Longitudinal Surveys of Australian Youth

Technical Paper 48

Estimating Attrition Bias in the Year 9 Cohorts of the Longitudinal Surveys of Australian Youth

Sheldon Rothman

April 2009

Australian Council *for* Educational Research

ACER

Published 2009 by
The Australian Council for Educational Research Ltd
19 Prospect Hill Road, Camberwell, Victoria, 3124, Australia.

Table of Contents

List of Tables

List of Figures

Estimating Attrition Bias in the Year 9 Cohorts of the Longitudinal Surveys of Australian Youth

1 INTRODUCTION

One of the major pitfalls of any survey research is the loss of members from the sample. No matter how well-planned a research project may be, some people will refuse to participate in the study when they are approached, or they may refuse to respond to an individual question. If the sample is drawn to be representative of a population, then *nonresponse* can cause problems for the survey.

> The most important consequence [of nonresponse] is that estimates may become biased, because part of the population that is not reached may differ from the part that is sampled. There is now ample evidence that these biases vary considerably from item to item and from survey to survey, being sometimes negligible and sometimes large. A second consequence is, of course, that the variances of estimates are increased because the sample actually obtained is smaller than the target sample. (Cochran, 1977, p. 396)

Longitudinal studies, in which study sample members are interviewed at regular intervals, often over a period of many years, face additional issues regarding nonresponse. Sample members may move between interviews and not have contacted the researchers to inform them of the change. They may not be available for a follow-up interview because of other commitments at the time. They may refuse to continue to participate for personal reasons. Thus the problem of nonresponse is increased for longitudinal studies, because there may be nonresponse for individual items as well as for an entire wave.

The permanent loss of sample members from a longitudinal survey is called *attrition*, and issues caused by this form of nonresponse may be exacerbated. Attrition can negatively affect the entire sample or specific subgroups only. In an effort to ensure the highest possible retention over the life of a longitudinal study, some studies are designed so that sample members who do not respond in one wave are contacted and encouraged to participate in subsequent waves, with values imputed for missing waves.

The full effect of attrition in surveys is impossible to quantify, because non-respondents have already indicated their unwillingness to respond to interviewers' questions. With no data from non-respondents, one cannot determine how much their nonresponse influences outcomes reported for respondents only. The best way to evaluate these effects would be to collect the information from non-respondents, then calculate the differences between findings for the complete sample and findings for the incomplete sample. This means, however, that non-respondents become respondents, and attrition is no longer a concern.[1]

It is possible to accommodate many of the problems caused by nonresponse and attrition. To compensate for missing responses, values may be imputed by estimating how a person would most probably respond to an individual item. To compensate for the loss of sample members in cross-sectional and longitudinal studies, weights can be assigned to remaining sample members to ensure that the distribution of the remaining sample resembles the distribution of the population that the sample was intended to represent. As a result, in a longitudinal study a person's weight may vary from year to year because of differential attrition between those subgroups on which the attrition weights are based.

---

[1] In some survey programs, late respondents—those who had initially refused to participate or were difficult to locate—are considered as a separate category of respondents and treated differently.

This technical paper examines the issue of attrition bias in two cohorts of the Longitudinal Surveys of Australian Youth (LSAY), based on an analysis using data from 1995 to 2002. Data up to 2002 provided eight years of information on members of the Y95 cohort and five years of information on members of the Y98 cohort. This amount of time was considered adequate to evaluate the extent of attrition bias and the performance of weights in correcting for bias. LSAY was designed to explore the transitions made by these cohorts of young people as they leave school and enter the labour force, engage in further study and become adults. It focuses on outcomes and how earlier factors may have influenced those outcomes. At the time each cohort was drawn, the sample represented the population of 15 year-old Australian students attending Australian schools, but like other longitudinal studies, LSAY experiences attrition of its respondents. The weighting schema was designed to ensure that remaining members of the cohort represented the original cohort, not to represent the population of young people in subsequent years.

The goals of this technical paper are:

- understanding the extent of attrition in the 1995 and 1998 Year 9 LSAY cohorts;
- calculating the amount of bias caused by attrition in these cohorts; and
- determining whether the current practice of calculating weights is appropriate or additional practices are required to ameliorate problems caused by attrition in LSAY.

2 ATTRITION IN LONGITUDINAL SURVEYS

Previous Studies of Sample Attrition

Many studies of attrition and bias in longitudinal surveys have already been undertaken, including analyses on the Michigan Panel Study on Income Dynamics (PSID) in the United States; the British Household Panel Survey (BHPS) in Great Britain; the European Community Household Panel (ECHP); and other major national studies of employment and income. Other researchers have examined the extent and effects of attrition in smaller cross-sectional and longitudinal studies. This section presents an overview of some of the issues that result from nonresponse and attrition, as discussed in reviews of other longitudinal studies.

Fitzgerald, Gottschalk and Moffitt (1998a) described patterns of attrition in the PSID. The study began in 1968 with a sample of 4,800 households, which included 17,807 people; by 1989, just over one-half of the original people from those households were still in the study. The largest incidence of attrition occurred in the first follow-up year (1969), when approximately 12 per cent of sample members (representing 10% of families) had not responded. Throughout the life of the study, a small proportion of the nonresponse was caused by death or a household move that could not be tracked; the greater proportion of nonresponse was due to family nonresponse. The authors determined background characteristics of those who responded to each survey over the life of the study.[2] When comparing results for male heads of household by attrition status ('always in' or 'ever out'), they found that 'nonattritors' (annual responders) were significantly more likely to be White, married and regularly employed, have more years of education, and own their home.

These same authors then examined the effects of attrition on a study of intergenerational relationships between parents and children in the PSID (Fitzgerald, Gottschalk, & Moffitt, 1998b). Once again they noted that one-half of their sample had been lost to attrition. They then compared the distribution of the remaining sample of children to the distribution of the population in the United States, based on the 1989 Current Population Survey (CPS). They found 'a close correspondence in characteristics for most demographic variables, especially when sample weights are used' (Fitzgerald et al., 1998b).

Other researchers examined the PSID for the effects of attrition. Lilliard (1998) concentrated on the influences of nonresponse on three outcomes: income, marriage formation and dissolution, and adult death. In one example, they concentrated on the effects of marriage on African-American males' income. They found that even though higher-income African-American males were more likely to become non-respondents in the PSID, only 'very mild' bias was introduced by their nonresponse. Similarly, Zabel (1998) found that even though there are differences in the labour market behaviour between respondents and non-respondents in the PSID, there was little evidence of serious attrition bias reported in the findings.

The National Center for Education Statistics (NCES) in the United States conducts a number of longitudinal studies on students. Two of these, based on young people engaged in post-secondary study, are the *Baccalaureate and Beyond* study, which follows young people after they complete their first university degree, and the *Beginning Postsecondary Students Longitudinal Study*, which follows young people after they enter any form of postsecondary education. Both programs have quantified in technical papers the effects of attrition (Charleston, Riccobono, Mosquin, & Link, 2003; Wine et al., 2002). These technical papers draw similar conclusions, as stated in Wine et al. (2002, p. 119): 'Note that while some variables do show statistically significant biases, the actual bias is generally very small.'

---

[2] In the PSID, it was possible for a household to rejoin the study after missing a year.

Since the analysis was completed for this technical report, papers describing attrition bias and the weighting methodology for two other Australian longitudinal surveys have been published. Interested readers should go to http://www.melbourneinstitute.com/hilda/hdps.html in relation to the Household Income and Labour Dynamics in Australia (HILDA) Survey and http://www.aifs.gov.au/growingup/pubs/technical.html in relation to the Longitudinal Study of Australian Children (LSAC).

Working with Attrition in Longitudinal Surveys

There are a number of ways to deal with attrition in surveys, particularly in longitudinal studies. Attrition can be reduced by retaining membership in the sample through extended tracking of participants if they fail to respond or by repeated attempts to contact sample members. This process can be very expensive, and must be considered feasible only in light of the study's purposes and overall value. For example, studies of men's and women's health—such as those that follow recipients of heart transplants—need to determine outcomes of medical procedures. In such situations, the expense of extended tracking is determined to be worthwhile.[3]

When extended tracking is not feasible, sample refreshment is one option to maintain sample size and representativeness. Sample refreshment involves the inclusion of new sample members who are similar to those who have left the sample. As each successive wave is interviewed, the sample is rebuilt to be representative of the original sample. For example, if attrition among Indigenous males in the study is higher than average, then Indigenous males who were not in the original sample are invited to join the study, provided that they are similar to the original sample members on other relevant characteristics. Sample refreshment is an important approach when the study is designed to provide annual estimates of the population from which the sample is drawn.

As expected, there are costs associated with this approach, in tracking current sample members and in recruiting new sample members. It can also be difficult to coordinate the timing of sample refreshment, as it is necessary first to determine which original sample members have dropped out of the study, then to recruit new members in the time available between preparations for a new study wave and conduct of the next round of interviews. This is possible with lists of the original populations (for example, all members of the appropriate grade levels in a school when only a proportion of the school was originally selected), although not all of the original population would have been tracked over the period of the study. Sample refreshment may also suffer from the lack of original, baseline data for new sample members, necessitating imputation on a number of earlier, unmeasured factors for those in the refreshment group. This may be a problem if these baseline data are important explanatory variables in the study.

It is also possible to adjust for sample attrition mathematically. Two common techniques used are the application of sample weights and the adjustment of estimates based on a group's likelihood of nonresponse. Researchers working with longitudinal data most frequently apply series of weights to their data: in the initial wave to adjust for sample design and in each subsequent wave to adjust for sample attrition. After the initial sampling weights have been calculated and applied, there are new weights to be determined. These new weights are calculated annually to account for differential attrition among groups within the sample.

Differential attrition occurs when subgroups of the sample respond to follow-up surveys at different rates.[4] The primary result of differential attrition is that some groups become

[3] In such situations, nonresponse as the result of death is a possible and informative outcome.

[4] Sample retention rates for selected subgroups in LSAY are shown in Table 1 and Table 2 in the following chapter.

overrepresented and others become underrepresented in these follow-ups. The weights compensate for differential attrition in the same way that the original post-stratification weights account for differential selection and initial responses. In some surveys, nonresponse weights are determined according to the reason for nonresponse, such as those conducted by NCES. Although small, there are differences in the characteristics of sample members classified as location non-respondents, refusal non-respondents, late respondents and refusal-to-nonrefusal respondents; these differences are incorporated into NCES's calculations of nonresponse weights (Charleston et al., 2003; Wine et al., 2002; Zahs, Pedlow, Morrissey, Marnell, & Nichols, 1995).

When used, weighting does introduce some new problems in the reporting and analysis of findings. The major problem is that standard errors of estimates, such as means and proportions, are larger than they would be if the data were not weighted. This occurs because the value of an individual's contribution to the overall statistic is adjusted by the weight so that it becomes lower or higher than originally recorded. To understand this, compare results for two students with test scores of 50; one is from a State that was oversampled and given a weight of .80 to reduce the value of his contribution, and the other is from a State that was undersampled and given a weight of 1.20 to increase the value of her contribution. When unweighted, these two scores have a mean of 50 and no standard error. When weights are applied, the scores are treated as if they were 40 and 60, respectively; the mean is 50, as it was when weights are not applied, but the standard error is 10. In large samples, such as those found in LSAY, the difference between the standard error of the unweighted mean and the standard error of the weighted mean is much smaller. Additionally, as the sample size decreases because of attrition, standard errors of estimates increase.

Another mathematical option for working with attrition is the sample selection method developed by Heckman (1979). This method is used to 'correct' for self-selection in samples, such as initial refusals and attrition, and has two steps. In the first step, the probability of being present for an interview is calculated, based on characteristics determined to be related to attrition. From this, a 'hazard' term is calculated, relating to the probability of being observed. In the second step, the hazard term is applied as a covariate in any modelling used to predict an outcome. It is also necessary that the variables used to predict attrition are not used as predictors to estimate an outcome.

An examination of three of these methods of attrition adjustments—extended tracking, weighting and sample selection models—was conducted by McGuigan, Ellickson, Hays and Bell (1997). The authors used data from a longitudinal study in the United States on substance use among secondary school students. They estimated results for extended tracking by separating follow-up respondents into two groups: those who responded when contacted and those who responded after tracking. Respondents and non-respondents were determined by their status in Year 10. The authors compared mean estimates of substance use in Year 8 using each of the three methods and found that weighting provided the least biased estimates, although with standard errors larger than those calculated after extended tracking. They also found that the sample selection model provided extremely inaccurate estimates, noting '... these results reflect the extreme sensitivity of the sample selection model to the underlying assumption of correct model specification' (McGuigan et al., 1997, p. 565). While tracking provided lower standard errors than weighting, it also provided more biased estimates of substance use at much greater expense, leading the authors to conclude that weighting was the best performer of the three methods (McGuigan et al., 1997, p. 565).

3 WEIGHTING AND ATTRITION IN THE 1995 AND 1998 YEAR 9 LSAY COHORTS

The LSAY program was developed as a successor to two earlier longitudinal studies: the Youth in Transition (YIT) program of the Australian Council for Educational Research (ACER); and the Australian Youth Survey (AYS), with its predecessor the Australian Longitudinal Survey (ALS), conducted by the Commonwealth government. In July 1995, YIT and AYS were brought together as part of LSAY, and a new longitudinal survey commenced, with the selection of a nationally representative sample of 13 000 Year 9 students. A second Year 9 cohort was selected in 1998, comprising more than 14 000 Year 9 students.

The 1995 and 1998 Year 9 LSAY samples (known as Y95 and Y98, respectively) began their participation in the program with 20-item tests in reading comprehension and mathematics and a brief questionnaire, providing information on achievement levels in literacy and numeracy, attitudes and aspirations, and family background. Annual surveys are then used in LSAY to determine young people's experiences in school and the labour force, changes in attitudes and aspirations, participation in social and community activities, and some aspects of their personal circumstances. Cohorts are followed until the young people reach their mid-twenties because it is then that they are fairly well-established themselves in the labour market and social relationships. Following the initial data collection in schools and mail surveys in the second wave, subsequent contact with the sample is by a telephone survey that averages 20 minutes in length.

The 1995 and 1998 Year 9 LSAY cohorts were drawn from the estimated Australian Year 9 population, as determined by the distribution of Year 8 students in 1994 and 1997, respectively. Detailed information on the sampling procedures used in LSAY is available in Long (1996) for the 1995 Year 9 LSAY cohort and in Long and Fleming (2002) for the 1998 Year 9 LSAY cohort. Weighting for the 1995 cohort is described in Marks and Long (2000); the same procedures were established for the 1998 cohort.

For each cohort, post-stratification weights are applied to adjust for sample selection procedures that allowed for oversampling in smaller States and Territories. For the first wave of each cohort, new weights are calculated to compensate for the changes in enrolments between the estimates based on Year 8 enrolments and actual Year 9 enrolments. For subsequent waves of LSAY, weights are also applied to adjust for differential attrition.

Earlier work by Marks and Long (2000) showed that attrition in the 1995 Year 9 LSAY cohort was most commonly associated with a combined measure of performance on tests of achievement in reading comprehension and mathematics, which were administered at the beginning of the survey, and that this attrition operates differently for males and females. They found that weights based on sex and achievement were providing sufficient adjustments in subsequent years, and that there was little change in the annual distribution of the sample on other first-wave variables, such as parent occupation and language background. The application of these weights ensures that individual students contribute to summary statistics only as much as their distribution in the Australian population—in the first wave of the study—would suggest.

Retention of Sample Members

Table 1 shows the retention rates for specific subgroups of the Y95 sample. Table 2 shows the same for the Y98 cohort. These tables are based on unweighted data, based on the sample members who were actually contacted each year. Raw numbers, weighted and unweighted, and annual response rates (percentage of those eligible who responded each year) are provided in the appendix. All characteristics are based on information gathered at first contact in Year 9.

Table 1 Numbers of Y95 sample members by selected characteristics in 1995, and overall retention rates (unweighted), 1996-2002

	1995	1996	1997	1998	1999	2000	2001	2002
All	13613	72.3%	75.7%	71.5%	64.5%	58.0%	50.5%	44.8%
Gender								
Male	6717	66.6%	74.5%	69.7%	62.1%	55.4%	47.7%	42.1%
Female	6896	77.8%	76.9%	73.4%	66.8%	60.5%	53.3%	47.4%
Indigenous background								
Indigenous	385	52.2%	59.2%	53.0%	44.7%	38.2%	31.4%	26.0%
Non-indigenous	12348	73.6%	76.7%	72.6%	65.6%	59.0%	51.7%	45.9%
Double response	4	50.0%	75.0%	75.0%	75.0%	75.0%	75.0%	50.0%
Home language								
English	11687	73.6%	77.1%	73.0%	65.9%	59.2%	51.8%	46.0%
Other language	1305	67.3%	69.0%	63.9%	57.3%	52.2%	44.5%	38.8%
Double response	116	66.4%	72.4%	67.2%	58.6%	51.7%	40.5%	36.2%
State (school)								
Australian Capital Terr.	599	76.1%	73.0%	69.3%	61.4%	55.9%	49.4%	46.1%
New South Wales	3090	68.0%	73.3%	68.0%	60.7%	54.6%	46.8%	40.9%
Victoria	2865	73.2%	77.9%	73.4%	65.4%	59.4%	52.7%	46.9%
Queensland	2524	71.9%	73.9%	69.8%	64.2%	56.8%	49.2%	43.2%
South Australia	1720	79.1%	81.8%	78.4%	72.7%	66.3%	58.6%	52.5%
Western Australia	1837	71.9%	76.7%	73.8%	66.6%	59.5%	51.3%	45.1%
Tasmania	582	76.6%	74.9%	70.8%	60.0%	53.6%	46.6%	42.3%
Northern Territory	396	60.6%	65.2%	60.9%	55.6%	47.7%	40.7%	36.6%
School sector								
Government	9081	70.2%	73.9%	69.3%	62.7%	56.1%	48.7%	43.3%
Catholic	2517	74.7%	79.1%	74.9%	67.8%	61.3%	53.3%	46.5%
Independent	2015	78.7%	79.7%	77.4%	68.4%	62.4%	55.3%	49.4%
Area								
Metropolitan	7564	71.9%	74.6%	70.7%	63.9%	58.0%	50.8%	44.9%
Regional	3378	71.6%	77.1%	72.5%	65.2%	58.3%	50.4%	44.8%
Rural/remote	2629	75.2%	78.2%	74.0%	66.5%	58.3%	50.7%	45.0%
Father's occupational group								
Managers/Farmers	2599	76.8%	79.4%	75.5%	67.5%	61.3%	53.2%	47.7%
Higher professionals	1282	79.9%	81.8%	78.3%	71.3%	66.1%	58.9%	53.2%
Lower professionals	906	81.0%	84.9%	81.9%	75.2%	70.8%	63.4%	57.6%
Other non-manual	1938	72.7%	76.6%	72.6%	65.4%	57.5%	51.2%	45.4%
Manual	4468	71.7%	75.7%	70.9%	64.1%	57.0%	49.6%	43.7%
Residual	2420	60.7%	64.4%	60.2%	53.8%	47.4%	39.5%	33.7%
Mother's occupational group								
Managers/Farmers	808	73.4%	78.1%	74.8%	64.9%	59.4%	51.4%	44.7%
Higher professionals	328	76.5%	77.7%	74.4%	69.2%	62.5%	54.9%	49.4%
Lower professionals	2185	80.5%	83.2%	80.1%	72.7%	67.5%	60.5%	55.0%
Other non-manual	3614	74.3%	78.8%	74.5%	68.2%	61.1%	53.3%	47.3%
Manual	1500	72.3%	75.1%	71.1%	64.5%	57.1%	50.2%	45.0%
Residual	5178	66.9%	70.1%	65.3%	58.2%	51.5%	44.1%	38.4%

Table 2 Numbers of Y98 sample members by selected characteristics in 1998, and overall retention rates (unweighted), 1999-2002

	1998	1999	2000	2001	2002
All	14117	65.8%	67.6%	62.2%	55.0%
Gender					
Male	7227	61.0%	65.8%	59.6%	52.2%
Female	6804	71.2%	70.1%	65.4%	58.3%
Indigenous background					
Indigenous	442	45.0%	48.0%	42.5%	36.2%
Non-indigenous	12917	67.4%	69.3%	63.9%	56.6%
Double response	11	36.4%	18.2%	9.1%	9.1%
Home language					
English	12078	67.4%	69.4%	64.0%	56.8%
Other	1191	59.9%	61.2%	55.5%	48.1%
Double response	253	56.9%	60.1%	52.6%	45.1%
State (school)					
Australian Capital Territory	558	65.4%	74.2%	69.4%	59.7%
New South Wales	3384	63.8%	67.1%	61.8%	54.1%
Victoria	2950	62.8%	66.8%	61.0%	54.5%
Queensland	3111	67.8%	67.8%	61.7%	54.4%
South Australia	1249	69.7%	69.3%	64.1%	57.4%
Western Australia	1689	65.5%	67.4%	61.9%	53.9%
Tasmania	715	77.2%	71.3%	68.0%	63.1%
Northern Territory	461	59.0%	58.6%	53.4%	47.9%
School sector					
Government	8887	64.4%	65.1%	59.0%	51.8%
Catholic	3122	66.9%	73.5%	69.3%	61.4%
Independent	2108	70.2%	69.8%	65.0%	58.9%
Area					
Metropolitan	7763	64.5%	67.0%	61.9%	54.8%
Regional	3169	68.9%	69.8%	63.7%	56.4%
Rural/remote	2474	71.2%	72.6%	66.4%	58.4%
Father's occupational group					
Professional	3740	70.6%	73.0%	67.7%	61.4%
Managers	1836	70.8%	71.8%	67.2%	59.5%
Clerical and personal service	1040	69.3%	73.1%	68.1%	59.7%
Trades	2538	65.2%	70.0%	64.3%	56.5%
Plant operators and labourers	1368	65.1%	65.6%	59.7%	51.5%
Unskilled manual	726	61.0%	64.0%	58.1%	52.2%
Mother's occupational group					
Professional	3514	72.0%	72.7%	68.2%	62.2%
Managers	419	68.5%	70.2%	65.6%	58.2%
Clerical and personal service	3849	68.1%	72.3%	67.4%	59.7%
Trades	422	61.4%	68.7%	60.2%	53.1%
Plant operators and labourers	236	66.5%	68.2%	60.2%	52.5%
Unskilled manual	904	65.7%	67.0%	61.4%	52.3%

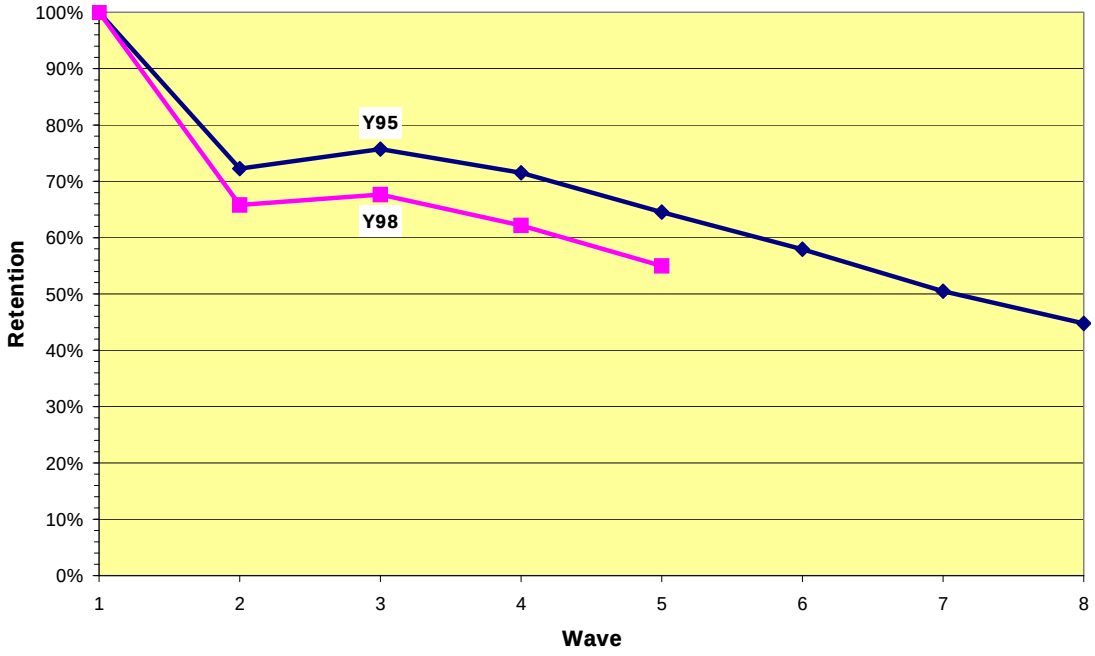

Figure 1 Annual retention of cohorts in LSAY samples, by wave

Figure 1 shows the overall retention of the original samples of the LSAY cohorts in each wave. The figure and tables indicate that the greatest attrition in both cohorts occurred at the second wave, when members were contacted with mail questionnaires. For the third wave, which was the first year of telephone interviews, the samples were rebuilt; attrition in the samples was highest in the first follow-up periods. Attrition is much higher among the Y98 cohort than the Y95 cohort. Further, young people were more likely to respond to the questionnaires while still at school; once they left school, however, some were more willing to respond and some were more difficult to locate.

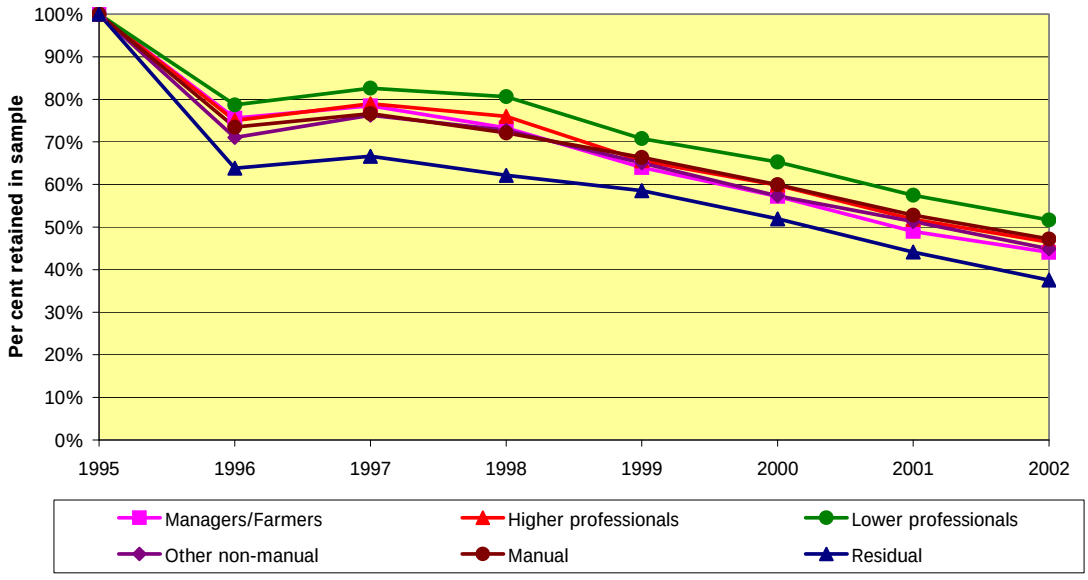

Figure 2 Annual retention of Y95 cohort, by father's occupational group

An example of differential attrition is evident in the rates by father's occupational group in the Y95 cohort. Two of these groups stand out in Figure 2: those whose fathers were in the Lower Professionals group, and those whose fathers were in the Residual group. For the other groups, there is little difference in the annual patterns of attrition, with some slight movement in the order among these groups. The Lower Professionals group has consistently higher retention in the sample, and the Residual group has consistently lower retention. For the Y98 cohort, different occupation groupings were used following changes to the Australian Standard Classification of Occupations. Nevertheless, patterns in differential attrition by father's occupation are similar to those seen for the Y95 cohort.

While these figures show that there is differential attrition among groups, not all members of a subgroup leave the survey (more than 100 Indigenous Australians remain in each cohort). Other events may occur and influence a cohort member to discontinue in the surveys. For example, young people in LSAY are not interviewed if they are overseas. Cohort members who miss the annual interview for any year—for any reason—are not included in any subsequent interviews.

An analysis of attrition should also include an understanding of the activities of both respondents and non-respondents in the last year of responses for the non-respondents to determine if annual events influence differences between the two groups. For the Y98 cohort, data were examined to determine the pre-attrition activities in years when most cohort members were still enrolled in school. Some of the differences that were found are listed in Table 3. This brief analysis suggests that non-respondents more often had left school without completing Year 12 or had plans to leave school before Year 12, and had no plans to attend university. Non-respondents' employment status varied by age; among those who had left school by 2000, fewer were in employment, but among those working, non-respondents were working longer hours and for a greater proportion of the year. Non-respondents in 2002 were more likely to have changed jobs between 2000 and 2001, and were more likely to be looking for work at the time of the 2001 interview.

Table 3 Differences between respondents and non-respondents in activities in the year before attrition among Y98 cohort members

2001 surveys and activity in 2000	2002 surveys and activity in 2001
More non-respondents had left school in 2000	More non-respondents had left school in 2001
Among those still at school, more non-respondents had changed schools since Year 9	
Among those still at school, more non-respondents were studying a TAFE subject	Among those still at school, more non-respondents were studying a TAFE subject
Among those still at school and in Year 11, fewer non-respondents planned to attend Year 12	not applicable
Among those planning to complete Year 12, fewer non-respondents planned to attend university	Among those in Year 12, fewer non-respondents planned to attend university
Among those who had left school by the end of Year 11, fewer non-respondents were working and more were looking for work as their main activity	Among those who left school during Year 12, more non-respondents were working and more were looking for work as their main activity
	Fewer non-respondents were doing further study or training
	More non-respondents had changed jobs since the previous interview
	More non-respondents were looking for work in the previous four weeks
More non-respondents moved out of their parents' home	More non-respondents moved out of their parents' home
Among those who had left school by the end of Year 11, non-respondents worked more hours per week and more weeks of the year	Hours worked about same; non-respondents worked for more weeks of the year and looked for work more weeks during the year

This information on non-respondents' activities—and how they differ from continuing participants' activities—is reflected in the information on all non-respondents shown in Table 1 and Table 2, as well as other information on annual non-respondents. LSAY research reports have described links between scores on the tests of achievement in reading comprehension and mathematics, which were administered to cohorts members when they entered the surveys in Year 9, and background factors, such as socioeconomic status (as determined by parent occupation), Indigenous status and language background. Figure 3 and Figure 4 show the mean achievement scores for respondents and non-respondents annually for the Y95 cohort. Figure 5 and Figure 6 show the same for the Y98 cohort. In Figure 3, for example, the mean reading score for all students was 49.8 in 1995. The following year, the mean for respondents was 50.9, and for non-respondents, 46.9. In 1997, after the sample was rebuilt, the means were 50.6 for respondents and 47.3 for non-respondents.

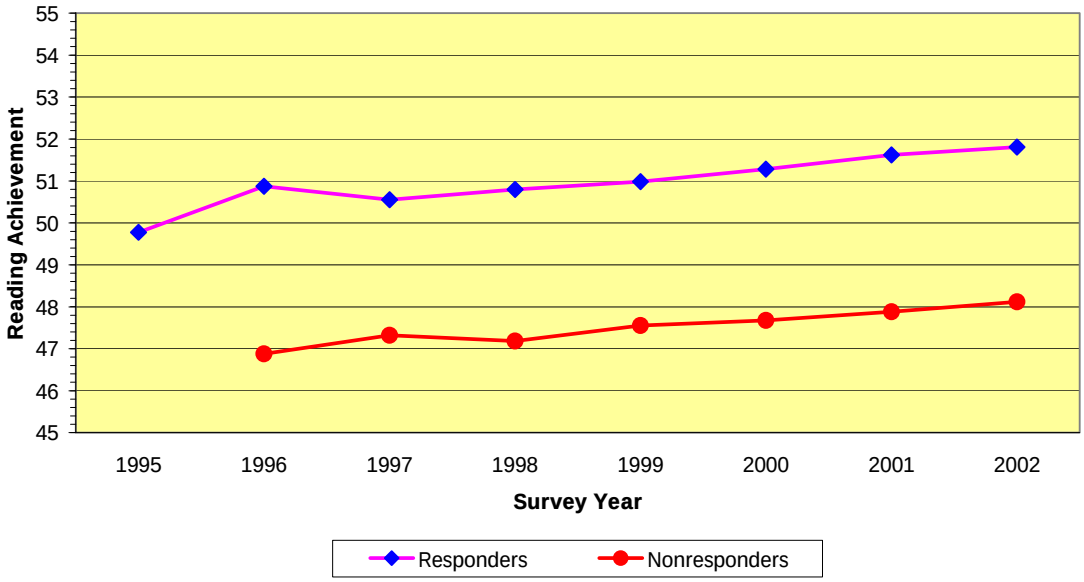

Figure 3 Mean scores on reading comprehension tests for respondents and non-respondents each year, Y95 cohort (unweighted)

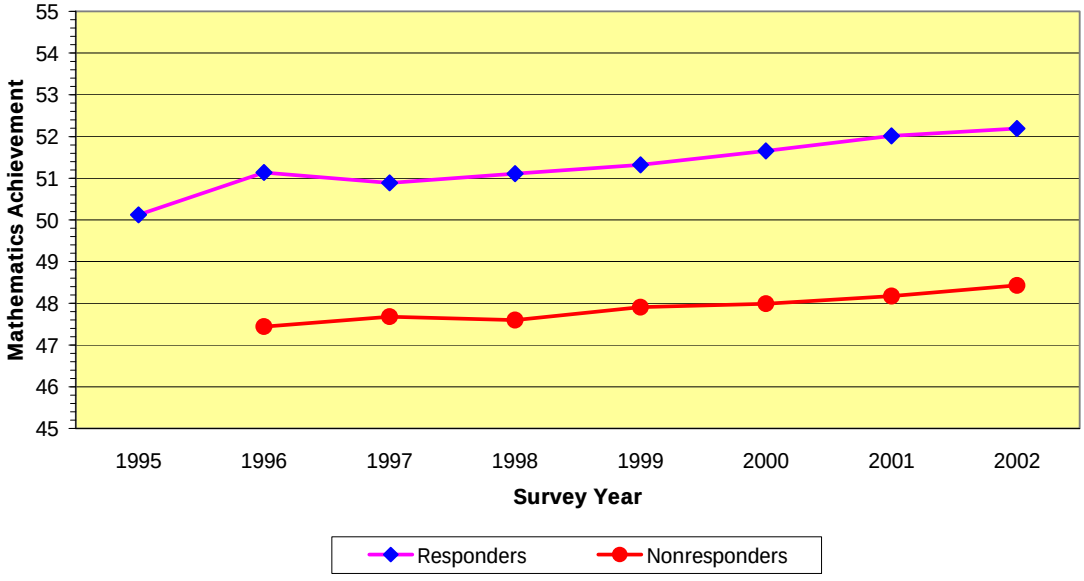

Figure 4 Mean scores on mathematics tests for respondents and non-respondents each year, Y95 cohort (unweighted)

The curve for non-respondents in the Y95 cohort suggests that lower achievers left the survey in the first few follow-up years, and that the achievement level of non-respondents increased in each subsequent year, so that the overall total mean achievement scores of non-respondents increased annually. The curve for non-respondents in the Y98 cohort is flat, indicating that achievement levels of non-respondents were similar from year to year. In the Y95 cohort, the difference between respondents and non-respondents has changed little, once attrition began. In the Y98 cohort, the difference between respondents and non-respondents has been increasing each year. For both cohorts, of course, as the lower achievers leave the survey, the mean achievement score increases for those remaining in the survey. That lower achievers are less likely to respond has already been recognised in LSAY, with the overall achievement quartile being used in the construction of annual attrition weights (see Marks & Long, 2000).

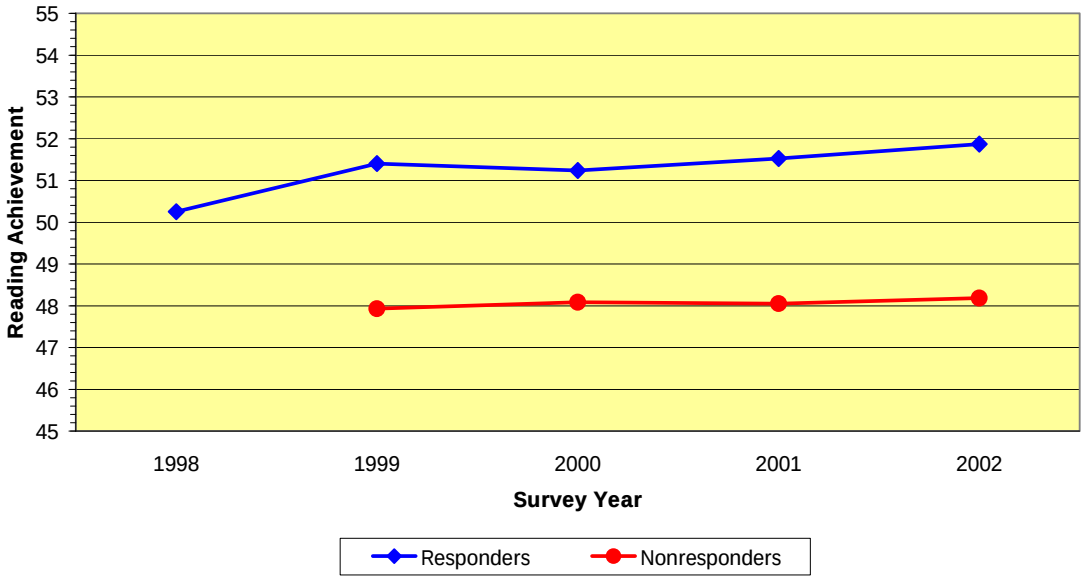

Figure 5 Mean scores on reading comprehension tests for respondents and non-respondents each year, Y98 cohort (unweighted)

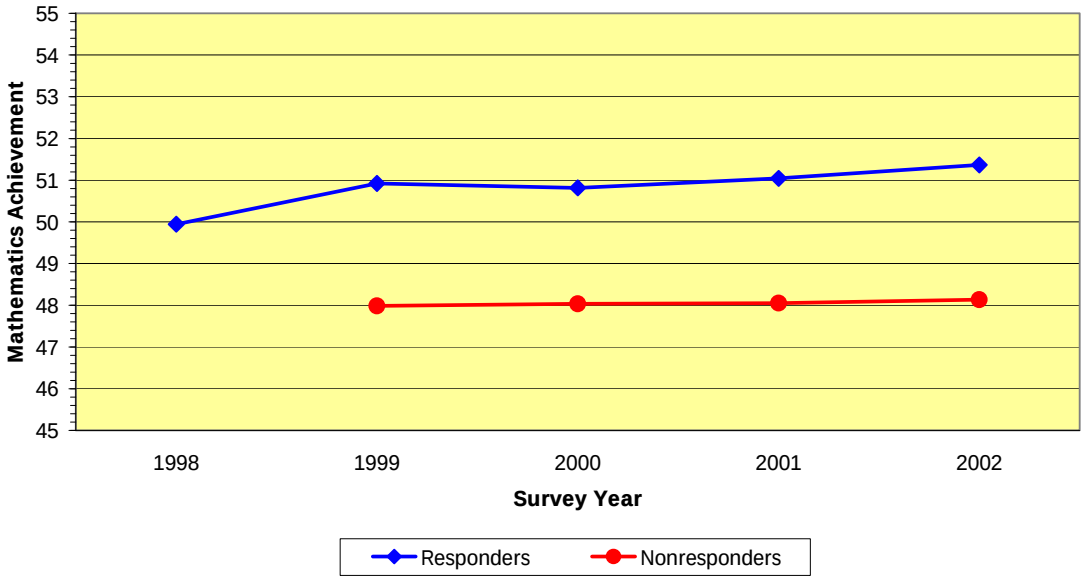

Figure 6 Mean scores on mathematics tests for respondents and non-respondents each year, Y98 cohort (unweighted)

Using the LSAY Cohorts to Represent the Annual Population

LSAY was designed as a longitudinal study that follows the transition from school of cohorts of young people. As such, the weighting schema used in LSAY results in an annual sample that is representative of the cohort as it was when first selected for LSAY. An alternative weighting schema could be developed so that remaining cohort members in any wave of the survey are assumed to represent a population of interest. For example, a researcher may be interested in using active members of the 1995 LSAY cohort to represent young people in the year 2000, because it is difficult to recruit a new representative sample. It would be possible to use the LSAY cohort to represent young people in a given year if an appropriate population could be identified.

It has been suggested that a weighting schema based on population data for Years 10, 11 and 12, which are available from the national schools census data published annually in *Schools Australia* (ABS catalogue no. 4221.0), could be used. This would account for young people from the cohort who are still attending school, with a minor adjustment for those who are not in the expected year level. There is not, however, a source of accurate population data for those who have left school.

There are no census data available to provide accurate counts of young people who are not attending school, particularly for the 1995 and 1998 LSAY cohorts, which are based on a year level at school. Other LSAY cohorts have been samples of young people of a specified age (for example, 14 year-olds), and there are reliable annual population estimates available for age-based cohorts. No such data exist for the grade-based cohorts used in LSAY in 1995 and 1998, as young people in Year 9 ranged in age from 12 to 18, with most split between ages 14 and 15.

Other factors preclude the use of external data to determine weights for the 1995 and 1998 LSAY cohorts in later waves. Data are not available to determine the rate of migration — inbound and outbound, domestic and overseas — by members of the population when the cohorts were selected. Deaths are also not considered in these calculations, although the death rate among the cohorts is generally small and would have a negligible statistical impact on any findings in LSAY.

4 EFFECTS OF ATTRITION ON THE 1995 AND 1998 YEAR 9 LSAY COHORTS

As noted above, the major difficulty in determining the effect of bias is that it is not possible to know how non-respondents would have responded. One approach to examining the effects of bias is to assign a score of zero, representing the lowest possible score on any item, to all non-respondents. Longitudinal surveys, however, tend to use categorical variables—in particular, dichotomous variables. In such cases, bias can be estimated by assuming that non-respondents have not attained a specific outcome. For example, if one is examining bias in a calculation of the proportion of young people who completed Year 12 or its vocational equivalent, it would be assumed that all non-respondents did not complete that level (see Cochran, 1977). As nonresponse increases and the size of the respondent group decreases, the standard error increases as well as the confidence interval around each estimate. Assigning a zero result to non-respondents, however, is an extreme approach, because the 'truth' is most likely between the zero option and the results reported for respondents.

Using sample attrition data for the Y98 LSAY cohort, Figure 7 demonstrates the effect of sample attrition on an item to which 50 per cent of cohort members respond positively. By the second wave—the mail questionnaire in 1999—34 per cent of the sample did not respond. In this situation, when 50 per cent of responses are positive, the lower limit of the 95 per cent confidence interval is 32 per cent, and the upper limit is 68 per cent. In 2000 (Wave 3), there was a slight reduction in attrition as a result of the sample being rebuilt, and only a slight narrowing of the confidence interval. By 2002 (Wave 5), the confidence interval ranges from 27 per cent to 73 per cent, which appears as a small change from Wave 2. For the Y95 cohort, the confidence interval in 2002 (Wave 8) ranges from 22 per cent to 78 per cent under a similar scenario.

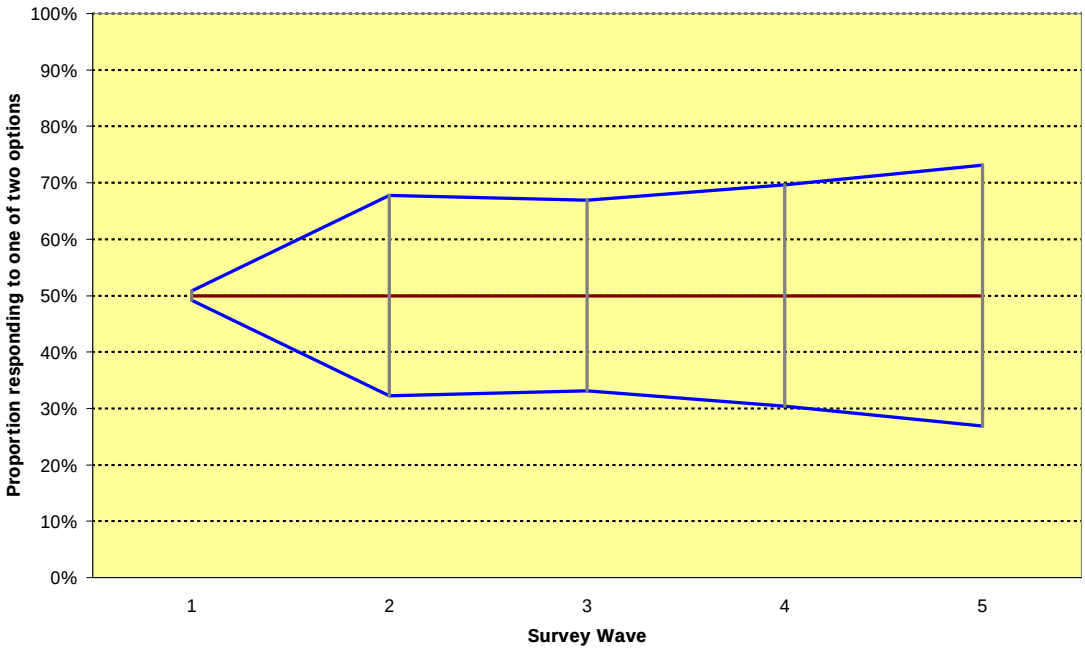

Figure 7 Estimate of confidence intervals when 50 per cent of the cohort respond positively to an item, based on attrition in the Y98 cohort, Waves 1 to 5

Estimates of Bias in the LSAY Cohorts

Charleston et al. (2003) and Wine et al. (2002) calculated the extent of bias caused by attrition of students participating in Baccalaureate and Beyond and the Beginning Postsecondary Students studies. They calculated bias twice, first using weights that were applied to the original sample, then using weights that were applied to those sampled in the year in question. This approach identifies how much bias is reduced by using nonresponse weights. For the present analysis, this approach was applied to the LSAY data.

Table 4 and Table 5 show the amount of bias in the Y95 cohort for distributions of cohort members by selected background characteristics, based on estimates of proportions. Table 4 calculates the amount of bias using the sampling weights applied in the original 1995 sample. Once nonresponse weights are applied for 1997 and each subsequent year, the bias is reduced, as shown in Table 5. In a small number of cases, the bias changes from negative to positive, or from positive to negative; however, in all cases, bias becomes less severe or remains nonsignificant. For the Y98 cohort, the effects of nonresponse weights on the overall bias are shown in Table 6; as a younger cohort, there are fewer time points. The numbers in the cells of these tables indicate the approximate shift to an estimated proportion caused by the calculated bias. The amount of bias for gender is approximately 1.7 (from Table 5; positive for females, negative for males). In estimating the gender distribution of young people in university study, for example, there would be small but significant bias of 1.7 percentage points, suggesting that the 'true' proportion of males in Year 9 in 1998 who later entered university study is up to 1.7 percentage points higher than estimates reported in LSAY.

In both LSAY cohorts, the bias caused by attrition is reduced for most subgroups of the total cohort after attrition weights are applied each year. In the Y95 cohort, gender (both males and females), home language (English and other), three States (Victoria, Queensland and South Australia), school sector (government and independent), father's occupational group (higher professionals, manual and residual), and mother's occupational group (lower professionals and residual) remain significantly biased after the application of post-stratification weights. In the Y98 cohort, gender, Indigenous background (Indigenous and double response), home language (other language and double response), State (Australian Capital Territory, South Australia, Tasmania and Northern Territory), school sector (government, Catholic and independent), and area (regional and rural/remote) remain significantly biased. Compared to the Y98 cohort at the end of Wave 5, there was no statistically significant bias after weighting in the Y95 cohort.

There is no single number that signifies a serious amount of bias, as the determination of significance is subject to the size of the sample.[5] In Table 6, for example, the most extreme level of bias in the Y98 cohort is –1.69, representing negative bias against young people who were in government schools in Year 9. There is less extreme, but also significant, positive bias in favour of young people who had attended independent schools in Year 9 (+0.58). For young people whose fathers were in professional occupations in 1998, the amount of bias was +0.73, greater than for independent schools but not significant.

The calculations in the tables above indicate that some subgroups in the LSAY samples have higher attrition rates and are subject to statistically significant bias. Table 5 and Table 6 also show, however, that significant bias does not occur uniformly across all groups, and it does not affect all subgroups within the same wave. Such differential changes support the idea that nonresponse is also related to activities that occur closer to the time of attrition, as examined in Table 3.

---

[5] Cochran (1977, p. 14) suggests that 10 per cent of one standard deviation is a reasonable cut-off for bias.

Table 4 Estimate of bias caused by attrition, before annual weighting for nonresponse, Y95 cohort

	1997	1998	1999	2000	2001	2002
Gender						
Male	-0.9432	-1.5861*	-2.0014*	-2.4938*	-3.0414*	-3.1619*
Female	0.9431	1.4439*	2.0006*	2.4939*	3.0378*	3.1599*
Indigenous background						
Indigenous	-0.9148*	-1.1335*	-1.2587*	-1.2503*	-1.3270*	-1.4408*
Non-indigenous	0.6053*	0.6990*	0.8755*	0.9209*	1.0366*	1.1501*
Double response	-0.0012	-0.0001	-0.0062	-0.0076	-0.0090	-0.0007
Home language						
English	0.9994*	1.2651*	1.3198*	1.1618*	1.5111*	1.6450*
Other	-1.0848*	-1.3200*	-1.2286*	-0.9842*	-1.2231*	-1.3602*
Double response	-0.1125	-0.1461	-0.1456	-0.1682	-0.2964*	-0.2716*
State (school)						
Australian Capital Terr.	-0.0428	-0.0092	-0.0472	-0.0022	0.0357	0.1593
New South Wales	-1.3768*	-1.7357*	-2.0269*	-2.1119*	-2.4610*	-2.7811*
Victoria	0.8121	0.9696	0.7786	1.0411*	1.5132*	1.6971*
Queensland	-0.2775	-0.2254	0.1557	-0.0662	-0.2329	-0.3598
South Australia	0.4972	0.7855*	1.0744*	1.1900*	1.3212*	1.4185*
Western Australia	0.2708	0.5279	0.5816	0.4792	0.3425	0.3508
Tasmania	-0.1619	-0.1595	-0.3810	-0.3276	-0.3021	-0.2671
Northern Territory	-0.2458	-0.1532	-0.1457	-0.1697	-0.1831	-0.1682
School sector						
Government	-1.9188*	-2.2549*	-2.0233*	-2.3428*	-2.5302*	-2.2044*
Catholic	0.7791	0.9221	1.0943*	1.2527*	1.1425*	0.7493
Independent	0.8476*	1.3328*	1.0951*	1.2661*	1.5646*	1.5924*
Area						
Metropolitan	-0.8949	-0.7354	-0.4199	0.0157	0.1224	-0.1323
Regional	0.4029	0.3008	0.1786	0.1515	0.0695	0.2455
Rural/remote	0.3852	0.4594	0.2427	-0.1741	-0.1937	-0.1159
Father's occupational group						
Managers/Farmers	0.9620*	1.1629*	0.9894*	1.1358*	0.9837*	1.2398*
Higher professionals	0.7337*	0.8223*	1.0490*	1.3709*	1.6995*	1.9120*
Lower professionals	0.6039*	0.7100*	0.9738*	1.2254*	1.4388*	1.6020*
Other non-manual	0.1329	0.1316	0.1307	-0.1629	0.1886	0.0752
Manual	-0.0051	-0.3172	-0.1656	-0.3536	-0.3929	-0.3157
Residual	-3.9925*	-4.0803*	-3.7318*	-4.0032*	-4.6918*	-5.3951*
Mother's occupational group						
Managers/Farmers	0.1559	0.2534	0.0235	0.1005	0.1072	0.0252
Higher professionals	0.0924	0.1085	0.1827	0.1776	0.2294	0.2434
Lower professionals	1.2672*	1.5247*	1.7962*	2.3460*	2.8219*	3.2232*
Other non-manual	1.1354*	1.1219*	1.6033*	1.6735*	1.7759*	1.8870*
Manual	0.0233	0.0539	0.0948	-0.0157	0.1781	0.2582
Residual	-3.6880*	-4.3115*	-4.4653*	-5.0855*	-5.8987*	-6.4288*

* Significant at $\alpha = .05$.

Table 5 Estimate of bias caused by attrition, after annual weighting for nonresponse, Y95 cohort

	1997	1998	1999	2000	2001	2002
Gender						
Male	-0.2200	-0.1624	-0.6999	-1.0387	-1.4818*	-1.7324*
Female	0.2200	0.1479	0.6997	1.0387	1.4801*	1.7313*
Indigenous background						
Indigenous	-0.0517	-0.0182	-0.1160	-0.1356	-0.2163	-0.1871
Non-indigenous	0.0356	0.0122	0.0857	0.1070	0.1796	0.1611
Double response	0.0003	0.0000	0.0025	0.0030	0.0034	-0.0041
Home language						
English	0.0794	0.1721	0.3763	0.5254	0.7932*	1.0172*
Other	-0.0912	-0.1924	-0.3869	-0.4985	-0.7413	-0.9355*
Double response	-0.0037	-0.0057	0.0024	-0.0121	-0.0339	-0.0597
State (school)						
Australian Capital Terr.	0.0122	-0.0026	0.0139	0.0396	0.0734	0.1589
New South Wales	0.0334	-0.4829	-0.4893	-0.6030	-0.7995	-1.0973
Victoria	0.0353	0.2697	0.4344	0.6370	0.9719	1.1869*
Queensland	-0.0675	-0.0627	-0.3752	-0.5857	-0.8630	-1.0696*
South Australia	0.0060	0.2185	0.4293	0.5659	0.7349*	0.8989*
Western Australia	-0.0029	0.1468	0.2650	0.2773	0.2605	0.3111
Tasmania	-0.0179	-0.0444	-0.1459	-0.1615	-0.1817	-0.1884
Northern Territory	0.0001	-0.0426	-0.0480	-0.0659	-0.0820	-0.0815
School sector						
Government	-0.1621	-0.6273	-1.0182	-1.4161*	-1.8570*	-1.9265*
Catholic	0.0395	0.2565	0.4770	0.6477	0.7325	0.6129
Independent	0.0975	0.3708	0.6298	0.8817*	1.2603*	1.4355*
Area						
Metropolitan	0.0426	-0.2740	-0.1456	0.0325	0.1113	0.0188
Regional	-0.0102	0.0516	-0.0106	-0.0913	-0.0807	0.0040
Rural/remote	-0.0271	0.2351	0.1593	0.0621	-0.0297	-0.0229
Father's occupational group						
Managers/Farmers	0.0767	0.1984	0.4043	0.5787	0.7771	0.8437
Higher professionals	0.0893	0.0931	0.3176	0.4522	0.7205*	0.8699*
Lower professionals	0.0357	0.0342	0.1540	0.2198	0.3179	0.3905
Other non-manual	0.0077	-0.0250	0.0054	-0.0031	-0.0086	0.0210
Manual	-0.0955	-0.1661	-0.3743	-0.6108	-0.9072	-1.1349*
Residual	-0.2033	-0.2341	-0.6110	-0.7582	-1.0210*	-1.1425*
Mother's occupational group						
Managers/Farmers	0.0265	0.0457	0.1068	0.1296	0.1600	0.1721
Higher professionals	0.0176	0.0109	0.0469	0.0861	0.1626	0.1713
Lower professionals	0.1090	0.1450	0.4596	0.6126	0.8412	1.0308*
Other non-manual	0.0478	-0.0018	0.1225	0.1626	0.3137	0.4166
Manual	-0.0428	-0.0865	-0.2173	-0.3134	-0.4090	-0.4787
Residual	-0.2216	-0.1701	-0.5798	-0.7738	-1.1854*	-1.4517*

* Significant at $\alpha = .05$.

Table 6 Estimate of bias caused by attrition, before and after annual weighting for nonresponse, Y98 cohort

	Before nonresponse adjustment			After nonresponse adjustment		
	2000	2001	2002	2000	2001	2002
Gender						
Male	-1.8671*	-2.6733*	-3.1318*	0.1314	-0.1351	-0.3639
Female	2.2712*	2.8191*	2.8153*	-0.1105	0.1174	0.3271
Indigenous background						
Indigenous	-2.0795*	-1.9571*	-1.7306*	-0.0582	-0.0973	-0.1513*
Non-indigenous	1.5830*	1.4947*	1.3165*	0.0362	0.0640	0.1092
Double response	-0.1998*	-0.1989*	-0.1643*	-0.0019	-0.0015	-0.0014*
Home language						
English	1.6851*	1.7241*	1.6336*	0.2994	0.3955	0.5516
Other	-1.3164*	-1.3184*	-1.2727*	-0.2514	-0.3389	-0.4716*
Double response	-0.4574*	-0.4989*	-0.4471*	-0.0622	-0.0799	-0.1054*
State (school)						
Australian Capital Terr.	0.2590	0.2601	0.1441*	0.0526	0.0782	0.0646*
New South Wales	-0.3745	-0.0837	-0.3107	-0.0597	0.0131	-0.0776
Victoria	-0.1007	-0.2407	-0.0390	-0.0366	-0.1045	-0.0709
Queensland	0.0798	-0.1946	-0.1999*	0.0267	-0.0502	-0.0746
South Australia	0.0694	0.0846	0.1966*	0.0127	0.0206	0.0818*
Western Australia	-0.0380	-0.0372	-0.1321*	-0.0169	-0.0226	-0.0745
Tasmania	0.2528	0.3327	0.4307*	0.0557	0.1051	0.1943*
Northern Territory	-0.1467	-0.1182	-0.0890*	-0.0348	-0.0388	-0.0420*
School sector						
Government	-3.1076*	-3.6256*	-3.3466*	-0.7681	-1.2364*	-1.6947*
Catholic	2.5468*	2.8935*	2.4402*	0.6086	0.9464*	1.1552*
Independent	0.6300	0.8141*	0.9866*	0.1742	0.3138	0.5816*
Area						
Metropolitan	-2.3470*	-1.7233*	-1.4631*	-0.4505	-0.4631	-0.6045
Regional	0.8469	0.6031	0.6663*	0.1600	0.1652	0.2553*
Rural/remote	1.4933*	1.1152*	0.7919*	0.2828	0.2923	0.3473*
Father's occupational group						
Professional	1.6501*	1.7084*	2.1266*	0.2434	0.4801	0.7328
Managers	0.2522	0.5039	0.4026	0.1702	0.2388	0.3712
Clerical/personal service	0.6345	0.5329	0.2866	-0.0156	-0.0271	-0.0378
Trades	-0.3273	-0.4710	-0.7393	-0.1651	-0.3024	-0.4584
Plant ops/labourers	-1.2540*	-1.3120*	-1.4116*	-0.1411	-0.2303	-0.3618
Unskilled manual	-1.0910*	-1.0825*	-0.7701*	-0.0999	-0.1638	-0.2545
Mother's occupational group						
Professional	0.8606	1.2927	1.9442*	0.3136	0.5055	0.7654
Managers	-0.1904	-0.1621	-0.1621	0.0314	0.0439	0.0834
Clerical/personal service	0.4868	0.5443	0.0703	-0.1301	-0.2232	-0.3908
Trades	-0.1901	-0.4828	-0.4256	-0.0166	-0.0212	-0.0209
Plant ops/labourers	-0.1695	-0.2506	-0.2088	-0.0237	-0.0448	-0.0727
Unskilled manual	-0.8967	-1.0769*	-1.3441*	-0.1569	-0.2417	-0.3637

* Significant at α = .05.

Another method to examine the effect of attrition is to calculate the bias associated with a variable that has an influence on other outcomes. Many LSAY reports have already highlighted the relationship between achievement in literacy and numeracy, as determined by the reading comprehension and mathematics tests administered in Year 9, and a number of outcomes, including Year 12 participation and university attendance. Bias is calculated as the change to estimates caused by attrition. In each follow-up survey for both cohorts, for example, bias increases positively, and more rapidly for the Y98 cohort. Models that estimate outcomes based on Year 9 reading comprehension or mathematics scores may be using biased estimates of achievement.

By using the combined literacy and numeracy quartile and sex as the basis for annual attrition weights, this bias is adequately reduced, as shown in Table 7.[6] Before attrition weights were applied to the Y95 data for 2002, reading comprehension scores were biased by 2.03 scaled score points. After the weights were applied, bias was reduced to –0.03, less than one per cent of a standard deviation. While there is increasing bias before weighting each year, this bias does not affect the general relationship between these test scores and predicted outcomes, and it is adequately adjusted by the attrition weights calculated after each year's collection of data. For the Y98 cohort, bias is greater at an earlier wave. For each follow-up wave, bias in reading comprehension ranges from –0.37 to –0.26 and in mathematics, from –0.26 to –0.21. At no time has the bias in the Y95 cohort reached these magnitudes.

Table 7 Estimate of bias in reading comprehension and mathematics test scores as a result of attrition, Y95 and Y98 cohorts

Y95 cohort		Before attrition weights		After attrition weights	
Wave	Year	Reading Comprehension	Mathematics	Reading Comprehension	Mathematics
2	1996	1.0963	1.0150	-0.0341	-0.1317
3	1997	0.7745	0.7666	-0.0503	-0.0985
4	1998	1.0195	0.9863	0.2121	0.1124
5	1999	1.2084	1.1984	-0.0156	-0.0526
6	2000	1.5084	1.5283	-0.0159	-0.0530
7	2001	1.8423	1.8884	-0.0214	-0.1434
8	2002	2.0303	2.0701	-0.0311	-0.1502

Y98 cohort		Before attrition weights		After attrition weights	
Wave	Year	Reading Comprehension	Mathematics	Reading Comprehension	Mathematics
2	1999	1.1544	0.9759	-0.2891	-0.2271
3	2000	0.9851	0.8706	-0.2623	-0.2074
4	2001	1.2743	1.1013	-0.3338	-0.2389
5	2002	1.6156	1.4214	-0.3654	-0.2630

[6] Similar reductions in bias by the use of attrition weights were noted in Marks and Long (2000).

5 SUMMARY AND DISCUSSION

Nonresponse is a major source of bias in sample surveys. In cross-sectional surveys, sample members randomly selected by researchers may choose not to respond to the interviewer's questions. If such nonresponse occurs among one group of the designed sample, bias has the potential to affect the results of the study. Bias due to nonresponse is compounded in longitudinal surveys, because participants may respond in the first year, then fail to respond in subsequent years. In LSAY, attrition increases throughout the life of the cohorts and is not random across the samples. Attrition is greater among some groups of young people in LSAY. While some factors are more highly associated with attrition, they share one common factor to a large extent: lower scores on the achievement tests taken in Year 9. At some point, the bias caused by attrition becomes too great and cannot be overcome by adjustments or replacement.

A review of practices to ameliorate the negative effects of attrition and associated bias suggests that the application of attrition weights, in addition to post-stratification weights that are applied after responses to each wave of the survey are received, can provide unbiased estimates of outcomes and relationships, although with slightly reduced precision. It is not possible to determine exactly the accuracy of bias reduction practices, because how non-respondents may have responded to any item is not knowable; however, it is possible to estimate how much bias is reduced. Weighting has also been found to be the most cost-effective of a number of alternative approaches available to ameliorate the effects of attrition.

Significance of Attrition in LSAY

Attrition in the LSAY cohorts is greater than that reported for other education-related longitudinal surveys, especially those conducted in the United States by the National Center for Education Statistics. It should be noted that those studies track participants through social security numbers, which is required information on university enrolment, student loan and income tax records. The LSAY cohorts experience nonrandom attrition, with some groups contributing to nonresponse more than other groups. In addition, attrition in Y98, the younger cohort, is increasing at a faster rate than it is in Y95.

In LSAY, one group experiences much greater attrition, and therefore is subject to greater problems of bias. Of the Indigenous Australians in the original samples, 26 per cent remained in the Y95 cohort after the eighth wave, and 36 per cent remained in the Y98 cohort after the fifth wave. These attrition rates are higher than for any other subgroup in the samples. Attrition of Indigenous Australians from longitudinal surveys has been noted elsewhere (Hunter & Smith, 2000), with the recognition that this group may require alternative methods of data collection. Because of these higher rates of attrition, data on Indigenous Australians in the LSAY Y95 and Y98 cohorts must be used with caution.

Effects of Weighting in LSAY

The attrition weights used in LSAY are calculated according to sex and four levels of combined scores on school achievement tests administered in Year 9. Attrition weights in LSAY differ from those used in other studies in two major respects. First, LSAY uses two variables to determine attrition weights; other studies tend to use more variables, based on their association with a number of different outcome variables of interest. Second, LSAY assumes that all types of attrition are similar, whether due to noncontact, refusal or other nonresponse; studies of attrition conducted by NCES show that different types of nonresponse are associated with different background characteristics (for example, see Wine et al., 2002). In this paper and in an earlier LSAY technical paper (Marks & Long, 2000), these weights as currently applied were shown to be adequate for reducing bias caused by attrition in the LSAY samples.

Research on sample attrition conducted for other longitudinal surveys has shown that there is no 'magic number' to indicate when a cohort becomes too small because of attrition bias, especially when attrition weights are applied. Problems do arise, however, when attrition occurs for some groups more than others, especially if those groups with higher attrition were smaller in the initial sample, as is the case for Indigenous Australians in the LSAY cohorts. For these smaller groups, high attrition means high standard errors of estimates, compromising the validity of findings for those groups. For the full cohort, however, standard errors remain reasonable for as long as the overall number of respondents has not fallen dramatically. Based on random samples selected from the Y95 cohort, standard errors are reasonable even when the remaining sample is as low as 25 per cent of the original sample (see Table 8). With a 10 per cent random sample, however, the estimates of mean scores in reading comprehension and mathematics have much larger standard errors than when first collected.

Table 8 Means, standard errors and 95% confidence intervals for random samples of LSAY Y95 cohort members on reading comprehension and mathematics tests

Reading Comprehension	n	Mean	se (mean)	CI (lower)	CI (upper)
100% sample	13367	49.78	0.0856	49.61	49.94
75% sample	10036	49.72	0.0991	49.53	49.92
60% sample	7975	49.86	0.1108	49.64	50.07
40% sample	5312	49.84	0.1358	49.57	50.11
25% sample	3327	49.90	0.1709	49.56	50.23
10% sample	1329	50.12	0.2763	49.57	50.66
Mathematics	**n**	**Mean**	**se (mean)**	**CI (lower)**	**CI (upper)**
100% sample	13356	50.12	0.0860	49.96	50.29
75% sample	10044	50.09	0.0991	49.89	50.28
60% sample	7964	50.12	0.1119	49.90	50.34
40% sample	5308	50.11	0.1357	49.84	50.37
25% sample	3316	50.22	0.1725	49.88	50.56
10% sample	1329	50.74	0.2659	50.22	51.26

It may appear that a sample that has lost more than one-half of its members has a problem of providing valid results. As demonstrated in this technical paper, however, the LSAY samples are robust and can still retain strong statistical properties, even down to 25 per cent of the original sample. This is because the LSAY samples were large when cohort members were first contacted, and because the weighting procedures reduce potential bias. Nevertheless, for some subgroups, such as Indigenous Australians, the size of the remaining sample can lead to unstable estimates of activities, which must be used with caution. For the full cohort, a sample size that is less than 25 per cent of the original sample size will also provide unstable estimates and therefore should be used with extreme caution.

Advice Regarding the Use of Weights with LSAY Data

If one LSAY cohort were to be used to represent a group of young Australians from a time other than when the cohort was selected, it should be acknowledged that the LSAY cohort will not be representative of that different population. For example, the population represented by the Y95 cohort at the time of selection was all young people in Year 9 at school in 1995, who were generally around 14 years old. Immigration and emigration are the most common reasons for differences, and internal migration—particularly interstate migration—affecting differences in the cohort. Researchers using the sample to represent 20 year-olds in 2001 should be aware that the population of interest has most likely changed, and they should re-weight the sample according to the 2001 population of 20 year-olds.

Researchers using the LSAY data sets can be confident that the post-stratification and attrition weights ensure that the remaining sample members represent the population from which they were selected. If comparisons are to be made with concurrent data—such as a comparison between Year 12 completion rates in LSAY and apparent retention rates published by the Australian Bureau of Statistics—then researchers are advised to explain how the results differ and why they are not expected to be the same. Such an approach recognises the design of LSAY and the limitations of using LSAY for point-in-time estimates.

REFERENCES

Charleston, S., Riccobono, J., Mosquin, P., & Link, M. (2003). *Baccalaureate and beyond longitudinal study: 2000–01 (B&B: 2000/01) methodology report* (No. 2003-156). Washington, DC: National Center for Educational Statistics.

Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). New York: John Wiley & Sons.

Fitzgerald, J., Gottschalk, P., & Moffitt, R. (1998a). An analysis of sample attrition in panel data: the Michigan Panel Study of Income Dynamics. *The Journal of Human Resources, 33*(2), 251-299.

Fitzgerald, J., Gottschalk, P., & Moffitt, R. (1998b). The impact of attrition in the Panel Study of Income Dynamics on intergenerational analysis. *The Journal of Human Resources, 33*(2), 300-344.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica, 47*(1), 153-161.

Hunter, B. H., & Smith, D. E. (2000). *Surveying mobile populations: lessons from recent longitudinal surveys of Indigenous Australians* (No. 203/2000). Canberra: Centre for Aboriginal Economic Policy Research, Australian National University.

Lilliard, L. A., & Panis, C. W. A. (1998). Panel attrition from the Panel Study of Income Dynamics: household income, marital status and mortality. *The Journal of Human Resources, 33*(2), 437-457.

Long, M. (1996). *Samples and sampling for the Y95 LSAY cohort* (Technical Paper No. 8). Melbourne: Australian Council for Educational Research.

Long, M., & Fleming, N. (2002). *The designed and achieved sample of the 1998 LSAY sample* (Technical Paper No. 16). Melbourne: Australian Council for Educational Research.

Marks, G. N., & Long, M. (2000). *Weighting the 1995 Year 9 cohort sample for differential response rates and sample attrition* (Technical Paper No. 15). Melbourne: Australian Council for Educational Research.

McGuigan, K. A., Ellickson, P. L., Hays, R. D., & Bell, R. M. (1997). Adjusting for attrition in school-based samples: bias, precision, and cost trade-offs of three methods. *Evaluation Review, 21*(5), 554-567.

Wine, J. S., Heuer, R. E., Wheeless, S. C., Francis, T. L., Franklin, J. W., & Dudley, K. M. (2002). *Beginning postsecondary students longitudinal study: 1996-2001 (BPS:1996/2001) methodology report* (No. 2002-171). Washington, DC: National Center for Educational Statistics.

Zabel, J. E. (1998). An analysis of attrition in the Panel Study of Income Dynamics and the Survey of Income and Program Participation with an application to a model of labor market behavior. *The Journal of Human Resources, 33*(2), 479-506.

Zahs, D., Pedlow, S., Morrissey, M., Marnell, P., & Nichols, B. (1995). *High school and beyond fourth follow-up methodology report* (No. 95426). Washington, DC: National Center for Education Statistics.

APPENDIX

Table A 1 Unweighted numbers of Y95 sample members interviewed from 1995 to 2002, by selected characteristics in 1995

	1995	1996	1997	1998	1999	2000	2001	2002
All	13613	9837	10307	9738	8783	7889	6876	6095
Gender								
Male	6717	4475	5001	4679	4174	3718	3201	2828
Female	6896	5362	5306	5059	4609	4171	3675	3267
Indigenous background								
Indigenous	385	201	228	204	172	147	121	100
Non-indigenous	12348	9093	9477	8962	8095	7282	6380	5673
Double response	4	2	3	3	3	3	3	2
Home language								
English	11687	8601	9005	8526	7701	6918	6055	5377
Other language	1305	878	901	834	748	681	581	506
Double response	116	77	84	78	68	60	47	42
State (school)								
Australian Capital Territory	599	456	437	415	368	335	296	276
New South Wales	3090	2102	2264	2100	1877	1686	1447	1263
Victoria	2865	2097	2231	2104	1874	1701	1510	1343
Queensland	2524	1814	1865	1761	1621	1433	1241	1090
South Australia	1720	1361	1407	1349	1250	1140	1008	903
Western Australia	1837	1321	1409	1356	1224	1093	942	829
Tasmania	582	446	436	412	349	312	271	246
Northern Territory	396	240	258	241	220	189	161	145
School sector								
Government	9081	6373	6712	6294	5698	5090	4420	3928
Catholic	2517	1879	1990	1885	1706	1542	1342	1171
Independent	2015	1585	1605	1559	1379	1257	1114	996
Area								
Metropolitan	7564	5442	5645	5344	4834	4386	3840	3397
Regional	3378	2419	2603	2448	2201	1968	1701	1514
Rural/remote	2629	1976	2057	1945	1747	1534	1334	1184
Father's occupational group								
Managers/Farmers	2599	1995	2063	1963	1755	1592	1383	1241
Higher professionals	1282	1024	1049	1004	914	848	755	682
Lower professionals	906	734	769	742	681	641	574	522
Other non-manual	1938	1409	1485	1407	1268	1114	993	880
Manual	4468	3205	3382	3166	2863	2546	2215	1954
Residual	2420	1470	1559	1456	1302	1148	956	816
Mother's occupational group								
Managers/Farmers	808	593	631	604	524	480	415	361
Higher professionals	328	251	255	244	227	205	180	162
Lower professionals	2185	1758	1819	1750	1589	1475	1321	1201
Other non-manual	3614	2685	2847	2693	2463	2207	1925	1708
Manual	1500	1085	1126	1067	967	857	753	675
Residual	5178	3465	3629	3380	3013	2665	2282	1988

Table A 2 Weighted numbers of Y95 sample members interviewed from 1995 to 2002, by selected characteristics in 1995

	1995	1996	1997	1998	1999	2000	2001	2002
All	13613	9837	10307	9738	8783	7889	6876	6095
Gender								
Male	6653	4814	5038	4669	4292	3856	3359	2978
Female	6960	5023	5269	5069	4491	4033	3517	3117
Indigenous background								
Indigenous	373	220	241	208	188	166	145	118
Non-indigenous	12359	9047	9435	8959	8051	7234	6330	5630
Double response	3	2	3	3	4	4	4	3
Home language								
English	11539	8431	8843	8371	7506	6726	5866	5189
Other language	1418	1015	1023	971	911	845	742	660
Double response	135	86	95	88	77	70	56	52
State (school)								
Australian Capital Territory	267	201	194	191	165	147	127	112
New South Wales	4557	3110	3316	3260	2887	2590	2246	1993
Victoria	3311	2449	2579	2368	2097	1882	1641	1455
Queensland	2499	1860	1893	1788	1736	1566	1377	1223
South Australia	1033	811	845	739	654	586	510	450
Western Australia	1439	1038	1120	1029	914	820	714	631
Tasmania	398	305	293	284	260	235	206	183
Northern Territory	110	64	68	79	70	63	54	48
School sector								
Government	9143	6535	6807	6540	5977	5383	4705	4173
Catholic	2744	2013	2145	1963	1748	1567	1363	1207
Independent	1726	1289	1355	1235	1058	939	808	715
Area								
Metropolitan	7489	5318	5588	5401	4848	4350	3788	3355
Regional	3244	2372	2512	2340	2119	1916	1660	1472
Rural/remote	2831	2147	2205	1996	1815	1621	1426	1268
Father's occupational group								
Managers/Farmers	2583	1951	2026	1893	1652	1477	1265	1139
Higher professionals	1239	930	978	941	812	741	642	576
Lower professionals	902	710	746	728	639	589	519	466
Other non-manual	1873	1331	1428	1362	1219	1074	960	841
Manual	4536	3333	3476	3272	3009	2720	2395	2142
Residual	2480	1583	1653	1542	1452	1288	1094	932
Mother's occupational group								
Managers/Farmers	784	567	598	569	483	438	381	334
Higher professionals	307	220	234	227	203	176	148	133
Lower professionals	2140	1649	1718	1649	1425	1315	1163	1047
Other non-manual	3591	2624	2824	2691	2437	2189	1895	1679
Manual	1477	1106	1140	1095	1025	920	823	737
Residual	5314	3670	3794	3507	3210	2851	2464	2164

Table A 3 Unweighted numbers of Y98 sample members interviewed from 1998 to 2002, by selected characteristics in 1998

	1998	1999	2000	2001	2002
All	14117	9289	9548	8777	7762
Gender					
Male	7227	4407	4753	4305	3776
Female	6804	4846	4770	4449	3967
Indigenous background					
Indigenous	442	199	212	188	160
Non-indigenous	12917	8702	8954	8252	7315
Double response	11	4	2	1	1
Home language					
English	12078	8140	8383	7732	6857
Other language	1191	713	729	661	573
Double response	253	144	152	133	114
State (school)					
Australian Capital Territory	558	365	414	387	333
New South Wales	3384	2160	2270	2092	1830
Victoria	2950	1854	1972	1800	1609
Queensland	3111	2108	2108	1921	1691
South Australia	1249	871	866	800	717
Western Australia	1689	1107	1138	1045	910
Tasmania	715	552	510	486	451
Northern Territory	461	272	270	246	221
School sector					
Government	8887	5722	5781	5244	4604
Catholic	3122	2088	2296	2162	1917
Independent	2108	1479	1471	1371	1241
Area					
Metropolitan	7763	5007	5198	4804	4258
Regional	3169	2185	2211	2020	1787
Rural/remote	2474	1761	1797	1643	1445
Father's occupational group					
Professional	3740	2639	2730	2533	2298
Managers	1836	1300	1318	1233	1093
Clerical and personal service	1040	721	760	708	621
Trades	2538	1654	1776	1633	1433
Plant operators and labourers	1368	890	897	817	705
Unskilled manual	726	443	465	422	379
Mother's occupational group					
Professional	3514	2529	2555	2397	2186
Managers	419	287	294	275	244
Clerical and personal service	3849	2622	2784	2595	2296
Trades	422	259	290	254	224
Plant operators and labourers	236	157	161	142	124
Unskilled manual	904	594	606	555	473

Table A 4 Weighted numbers of Y98 sample members interviewed from 1998 to 2002, by selected characteristics in 1998

	1998	1999	2000	2001	2002
All	14117	9289	9548	8777	7762
Gender					
Male	7208	4532	4695	4309	3808
Female	6829	4721	4826	4443	3933
Indigenous background					
Indigenous	447	209	221	198	175
Non-indigenous	12926	8695	8937	8230	7293
Double response	11	4	2	0	0
Home language					
English	12024	8016	8242	7593	6720
Other language	1235	809	835	763	674
Double response	262	157	169	150	132
State (school)					
Australian Capital Territory	272	179	185	170	150
New South Wales	4627	3032	3122	2868	2534
Victoria	3311	2186	2243	2065	1830
Queensland	2833	1866	1914	1759	1555
South Australia	1075	710	728	670	592
Western Australia	1490	981	1010	929	822
Tasmania	388	255	262	241	213
Northern Territory	121	80	82	76	67
School sector					
Government	9373	6184	6359	5852	5190
Catholic	2879	1893	1937	1780	1571
Independent	1865	1212	1252	1145	1001
Area					
Metropolitan	7491	4985	5110	4700	4165
Regional	3368	2238	2312	2124	1881
Rural/remote	2841	1905	1955	1797	1581
Father's occupational group					
Professional	3604	2472	2553	2336	2095
Managers	1865	1264	1280	1188	1043
Clerical and personal service	1032	718	764	707	619
Trades	2544	1691	1815	1686	1488
Plant operators and labourers	1460	992	999	919	809
Unskilled manual	758	481	500	461	427
Mother's occupational group					
Professional	3375	2313	2360	2188	1968
Managers	406	260	274	253	220
Clerical and personal service	3826	2635	2763	2578	2287
Trades	420	254	294	259	228
Plant operators and labourers	246	168	173	158	142
Unskilled manual	932	642	660	606	523

Table A 5 Annual attrition rates for Y95 cohort telephone interviews, 1997-2002, by selected characteristics in 1995 (unweighted)

	1997	1998	1999	2000	2001	2002
All	24.3%	5.5%	9.8%	10.2%	12.8%	11.4%
Gender						
Male	25.5%	6.4%	10.8%	10.9%	13.9%	11.7%
Female	23.1%	4.7%	8.9%	9.5%	11.9%	11.1%
Indigenous background						
Indigenous	40.8%	10.5%	15.7%	14.5%	17.7%	17.4%
Non-indigenous	23.3%	5.4%	9.7%	10.0%	12.4%	11.1%
Double response	25.0%	0.0%	0.0%	0.0%	0.0%	33.3%
Home language						
English	22.9%	5.3%	9.7%	10.2%	12.5%	11.2%
Other language	31.0%	7.4%	10.3%	9.0%	14.7%	12.9%
Double response	27.6%	7.1%	12.8%	11.8%	21.7%	10.6%
State (school)						
Australian Capital Territory	27.0%	5.0%	11.3%	9.0%	11.6%	6.8%
New South Wales	26.7%	7.2%	10.6%	10.2%	14.2%	12.7%
Victoria	22.1%	5.7%	10.9%	9.2%	11.2%	11.1%
Queensland	26.1%	5.6%	8.0%	11.6%	13.4%	12.2%
South Australia	18.2%	4.1%	7.3%	8.8%	11.6%	10.4%
Western Australia	23.3%	3.8%	9.7%	10.7%	13.8%	12.0%
Tasmania	25.1%	5.5%	15.3%	10.6%	13.1%	9.2%
Northern Territory	34.8%	6.6%	8.7%	14.1%	14.8%	9.9%
School sector						
Government	26.1%	6.2%	9.5%	10.7%	13.2%	11.1%
Catholic	20.9%	5.3%	9.5%	9.6%	13.0%	12.7%
Independent	20.3%	2.9%	11.5%	8.8%	11.4%	10.6%
Area						
Metropolitan	25.4%	5.3%	9.5%	9.3%	12.4%	11.5%
Regional	22.9%	6.0%	10.1%	10.6%	13.6%	11.0%
Rural/remote	21.8%	5.4%	10.2%	12.2%	13.0%	11.2%
Father's occupational group						
Managers/Farmers	20.6%	4.8%	10.6%	9.3%	13.1%	10.3%
Higher professionals	18.2%	4.3%	9.0%	7.2%	11.0%	9.7%
Lower professionals	15.1%	3.5%	8.2%	5.9%	10.5%	9.1%
Other non-manual	23.4%	5.3%	9.9%	12.1%	10.9%	11.4%
Manual	24.3%	6.4%	9.6%	11.1%	13.0%	11.8%
Residual	35.6%	6.6%	10.6%	11.8%	16.7%	14.6%
Mother's occupational group						
Managers/Farmers	21.9%	4.3%	13.2%	8.4%	13.5%	13.0%
Higher professionals	22.3%	4.3%	7.0%	9.7%	12.2%	10.0%
Lower professionals	16.8%	3.8%	9.2%	7.2%	10.4%	9.1%
Other non-manual	21.2%	5.4%	8.5%	10.4%	12.8%	11.3%
Manual	24.9%	5.2%	9.4%	11.4%	12.1%	10.4%
Residual	29.9%	6.9%	10.9%	11.5%	14.4%	12.9%

Table A 6 Annual attrition rates for Y95 cohort telephone interviews, 1997-2002, by selected characteristics in 1995 (weighted)

	1997	1998	1999	2000	2001	2002
All	24.3%	5.5%	9.8%	10.2%	12.8%	11.4%
Gender						
Male	24.3%	7.3%	8.1%	10.2%	12.9%	11.3%
Female	24.3%	3.8%	11.4%	10.2%	12.8%	11.4%
Indigenous background						
Indigenous	35.5%	13.5%	9.6%	11.7%	13.1%	18.2%
Non-indigenous	23.7%	5.1%	10.1%	10.2%	12.5%	11.1%
Double response	-5.8%	5.0%	-20.5%	-0.9%	0.0%	33.6%
Home language						
English	23.4%	5.3%	10.3%	10.4%	12.8%	11.5%
Other language	27.8%	5.1%	6.2%	7.2%	12.2%	11.0%
Double response	29.4%	7.7%	12.3%	9.5%	19.6%	6.8%
State (school)						
Australian Capital Territory	27.4%	1.4%	13.5%	10.7%	13.6%	12.1%
New South Wales	27.2%	1.7%	11.4%	10.3%	13.3%	11.3%
Victoria	22.1%	8.2%	11.4%	10.3%	12.8%	11.3%
Queensland	24.2%	5.6%	2.9%	9.8%	12.1%	11.2%
South Australia	18.3%	12.5%	11.5%	10.4%	12.9%	11.8%
Western Australia	22.1%	8.1%	11.2%	10.3%	12.9%	11.6%
Tasmania	26.3%	3.0%	8.7%	9.4%	12.6%	11.1%
Northern Territory	38.3%	-16.0%	10.8%	10.5%	13.3%	11.8%
School sector						
Government	25.5%	3.9%	8.6%	9.9%	12.6%	11.3%
Catholic	21.8%	8.5%	11.0%	10.3%	13.0%	11.4%
Independent	21.5%	8.9%	14.3%	11.2%	13.9%	11.6%
Area						
Metropolitan	25.4%	3.4%	10.2%	10.3%	12.9%	11.4%
Regional	22.6%	6.8%	9.4%	9.6%	13.4%	11.3%
Rural/remote	22.1%	9.4%	9.1%	10.6%	12.1%	11.1%
Father's occupational group						
Managers/Farmers	21.5%	6.6%	12.7%	10.6%	14.4%	10.0%
Higher professionals	21.0%	3.8%	13.7%	8.7%	13.4%	10.4%
Lower professionals	17.4%	2.4%	12.2%	7.8%	11.9%	10.2%
Other non-manual	23.7%	4.7%	10.5%	11.9%	10.5%	12.4%
Manual	23.4%	5.9%	8.0%	9.6%	11.9%	10.6%
Residual	33.4%	6.7%	5.8%	11.3%	15.0%	14.9%
Mother's occupational group						
Managers/Farmers	23.7%	4.7%	15.2%	9.2%	13.0%	12.4%
Higher professionals	23.7%	3.1%	10.7%	13.3%	15.5%	10.1%
Lower professionals	19.7%	4.0%	13.6%	7.8%	11.5%	10.0%
Other non-manual	21.4%	4.7%	9.4%	10.2%	13.4%	11.4%
Manual	22.8%	4.0%	6.4%	10.2%	10.5%	10.5%
Residual	28.6%	7.6%	8.5%	11.2%	13.6%	12.2%

Table A 7 Annual attrition rates for Y98 cohort telephone interviews, 2000-2002, by selected characteristics in 1998 (unweighted and weighted)

	Unweighted			Weighted		
	2000	2001	2002	2000	2001	2002
All	32.4%	8.1%	11.6%	32.4%	8.1%	11.6%
Gender						
Male	34.2%	9.4%	12.3%	34.9%	8.2%	11.6%
Female	29.9%	6.7%	10.8%	29.3%	7.9%	11.5%
Indigenous background						
Indigenous	52.0%	11.3%	14.9%	50.5%	10.5%	11.3%
Non-indigenous	30.7%	7.8%	11.4%	30.9%	7.9%	11.4%
Double response	81.8%	50.0%	0.0%	84.7%	72.0%	5.4%
Home language						
English	30.6%	7.8%	11.3%	31.5%	7.9%	11.5%
Other language	38.8%	9.3%	13.3%	32.4%	8.6%	11.7%
Double response	39.9%	12.5%	14.3%	35.3%	11.7%	11.6%
State (school)						
Australian Capital Territory	25.8%	6.5%	14.0%	31.9%	8.5%	11.9%
New South Wales	32.9%	7.8%	12.5%	32.5%	8.2%	11.6%
Victoria	33.2%	8.7%	10.6%	32.2%	7.9%	11.4%
Queensland	32.2%	8.9%	12.0%	32.5%	8.1%	11.6%
South Australia	30.7%	7.6%	10.4%	32.3%	8.0%	11.6%
Western Australia	32.6%	8.2%	12.9%	32.2%	8.0%	11.5%
Tasmania	28.7%	4.7%	7.2%	32.3%	8.3%	11.7%
Northern Territory	41.4%	8.9%	10.2%	31.8%	8.2%	11.6%
School sector						
Government	34.9%	9.3%	12.2%	32.2%	8.0%	11.3%
Catholic	26.5%	5.8%	11.3%	32.7%	8.1%	11.7%
Independent	30.2%	6.8%	9.5%	32.9%	8.5%	12.6%
Area						
Metropolitan	33.0%	7.6%	11.4%	31.8%	8.0%	11.4%
Regional	30.2%	8.6%	11.5%	31.4%	8.1%	11.4%
Rural/remote	27.4%	8.6%	12.1%	31.2%	8.1%	12.1%
Father's occupational group						
Professional	27.0%	7.2%	9.3%	29.2%	8.5%	10.3%
Managers	28.2%	6.4%	11.4%	31.4%	7.1%	12.2%
Clerical and personal service	26.9%	6.8%	12.3%	25.9%	7.5%	12.5%
Trades	30.0%	8.1%	12.2%	28.7%	7.1%	11.7%
Plant operators and labourers	34.4%	8.9%	13.7%	31.6%	8.0%	12.0%
Unskilled manual	36.0%	9.2%	10.2%	34.0%	7.9%	7.3%
Mother's occupational group						
Professional	27.3%	6.2%	8.8%	30.1%	7.3%	10.0%
Managers	29.8%	6.5%	11.3%	32.5%	7.9%	13.1%
Clerical and personal service	27.7%	6.8%	11.5%	27.8%	6.7%	11.3%
Trades	31.3%	12.4%	11.8%	30.0%	12.1%	11.9%
Plant operators and labourers	31.8%	11.8%	12.7%	29.9%	8.7%	9.6%
Unskilled manual	33.0%	8.4%	14.8%	29.2%	8.2%	13.6%