

# Using Longitudinal Data for Research on VET

**Phillip McKenzie**  
**Australian Council for Educational Research (ACER) &**  
**Monash University-ACER**  
**Centre for the Economics of Education and Training (CEET)**

## INTRODUCTION

Longitudinal studies can provide insights on young people's transition from education to work that other forms of data cannot. The Longitudinal Surveys of Australian Youth (LSAY) program, which is managed jointly by ACER and the Commonwealth Department of Education, Training and Youth Affairs (DETYA), has now accumulated more than 20 years of data that follow successive cohorts of young Australians as they move through education and training and into the labour market. The data are added to every year. This paper explores the potential of LSAY data for research on VET, and also some of the challenges that VET poses for longitudinal analyses. Illustrations are provided from recent research on the backgrounds of young people participating in VET, and the links between VET and labour market outcomes. The LSAY data are publicly available for use by researchers, and the paper concludes by indicating how the database can be accessed.

## FEATURES OF THE LSAY PROGRAM

LSAY has essentially been designed to inform policy aimed at improving young people's transition from school to work. Its origins in the late 1970s arose largely from concerns about the rapid rise in youth unemployment and social exclusion that first became evident in Australia and most other OECD countries in the mid-1970s. The focus on school-to-work transition has led to certain design features in LSAY.

The sample is first contacted when respondents are in Year 9, which is just before the end of compulsory schooling. Since school background and achievement in literacy and numeracy play key roles in shaping post-compulsory education and training opportunities and access to jobs, LSAY starts by collecting extensive information while young people are still in school, and from their teachers and school principals as well.

The focus on school-to-work transition means that the LSAY data collections are annual. Annual surveys are necessary until young people reach their mid-twenties because of the rapidly changing education and labour market circumstances, and high mobility, of teenagers and young adults during the transition process. Annual contact with the sample also reduces sample attrition.

LSAY has a strong concentration on education-to-work linkages so it collects extensive information on educational and occupational aspirations, part-time work while in full-time study, VET in schools, and a wide variety of labour market outcomes (extent and duration of unemployment; occupational status; hours of work; earnings; access to job training; job satisfaction; job mobility and so on).

The LSAY samples are now relatively large (around 13,500 in the original Year 9 sample and about 11,500 when the telephone interviewing commences two years later) because of the interest in educational and labour market outcomes classified by social background factors, State and school sector.

Following initial data collection in school, and a mail survey in the second year, subsequent contact with the sample is via a telephone survey that averages about 20 minutes in length. Using computer-aided telephone interviewing for the subsequent data collections allows more extensive information to be collected than is possible through mail surveys, and at a lower unit cost than would be incurred through face-to-face interviewing.

The program is multi-wave as well as longitudinal. LSAY adds a new cohort at regular intervals, the two most recent being the Year 9 classes of 1995 and 1998. New cohorts are added at reasonable time intervals to allow for monitoring of the effects of contextual changes as well whether the relationships among key variables change over time. Much of the background to LSAY resides in two

earlier programs of longitudinal studies: the ACER program called Youth in Transition (YIT); and the Australian Youth Survey (AYS) (and its predecessor the Australian Longitudinal Survey) conducted by the-then DEETYA. The earlier data collections enable the current experiences of young people as they move through education into the labour market to be compared with those experienced by other cohorts of young Australians over the past 20 years.

## STRUCTURE OF THE LSAY DATA COLLECTIONS

Over time the LSAY data collections from each cohort build up a comprehensive picture of the social and educational backgrounds of young people, their participation in various forms of education, training and work, and their attitudes to education, work and life more generally. Not all of these data are collected each year, and the data collection changes somewhat in coverage as the cohorts gradually get older, although there is a common core of data items. The longitudinal nature of the LSAY data collections means that new surveys are closely linked to, are comparable with, and build on the previous corresponding surveys.

The common areas covered each year are as follows:

- Education experiences (program, institution, type of enrolment, performance)
- Labour market experiences (employment, type of job, occupation, industry, earnings, job training, job history, job search activity)
- Non-work and education activities
- Health, living arrangements and financial support
- Attitudes and aspirations

Table 1 shows the structure of the LSAY data collections, and how the variables collected change over time as the cohorts age. For example, from year 5 of the data collections onwards much more extensive data are collected on labour market experiences than in earlier years since by that stage almost all of the sample have left school. The data collections also change in response to emerging policy concerns and extensive consultative processes.

**Table 1: Structure of the LSAY cohort data collections**

Year data collected	Modal age	Young people's activities	Data collection method	Main data collected
1	14	All in Year 9	In-school tests and survey	<ul style="list-style-type: none"> <li>• Social background</li> <li>• Literacy &amp; numeracy</li> <li>• Attitudes to school</li> <li>• Aspirations</li> </ul>
2	15	Almost all in Year 10; small number of early school leavers	<ul style="list-style-type: none"> <li>• At-home mail survey</li> <li>• In-school mail surveys (school principal &amp; Year 10 teachers)</li> </ul>	<ul style="list-style-type: none"> <li>• Schooling &amp; labour market activities</li> <li>• School structure, programs and environment</li> </ul>
3	16	Most in Year 11; some in VET or the labour market	Phone survey	<ul style="list-style-type: none"> <li>• School activities</li> <li>• Transition from school</li> <li>• Post-school education &amp; training</li> <li>• Employment</li> <li>• Job search</li> <li>• Not in the labour market</li> <li>• Living arrangements, health, general attitudes</li> </ul>
4	17	Most in Year 12; some in VET, the labour market, or outside the labour force	Phone survey	<ul style="list-style-type: none"> <li>• As above, with more detailed questions on Year 12 courses</li> </ul>
5	18	Almost all have left school; wide variety of higher education, VET, labour market and non-work	Phone survey	<ul style="list-style-type: none"> <li>• As above, with new questions on post-school study and training, and participation in non-formal learning</li> </ul>

		activities		• Final school certification and results
6	19	Wide variety of higher education, VET, labour market and non-work activities	Phone survey	As above, with new questions on family formation
7 and onwards	20 plus	Increasingly differentiated education, training and employment pathways and outcomes	Phone survey	As above, with new questions to reflect increasing variety of experiences and outcomes

## WHAT CAN LONGITUDINAL DATA PROVIDE?

The essence of longitudinal data is that the same people are surveyed on a regular basis over time. In the case of LSAY the surveys start at around age 14 or 15 and, resources permitting, continue until the young people are aged at least in their mid-twenties. In the case of the cohort born in 1961 the surveys continued until they were aged 33 years. By surveying the same young people over time the program enables an understanding of the changes taking place in their lives -- and the ways that previous experiences influence what is happening to them now.

### Mapping pathways

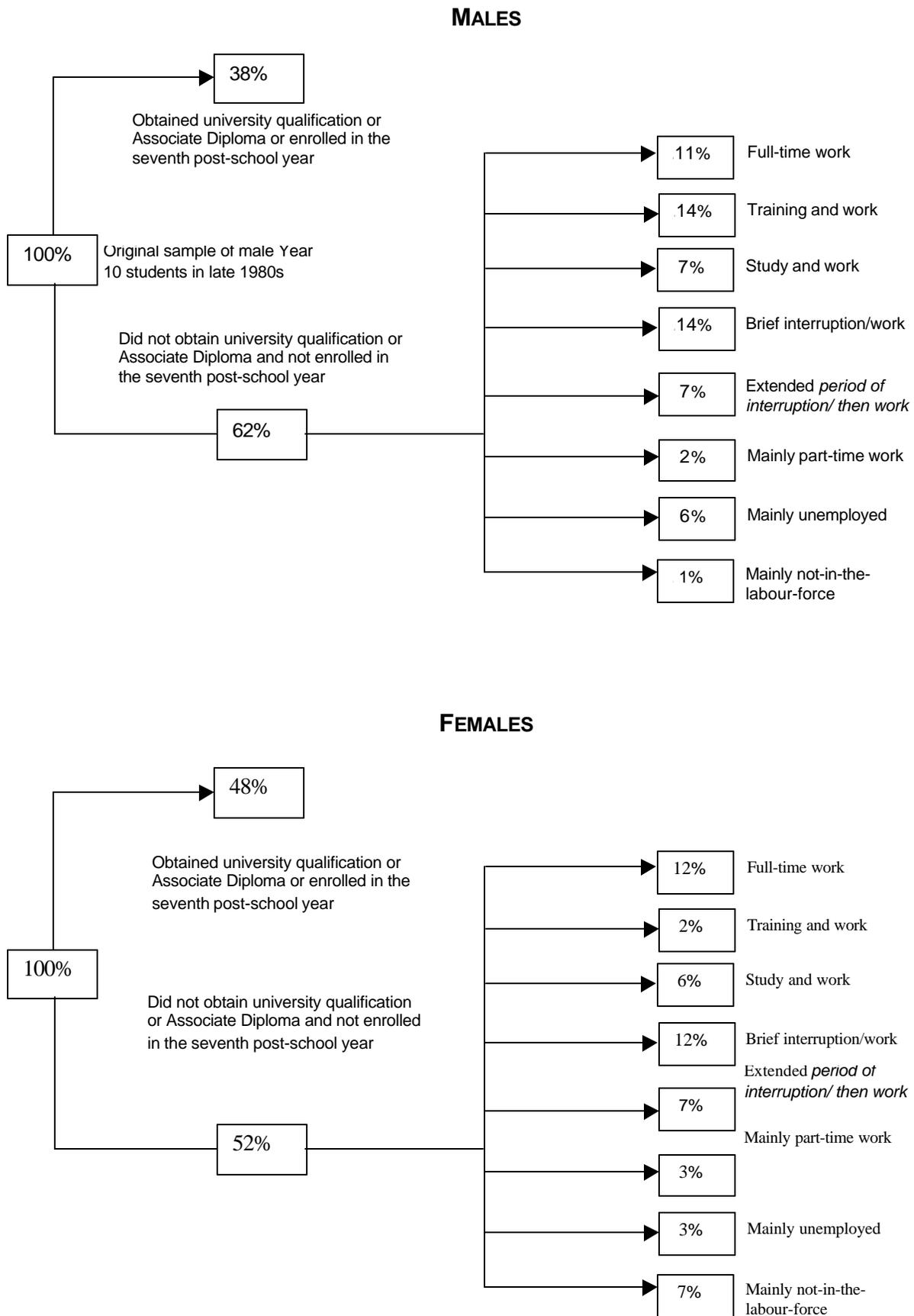
The recent OECD comparative review of education-to-work transition showed that, in comparison with many OECD countries, Australia places an increasing emphasis on individuals constructing their own pathways through education and training and into work (OECD, 1999). As the LSAY database tracks individuals on an annual basis from around the age of 14 years, it is able to identify the range of pathways followed by young Australians, their relative significance in terms of the numbers and types of people involved, and the destinations to which they lead.

Lamb and McKenzie (in press) used LSAY data to analyse the variety of post-school pathways followed by young people from different social and educational backgrounds. They examined the first seven years after leaving school, and identified almost 500 different patterns of activity in the transition from school (in terms of participation each year in various forms of education and training, full-time or part-time employment, unemployment, or being outside the labour force altogether). The diverse patterns of activities evident from these data reinforce the point that young people's post-school pathways in Australia are highly individualised. It is unlikely, for example, that an equivalent analysis in a country such as Germany, where pathways are more institutionally structured, would reveal that around 80 per cent of the sample engage in almost 500 different patterns of activity over the first 7 post-school years, most of them involving fewer than 10 people following the same pattern.

Figure 1 abstracts from the mass of detailed individualised data by grouping patterns of activity that are in fact very similar to each other. It documents the proportions of males and females who had obtained university or TAFE Associate Diploma qualifications by their early twenties and the principal pathways followed by those who did not obtain such qualifications. The data show that female school leavers were more likely to obtain university or advanced TAFE qualifications than young men, but that young women were also more likely to be on post-school pathways involving mainly part-time work or being outside of the labour force altogether. The "training and work" pathway - essentially apprenticeships - was much more significant for male school leavers than for females.

The analyses showed that young people whose principal activity in the first year after leaving school was either an apprenticeship/traineeship, full-time employment, full-time study, or part-time work and study, were much more likely to experience a successful pathway over the first seven post-school years (defined as spending the majority of that time in full-time employment), than were young people whose principal activity in the first post-school year was either part-time work, being unemployed, or outside the labour force. A good early start - in the sense of being in full-time education, training or employment -- seemed to be particularly important for female school leavers. Findings such as these reinforce the need for tracking the experiences of school leavers and early intervention to assist those at risk in the transition process.

**Figure 1: Pathways of school leavers over the first seven years after leaving school in the late 1980s, by gender (from Lamb & McKenzie)**



## Facilitating causal analyses

Unlike cross-sectional or one-off data collections, longitudinal studies are important for policy analysis not only because they document change over time but also because they enable the influence of policies and practice to be isolated from confounding influences such as social background and context. Each longitudinal record contains information about the past social and educational background of that individual as well as their current occupational or educational status. Allowance can therefore be made for relevant aspects of background when investigating the impact of policy or practice on outcomes.

To understand the processes involved in life histories we need to collect data from the same individuals across time and over an extended period of time ... Cross sectional data collected on repeated occasions enables us to monitor the effects of societal change on the prevalence of population characteristics ... [but] longitudinal data is essential to measure changes in individuals within the population as well ... these incorporate the information essential to gain any purchase on causal processes: we need to know about sequences of life experiences and events, and which individuals are affected by environmental changes, while others remain impervious to them. (Bynner 1996, 6)

The capacity to control for the impact of background influences, and to measure growth or change over time, are key features of longitudinal data.

## Controlling for Background Influences

When there are a number of factors affecting an outcome it is necessary to isolate the effect of each factor holding constant the effects of all the other factors so that the influence of several influences are not confounded. Because in practice individuals are not randomly assigned to different policies, practices and programs any assessment of the effect of these on outcomes must make statistical allowance for the effect of differences in background. Although cross-sectional studies can and do make use of analyses in which statistical controls for the influence of background are invoked the potential is much greater in longitudinal designs because they gather background data at an earlier time rather than when outcomes are measured.

Figure 2 outlines the model used by Long et al (1999) to examine the impact of different factors on participation in post-school education and training. Their work used YIT data to analyse the way in which the family and personal characteristics of respondents, their educational experiences and attitudes and expectations while at school, and Year 12 completion and immediate post-school educational participation are related. The first block of variables in the model, Family & Personal Characteristics, refer to factors that young people in effect bring to school with them, and which are not amenable to policy intervention, at least not in the short-run. The second block of variables, School Experiences & Post-school Expectations, attempts to map the mechanisms by which any effect of family and personal characteristics on educational participation is transmitted. Such variables are in principle more open to policy influences.

Figure 2: Model of Influences on Educational Participation (Long et al, 1999)

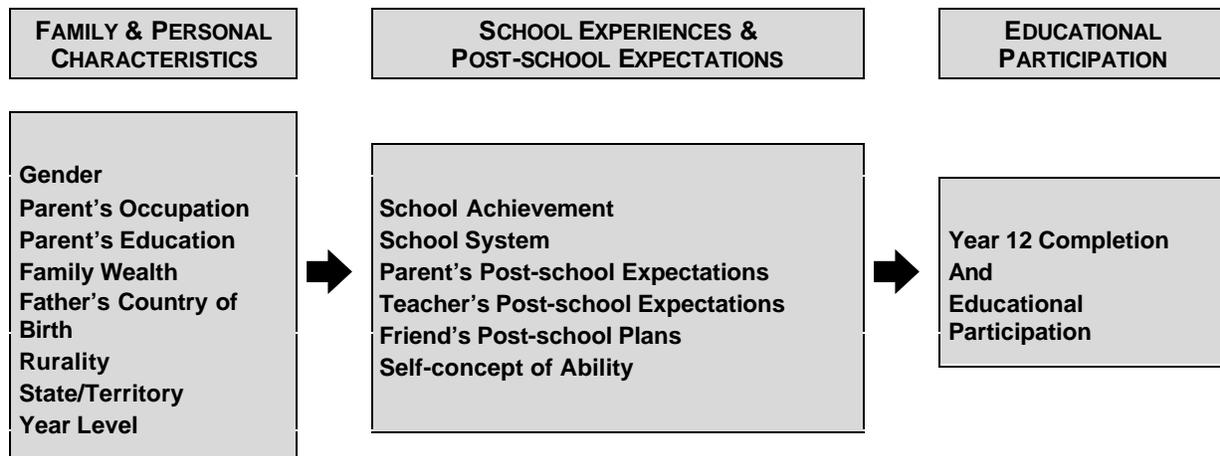


Table 2 applies this general model to the study of participation in apprenticeships at age 19 by the cohort born in 1975. Column 1 (the observed participation rate) shows that young people from homes in which the parent's occupation is classified as *Skilled* are more likely to participate in apprenticeships than other young people. Given that *Skilled* occupations are often those that require an apprenticeship for entrance, and that apprenticeships are a predominantly male activity, these results suggest that sons may be following in their father's footsteps. Young people from semi-skilled and, to a lesser extent, clerical families have participation rates towards the upper end of the range. Apprenticeship participation rates for the *Professional* category are relatively low in each cohort.

**Table 2. Participation in an apprenticeship by age 19 for the cohort born in 1975, by parental occupation**

Parental occupation	Observed participation rate (%)	Rate adjusted for family & personal characteristics (%)	Rate adjusted for column 2 plus school experiences & post-school expectations (%)
Professional	8	8	10
Managerial	14	14	14
Clerical	11	11	11
Skilled	19	19	17
Semiskilled	16	16	14
Unskilled	17	17	15

Source: Long et al (1999).

Statistical adjustment based on multivariate techniques to try to isolate the relative contributions of other factors to the observed differences appear to have little effect on the differences between social groups (see columns 2 and 3). As expected, the statistical adjustments narrow the differences in apprenticeship participation among young people from different social groups which suggests that it is not just parental background which accounts for the differences (eg young people who live in rural areas tend to have a higher participation rate in apprenticeships). However, the fact that the adjusted percentages differ in only small ways from the observed percentages suggest that parental occupation plays a uniquely powerful role in participation in apprenticeships.

### Identifying Value Added

Another illustration of the analytical perspective facilitated by longitudinal studies is that data on early school achievement allow identification of the separate influence of post-school educational attainment on labour market outcomes such as earnings. Not all of the observed higher earnings of university graduates, for example, can be attributed to their degree. Those who enter higher education typically have relatively high levels of early school achievement, are more likely to have completed Year 12, and to have come from more privileged home backgrounds - all of which are associated with lower unemployment rates and higher earnings. The application of analytical models similar to those in Figure 2 by Marks and Fleming (1998) found that after controlling for such factors the earnings advantage of a university qualification does decline, but that it is still positive, adding around 8 per cent to earnings on average, other factors equal. However, the same analyses found that while having an apprenticeship qualification is certainly associated with higher earnings, especially in the early years of employment, the qualification itself appears to add relatively little on its own to earnings once young people reach their late twenties.

## THE CHALLENGES OF VET FOR LONGITUDINAL ANALYSES

Longitudinal data are not easy to analyse. The data sets are typically large, involving a number of years of data collections with varying response rates, requiring a considerable investment of time in data checking and variable construction, and necessitating complex weighting procedures to allow for the problems of sample attrition. Although the attrition rate of the LSAY samples is relatively low, averaging less than 10 per cent per year, it is generally the case that those most likely to drop out of the sample are those who are less successful in educational and employment terms. Weighting can overcome many of the problems due to attrition - since a great deal is known about those who are no longer in the sample - it is an issue that requires on-going attention.

To these general challenges of conducting longitudinal studies can be added some particular features of the Australian VET system and indeed post-school education and training more generally. As is

well known VET provision is becoming increasingly flexible, modularised, provided in a greater variety of settings, and leading to a great diversity of qualifications - or in many cases no completed qualifications although other measures (such as module completions) indicate successful outcomes. Accordingly it is necessary in an annual survey such as LSAY to ask not just about educational participation at the time of the survey (which is typically in the October to December period), but also to seek recall data since the time of the last survey. It is also necessary to seek data about educational participation in a variety of ways (eg name of course, name of institution, main area of study, type of enrolment) to check the consistency of the responses both internally and against other data sources. The fluidity of participation in post-school education and training also means that many analyses are best conducted in terms of ever enrolled in (say) TAFE by age 22 rather than enrolled in TAFE at age 22.

## **ACCESSING LSAY DATA AND RESULTS**

LSAY data are deposited with the Social Science Data Archives (SSDA) at the Australian National University in Canberra in the year after collection. Data from the earlier ACER and DETYA longitudinal surveys are also available through the SSDA. The data are deposited in a form that does not permit the identification of individual sample members or participating schools. The data can be purchased from SSDA in SPSS format for a nominal charge.

ACER maintains an extensive LSAY analytical program organised around three main themes:

- participation in different forms of education and training;
- school and educational effects; and
- transitions from education and training to the labour market and adult life.

The LSAY program has also produced an extensive range of Technical Papers, conference papers, articles, briefings for education and training authorities, and special data analyses on request. Details of the program and its output are available on the ACER Website: [www.acer.edu.au](http://www.acer.edu.au)

## **References**

- Bynner, J. (1996) *Use of Longitudinal Data in the Study of Social Exclusion*, London, Social Statistics Research Unit, City University.
- Lamb, S. & McKenzie, P. (in press) *Patterns of Success and Failure in the Transition from School to Work in Australia*, Melbourne, ACER.
- OECD (1999). *Thematic Review of the Transition from Initial Education to Working Life: Final Comparative Report*. DEELSA/ED 99(11), October 1999. Paris: OECD.
- Marks, G. & Fleming, N. (1998) *Youth Earnings in Australia: 1980-1994: A comparison of three youth cohorts*, LSAY Research Report No. 8, Melbourne, ACER.
- Long, M., Carpenter, P. & Hayden, M. (1999) *Participation in Education and Training in Australia 1980-1994*, LSAY Report No. 13, Melbourne, ACER.