say

Longitudinal Surveys of Australian Youth

Technical Paper Number 16

The Designed and Achieved Sample of the 1998 LSAY Sample

Mike Long
Nicole Fleming

February 2002

ACER

Australian Council for Educational Research

Longitudinal Surveys of Australian Youth

The 1998 Sample

This paper outlines the intended and achieved samples for the second cohort in the program of *Longitudinal Surveys of Australian Youth*. Some earlier discussions that relate to the sample design are available in LSAY Technical Reports Numbers 2, 7 and 8.

The sample design

The design was intended to provide a national stratified sample of Year 9 students which would permit a sample of some 10,000 young people to be interviewed by phone in late 2000. The major stratum considered in the design was State of schooling. Students from smaller states were to be over-sampled and, correspondingly, students from larger states were under-sampled.

However, in the 1998 sample, the over-sampling of the smaller states and the territories was undertaken at a lower rate than for the 1995 sample. There were two reasons for this:

1. Continued over-sampling of schools in the smaller states would place a greater burden on schools within those States. As the proportion of schools required approaches 100 per cent of all the schools (allowing for refusal, non-return of materials, and so on), sampling becomes more problematic and school and system-level resistance increases.

2. National estimates are produced by weighting the sample. Greater variation in sampling fractions between States (that is, higher levels of over-sampling) requires greater variation in those weights -- over-sampled States require smaller weights and under-sampled States require smaller weights. Disparity in the weights can contribute to instability in analyses over the years of the survey.

Reducing the extent of over-sampling of the smaller States would address these problems somewhat. It is, however, a question of balance between these concerns and the desire to provide stable State estimates from the survey.

Table 1 shows the proposed and actual sampling fractions for the 1995 survey and the proposed sampling fractions for the 1998 survey. It also shows, on the assumption of 10,000 respondents, the resultant sample sizes for the States and the change in sample size between the 1995 and proposed 1998 designs. The effect of the proposed change is evident -- sample sizes for the smaller States and the territories would be reduced in 1998.

Table 1 Sampling fractions and sample sizes by State and sample design

| | *Sampling fraction* | | | | | *Sample size* | | | | *Change* | *Std* |
| | *Schools* | | *Students* | | | *Students* | | | | *in Std* | *Error* |
	1995	*1998*	*1995*	*1995*	*1998*	*1995*	*1995*	*1998*	*Change*	*Error %*	*1998*
	Designed	*Designed*	*Designed*	*Achieved*	*Designed*	*Designed*	*Achieved*	*Designed*			
NSW	13.0	14.5	0.70	0.69	0.78	2279	2246	2560	+282	-7	1.29
VIC	15.5	17.0	0.81	0.90	0.90	1887	2097	2109	+222	-5	1.33
QLD	19.1	20.0	0.93	0.95	1.00	1852	1892	2003	+151	-4	1.44
SA	23.7	20.5	1.49	1.65	1.30	1143	1266	1003	-140	+8	1.48
WA	21.3	18.9	1.36	1.28	1.20	1414	1331	1255	-159	+5	1.52
TAS	31.7	22.8	2.38	1.50	1.80	655	413	498	-157	+17	2.06
NT	24.2	14.5	3.17	3.62	2.00	293	335	186	-107	+22	5.57
ACT	54.5	40.9	2.49	2.20	2.00	477	422	386	-92	+13	2.18
AUS	18.8	18.8	1.00	1.00	1.00	10000	10000	10000	0	0	

Notes

1. Sampling fractions for schools are fractions of the schools in a State that will be contacted. This is an under-estimate, particularly for the Northern Territory, because an allowance needs to be made for schools that will yield fewer than 44 students in Year 9.

2. Sampling fractions for schools assume 66.7 response rate from schools and 21 per cent attrition between contact and successful interview.

3. Sampling fractions for students show the variation from the overall sampling fraction of the survey -- not the sampling fraction *per se*. A value of 2.00, for instance, indicates twice the overall sampling fraction.

4. Sample sizes for 1995 and 1998 have been standardised to 10,000 to assist comparison.

5. Values are for the main sample only and therefore exclude consideration of the pilot sample.

6. Designed sample sizes are based on enrolments in Year 8 in ABS, *Schools Australia 1996 (Cat. No. 4221.0)* -- the latest information to hand at the time of design.

7. Change in standard error is the percentage change from the 1995 design to the 1998 design for an overall measure of achievement based on reading and mathematics. Standard errors take into account the two-stage nature of the design and consequently are based on the number of schools rather than the number of students.

8. Standard errors in 1998 are for the overall measure of achievement. The achievement variable is likely to be a worst-case choice because of the greater homogeneity within schools (and hence greater variance between schools). The weighted student mean and standard deviation are 49.48 and 17.70 respectively.

The effect of the change in sampling fractions on the standard errors of a combined reading-maths test measure is shown in Table 1. The standard errors for the five mainland states are similar and small relative to the standard deviation. The standard errors for Tasmania and the ACT are somewhat larger and markedly larger for the Northern Territory. The proposed changes in sampling fractions maintain reasonable standard errors for most States, but take into account administrative practicalities and the need for more stable weights.

Selection of students within States was to be proportional by Sector. Three sectors were used as strata: Government schools, Catholic schools and Non-government, Non-Catholic (referred to as *independent*) schools. The population data for strata were taken from the *Schools Australia* series (ABS). Within strata, schools were to be selected proportional to their size. Information on the number of Year 9 students in each school came from ACER's Sampling Frame which, in turn, was based on information provided by the relevant State authorities and, in the case of non-government schools, by DEETYA (as it was then). These figures were from the 1997 annual school census. Within schools two classes were to be randomly selected (again, proportional to their size). Schools were asked for a list of the number of students enrolled in each of their Year 9 classes for a subject studied by all Year 9 students in the school (usually English classes). Responses would be weighted to correct for the disproportionate sampling between strata and to correct for the variation between strata due to differential response rates and variable class sizes.

Based on an achieved sample of 13,513 in 1995, the designed sample size for 1998 was 14,000, so as to yield approximately 10,000 students for the telephone interviews, after sample attrition (including refusal and inability to contact). While the designed sample for 1995 was based on a minimum of 35 students per school, the design was based on an average 25 students per class, yielding 50 students per school.

An additional sample of some 500 Year 9 students was to be selected by sampling a further class from some schools. This group was to form the basis for a pilot sample in subsequent contacts with the cohort.

The sampling process

Lists of schools for each sector within each State were sorted by postcode. The cumulative total of Year 9 students was calculated. The interval required to yield the designed number of schools was determined, a random start made within that interval, and then the interval was applied to select schools from the list. This process yields selection proportional to the size of the school and an implicit stratification by geography because of the postcode-order of the list. If constant numbers are selected from each school, the sample within each State becomes self-weighting.

Schools with less than 50 Year 9 students were combined with other small schools to form a pseudo-school with more than 50 Year 9 students. The selection of one school within a pseudo-school implied the selection of other schools within the pseudo school. Without the formation of pseudo-schools, there was the likelihood that the achieved sample would be less than the designed sample because 50 Year 9 students would not be able to be selected from smaller schools.

In practice, the procedure for selection of schools was modified from that outlined above. The *Third International Mathematics and Science Study* (TIMSS) -- another ACER project -- had already approached some schools in 1998. These schools were removed from the list before selection began. This had a relatively small effect on sample selection. This study focused on schools with Year 12 students. Selection in states with senior colleges (mainly Tasmania and the Australian Capital Territory) was therefore relatively unaffected and few schools were involved in the other small States. In the larger States, proportionately fewer schools were involved.

Due to problems with replacing schools that refused to take part in the 1995 sample, it was proposed, for the 1998 sample, to approach more schools than would be required for the designed sample so as to allow for a proportion of refusals. This is instead of contacting the replacement school on the refusal of the selected school, since often it was difficult to know when to determine a refusal, if the school delayed their response. Thus, each sector was over-contacted by approximately 50%, to allow for refusal.

The first contact with a school was by a letter to the principal. The letter was accompanied by a form and a reply-paid envelope. The form allowed schools to indicate whether or not they would participate in the project and, if they agreed, the class structure of their Year 9 students. Several weeks was allowed for the request to be considered by internal staff committees or school councils. If, after several weeks, no reply was received, a follow-up fax was sent to the principal. Again, if no response was received, a second fax was sent. In some cases a further series of phone calls was required. Frequently the first or second fax elicited a request that the initial letter be sent again.

In New South Wales, due to requirements of the Department of Education and Training, we were to seek explicit parental permission for each student's participation. This situation had the potential to significantly reduce response rates within New South Wales government schools. Therefore, with the permission of the Department, we selected 3 classes of Year 9 students from those schools that had 3 or more Year 9 classes, in order to compensate for any increase in refusal rates.

Selection within schools proceeded as intended. In schools where selection of two classes did not yield 50 students, a third class was selected. In several schools, the principal requested that all students in Year 9 be tested. Table 2 outlines the designed and achieved samples.

Table 2   Designed and Preliminary Achieved Samples and Response Rates

State	Sector	Designed Schools	Approached Schools	Achieved Schools	Response Rate	Designed Students	Achieved Students	Over-sampling
ACT	GOV	6	9	8	88.9	300	348	16
	CATH	3	4	3	75.0	150	122	-18.7
	IND	1	2	2	100.0	50	88	76.0
	Total	10	15	13	86.7	500	558	11.6
NSW	GOV	47	71	40	56.3	2350	2323	-1.1
	CATH	16	23	13	56.5	800	651	-18.6
	IND	8	11	7	63.6	400	410	2.5
	Total	71	105	60	57.1	3550	3384	-4.7
VIC	GOV	39	58	43	74.1	1950	1715	-12.1
	CATH	14	21	15	71.4	700	741	5.9
	IND	10	14	10	71.4	500	494	-1.2
	Total	53	93	68	73.1	3150	2950	-6.3
QLD	GOV	36	56	39	69.6	1800	1879	4.4
	CATH	10	15	13	86.7	500	762	52.4
	IND	9	13	10	76.9	450	470	4.4
	Total	55	84	62	73.8	2750	3111	13.1
SA	GOV	19	30	18	60.0	950	733	-22.8
	CATH	5	8	6	75.0	250	284	13.6
	IND	5	7	5	71.4	250	232	-7.2
	Total	29	45	29	64.4	1450	1249	-13.9
WA	GOV	26	39	22	56.4	1300	1144	-12.0
	CATH	8	11	7	63.6	400	336	-16.0
	IND	6	8	5	62.5	300	209	-30.3
	Total	40	58	34	58.6	2000	1689	-15.6
TAS	GOV	10	16	10	62.5	500	432	-13.6
	CATH	2	3	3	100.0	100	184	84.0
	IND	1	2	2	100.0	50	99	98.0
	Total	13	21	15	71.4	650	715	10.0
NT	GOV	4	6	5	83.3	200	313	56.5
	CATH	1	1	1	100.0	50	42	-16.0
	IND	1	2	2	100.0	50	106	112.0
	Total	6	9	8	88.9	300	461	53.7
AUS	GOV	187	285	185	64.9	9350	8887	-4.9
	CATH	59	86	61	70.9	2950	3122	5.8
	IND	41	59	43	72.8	2050	2108	2.8
	Total	287	430	289	67.2	14350	14117	-1.6

Notes:
1 *Schools* refers to pseudo-schools.
2 *Response rate* is response rate for schools.
3 *Designed students* is based on 50 students per school.
4 Includes students to be assigned to the pilot sample.