Weighting for the Australian Study of Health and Relationships

January 2003

Nicole Watson and Louise Gates

Prepared for

Assoc Prof Anthony Smith and Dr Richard de Visser
Australian Research Centre in Sex, Health and Society
La Trobe University
1st Floor, 215 Franklin St
Melbourne VIC 3000

Contents

Contents *2*
Introduction *3*
Survey design of the ASHR *3*
Sample design **3**
Sample selection **3**
Fieldwork **4**
Data issues *4*
Data problems **4**
Modified strata **5**
Non-response Analysis *6*
Response rates **6**
Respondent characteristics **7**
Weighting methodology *8*
Design weights **8**
Adjustment to weights **9**
Comparison to independent sources *11*
Conclusion *14*
Appendix 1: Final outcomes for the male and female samples *15*

Introduction

This report details the work undertaken in weighting the Australian Study of Health and Relationships (ASHR). There are two components to this work:

- Analysis of non-response – To determine what effect non-response may have on estimates from the survey, the non-response rates were considered across geographic regions and comparisons of demographic variables were made against external data sources.

- Construction of the weights – Weights are used in surveys to adjust the contribution to estimates of certain respondents – those with a high probability of being included in the survey have a low weight and those with a low probability of being included have a high weight. For the ASHR, the weights reflect both the design of the study (which over-sampled some geographical areas and under-sampled people at low risk) and likelihood of response (those less likely to respond are given a higher weight and those more likely to respond are given a lower weight).

This report first outlines our understanding of the survey design features for the ASHR and discusses some of the data problems encountered. The design of the survey determines the initial weights that are calculated and these are then adjusted, where possible, for non-response biases. Adjusting the weights for non-response assumes that respondents are like non-respondents – this is a fairly strong assumption, but it is frequently adopted in weighting procedures. To help understand the characteristics of the non-respondents, the response rates across various regions in Australia are considered and estimates using the design weight are compared to data from external sources. To understand the improvements made in adjusting the weights, the weighted estimates are also compared to the external data sources.

Survey design of the ASHR

Sample design

Over 19,000 telephone interviews were conducted with a disproportionate stratified two-phase sample of men and women aged 16 to 59 from across Australia.

The sample was stratified by sex and then into 14 geographic regions for the female sample and 15 geographic regions for the male sample – resulting in 29 strata all together. Separating the male sample from the female sample allowed the interviewers to market the survey to respondents as a survey of men's health or a survey of women's health. The basic geographic stratification was by State, and, as funding partners were identified, extra sample was added to certain parts of some States.

During the pilot test for the survey, it was identified that significant interviewing effort was expended on people with one sexual partner of the opposite sex in the last 12 months. This led to the introduction of a second phase of sampling after the respondent had answered some initial questions about their sexual practice in the last 12 months.

Sample selection

The telephone numbers were generated using a form of random digit dialing. Area codes together with 4 digit prefixes that were most likely to be in use were identified from the Electronic White Pages directory listing. These prefixes formed the basis of the frame of telephone numbers from which the sample was drawn (where all combinations of the remaining 4 digits were possible).

On making contact with a household, the number of eligible people in the household was determined. For households in the male sample, the number of eligible people was the number of males aged 16 to 59. In households in the female sample, the number of eligible people was the number of females aged 16 to 59. If no-one in the household was eligible, the household was coded as such and another number was called. When identifying a household with more than one eligible person, the computer randomly selected one person.

If, after completing the introductory sections of the questionnaire, the respondent only had one sexual partner of the opposite sex in the last 12 months, then they were randomly assigned to receive a long form or a short form.[1] Approximately 20 per cent of these people received the long form. In all other cases, the respondent went through all of the sections that were relevant to them.

Fieldwork

The Hunter Valley Research Centre conducted the interviews between May 2001 and June 2002. The male and female samples were generally worked on in a similar manor throughout this period, with the exception that the male sample took three extra weeks to finalise at the end of the fieldwork period.[2]

An extensive call back routine was adopted to maximize the chances of making contact with the household and then following up an interview with the selected individual.

Data issues

Data problems

There were a number of problems identified with the data available for the construction the weights. This either caused significant work to circumvent the problem or limited the options available in adjusting for non-response. These issues are briefly discussed below:

- *Overlapping strata* – As mentioned earlier, the first stratification of the sample was by State. When additional funding became available to top up the sample in some areas, additional sample was drawn for these areas. However, the initial strata were not reconstructed at this stage. Prefixes had to be obtained from La Trobe University for the creation of independent strata (described in more detail below). This means that, for example, the Victorian sample had to be split into a Melbourne component and a rural Victoria component. The rural component of the sample was then combined with the top up sample selected for rural Victoria to form a separate rural Victoria stratum. The Melbourne component forms a separate Melbourne stratum.

- *Geographic area identifier was not linked to Local Government Area (LGA)* – Respondents provided their postcode at the end of the interview. This could have been linked to LGA, which could have allowed the age and sex adjustment to the weights to occur at a lower geographic level than State. (For example, Central Sydney could have been adjusted to the age and sex profile of the group of LGAs that make up Central Sydney.) It is noted, however, that the stratification boundaries used (based on telephone prefixes) may not have lined up particularly well with LGA boundaries, which may have introduced additional problems into the weighting methodology.

- *Strata crossing State boundaries* – State boundaries were identified by the telephone prefixes used in the area. As a result, there are some people selected in one State, but actually live in another State. (For example, people living near the border of NSW and Victoria may have been selected in NSW as they have an 02 area code, but when interviewed, actually said that they were in Victoria.) This problem is fairly minor, and has been ignored in the weighting process.[3] The State variable rather than strata, however, should be used when producing estimates by State.

---

[1] Respondents were randomly assigned the form type based on the modulus of their unique identifier when divided by 5. The unique identifier is assigned to the individual by the computer at the beginning of the interview.

[2] It has been assumed in the weighting strategy adopted that the variables of interest from the survey are not affected by the time of the year that the respondent was interviewed. Therefore, no account has been made of any lumpiness in interviewing during the fieldwork period.

[3] There are 0.7 per cent of the respondents who were selected in one state and report living in another. The majority of these cases are from cross-overs between NSW and Victoria or between NSW and ACT.

- *Frame information incorrect* – Information about the number of telephone numbers from which the sample was selected had to be regenerated from the selection programs. Unfortunately, the staff involved with the ASHR had changed and some of the programs did not run properly. As a result, the quality of some of this information was relatively suspect. This information was very important to the construction of the design weights, so had to be corrected based on the observed characteristics of the original and top-up samples.[4]

- *Unlikely combinations of number of phone lines and number of people in the household* – There were 33 respondents that reported having five or more separate telephone lines in their household. Generally these households were not particularly large in size. The number of phone lines in the selected household was used in the calculation of the design weights, so it is important that this information is correct. This information has been treated as correct in the weighting procedure.

- *Age missing or outside of age range* – The age of each respondent (as at the end of 2001) was constructed from their year of birth. This is inexact, as we do not know if the interview occurred before or after their birthday. In fact, for a quarter of cases, we do not even know the date of interview. People whose age was calculated as 15 were recoded to age 16 and people whose age was calculated as 60 were recoded to age 59. Where year of birth was not provided, age bands were usually available. There were 14 cases for which either the age group could not be calculated or they were aged 61 (and should not have participated in the interview). As a result, weights were not assigned to these 14 cases.

Modified strata

As mentioned earlier, the stratification needed to be revised so that strata did not overlap and the selection probabilities could be calculated. For example, a household in rural Victoria could have been selected in the original Victorian sample or in the extra rural Victorian top up sample.

Telephone prefixes were used to identify the section of the original State strata for which top-up samples were selected. As a result of this, the following changes were made:

- 105 men and 134 women were moved from the NSW strata (region=1) into Central Sydney;
- 48 men were moved from the NSW strata (region=1) into Gay Sydney;
- 683 men and 734 women were moved from the NSW strata (region=1) into Rural NSW;
- 717 men and 726 women were moved from the Vic strata (region=2) into Rural Vic;
- 722 men and 753 women were moved from the Qld strata (region=2) into Rural Qld;
- 4 men and 3 women were moved from the WA strata (region=5) into WA rural 1;
- 10 men and 10 women were moved from the WA strata (region=5) into WA rural 2.

As the meaning attached to the original stratification codes had changed, it was thought better to construct a revised stratification labeling convention to avoid confusion. The revised stratification uses two digits – the first digit being a State code (1 through to 8) and the second being a region identifier: 0 = most of State, 1 = metropolitan, 2 = rural, 3 = special sub-sample 1, 4 = special sub-sample 2. The following table shows how the original stratification and the revised stratification are connected.

[4] For example, the proportion of the responding male sample in NSW (region==1) falling into Sydney, Central Sydney, Gay Sydney and rural NSW was used together with the interview to selected sample ratio in top-up samples for Central Sydney (region==9), Gay Sydney (region==10) and rural NSW (region==11) to divide the population of telephone numbers for NSW into Sydney, Central Sydney, Gay Sydney and rural NSW.

Table 1: Original and revised stratification

Original Stratification	Revised Stratification	Original Stratification	Revised Stratification
1 = NSW	11 = Sydney 12 = Rural NSW 13 = Central Sydney 14 = Gay Sydney	9 = Central Sydney	13 = Central Sydney
2 = Vic	21 = Melbourne 22 = Rural Vic	10 = Gay Sydney	14 = Gay Sydney
3 = Qld	31 = Brisbane 32 = Rural Qld	11 = Rural NSW	12 = Rural NSW
4 = SA	40 = SA	12 = Rural Vic	22 = Rural Vic
5 = WA	50 = WA (except rural 1 and 2) 53 = WA rural 1 54 = WA rural 2	13 = Rural Qld	32 = Rural Qld
6 = Tas	60 = Tas	14 = WA rural 1	53 = WA rural 1
7 = NT	70 = NT	15 = WA rural 2	54 = WA rural 2
8 = ACT	80 = ACT		

Non-response Analysis

The first step in considering the impact of non-response on a sample is to determine the size of the group of non-respondents. The second step is to consider how different the non-respondents are from the respondents – this is often difficult to do directly as we do not have any information about the non-respondents. The usual approach, therefore, is to consider how different the respondent characteristics are from accepted population characteristics.

Response rates

The number of interviews and response rates obtained in each strata for the male and female sample are provided in Table 2. The original stratification has been used in this table, as prefixes were not provided for all the selected telephone numbers (only for the responding sample). Further information on the final outcome status is given in Appendix 1.

Two response rates are provided in Table 2. The numerator in both response rates is the same – that is, the number of interviews obtained. The denominator in the response rate calculated by the La Trobe University is the sum of interviews and refusals (at the individual level and household level) and incomplete interviews. It can be argued that this provides an overestimate of the true response rate as it does not include all households where there was an eligible person or where it can be expected that their was an eligible person. Households with an eligible person also include those that were scheduled for a callback (either at individual or household level), sent a letter, or where the individual was unavailable for the duration of interviewing. A portion of households with language difficulties would have also had an eligible person – an estimate of the number of such households was obtained by applying the same rate of eligibility in households where there weren't English difficulties. The alternative response rate is approximately 4 percentage points lower than that calculated by La Trobe University.[5] Most of this difference is due to the inclusion of an eligible proportion of households with language difficulties.

[5] It may also be argued that some of the numbers that have been classified as out of scope (as they always rang out, or was always engaged) may actually contain an eligible person. However, it is unlikely that there would be many such cases due to the extensive call back schedule employed. All of these types of numbers have been treated as non-household numbers.

Table 1: Interviews achieved and response rates for the ASHR male and female samples

	Male sample			Female sample		
Location	**Interviews**	**Response rate (La Trobe)**	**Response rate (Alternative)**	**Interviews**	**Response rate (La Trobe)**	**Response rate (Alternative)**
NSW	1456	67.0%	63.0%	1454	76.0%	71.3%
VIC	2099	67.7%	62.5%	2086	77.4%	71.7%
QLD	1257	66.5%	64.5%	1295	76.6%	73.6%
SA	501	70.1%	67.8%	531	75.7%	72.8%
WA	517	68.9%	66.1%	505	74.3%	72.7%
TAS	525	73.1%	70.7%	509	81.7%	79.9%
NT	510	74.3%	70.1%	503	80.1%	76.1%
ACT	503	75.8%	73.1%	506	83.2%	80.7%
Central Sydney	503	71.0%	62.3%	514	78.5%	66.3%
Gay Sydney	1022	70.2%	63.3%	-	-	-
NSW Rural	338	67.2%	65.4%	335	76.7%	75.3%
VIC Rural	347	71.8%	69.1%	357	79.0%	76.2%
QLD Rural	386	74.4%	73.2%	335	77.4%	75.7%
WA rural 1	115	70.6%	68.0%	108	80.6%	77.0%
WA rural 2	104	74.3%	73.1%	104	84.6%	84.1%
Australia	10183	69.4%	65.4%	9142	77.6%	73.4%

The response rates are very good for a voluntary telephone survey and differ largely as expected across the areas sampled. NSW and Victoria had lowest response rates, and Tasmania, NT and ACT had higher response rates. The rural areas outperformed the response rate figure for the State, though often not by much. The female sample had higher response rates than the male sample. The Central Sydney and Gay Sydney areas performed fairly well given the usual expectation that areas with greater population density are likely to have lower response rates.

With approximately one in every third selected male and one in every fourth selected female not participating in the ASHR, there is certainly scope for non-response bias.

Respondent characteristics

Comparison of respondent characteristics to population estimates provides further information on the likely characteristics of non-respondents. As the ASHR has a disproportionate stratified sample, estimates produced from the sample should be weighted. The appropriate weight to use at this point is one that accounts for the design of the sample – this is called the design weight and the calculation of this weight is discussed in the next section. Estimates of key demographic variables based on the sample are provided in Table 5 later in this report. It is sufficient to point out at this stage several features of estimates from the weighted sample compared to estimates from the Australian Bureau of Statistics:

- *State* – The distribution of the ASHR sample appears skewed towards NSW, SA, WA and ACT. Given the response rates are not hugely different across state, it is expected that this problem is likely to be a result of inaccurate information about the size of the frame of telephone numbers.

- *Age* – People aged 20-29 were less likely to participate in the ASHR and people aged 30-49 were more likely to participate in the survey. This is not an uncommon finding.

- *Birthplace* – People born in countries that do not mainly speak English are under-represented in the ASHR. Provision of interpreters would have helped avoid this bias.

- *Indigenous status* – The proportion of people who are of indigenous origin in the ASHR is consistent with the proportion expected based on ABS information.

- *Household size* – The proportion of people living alone in the ASHR is very close to what was expected from ABS data.

This analysis suggests that the weights for the male and female sample would benefit from state and age adjustments. It would be unwise to adjust for birthplace as it is unlikely that people with language difficulties would be similar to people without language difficulties. This should be treated as a recognized bias of the survey.

Weighting methodology

Design weights

The basic weight is the design weight, which is calculated as the inverse of the probability of selection. For the ASHR, the design weight needs to reflect the following:

- increased probability of selection in the smaller States compared to the larger States;
- increased probability of selection in certain geographic regions where the sample was topped up;
- increased probability of selection for people with multiple phone lines (as we could have called on any one of their numbers and reached the same household);
- decreased probability of selection for people living in households with more eligible people (as we only selected one per household); and
- for people with only one sexual partner of the opposite sex in the last 12 months, decreased probability of selection for completing the long form.

The second phase selection process can be thought of dividing the sample in each strata into two substrata as shown in Table 3. Approximately 20 per cent of the people in substrata 1 received the long form, and the remainder received the short form. All people in substrata 2 received the long form. Therefore, to calculate estimates for variables in the short form, all people selected into the sample can be used. However, to calculate estimates for variables in the long form, only those people receiving the long form from substrata 1 and everyone in substrata 2 can be used.

Table 3: Second phase stratification

	Sub-strata 1	Sub-strata 2
Type of people	➢ one heterosexual partner in last 12 months	➢ no sexual experience ➢ heterosexual partner, but not in last 12 months ➢ two or more heterosexual partners in last 12 months ➢ homosexual partner
Sample 1 (Short Form items)	All	All
Sample 2 (Long Form items)	~ 20 per cent	All

Therefore, the design weight in each strata h is calculated as:

$$W_{design,\,h} = \frac{N_{ph\ on\ frame,\,h} \times n_{ph\ lines\ in\ HH} \times F}{n_{ph\ selected,\,h} \times n_{eligile\ people\ in\ HH}}$$

$$\text{Where } F = \begin{cases} 1 & \text{, if person is in sub-strata 2} \\ N_{\text{sample 2, h}} / N_{\text{sample 1, h}} & \text{, if person is in sub-strata 1} \end{cases}$$

The number of eligible people and the number of phone lines differs for each selected household, but information on the number of phone numbers on the frame, the number of phone numbers selected and the number in sub-strata 1 and sub-strata 2 for each strata is provided in Table 4. The initial design weights are provided for indicative purposes – they provide information about the relativities between the design weights in different strata, but they do not sum to the population as not every selection resulted in an interview with an eligible person. For the design weights to sum to the total male population for Australia aged 16 to 59, a factor of 3.2 needs to be applied. The equivalent factor for the female sample is 2.5.

While the intention was to sample 20 per cent of the people in substrata 1, this actually varied from 8 per cent to 30 per cent, with much of the variability coming from the strata with small samples.

Adjustment to weights

The design weights were adjusted to age and sex benchmarks obtained from the Australian Bureau of Statistics. The benchmarks used were the 2001 Estimated Residential Population by State, sex and 5 year age groups (with the first age group being only 4 years – 16 to 19).

The design weights within each State, sex and 5 year age group were adjusted by a factor such that the sum of the adjusted weights equaled the population in this category. For the people in sub-strata 2, the weights for sample 1 (short form items) and sample 2 (long form items) do not exactly match due to the adjustments made to the external benchmarks. This should not be a problem for users, given La Trobe University will be providing users with two datasets (one for short form items and one for long form items).[6]

[6] A method to ensure an individual in sub-strata 2 received the same weight for sample 1 and 2 was investigated, but resulted in significant variability in the weights, which was unwarranted. Should this be a desirable feature of the weights, then the use of a SAS macro, GREGWT, should be investigated as it may be able to provide such adjustments to the weights. (GREGWT can be obtained from the Methodology Division of the Australian Bureau of Statistics.)

Table 4: Design information for the ASHR

Location	Sydney	Melb	Bris	SA	WA	TAS	NT	ACT	Central Sydney	Gay Sydney	NSW Rural	VIC Rural	QLD Rural	WA rural 1	WA rural 2
Strata (digits 1&2)	11	21	31	40	50	60	70	80	13	14	12	22	32	53	54
MALE SAMPLE															
Phone numbers on frame (N_h)	2,029,880	1,566,391	480,566	1,110,302	1,402,071	266,641	42,689	238,972	460,405	228,994	2,048,578	667,872	650,571	10,733	28,152
Selected phone numbers (n_h)	4,255	8,989	3,132	3,017	3,214	2,936	3,097	2,535	4,850	10,701	6,297	5,659	6,238	825	665
Substrata 1															
Long form	98	201	88	83	77	75	69	65	66	97	163	170	153	12	7
Short form	340	836	335	290	312	319	310	303	339	496	616	679	660	79	85
Actual long form sampling fraction	0.22	0.19	0.21	0.22	0.20	0.19	0.18	0.18	0.16	0.16	0.21	0.20	0.19	0.13	0.08
Substrata 2															
Long form	182	345	154	129	114	130	132	135	203	477	242	214	253	28	22
Initial design weights															
Long form strata 1	2132	899	737	1654	2204	477	76	534	582	131	1555	589	554	99	557
Strata 2, Short form strata 1	477	174	153	368	436	91	14	94	95	21	325	118	104	13	42
FEMALE SAMPLE															
Phone numbers on frame (N_h)	2,258,874	1,566,391	480,566	1,110,302	1,402,071	266,641	42,689	238,972	460,405	-	2,048,578	667,872	650,571	10,733	28,152
Selected phone numbers (n_h)	3,650	7,091	2,551	2,492	2,523	2,198	2,497	2,024	4,132	-	4,920	4,536	4,999	618	649
Substrata 1															
Long form	75	178	88	76	74	80	80	66	78	-	156	148	162	16	26
Short form	371	835	326	333	287	324	297	299	358	-	689	703	704	79	62
Actual long form sampling fraction	0.17	0.18	0.21	0.19	0.20	0.20	0.21	0.18	0.18	-	0.18	0.17	0.19	0.17	0.30
Substrata 2															
Long form	138	346	128	121	131	105	126	141	212	-	224	232	222	16	26
Initial design weights															
Long form strata 1	3680	1257	886	2398	2711	613	81	653	623	-	2256	847	696	103	147
Strata 2, Short form strata 1	619	221	188	446	556	121	17	118	111	-	416	147	130	17	43

Comparison to independent sources

The impact of the weight adjustments can be seen though a comparison of ASHR estimates to external data sources using the alternative ASHR weights. Table 5 provides ASHR estimates for Sample 1 and Sample 2 and compares these against ABS estimates for several variables. Significant differences between the ASHR and ABS estimates are marked with an asterisk.[7]

It is difficult to find information readily available from ABS publications that are applicable to people aged 16 to 59. The information compiled for the comparison purposes in this report come from a variety of surveys and relate to a variety of age ranges (see the footnote of Table 5 for more details). It is recommended that La Trobe University undertake a similar exercise using CDATA from the 2001 Census where the appropriate age range can be specified.

The following observations are made about the comparison of ASHR estimates (calculated using the adjusted weights) to ABS estimates:

- *Age, Sex and State* – As the design weights are adjusted to age, sex and State benchmarks, it follows that the ASHR estimates will closely match those from the ABS. There are some small differences by state and these are due to the sample crossing over State boundaries.

- *Legal marital status* – The ASHR estimates are higher than the ABS estimates for married and lower for never married. It is expected that these differences are observed due to the question asked in the ASHR. Interviewers were instructed not to accept 'defacto' as a response and to reallocate these people to other categories. It is likely that some respondents who were not married, but were in a defacto relationship, may have called themselves 'married'. The other categories of legal marital status line up well with the ABS source.

- *Birthplace* – People born in countries where the main language was not English are under-represented. Interviews were not conducted in languages other than English, so this was a known bias introduced by the survey methodology. Note that the ABS estimates include people aged 60 and over, so may overstate the proportion of people born in another country, but it is not expected that would change the estimates very much. This is one variable that should be reconsidered when CDATA 2001 is available.

- *Indigenous status* – The proportion of the ASHR sample identifying as indigenous is consistent with ABS figures.

- *Labour force status* – The ASHR appears to over-represent people who work part time and under-represent people who are not in the labour force. The scope, timing and age group of the ASHR and ABS surveys are similar. It is possible that the definitions used do not line up well – the ABS ask a large number of questions to work out the correct labour force status. If the difference between the estimates is considered a problem, then La Trobe University should investigate the definitional issue further with the ABS.

- *Educational attainment and occupation* – It appears that there is an over-representation of people with higher educational qualifications and more skilled occupations than

[7] The formulae used for sample 1 and sample 2 are provided in section 3.5 of Warszawsi et al (1997) with the extension to a stratified sample. [Warszawski, J., Messiah, A., Lellouch, J., Meyer, L., Deville, J., (1997), 'Estimating means and percentages in a complex sampling survey: Application to a French National Survey on Sexual Behaviour (ACSF)', *Statistics in Medicine*, Vol 16, p397-423.]

compared with ABS figures. It is, however, possible that respondents inflated their achievements to provide more socially desirable responses. Again, the questions used in the ASHR are not as extensive as the ABS questions. Further, the question used in the ASHR to obtain occupation relied on the interviewer and the respondent classifying the respondent's occupation into the list of relatively unfriendly and incomprehensible categories that correspond to the top level of the Australian Standard Classification of Occupation. It is not expected that this method would have resulted in very good occupation data.

- *Lone person households* – The proportion of lone person households in the ASHR compares favorably with the ABS figure. With telephone surveys, it is expected that contact with larger households would be more likely than contact with smaller households as there is more chance of catching someone at home. For the ASHR, it appears that the extensive call back routine adopted negated any such effect.

- *Smoking status* – The ASHR has a higher proportion of current smokers than identified by the ABS (admittedly in 1995). If the size of this difference is a concern to La Trobe University, then it is recommended that this be further investigated.

Overall, it is possible that some of the differences identified between the ASHR estimates and the ABS estimates are due to the different time periods, different scope of the sample, and different questions that were asked. However, it is also likely that some of the differences identified are real and are suggestive of biases in the ASHR data. While some differences may be significant, they may not be of practical importance, and La Trobe University will need to consider which differences are of importance to their research.

Table 5: Comparison of ASHR estimates to ABS estimates[(a)]

		ASHR[(h)]			
		Sample 1 (Short form items)		Sample 2 (Long form items)	
	ABS	Design Weights	Weights adjusted for location, age and sex	Design Weights	Weights adjusted for location, age and sex
State[(b)]					
New South Wales	33.5%	39.6%**	33.4%	39.9%**	33.4%
Victoria	24.9%	20.5%**	24.9%	20.2%**	24.9%
Queensland	18.7%	11.0%**	18.6%	11.1%**	18.7%
South Australia	7.6%	10.7%**	7.7%	10.4%**	7.7%
Western Australia	10.0%	12.4%**	10.0%	12.6%**	10.0%
Tasmania	2.4%	2.7%**	2.4%*	2.7%**	2.4%
Northern Territory	1.1%	0.4%**	1.1%	0.4%**	1.1%
Australian Capital Territory	1.7%	2.7%**	1.8%	2.6%**	1.8%
Sex[(b)]					
Males	50.4%	50.4%	50.4%	50.4%	50.4%
Females	49.6%	49.6%	49.6%	49.6%	49.6%
Age group[(b)]					
16-19	9.2%	9.2%	9.2%	9.4%	9.2%
20-24	11.5%	10.0%**	11.5%	10.0%*	11.5%
25-29	12.5%	10.8%**	12.5%	10.4%**	12.5%
30-34	12.0%	13.0%**	12.0%	13.3%*	12.0%
35-39	12.6%	13.3%*	12.6%	13.6%	12.6%
40-44	12.2%	14.0%**	12.2%	13.7%*	12.2%
45-49	11.4%	11.7%	11.4%	12.3%	11.4%
50-54	10.5%	10.3%	10.5%	9.5%*	10.5%
55-59	8.1%	7.6%*	8.1%	7.8%	8.1%

		ASHR[h]			
		Sample 1 (Short form items)		Sample 2 (Long form items)	
	ABS	Design Weights	Weights adjusted for location, age and sex	Design Weights	Weights adjusted for location, age and sex
Legal marital status[c]					
Married (excluding de-facto)	49.1%	53.3%**	52.0%**	53.2%**	51.9%**
Divorced	7.4%	7.5%	7.3%	7.7%	7.4%
Separated	3.7%	3.4%	3.4%*	3.3%	3.3%
Widowed	1.0%	0.8%	0.9%	0.6%**	0.7%**
Never married	38.7%	34.9%**	36.4%**	35.1%**	36.6%*
Birthplace[d]					
Born in Australia	72.4%	75.4%**	76.2%**	75.7%**	76.2%**
Born outside Australia	27.6%	24.5%**	23.8%**	24.0%**	23.8%**
Main English-speaking countries	10.2%	11.1%**	10.6%	11.9%*	11.5%*
Other countries	17.4%	13.5%**	13.2%**	12.4%**	12.3%**
Indigenous status[e]					
Indigenous	1.7%	1.5%	1.6%	1.7%	1.8%
Non-indigenous	98.3%	97.9%	97.9%	97.9%	97.8%
Labour force status[f]					
Employed	69.8%	74.7%**	75.5%**	74.4%**	75.3%**
Full-time	50.8%	52.3%**	52.4%**	51.4%	51.9%
Part-time	19.0%	23.2%**	23.1%**	23.7%**	23.4%**
Unemployed	5.2%	4.2%**	4.2%**	4.0%**	4.1%**
Not in the labour force	25.0%	20.3%**	20.3%**	20.9%**	20.6%**
Educational attainment[g]					
Postgraduate degree (inc grad.dip. & grad.cert.)	4.6%	9.4%**	9.0%**	9.7%**	9.1%**
Undergraduate degree (ie bachelor degree)	12.5%	13.7%**	13.5%**	14.5%**	14.1%*
College certificate or diploma (ie adv.dip or dip)	6.7%	11.8%**	11.8%**	11.8%**	12.1%**
Higher secondary school (ie Yr11 & Y12)	28.3%	21.7%**	22.2%**	21.3%**	21.4%**
Technical or trade certificate (ie Cert I, II, III or IV)	15.1%	11.9%**	12.0%**	10.7%**	11.2%**
Lower secondary school or lower (Y10 and below)	31.4%	31.3%	31.3%	32.0%	32.1%
Level not determined	1.1%	0.1%**	0.1%**	0.0%**	0.0%**
Occupation (employed persons) [f]					
Managers & administrators	7.3%	17.0%**	16.4%**	16.9%**	16.5%**
Professionals	18.8%	20.3%**	19.9%**	22.7%**	22.0%**
Associate professionals	11.6%	8.6%**	8.7%**	8.2%**	8.2%**
Tradespersons and related workers	12.9%	13.2%	13.4%	12.7%	12.9%
Advanced clerical, sales and service workers	4.7%	7.4%**	7.3%**	6.7%**	6.8%**
Intermediate clerical, sales and service workers	17.1%	13.3%**	13.6%**	13.4%**	13.5%**
Intermediate production and transport workers	8.6%	3.9%**	3.9%**	4.0%**	3.9%**
Elementary clerical, sales and service workers	9.9%	7.3%**	7.6%**	6.5%**	6.8%**
Labourers and related workers	9.1%	8.9%	9.1%	8.9%	9.4%
Occupation not determined					

		ASHR[h]			
		Sample 1 (Short form items)		Sample 2 (Long form items)	
	ABS	Design Weights	Weights adjusted for location, age and sex	Design Weights	Weights adjusted for location, age and sex
Relationship in Household[c]					
Lone person	8.8%	8.7%	8.7%	8.9%	9.1%
Smoking status[g]					
Never smoked	48.7%			46.0%**	46.1%**
Past smoker	25.3%			26.0%	25.2%
Current smoker	26.0%			27.5%	28.3%**

Notes:

a. ** and * denotes statistically significant values from ABS estimates at the 99 and 95 per cent confidence levels, respectively.

b. The ABS estimates for State, sex and age group were the preliminary Estimated Residential Population as at June 2001 as reported in publication Population by Age and Sex (ABS Cat. No. 3201.0). These estimates are based on the 2001 Census and have been updated for known births, deaths, immigration and emigration. The intrastate migration has been estimated based on Medicare data. The ABS does not provide standard errors for these estimates and regards the information as very accurate.

c. The ABS figures for marital status and lone households were obtained from 2001 Census of Population and Housing, Release 1 (ABS Cat. No. 2001.0). The figures relate to individuals aged 15 to 59 and include people in non-private dwellings.

d. The ABS estimates for country of birth were obtained from the Labour Force Survey conducted in January 2001 (ABS Cat. No. 6203.0). These estimates relate to individuals aged 15+.

e. The ABS estimates for indigenous people were obtained from the Labour Force Survey conducted in August 2001 (ABS Cat. No. 6203.0). These estimates relate to individuals aged 15+ and includes people in non-private dwellings.

f. The ABS estimates for labour force status, educational attainment and occupation were obtained from the Survey of Education and Work conducted in May 2001 (ABS Cat. No. 6227.0). These estimates relate to individuals aged 15 to 64 and excludes people in remote parts of Australia (approx 80,000 people), people in most non-private dwellings (but includes boarding schools), and people who are permanently unable to work.

g. ABS estimates for smoking status were obtained from the 1995 National Health Survey (ABS Cat. No. 4364.0). These estimates relate to individuals aged 18 to 64 and includes people in non-private dwellings.

h. The ASHR was conducted between May 2001 and June 2002. The estimates relate to individuals aged 16 to 59, and exclude people in non-private dwellings. These estimates include people in remote parts of Australia along with those who are permanently unable to work.

Conclusion

This report has considered the magnitude and likely composition of the non-respondents to the ASHR. Based on this information, the design weights for the survey have been adjusted to external benchmarks. Even though the response rates achieved for the ASHR are relatively high for a voluntary telephone survey, there is still considerable scope for non-response bias. Comparison of ASHR estimates with ABS estimates suggest that the non-respondents are more likely to be young, born in non-English speaking countries, not in the labour force, and live in more populous areas. Adjusting the design weights to external benchmarks for State, sex and age seeks to correct for some of these non-response problems. It will still be important for users of the ASHR data be aware of the potential non-response bias and make appropriate allowances for this.

Appendix 1: Final outcomes for the male and female samples

Location	NSW	VIC	QLD	SA	WA	TAS	NT	ACT	Central Sydney	Gay Sydney	NSW Rural	VIC Rural	QLD Rural	WA rural 1	WA rural 2	Australia
MALE SAMPLE																
Summary outcomes																
Interviews	1456	2099	1257	501	517	525	510	503	503	1022	338	347	386	115	104	10183
Refusals/Incomplete ivws	718	1000	634	214	233	193	176	161	205	434	165	136	133	48	36	4486
Other non-response	64	153	37	10	15	16	30	12	39	118	10	14	4	6	1	529
Ineligible	2188	3024	1584	744	671	729	383	474	736	1784	487	512	402	60	67	13845
Eligibility not determined - language	144	201	37	28	33	17	18	21	119	86	7	10	8	0	2	731
Assumed out of scope (not HH)	1288	1932	773	320	389	292	528	349	586	2177	199	226	204	119	64	9446
Definite out of scope (not HH)	4137	4412	3051	1200	1445	1164	1452	1015	1697	4600	797	582	860	452	326	27190
TOTAL	**9995**	**12821**	**7373**	**3017**	**3303**	**2936**	**3097**	**2535**	**3885**	**10221**	**2003**	**1827**	**1997**	**800**	**600**	**66410**
Response rate – La Trobe	67.0%	67.7%	66.5%	70.1%	68.9%	73.1%	74.3%	75.8%	71.0%	70.2%	67.2%	71.8%	74.4%	70.6%	74.3%	69.4%
Response rate – Alternative	63.0%	62.5%	64.5%	67.8%	66.1%	70.7%	70.1%	73.1%	62.3%	63.3%	65.4%	69.1%	73.2%	68.0%	73.1%	65.4%
Non-HH rate	54.3%	49.5%	51.9%	50.4%	55.5%	49.6%	63.9%	53.8%	58.8%	66.3%	49.7%	44.2%	53.3%	71.4%	65.0%	55.2%
Eligibility yield	22.4%	25.4%	26.1%	24.0%	23.2%	25.0%	23.1%	26.7%	19.2%	15.4%	25.6%	27.2%	26.2%	21.1%	23.5%	22.9%
Ivw yield	14.6%	16.4%	17.0%	16.6%	15.7%	17.9%	16.5%	19.8%	12.9%	10.0%	16.9%	19.0%	19.3%	14.4%	17.3%	15.3%
FEMALE SAMPLE																
Summary outcomes																
Interviews	1454	2086	1295	531	505	509	503	506	514	-	335	357	335	108	104	9142
Refusals/Incomplete ivws	459	610	396	170	175	114	125	102	141	-	102	95	98	26	19	2632
Other non-response	32	82	34	10	4	8	10	8	33	-	4	11	3	4	0	243
Ineligible	1440	2179	1156	508	437	482	267	331	583	-	355	417	275	39	54	8523
Eligibility not determined - language	162	231	58	31	17	10	33	17	162	-	7	11	11	3	1	754
Assumed out of scope (not HH)	995	1415	621	274	322	225	368	240	527	-	133	185	165	68	56	5594
Definite out of scope (not HH)	3163	3511	2445	968	1133	850	1191	820	1428	-	673	437	658	351	364	17992
TOTAL	**7705**	**10114**	**6005**	**2492**	**2593**	**2198**	**2497**	**2024**	**3388**	**-**	**1609**	**1513**	**1545**	**599**	**598**	**44880**
Response rate – La Trobe	76.0%	77.4%	76.6%	75.7%	74.3%	81.7%	80.1%	83.2%	78.5%	-	76.7%	79.0%	77.4%	80.6%	84.6%	77.6%
Response rate – Alternative	71.3%	71.7%	73.6%	72.8%	72.7%	79.9%	76.1%	80.7%	66.3%	-	75.3%	76.2%	75.7%	77.0%	84.1%	73.4%
Non-HH rate	54.0%	48.7%	51.1%	49.8%	56.1%	48.9%	62.4%	52.4%	57.7%	-	50.1%	41.1%	53.3%	69.9%	70.2%	52.6%
Eligibility yield	25.2%	27.5%	28.7%	28.5%	26.4%	28.7%	25.6%	30.4%	20.3%	-	27.4%	30.6%	28.2%	23.0%	20.6%	26.8%
Ivw yield	18.9%	20.6%	21.6%	21.3%	19.5%	23.2%	20.1%	25.0%	15.2%	-	20.8%	23.6%	21.7%	18.0%	17.4%	20.4%