of whether or not a respondent replied to the survey in a given year was regressed against a series of factors that may contribute to non-response. The inverse of the predicted probability of responding then forms the attrition weights.

The final LSAY weights for each wave combine both the sample and attrition weights. Two sets of final weights are produced. The first reproduces the sample sizes in each wave, and the second reproduces the population size (240 851) at each wave. In both cases, the distributions in each wave match those obtained in the original population.

Users must be aware that bias resulting from survey attrition may not be fully accounted for in the weighting strategies used. To allow users to determine the effectiveness of the attrition weights, data in the cohort report demographic tables are presented both weighted and unweighted. The Y09 cohort reports can be accessed at: <http://www.lsay.edu.au/cohort/2009/101.html>. Researchers are encouraged to determine their own weighting or analysis methodology to counteract attrition; this may include using methods of multiple imputations for missing values.

Table 10 shows the three different types of available weights and the variable naming convention for each, where *YY* or *YYYY* denotes the survey year at two or four digits respectively. Weights that sum to the population size are denoted by '_P' at the end of the weight variable.

Table 10 Weight variables

Weight	Variables	Sum
Sample weight	WTYYGEN	Sample size in YY
Sample weight (N)	WTYYGEN_P	Population size (240 851)
Attrition weight	ACHYYWT	Sample size in YY
Attrition weight (N)	ACHYYWT_P	Population size (240 851)
Final weight	WTYYYY	Sample size in YYYY
Final weight (N)	WTYYYY_P	Population size (240 851)

Sampling error

Users of the LSAY data must consider the size of the sampling error when deriving or interpreting estimates obtained from LSAY. Sampling error arises because estimates are obtained from the use of a sample rather than from measuring the entire population. It is possible to select many different individual samples from a single population; each of these would provide a different population estimate. An estimate obtained from a sample is subject to sample-to-sample variation (sampling error). In random (probability) sampling, the size of the sampling error (for a given sample) is measured using the standard error of the estimate.

It is important that users take into consideration the reliability of estimates obtained from survey data. Standard errors, confidence intervals and relative standard errors (RSEs) can be calculated to determine the reliability of the estimate(s).

The greatest contributor to standard error is the sample size. Small sample sizes generally result in higher standard errors and wider confidence intervals. The relative standard error enables a comparison of the accuracy between two different estimates. An estimate with a high relative standard error or wide confidence interval should be used with caution, and users are advised against relying on estimates obtained from sample sizes of fewer than five, or estimates that have a relative standard error of greater than 25%.

Standard errors

The standard error of an estimate indicates the accuracy to which that estimate approximates the true population parameter. There are multiple methods for calculating the standard errors in complex surveys. One method commonly used is the Taylor series expansion.[3] This technique has been applied to obtain estimates of standard errors for the LSAY cohort reports. These standard errors can then be used to calculate confidence intervals and relative standard errors.

Confidence intervals

The confidence interval is an interval estimate of the population parameter. Sample estimates which have high standard errors will have wide confidence intervals.

The mathematical derivation of a 95% confidence interval for a proportion is:

$$\hat{p} \pm 2 \times se(\hat{p})$$

where $\hat{p}$ is the estimate obtained from the sample, and $se(\hat{p})$ is the standard error of the estimate (typically obtained from a statistical analysis package).

Relative standard errors

The relative standard error is a standardised measure that enables the comparison between different estimates in terms of their reliability. The relative standard error is derived by dividing the standard error of the estimate by the estimate itself, expressed as a percentage:

$$RSE(\hat{p}) = \frac{se(\hat{p})}{\hat{p}} \times 100$$

[3] For further information on this technique, users should consult William Cochran, *Sampling techniques*, 3rd edn, John Wiley and Sons, New York, 1977, sections 11.18, 11.19, 11.20.

Examples

Consider the following estimates of highest school level completed (XHSL2010) to 2010 taken from the Y09 cohort reports. In this example, estimates obtained from a large sample are compared with estimates obtained from a small sample. Table 10 presents the highest school level for all respondents (large sample), while table 11 presents the highest school level obtained for those from remote areas (small sample).

Table 11 Estimates, standard errors, RSEs and confidence limits for highest school level completed, Y09 cohort in 2010 for a large sample (all respondents)

Year level	Frequency	%	Standard error of %	RSE (%)	95% confidence interval	
					Lower limit	Upper limit
Year 12	63	0.75	0.10	13.84	0.55	0.95
Year 11	1817	21.45	0.51	2.37	20.45	22.44
Year 10	5990	67.08	0.61	0.91	65.88	68.28
Year 9 or below	889	10.72	0.44	4.14	9.85	11.59
Total	**7626**	**100**				

Table 12 Estimates, standard errors, RSEs and confidence limits for highest school level completed, Y09 cohort in 2010 for a small sample (remote respondents)

Year level	Frequency	%	Standard error of %	RSE (%)	95% confidence interval	
					Lower limit	Upper limit
Year 12	4	2.46**	1.40	56.83	0.00	5.21
Year 11	47	31.86	5.01	15.73	21.99	41.73
Year 10	182	62.34	5.08	8.16	52.33	72.36
Year 9 or below	10	3.34*	1.44	43.08	0.51	6.17
Total	243	100.00				

Notes: * Estimate has a relative standard error greater than 25%.
** Estimate has a sample size of fewer than five.

Using this example, we see the estimate for all respondents who finished Year 11 is 21.45%, with a relative standard error of 2.37%. The estimate for remote respondents who finished Year 11 is 31.86%, with a relative standard error of 15.73%. Both estimates have a relative standard error of less than 25%, so are considered reliable; however, the estimate for remote respondents is much less reliable than the estimate for all respondents, given that the relative standard error for remote respondents (15.73%) is considerably higher than that of all respondents (2.37%).

In addition, we would not recommend using estimates obtained from respondents from remote areas who have completed Year 12 or Year 9 or below, as the relative standard errors are higher than 25%.

The interpretation of the confidence intervals for all respondents (table 10) is: we are 95% confident that the true population estimate of Year 11 completion lies between 20.45 and 22.44%.

Classifications and code frames

There are a number of variables contained in the LSAY datasets that are coded using standard classifications. The information for these variables is collected using open-ended questions, and verbatim responses are recorded. These responses are then coded using standard classifications.

The details of these classifications are not provided in the data elements documents because they are very lengthy and can be summarised in a variety of ways. This section provides a summary of the classifications and code frames used for each survey wave.

Table 13 Summary of classifications and code frames used in the LSAY Y06 dataset

Wave/year	Education	Occupation	Industry	Institution
1/2006	ISCED 97	ISCO 88 ANZSCO 1st edition	Not applicable	Not applicable
2/2007	ASCED	ANZSCO 1st edition	ANZSIC 2006	Institution code frame
3/2008	ASCED	ANZSCO 1st edition	ANZSIC 2006	Institution code frame

Notes: ISCED – International Standard Classification of Education
ASCED – Australian Standard Classification of Education
ISCO – International Standard Classification of Occupations
ANZSCO – Australian and New Zealand Standard Classification of Occupations
ANZSIC – Australian and New Zealand Standard Industrial Classification.

Education

The International Standard Classification of Education (ISCED) 1997 is used to code parental education levels and expected student educational levels in the first wave of the 2009 cohort as part of PISA.

The ISCED has the following categories:

- ISCED 1 (primary education)
- ISCED 2 (lower secondary e.g. up to Year 10)
- ISCED 3B or 3C (vocational/pre-vocational upper secondary e.g. Year 11 with Certificate III)
- ISCED 3A (upper secondary e.g. Year 12)
- ISCED 4 (non-tertiary post-secondary e.g. certificate IV)
- ISCED 5B (vocational tertiary e.g. diploma)
- ISCED 5A or 6 (theoretically oriented tertiary and postgraduate e.g. bachelor degree, postgraduate degree).

Further information about ISCED is available at:
<http://www.uis.unesco.org/Education/Pages/international-standard-classification-of-education.aspx>.

The Australian Standard Classification of Education[4] (ASCED) is used to code the area of study from wave 2 (2010).

[4] ABS (Australian Bureau of Statistics), *Australian Standard Classification of Education (ASCED)*, cat.no.1272.0, Canberra, 2001.

Occupation

The International Standard Classification of Occupations (ISCO) 88 is used to code parental occupation in the first wave of the 2009 cohort as part of PISA.

Further information about ISCO is available at:
<http://www.ilo.org/public/english/bureau/stat/isco/isco88/index.htm>

The Australian and New Zealand Standard Classification of Occupations[5] (ANZSCO) first edition is used to code the remaining occupational data. The 'national options' questions relating to occupation asked at wave 1 as part of PISA have been coded using ANZSCO. This includes information about respondents' jobs while at school, and the type of job the respondent expects to have at age 30.

Industry

The Australian and New Zealand Standard Industrial Classification[6] (ANZSIC) 2006 is used to code industries for all waves of the 2009 cohort.

Institution

Non-standard institution code frames have been developed specifically for LSAY to enable consistent coding of education institutions. The code frame incorporates information about the institution campus and uses six digits to code institutions (including campus) from wave 2 (2010).

The institution code frames can be accessed under the 'supporting documents' tab at:
<www.lsay.edu.au/publications/2547.html>.

[5] ABS, *Australian and New Zealand Standard Classification of Occupations*, 1st edn, cat.no.1220.0, ABS, 2006.
[6] ABS, *Australian and New Zealand Standard Industrial Classification,* cat.no.1292.0, Canberra, 2006.

Topic maps

The following series of topic maps list the data elements for each sub-major topic area by minor topic area. The digits within the tables indicate the:

- survey waves in which this data element exists
- number of times the data element appears within a wave. This is equivalent to the number of variables that correspond to the data element in a single wave.

'Topic map 1: Demographics — Student' contains demographic information relating to respondents' place of residence, gender, Indigenous status, date of birth and age, country of birth, language spoken at home, and socioeconomic status.

'Topic map 2: Demographics — Parent' contains demographic information relating to the country of birth, occupation and education levels of a respondent's mother and father.

'Topic map 3: Education — School' contains school education information relating to respondents' school characteristics, student characteristics, student achievement, time spent learning, perceptions about self and school, reading activities, use of computers, teaching and learning, subjects and courses, school plans, careers advice, work experience, workplace learning, qualifications and results, and government payments and income.

'Topic map 4: Education — School transition' contains school transition information about respondents' post-school plans, school leavers, reasons for leaving school, and school leavers' main activity since leaving school.

'Topic map 5: Education — Post-school' contains post-school education information relating to study (including current and past study, apprenticeships and traineeships), qualifications obtained, reasons for withdrawing/deferring from study, changes in study status and/or details (including changes to course, institution, employer, and apprenticeship or traineeship), satisfaction with study, careers advice, and government payments and income.

It is worth noting that within the following minor topic areas:

- 'Study' may refer to past and/or current study as well as apprenticeships and traineeships (for some waves).
- 'Current study' may refer to apprenticeships and traineeships (for some waves).
- 'Past study' may refer to apprenticeships and traineeships (for some waves).
- 'Apprenticeship/traineeships' may refer to past and/or current apprenticeships (for some waves).

'Topic map 6: Employment — Current' contains respondents' current employment including: employment characteristics, time worked, wages and benefits, when started and left work, reasons for leaving work, employment while at school, post-school employment, job training, and job satisfaction.

'Topic map 7: Employment — Job history and training' contains respondents' job history and training information (including any other employment currently undertaken by the respondent)

relating to employment characteristics, time worked, wages and benefits, job training undertaken, reasons for leaving work, and perceptions about work.

'Topic map 8: Employment — Seeking employment' contains information about respondents' job-seeking behaviour, including whether they were looking for work, job search activity details and problems looking for work.

'Topic map 9: Employment — Not in the labour force' contains respondents' main activity while not in the labour force and their intentions for seeking employment or commencing study.

'Topic map 10: Social — Health, living arrangements and finance' contains information about respondents' living arrangements, household possessions, children, marriage, disability and health, government payments, housing payments and financial circumstances.

'Topic map 11: Social — General attitudes' contains information about what respondents do in their leisure time, their life satisfaction and aspirations.

Topic map 1: Demographics – Student

		Wave/year		
Minor topic area	**Data element**	**1/2009**	**2/2010**	**3/2011**
Place of residence	State		1	1
	Postcode		1	1
Gender	Gender	1	1	
Indigenous status	ATSI	1		
Date of birth/ age	Age	1		
	Date of birth: month	1		
	Date of birth: year	1		
	Date of birth	1		
	Date of birth: SAS date	1		
Country of birth	Country of birth	1		
	Country of birth: other	1		
	Country of birth: all	1		
	Immigration status	1		
	Immigration status: Australian definition	1		
	Age of arrival	1		
Language spoken at home	Language spoken at home	2		
	Language spoken at home: all	2		
Socioeconomic status	Cultural possessions (index)	1		
	Educational resources (index)	1		
	Household possessions (index)	1		
	Wealth	1		
	Economic social and cultural status (index)	1		

Topic map 2: Demographics – Parent

		Wave/year		
Minor topic area	**Data element**	**1/2009**	**2/2010**	**3/2011**
Country of birth	Mother's country of birth	1		
	Mother's country of birth: other	1		
	Mother's country of birth: all	1		
	Father's country of birth	1		
	Father's country of birth: other	1		
	Father's country of birth: all	1		
Occupation	Mother's occupation (ISCO)	1		
	Mother's occupation: white/blue collar classification	1		
	Mother works in job/business		1	
	Mother works full/part-time		1	
	Mother's occupation (ANZSCO)		1	
	Mother's main activity	1	1	
	Father's occupation (ISCO)	1		
	Father's occupation: white/blue collar classification	1		
	Father works in job/business		1	
	Father works full/part-time		1	
	Father's occupation (ANZSCO)		1	
	Father's main activity	1	1	
	Parents' occupation: white/blue collar classification	1		
Education	Mother's schooling	1		
	Mother's qualifications: post-secondary training certificate	1		
	Mother's qualifications: post-secondary training qualification	1		
	Mother's qualifications: university degree	1		
	Mother's qualifications: doctorate (PhD) or equivalent	1		
	Mother's highest education level (ISCED)	1		
	Mother's qualifications: post-secondary qualification		1	
	Mother's qualifications: post-secondary qualification (type)		1	
	Father's schooling	1		
	Father's qualifications: post-secondary training certificate	1		
	Father's qualifications: post-secondary training qualification	1		

Minor topic area	Data element	Wave/year 1/2009	2/2010	3/2011
	Father's qualifications: university degree	1		
	Father's qualifications: doctorate (PhD) or equivalent	1		
	Father's highest education level (ISCED)	1		
	Father's qualifications: post-secondary qualification		1	
	Father's qualifications: post-secondary qualification (type)		1	
	Parents' highest education level (ISCED)	1		
	Parents' highest education level (years)	1		
Socioeconomic status	Mother's ISEI score	1		
	Father's ISEI score	1		
	Parents' ISEI score	1		

Topic map 3: Education – School

		Wave/year		
Minor topic area	**Data element**	**1/2009**	**2/2010**	**3/2011**
School characteristics	School offers IB		1	1
	Geographic location	1		
	School state	1	1	2
	Postcode	1		
	School sector	1	1	1
	School identifier	1		
Student characteristics	Student identifier	2		
	At school		1	2
	At school (at last interview)			1
	Year level	1	1	1
	Full-time or part-time study		1	1
	Study program	1		
	ISCED level	1		
	ISCED program	1		
	ISCED orientation	1		
	Studying for IB	1	1	1
	Studying for IB (at last interview)			1
	Attended kindergarten/pre-school	1		
	Primary school commencing age	1		
	Repeated year level: primary	1		
	Repeated year level: lower secondary	1		
	Repeated year level: Year 11 or 12	1		
	Missed primary school	1		
	Missed secondary school	1		
	Changed schools (primary)	1		
	Changed schools (secondary)	1		
	Changed schools		1	1
	Changed schools: month		1	1
	Changed schools: year		1	1
	Current school level (derived variable)	1	1	1
Student achievement	Plausible value in maths	5		

		Wave/year		
Minor topic area	**Data element**	**1/2009**	**2/2010**	**3/2011**
	Plausible value in reading	5		
	Plausible value in science	5		
	Plausible value in reading (access and retrieve)	5		
	Plausible value in reading (integrate and interpret)	5		
	Plausible value in reading (reflect and evaluate)	5		
	Plausible value in reading (continuous text)	5		
	Plausible value in reading (non-continuous text)	5		
	English mark	1		
Time spent learning	Minutes per class	3		
	Minutes per week	3		
	Classes	4		
	Out-of-school	13		
	Out-of-school (primary)	3		
	Science	1		
	Maths	1		
	Other	1		
Perceptions about self and school	Subject: English	1		
	Subject: maths	1		
	Subject: science	1		
	Subjects: overall	1		
	Attitudes towards school	5		
	Student teacher relations	6		
	Life at school	30		
	Coping		5	
Reading activities	Time spent	1		
	Enjoyment	13		
	Diversity	6		
	Online	7		
	Study strategies	16		
Libraries	Use of libraries	8		

		Wave/year		
Minor topic area	**Data element**	**1/2009**	**2/2010**	**3/2011**
	School has library	1		
Reading tasks	Understanding and memorising	7		
	Summarising	6		
Reading for school	Texts and tasks	21		
Use of computers	At school	10		
	Resources: home	10		
	Resources: school	6		
	Entertainment	10		
	At home for school work	6		
	Self-confidence	6		
	Time spent in class	4		
	Time spent outside class	1		
	Attitudes	5		
	Used computer	1		
Teaching and learning English	Class size	1		
	Disciplinary climate	6		
	Stimulating	8		
	Strategies	10		
Science career	Future	3		
Subjects/ courses	Provided school subject information			1
	English		1	1
	English subject		4	4
	LOTE		1	1
	LOTE subject		4	4
	Maths		1	1
	Maths subject		4	4
	Science		1	1
	Science subject		4	4
	Business		1	1

Minor topic area	Data element	Wave/year 1/2009	2/2010	3/2011
	Business subject		4	4
	Humanities/SOSE		1	1
	Humanities/SOSE subject		4	4
	Arts		1	1
	Arts subject		4	4
	Health/PE		1	1
	Health/PE subject		4	4
	Computing		1	1
	Computing subject		4	4
	Home economics		1	1
	Home economics subject		4	4
	Technology		1	1
	Technology subject		4	4
	Other		1	1
	Other subject		4	4
	Other study		1	1
	Qualification		1	1
Subjects/ courses: VET	VET subjects	1	1	1
	Number of VET subjects		1	1
	VET subjects part of apprenticeship/traineeship	1	1	1
	VET subjects at school		1	1
	VET subjects at TAFE		1	1
	VET subjects at ACE		1	1
	VET subjects at other training organisation		1	1
	TAFE subjects	1		
	TAFE subjects part of apprenticeship/traineeship	1		
	English subject is VET		4	4
	LOTE subject is VET		4	4
	Maths subject is VET		4	4
	Science subject is VET		4	4
	Business subject is VET		4	4
	Humanities/SOSE subject is VET		4	4
	Arts subject is VET		4	4
	Health/PE subject is VET		4	4

Minor topic area	Data element	Wave/year 1/2009	2/2010	3/2011
	Computing subject is VET		4	4
	Home economics subject is VET		4	4
	Technology subject is VET		4	4
	Other subject is VET		4	4
School plans	Educational expectations	6		
	Plan to complete Year 12	1	1	1
Careers advice	Helped to develop formal plan		1	1
	Helped to make career decision		1	1
	Helped to make subject/course decision		1	1
	Helped to prepare to apply for job		1	1
	Helped to prepare post-school study application		1	1
	Helped to develop formal plan (at school)		1	1
	Helped to make career decision (at school)		1	1
	Helped to make subject/course decision (at school)		1	1
	Helped to prepare to apply for job (at school)		1	1
	Helped to prepare post-school study application (at school)		1	1
	Talked to a teacher/careers guidance officer		1	
	Talked to careers guidance officer			1
	Talked to a teacher			1
	Talked with person in desired job		1	1
	Questionnaire		1	
	Visited workplace		1	1
	University information session		1	1
	TAFE information session		1	1
	Careers expo/fair		1	1
	Used internet site/computer program		1	1
	Talked with family		1	1
	Talked with friends		1	1
	Most useful careers advice		1	1
Work experience	Work experience	1	1	
	Work experience (undertaken)		1	
	Number of days	1	1	

		Wave/year		
Minor topic area	**Data element**	**1/2009**	**2/2010**	**3/2011**
	Teaches what work is really like	1	1	
	Teaches about people	1	1	
	Teaches about instructions	1	1	
	Teaches about thinking for yourself	1	1	
	Teaches about confidence	1	1	
	Teaches about job skills	1	1	
	Teaches about work conditions	1	1	
	Teaches about your future career	1	1	
Workplace learning (TAFE)	Workplace learning	3		
	Number of days (planned)	1		
	Number of days (actual)	1		
	Teaches what work is really like	1		
	Teaches about people	1		
	Teaches about instructions	1		
	Teaches about thinking for self	1		
	Teaches about confidence	1		
	Teaches about job skills	1		
	Teaches about work conditions	1		
	Teaches about your future career	1		
Workplace learning (VET)	Workplace learning	3	1	1
	Workplace learning (undertaken)		1	1
	Time specified		1	1
	Number of days (planned)	1		
	Number of days (actual)	1		
	Number of days		1	1
	Number of hours		1	1
	Teaches what work is really like	1	1	1
	Teaches about people	1	1	1
	Teaches about instructions	1	1	1
	Teaches about thinking for yourself	1	1	1
	Teaches about confidence	1	1	1
	Teaches about job skills	1	1	1
	Teaches about work conditions	1	1	1

Minor topic area	Data element	Wave/year 1/2009	2/2010	3/2011
	Teaches about your future career	1	1	1
	Certificate of attainment		1	1
Qualifications and results	Awarded certificate			1
	Received any other certificate			1
	Certificate name			2
	Received (state-specific) score			1
	Result			1
	Highest school level completed (derived variable)	1	1	1
	Completed Year 12 or certificate II or higher (derived variable)	1	1	1
	Completed Year 12 or certificate III or higher (derived variable)	1	1	1
Government payments and income	Receive Youth Allowance/ABSTUDY		1	1
	Fortnightly Youth Allowance/ABSTUDY payment		1	1
	Stay on at school without Youth Allowance/ABSTUDY		1	1

Topic map 4: Education – School transition

		Wave/year		
Minor topic area	**Data element**	**1/2009**	**2/2010**	**3/2011**
Post-school plans	Student plans	1		
	Student plans (immediate)		1	1
	Student plans (eventual)		1	1
	Parents' plans	1		
	Friends' plans	1		
	Study plans	1	1	1
	Study plans (verbatim)	1		
	Study plans (type)		1	1
	Study plans (timeframe)		1	1
	Influence: family		1	1
	Influence: friends		1	1
	Influence: school teachers		1	1
	Influence: university/tafe representatives		1	1
	Influence: media		1	1
	Influence: career advisors		1	1
	Influence: job		1	1
	Influence: information from employers		1	1
	Influence: school work experience		1	1
	Influence: community groups		1	1
	Prepared to make decisions about future		1	1
	Better prepared: school subject information		1	1
	Better prepared: work experience/on-the-job training		1	1
	Better prepared: career options information/tools		1	1
	Better prepared: future study options		1	1
	Better prepared: support		1	1
	Better prepared: other		1	1
School leavers	Left school before completing Year 12		1	1
	Month left school		1	1
	Year left school		1	1
	Year level left school		1	2
	Feelings about having left school		1	1

Minor topic area	Data element	Wave/year 1/2009	2/2010	3/2011
	Main activity		1	1
	Prepared to make decisions about future career		1	1
	Reason: have job/apprenticeship		1	1
	Reason: to get job/apprenticeship		1	1
	Reason: not good at school		1	1
	Reason: study/training not available		1	1
	Reason: didn't like school		1	1
	Reason: financially difficult		1	1
	Reason: teachers		1	1
	Reason: earn own money		1	1
	Reason: parents		1	1
	Reason: Year 12 wouldn't help get a job		1	1
	Reason: Year 12 wouldn't help with further study/training		1	1
	Reason: main reason		1	1
	Received study/training advice: university		1	1
	Received study/training advice: TAFE		1	1
	Received study/training advice: other educational organisation		1	1
	Received study/training advice: none		1	1
	Study/training advice: on-campus (university)		1	1
	Study/training advice: on-campus (TAFE)		1	1
	Study/training advice: on-campus (other)		1	1
	Study/training advice: mentoring		1	1
	Study/training advice: summer school/short course		1	1
	Study/training advice: Youth Allowance		1	1
	Main reason returned to school			1

Topic map 5: Education – Post-school

		Wave/year		
Minor topic area	**Data element**	**1/2009**	**2/2010**	**3/2011**
Study	Study status (at last interview)			1
	Still studying		4	6
	Confirmation of study			1
	Confirmation of deferred study			1
	Resumption of deferred study			1
	Commenced study		1	1
	Study type		1	1
	Qualification		2	2
	Qualification (at last interview)			1
	Main area of study		1	2
	Institution		2	4
	Month started study		1	1
	Year started study		1	1
	Applied for university place		1	1
	Intend to apply for university place		1	1
	Intend to reapply for university place		1	2
	First preference: offered place		1	1
	First preference: institution		1	1
	First preference: accepted place			1
	First preference: reason did not take up place (taking break/holiday/travel)		1	1
	First preference: reason did not take up place (required leaving home)		1	1
	First preference: reason did not take up place (need Youth Allowance)		1	1
	First preference: reason did not take up place (considering options)		1	1
	First preference: reason did not take up place (course costs)		1	1
	First preference: reason did not take up place (financial)		1	1
	First preference: reason did not take up place (prefer to work)		1	1
	First preference: reason did not take up place (prefer to study at TAFE)		1	1
	First preference: reason did not take up place (other)		1	1

Minor topic area	Data element	Wave/year 1/2009	2/2010	3/2011
	First preference: reason did not take up place (main reason)		1	1
	University: offered place		1	1
	University: institution		1	1
	University: accepted place		1	1
	University: reason did not take up place (taking break/holiday/travel)		1	1
	University: reason did not take up place (required leaving home)		1	1
	University: reason did not take up place (need Youth Allowance)		1	1
	University: reason did not take up place (considering options)		1	1
	University: reason did not take up place (course costs)		1	1
	University: reason did not take up place (financial)		1	1
	University: reason did not take up place (prefer to work)		1	1
	University: reason did not take up place (prefer to study at TAFE)		1	1
	University: reason did not take up place (other)		1	1
	University: reason did not take up place (main reason)		1	1
	Study status in bachelor degree or higher (derived variable)	1	1	1
	Study status in VET (derived variable)	1	1	1
Current study	Study type		1	1
	Qualification		1	1
	Main area of study		1	2
	Institution		3	10
	Full-time or part-time study		3	4
	Month started study		1	2
	Year started study		1	2
	Month expect to complete study		1	1
	Year expect to complete study		1	1
	Current qualification level (derived variable)	1	1	1
	Full-time or part-time study status (derived variable)	1	1	1
Past study	Study completed/withdrawn/deferred/changed		2	2
	Main area of study		1	2

Minor topic area	Data element	Wave/year 1/2009	2/2010	3/2011
	Institution		2	4
	Full-time or part-time study		3	3
	First preference		1	1
	Month stopped study		3	3
	Year stopped study		3	3
	Highest qualification level completed (derived variable)	1	1	1
Apprenticeships/traineeships	Still studying		1	2
	Confirmation of apprenticeship/traineeship			1
	Qualification		1	1
	Main area of study		1	1
	Employer type		1	1
	Classes/off-the-job training at TAFE		1	1
	Provider of off-the-job training		1	1
	Month started study		1	1
	Year started study		1	1
	Status in apprenticeship/traineeship (derived variable)	1	1	1
Current apprenticeships/ traineeships	Employer type		2	2
	Classes/off-the-job training at TAFE		1	1
	Provider of off-the-job training		1	1
	Full-time or part-time study		2	2
	Month expect to complete study		1	1
	Year expect to complete study		1	1
Past apprenticeships/ traineeships	Study completed/withdrawn/time out/other		1	1
	Employer type		1	1
	Reason apprenticeship/traineeship ended		1	1
	Month stopped study		1	1
	Year stopped study		1	1
Deferred/withdrew from study	Reason: problems juggling study and work commitments		1	1
	Reason: wanted job/apprenticeship/traineeship		1	1

Minor topic area	Data element	Wave/year 1/2009	2/2010	3/2011
	Reason: financially difficult		1	1
	Reason: lost interest		1	1
	Reason: never wanted to study		1	1
	Reason: course was not what you wanted		1	1
	Reason: wouldn't have led to good job/career		1	1
	Reason: poor results		1	1
	Reason: study load		1	1
	Reason: never intended to complete the course		1	1
	Reason: access/transport		1	1
	Reason: health/personal reasons		1	1
	Reason: main reason		1	1
Changed institutions	Same institution		5	7
	Reason: not first choice		5	7
	Reason: better quality education		5	7
	Reason: poor results		5	7
	Reason: course was not what you wanted		5	7
	Reason: course not available at first institution		5	7
	Reason: access/transport		5	7
	Reason: health/personal reasons		5	7
	Reason: main reason		5	7
Changed course	Same course			2
	Reason: course costs		1	2
	Reason: pre-requisite		1	2
	Reason: didn't like course		1	2
	Reason: course was not what you wanted		1	2
	Reason: better career prospects		1	2
	Reason: poor results		1	2
	Reason: study load		1	2
	Reason: preferred to do second course		1	2
	Reason: health/personal reasons		1	2
	Reason: main reason		1	2
Changed/left employer	Same employer		2	2

Minor topic area	Data element	Wave/year 1/2009	2/2010	3/2011
	Circumstances of changing employer		2	2
	Reason: offered better job		2	2
	Reason: boss/other people at work		2	2
	Reason: on-the-job training		2	2
	Reason: travelling/transport		2	2
	Reason: health/personal reasons		2	2
	Reason: main reason		2	2
	Way in which next job was better			2
	Month changed employer		2	2
	Year changed employer		2	2
Changed/stopped apprenticeship/traineeship	Reason: offered better job		1	1
	Reason: pay		1	1
	Reason: job prospects		1	1
	Reason: type of work		1	1
	Reason: boss/other people at work		1	1
	Reason: on-the-job training		1	1
	Reason: off-the-job training		1	1
	Reason: study/training too difficult		1	1
	Reason: travelling/transport		1	1
	Reason: health/personal reasons		1	1
	Reason: main reason		1	1
Satisfaction with study	Problem-solving skills		1	1
	Analytic skills		1	1
	Ability to work as a team member		1	1
	Confidence in tackling unfamiliar problems		1	1
	Communication skills		1	1
	Work planning		1	1
	Overall satisfaction		1	1
	Improved career prospects		1	1
	Helped make contacts		1	1
	Impressions: like being tertiary student		1	1
	Impressions: student life suits you		1	1

Minor topic area	Data element	Wave/year 1/2009	2/2010	3/2011
	Impressions: like campus atmosphere		1	1
	Impressions: student life meets expectations		1	1
	Impressions: made close friends		1	1
	Problems: paying fees		1	1
	Problems: juggling study and work commitments		1	1
	Problems: course more difficult than expected		1	1
	Problems: conflict between family and study		1	1
	Problems: caring for children or other family members		1	1
	Problems: balancing personal relationships		1	1
	Problems: fitting in with other students and making friends		1	1
	Problems: finding time for other commitments		1	1
	Problems: other		1	1
	Problems: none		1	1
	Problems: main problem		1	1
Careers advice	Careers guidance officer			1
	Questionnaire			1
	Job application assistance			1
	Information about further study			1
	Online tool			1
	Source: educational institution			1
	Source: government agency			1
	Source: employer program			1
	Source: private provider (you paid)			1
	Source: internet			1
	Source: other			1
	Usefulness			1
	Reason for not accessing careers advice			1
Government payments and income	Sources of income: Youth Allowance/ABSTUDY		1	1
	Amount of Youth Allowance/ABSTUDY		1	1
	Youth Allowance/ABSTUDY (independent/dependent)		1	1
	Sources of income: paid work		1	1
	Sources of income: parents or family		1	1

Minor topic area	Data element	Wave/year 1/2009	2/2010	3/2011
	Sources of income: scholarship or cadetship		1	1
	Sources of income: other government allowance		1	1
	Sources of income: other		1	1
	Sources of income: none		1	1
	Course fees: none		1	1
	Course fees: respondent		1	1
	Course fees: parents/family		1	1
	Course fees: employer		1	1
	Course fees: government		1	1
	Course fees: other		1	1
	Commonwealth supported (HECS)		1	1
	Commonwealth supported (HECS)/full-fee paying		1	1
	Full-fee paying		1	
	Full-fee paying: FEE-HELP			1
	Full-fee paying: up-front			1
	Full-fee paying: payment scheme			1
	Full-fee paying: employer			1
	Full-fee paying: scholarship			1

Topic map 6: Employment – Current

		Wave/year		
Minor topic area	**Data element**	**1/2009**	**2/2010**	**3/2011**
Employment characteristics	Work in job/business/farm		1	1
	Still have job (reported at last interview)			1
	Away from job		1	1
	School holiday job		1	1
	More than one job		1	1
	Number of other jobs had		1	1
	Wages/salary/self-employed		1	1
	Kind of work (ANZSCO)	1	1	1
	Employer's main kind of business (ANZSIC)		1	1
	Change of work conditions: pay			1
	Change of work conditions: skills			1
	Change of work conditions: responsibility			1
	Change of work conditions: promotion			1
	Part-time/casual	1		
	Workplace learning job		1	1
	Labour force status (derived variable)	1	1	1
	Permanent or casual employment (derived variable)	1	1	1
	Occupation (1 digit ANZSCO First Edition) (derived variable)	1	1	1
	In full-time employment or full-time education (derived variable)	1	1	1
	Job mobility during last year (derived variable)	1	1	1
	Any spell of unemployment during the year (derived variable)	1	1	1
Time worked	Hours worked per week (present job)	1	1	1
	Hours worked per week (main job if more than one)		1	1
	Hours worked per week (all jobs if more than one)		1	1
	Hours worked per week (job reported at last interview)			1
	Hours worked per week (weekdays)	1		
	Hours worked per week (weekend)	1		
	Prefer different hours		1	1
	Preferred weekly hours		1	1
	Months worked		13	20

Minor topic area	Data element	Wave/year 1/2009	2/2010	3/2011
	Full-time or part-time employment status (derived variable)	1	1	1
	Average weekly hours worked (derived variable)	1	1	1
Wages and benefits	Frequency of pay	1	1	1
	Gross pay		1	1
	Annual salary		1	1
	Hourly rate		1	1
	Average weekly earnings		1	1
	Take-home pay (dollars)	1		
	Take-home pay (cents)	1		
	Take-home pay		1	1
	Pay type		1	
	Annual/sick leave		1	1
	Average weekly pay (derived variable)	1	1	1
	Average hourly pay (derived variable)	1	1	1
Starting work	Month began job		1	1
	Year began job		1	1
	How found job		1	1
Looking for work	Looking for work		1	1
	Looking for work (additional or to change jobs)		1	1
Working in a job while at school	Kind of work want as career	1		
	Enjoy work	1		
	Family needs money	1		
	Independence	1		
	Help get job	1		
	Family business	1		
	Support myself	1		
	Spending money	1		
	Counts towards school leaving certificate	1		
Working in a job post-school	Full-time job since leaving school		1	1

Minor topic area	Data element	Wave/year 1/2009	2/2010	3/2011
	Full-time job since leaving full-time study		1	1
	Time taken to find full-time job		1	1
	Still have job		1	1
Job training	Classroom-based training		1	1
	Hours of classroom-based training		1	1
	Training outside workplace		1	1
	Hours of training outside workplace		1	1
	On-the-job training		1	1
	Training helped get promotion or pay rise		1	1
	Training could help to get promotion or pay rise		1	1
	Training could help to get more responsibility		1	1
	Training could help to get a different type of job		1	1
	Use of training		1	1
	Suitable amount of training received		1	1
Job satisfaction	Like job as career		1	1
	Kind of work		1	1
	Utilise skills/experience			1
	Immediate boss/supervisor		1	1
	Other people		1	1
	Pay		1	1
	Opportunities for training		1	1
	Tasks assigned		1	1
	Recognition		1	1
	Opportunities for promotion		1	1
Perceptions about work	Teaches what work is really like		1	1
	Teaches about people		1	1
	Teaches about instructions		1	1
	Teaches about thinking for yourself		1	1
	Teaches about confidence		1	1
	Teaches about work conditions		1	1
	Teaches about future career		1	1

Topic map 7: Employment – Job history and training

		Wave/year		
Minor topic area	**Data element**	**1/2009**	**2/2010**	**3/2011**
Employment characteristics	Work in job/business/farm (at last interview)			1
	Re-definition of second job as main job			1
	Kind of work (ANZSCO)		1	1
	Kind of work: other/second job (ANZSCO)		1	1
	Kind of work: other/third job (ANZSCO)		1	1
	Employer's main kind of business (ANZSIC)		1	1
	Employer's main kind of business: other/second job (ANZSIC)		1	1
	Employer's main kind of business: other/third job (ANZSIC)		1	1
	Wages/salary/self-employed: other/second job		1	1
	Wages/salary/self-employed: other/third job		1	1
Time worked	Hours worked per week: other/second job		1	1
	Hours worked per week: other/third job		1	1
Wages and benefits	Pay type: other/second job		1	1
	Pay type: other/third job		1	1
	Gross weekly pay: other/second job		1	1
	Gross weekly pay: other/third job		1	1
	Average weekly earnings: other/second job		1	1
	Average weekly earnings: other/third job		1	1
	Hourly rate: other/second job		1	1
	Hourly rate: other/third job		1	1
	Annual salary: other/second job		1	1
	Annual salary: other/third job		1	1
Job training	Classroom-based training		1	1
	Hours of classroom-based training		1	1
	Training outside workplace		1	1
	Hours of training outside workplace		1	1

Minor topic area	Data element	Wave/year 1/2009	2/2010	3/2011
	On-the-job training		1	1
	Training helped get promotion or pay rise		1	1
	Training could help to get more responsibility		1	1
	Training could help to get different type of job		1	1
	Suitable amount of training received		1	1
Leaving work	Main reason left job		1	3
	Month left job			2
	Year left job			2
	Way in which next job was better			1

Topic map 8: Employment – Seeking employment

		Wave/year		
Minor topic area	**Data element**	**1/2009**	**2/2010**	**3/2011**
Looking for work	Looking for work (in the last 4 weeks)		1	1
	Looking for full-time or part-time work		1	1
	Prefer full-time work		1	1
	Available for work last week		1	1
Job search activity	Looked for work		1	1
	Number of weeks looking for work		1	
	Months looking for work		13	20
	Registered with Centrelink		1	1
	Checked Centrelink touch screens/computers		1	1
	Checked/registered with Job Network/Job Services member		1	1
	Checked/registered with any other employment agency		1	1
	Looked at advertisements in newspaper/on the internet		1	1
	Answered advertisements in newspapers/on the internet		1	1
	Contacted friends or relatives		1	1
	Written/phoned/approached an employer about a job		1	1
	Checked workplace noticeboards		1	1
	Asked school or another organisation for advice		1	1
	Advertised/tendered for work		1	1
Problems looking for work	Health problems or some disability		1	1
	Problems with childcare		1	1
	Don't have suitable transport		1	1
	Not enough of the right kind of education		1	1
	Don't have enough work experience		1	1
	Not enough jobs available		1	1
	Gender		1	1
	Racial/ethnic background		1	1
	Need better reading and writing skills		1	1
	Don't have good interview skills		1	1
	Lack of skills in writing job applications		1	1
	Lack confidence		1	1

Minor topic area	Data element	Wave/year 1/2009	2/2010	3/2011
	Not good with numbers		1	1
	Poor language or communication skills		1	1
	Age, gender or other discrimination		1	1
	Age discrimination		1	1
	Other discrimination		1	1

Topic map 9: Employment – Not in the labour force

		Wave/year		
Minor topic area	**Data element**	**1/2009**	**2/2010**	**3/2011**
Main activity	Main activity		1	1
Education	Likelihood of beginning full-time study		1	1
	Timeframe for beginning study		1	1
Employment	Likelihood of seeking employment		1	1
	Timeframe for seeking employment		1	1

Topic map 10: Social – Health, living arrangements and finance

		Wave/year		
Minor topic area	**Data element**	**1/2009**	**2/2010**	**3/2011**
Living arrangements	Family structure	1		
	Live with parents		1	1
	Number of (other) people in household		1	1
	Father/step-father	1	1	1
	Mother/step-mother	1	1	1
	Brother/step-brother	1	1	1
	Sister/step-sister	1	1	1
	Grandparents	1		
	Other	1		
	Husband/wife/de facto		1	1
	Partner		1	1
	Boyfriend/girlfriend		1	1
	Own children		1	1
	Other relatives		1	1
	Non-relatives		1	1
	Living with parent(s) (derived variable)	1	1	1
	Living in own home (derived variable)	1	1	1
	Number of dependent children (derived variable)	1	1	1
Household possessions	Desk	1		
	Own room	1		
	Quiet study place	1		
	Computer	1		
	Software	1		
	Internet	1		
	Literature	1		
	Poetry	1		
	Art	1		
	Textbooks	1		
	Technical reference books	1		
	Dictionary	1		
	Dishwasher	1		
	DVD/VCR	1		

		Wave/year		
Minor topic area	**Data element**	**1/2009**	**2/2010**	**3/2011**
	Cable/pay TV	1		
	Digital camera	1		
	Plasma TV	1		
	Number of mobile phones	1		
	Number of TVs	1		
	Number of computers	1		
	Number of cars	1		
	Number of rooms with bath/shower	1		
	Number of books	1		
Marriage	Marital status (derived variable)	1	1	1
Disability and health	General health		1	
Government payments	Youth Allowance/Newstart Allowance		1	1
	Parenting Payment		1	1
	Sickness Allowance		1	1
	Disability Support Pension		1	1
	Family Tax Benefit		1	1
	Other		1	1
	None of these		1	1
	Amount per fortnight received in government payments		1	1
	Amount per year received in government payments			1

Topic map 11: Social – General attitudes

		Wave/year		
Minor topic area	**Data element**	**1/2009**	**2/2010**	**3/2011**
Leisure	Hours spent watching TV	1		
	Hours spent listening to music	1		
	Hours spent playing sport	1		
	Hours spent reading for pleasure	1		
	Hours spent doing unpaid/volunteer work	1		
	Hours spent using the internet for pleasure	1		
	Go to the library			1
	Read books			1
	Read newspapers or magazines			1
	Use the internet			1
	Play computer/video games			1
	Play sport or do exercise			1
	Community activities			1
	Go to church/place of worship			1
	Volunteer			1
Life satisfaction	The work you do		1	1
	What you do in your spare time		1	1
	How you get on with people		1	1
	The money you get each week		1	1
	Your social life		1	1
	Your independence		1	1
	Your career prospects		1	1
	Your future		1	1
	Your life at home		1	1
	Your standard of living		1	1
	Where you live		1	1
	Your life as a whole		1	1
Job aspirations	Type of job expect at age 30 (ANZSCO)	1		
Aspirations	Personal goal			1

Appendix A: Updates to the Y09 data file

The following table tracks updates made to the Y09 data file deposited with the Australian Data Archive. Users are encouraged to download the most recent version of the dataset to ensure all updates are included.

Table 14 Summary of changes made to the Y09 data file

Wave/year	Version	Date published	Description	Variables and number of observations affected
Waves 1 to 3 (2009 to 2011)	Version 1	October 2012	Data file created incorporating data from waves 1, 2 and 3 (2009, 2010 and 2011).	na

Australian Government

Department of Education, Employment and Workplace Relations

National Centre for Vocational Education Research Ltd
Level 11, 33 King William Street, Adelaide, South Australia
PO Box 8288, Station Arcade, SA 5000 Australia
Telephone +61 8 8230 8400 Facsimile +61 8 8212 3436
Website www.ncver.edu.au Email ncver@ncver.edu.au