The Australia Institute
Research that matters.

What does $50,000 buy in a population survey?

Characteristics of internet survey participants compared with a random telephone sample

Technical Brief No.4
October 2009
ISSN 1836-9014

Hilary Bambrick, Josh Fear and Richard Denniss

TAI

Technical Brief

Summary

Traditional methods of obtaining a representative sample for survey research are expensive and time consuming, and may be no more representative of the general population than relatively new and cost-effective internet-based sampling techniques. As access to and use of the internet becomes near universal in Australia, internet surveys are likely to become the method of choice for researchers undertaking population-based surveys.

This paper compares key social and demographic characteristics of respondents recruited using two different survey methods: random-digit dialling (RDD) followed by a telephone interview (1024 participants), and internet panel recruitment followed by an online survey (1000 participants). Both surveys were of adults and were national in scope.

The phone survey cost approximately $56 000 and took more than a month to complete, while the internet survey cost approximately $6000 and took five days to complete.

Participants in the two surveys were compared by age, sex, state/territory, highest level of education attained, country of birth (Australia or overseas), employment status, household income, and place of residence (metropolitan, rural or remote). 'Community mindedness' was also assessed via questions asking whether respondents had donated blood and were registered organ donors.

The two samples were near identical on each of these characteristics. When compared with ABS Census statistics, participants in the internet survey tended to be younger than the national age profile, while participants in the phone survey were older. However, internet participants appeared more nationally representative by age than the phone participants, with many more having been born overseas, probably reflecting the slightly younger age profile of this sample.

We conclude that using the internet to recruit participants for survey research can produce a sample that is at least as representative as traditional phone surveys, and at a much lower cost.

Nevertheless, researchers should take care when selecting a sampling method. Generally speaking, internet-based recruitment is less appropriate in the following circumstances:

- when it is especially important to include the views of people belonging to certain population groups that tend to have lower levels of internet access (for example, welfare recipients, Indigenous people, the aged) or lower levels of literacy in written English (for example, people born overseas). These groups would, however, also be underrepresented in surveys using RDD.
- when survey topics or questions are likely to correlate with certain kinds of online behaviour, such as joining an online panel or spending unusually large amounts of time on the internet (for example, a survey on the uptake of social networking or downloading music).

In these situations, it may be advisable for researchers to source some or all respondents using alternative 'offline' methods. However, when it is less critical that the survey include population representation from lower-use groups, and when the topic is likely to be unrelated to internet use, online surveys are an affordable and effective research tool.

1. Introduction

Australia has strict privacy laws regarding the proper use of personal information, and it is becoming increasingly difficult to obtain a representative sample of the population for survey research. For example, the current national electoral roll is no longer readily available to researchers. To maximise population representativeness, large research surveys often recruit and survey participants with RDD, thus bypassing the need to use personal information to contact potential respondents.

Increasingly, the widespread availability and affordability of new communication technology such as mobile phones and the internet mean that fewer people are available on a landline phone number. This is especially true for younger people and those living in rental accommodation; landline penetration is now as low as 60 per cent among some groups.[1] Intensive telemarketing and overly lengthy phone surveys have also probably contributed to reduced rates of response to requests to participate in research by phone.

RDD is expensive and time-consuming, involving additional training for interviewers and numerous call-backs to unanswered phones. In any phone survey, the time taken to train interviewers and implement the survey is a major expense for researchers. Despite these limitations, RDD is considered the favoured method for producing a survey sample most representative of the population.

In contrast with telephone-based sampling, the internet has the potential to reach particular target populations readily and at lower cost, especially in high take-up countries such as Australia. In 2007–2008, 67 per cent of Australian households were connected to the internet, and 78 per cent of these were broadband connections.[2] If current trends continue, approximately 90 per cent of households will have an internet connection by 2011, and the vast majority of these will be broadband.

Commercial market researchers in Australia and overseas have wholeheartedly embraced internet-based surveys, but academics and other 'traditional' researchers (such as those working on behalf of government) continue to rely on older methods. This is perhaps as much the result of inertia as a deliberate and well-considered choice.

This paper compares the characteristics of respondents to a phone survey using RDD and respondents to an internet panel survey on the same topic. While the survey questions and the sample size were virtually identical, the cost differences were stark: around $6000 for the internet survey and $56 000 for the phone survey. If key sociodemographic characteristics of survey participants and the results of comparable attitudinal questions are sufficiently similar, there would appear to be little reason to opt for the more expensive approach, and a strong case for pursuing online surveys as a reliable and affordable option for social researchers.

This paper is not intended as a critique of any particular survey-fieldwork provider. The methods described and the costs involved will be comparable across the industry and even in other countries. Instead, we are using a specific survey—in this case on the topic of blood and organ donation—to make a more general point about sample representativeness and the time- and cost-effectiveness of different survey-sampling methods in today's internet-saturated society.

---

1 ACMA (Australian Communications and Media Authority), *Convergence and Communications Report 1: Australian household consumers' take-up and use of voice communications services*, Australian Government, Canberra, 2008.

2 ABS (Australian Bureau of Statistics), *Household Use of Information Technology, Australia, 2007–08: Household access to computers and the internet*, Cat. 8146.0, Canberra, 2008.

2. Methods

We recruited two national samples of adults into a survey of community attitudes towards blood and organ donation; 1024 participants were recruited through RDD and 1000 through an internet panel. The two surveys were conducted within two months of each other in late 2008.

Phone sampling

A professional survey company[3] contacted and surveyed the RDD phone sample using a method based on existing prefixes (the first six digits of a ten-digit phone number, including area code) from the electronic white pages in approximately the same proportion as the national distribution of phone numbers. Suffixes (the last four digits of each phone number) were then randomly generated. The sample of RDD numbers was compared with the electronic yellow pages, and numbers belonging to businesses were removed. The remaining sample contained silent and unlisted numbers, which would have been missed had the sample been based solely on the electronic white pages.

If a household contained more than one eligible person (a person aged over 18 years and living in the household), the potential respondent was randomly selected from all eligible participants living in that household; this selection process was guided by a Computer Assisted Telephone Interview (CATI) system.

When approximately half of the sample had been obtained, the survey company observed that young males were not participating at a level comparable to other demographic groups. Thus, for the remainder of the survey period the company over-sampled young males by changing the way the CATI system selected respondents in each household. The new sampling regime selected a respondent aged under 60 years at the rate of one in two interviews; a male respondent at the rate of one in four interviews and a respondent under 40 at the rate of one in seven interviews. The overall response rate was 65 per cent. The survey cost $51 364, used 26 interviewers, and took one month to complete.

Online sampling

Participants in the online survey sample of 1000 people were sourced from an online panel, in essence a collection of 'pre-recruited' respondents who have agreed to take part in online surveys from time to time in exchange for a small incentive.[4] When they join the panel, panellists are asked to answer a series of questions about themselves, including demographic topics, so that particular sub-groups can be targeted in any subsequent invitation to participate. Panellists are invited by email to take part in surveys and, if they fit the eligibility criteria and complete the survey, the incentive is credited to their account and can be redeemed for vouchers, movie tickets, competition entries and other small rewards.[5]

---

[3] The names of the companies used to carry out the surveys reported here are deliberately withheld as the focus of this paper is on comparing two survey methods rather than the actual results. Both companies (one phone-survey company and one online-panel company) carried out their work promptly and professionally and to an excellent standard.

[4] The level of incentive varies according to the length of the survey and the difficulty in finding members of the target group. Incentives for a standard population-based survey typically range from $1 to $5 for each complete questionnaire. Incentive costs (which accrue to respondents) are in addition to sampling costs (which accrue to the online-panel provider).

[5] There are now a number of high-quality online-panel providers operating in Australia, along with some poor-quality providers. Quality in this context refers to the way that panellists are managed (to ensure good response rates and engagement with the questionnaire), the way they are recruited (both online and offline, to

Participants in this survey were selected based on quotas corresponding to age, sex and state/territory. The quotas were derived from the adult Australian population, so the survey was by definition representative of the adult Australian population by age, by sex and by state/territory (but not all three in combination). It is not possible with this method of sampling to calculate a response rate, the proportion of people who received the invitation who completed the survey. However, because participants have pre-elected to be contacted to participate in surveys, the response rate is expected to be high.

While respondents were sourced externally from an online panel, The Australia Institute scripted and hosted the online survey using commercially available software. This brought the costs of fieldwork down considerably, to around $6000, or just over one-tenth of the cost of the telephone survey. It took approximately one week to format the survey questions appropriately, and another five days to collect the required number of responses.

The questionnaire

Each survey was roughly 10 minutes in length. While the majority of the questions were identical between the two surveys, there were some unavoidable small differences due to the manner in which questions are presented to telephone and online respondents.[6] For the sake of brevity, the results presented in this paper are confined to the most readily comparable, certain key sociodemographic characteristics elicited from questions that did not change between the two survey types. Selected attitudinal and behavioural questions that could be interpreted as relating to 'community-mindedness' are also compared in order to assess the likely degree of similarity between the two samples on non-sociodemographic factors.

3. Results

The two survey methods produced samples of participants that were very similar across a range of key sociodemographic criteria.

Sex

Both survey methods produced near identical proportions of male and female participants (Table 1), which held across the different states (Figure 1), demonstrating that the random phone survey, with its over-sampling among men, was representative by sex since the internet survey used quotas based on the actual population (proportion of adult men and women in each state).

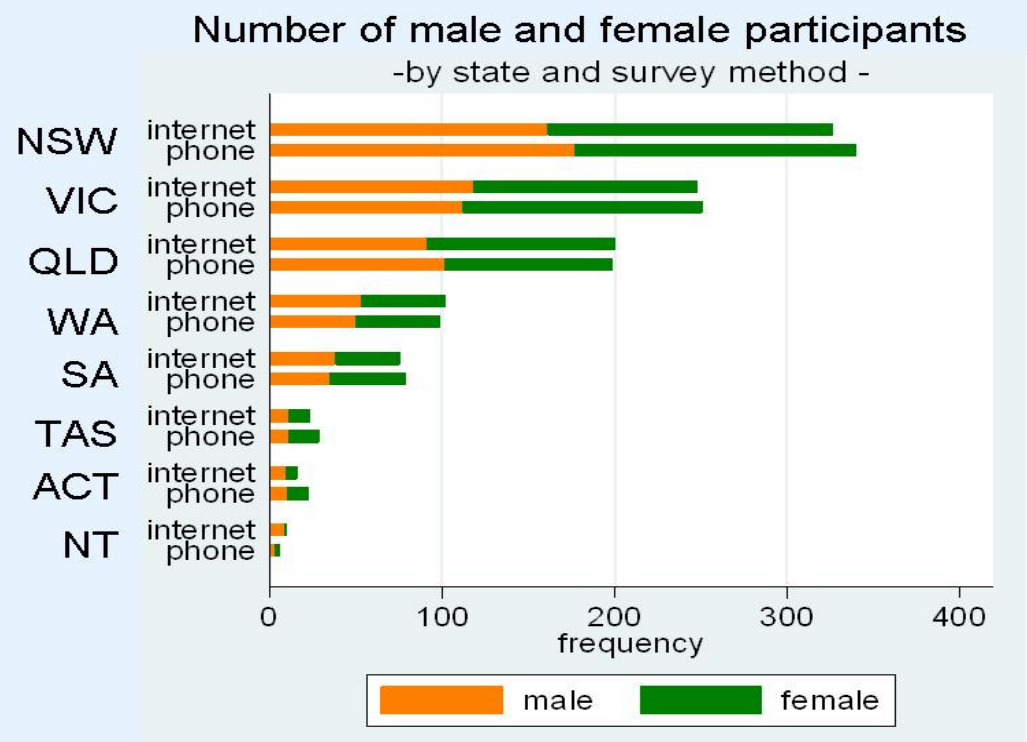
encourage greater representativeness), and whether the panel is used for any non-research activities (such as marketing).

[6] For example, a 'don't know' or 'refuse' option is usually not read out to respondents, even though it is available. If a 'don't know' option is available in an online survey, it needs to be presented to respondents.

Table 1: Participants by sex and survey method

Sex	Method						
	Internet		Phone		Total		Census[7]
	Number	%	Number	%	Number	%	%
Males	489	48.9	499	48.73	988	48.81	48.7
Females	511	51.1	525	51.27	1036	51.19	51.3
Total	1000		1024		2024		

Figure 1: Number of participants by state and survey method



Age

As might be expected, the youngest age group (18 to 24 years) was slightly overrepresented in the internet sample, while the oldest age group was slightly underrepresented (Figure 2). By contrast, those aged 35 to 54 years were overrepresented in the phone survey.[8,9]

---

[7] ABS (Australian Bureau of Statistics), *2006 National Census*, ABS, Canberra, 2006.

[8] Because quotas were used to achieve representativeness in the online survey, it might be expected that the final sample would exactly reflect the intended sample (which in this case is based on the population profile). In practice, quotas always fill up unevenly, and it becomes difficult to find respondents with the right characteristics to fill the remaining quotas as the survey draws to a close. Sometimes, the most efficient way to complete the survey is to relax incrementally some of the already-full quotas, which results in a final sample that is very close, but not identical, to the original quotas. The differences in intended versus actual quotas are usually only one or two respondents per quota; this means that the bigger the overall sample, the less of an influence such

Although quotas were used in the online survey to capture a predetermined number of respondents of each sex and age group, there were no additional quotas to control how many men and women *within* each age group took part.[10] Notwithstanding, the numbers of men and women in each age group in the respondent sample provided a good indication of the reach of the survey and the representativeness of each sample.

Figure 2: Number of participants by age and survey method

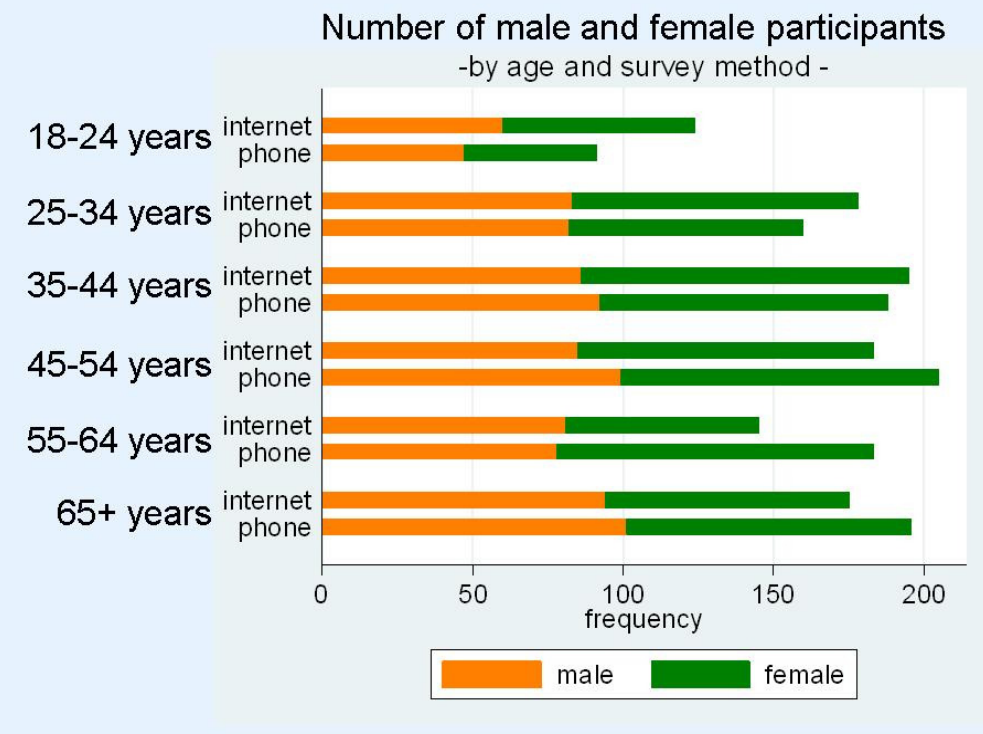


Table 2 shows the numbers of males and females by age group achieved in each sample and the numbers that would be expected if each survey had been truly representative of the Australian population as a whole. Figure 3 illustrates the differences between the numbers that were achieved and the numbers that would have been expected by age and sex in each survey method.

---

quota management techniques will have on representativeness. These observations apply equally to any survey method that employs a quota system, whether the sampling is conducted online, by phone or face-to-face.

9 *Overrepresentation* means having more respondents in a particular category than would be expected if the sample were truly representative of the wider population, while *underrepresentation* means having fewer respondents than expected. These differ from *over-* and *under-sampling*, which refer to the techniques used to select survey respondents prior to their participation.

10 In other words, the quotas were not 'interlocking'. Interlocking quotas in this context would mean recruiting a certain number of males and females *within* each age group, repeated in each state and territory. By contrast, the quotas in this survey were non-interlocking and involved recruiting a certain number of males and females, a certain number in each age group and a certain number in each state/territory, with no link between the different quotas.

Table 2: Expected participant numbers by age and sex based on the total Australian population compared with achieved numbers for both survey methods

		Internet sample		Phone sample	
Sex	Age group	Expected	Achieved	Expected	Achieved
Female	18 to 24	61	64	63	45
	25 to 34	90	95	92	78
	35 to 44	99	109	102	96
	45 to 54	93	98	95	106
	55 to 64	73	64	74	105
	65+	97	81	99	95
Male	18 to 24	63	60	65	47
	25 to 34	88	83	90	82
	35 to 44	95	86	98	92
	45 to 54	90	85	92	99
	55 to 64	73	81	74	78
	65+	79	94	80	101

Figures 3 and 4 below show how the patterns of over- and under-sampling differed between male and female respondents. Both surveys tended to underrepresent younger men and to overrepresent older men, with the degree of underrepresentation among younger men lower for the internet survey than for the phone survey, as was the degree of oversampling among older men.

The internet survey overrepresented younger women and underrepresented older women, while the opposite was true of the phone survey. As with the male sample, the degree of over and underrepresentation appears lower for the internet survey than for the phone survey.

In other words, differences between expected and achieved numbers by age and sex were generally less for the internet survey than for the phone survey. These comparisons indicate that *internet-based methods can provide a more representative sample by age and sex than phone surveys*, even in the absence of interlocking quotas to control the numbers of male and female respondents within in each age group.[11]

[11] See Footnote 7.

Figure 3: The difference for female participants between the achieved number of participants in each age and sex group and the expected number (based on the total Australian population)

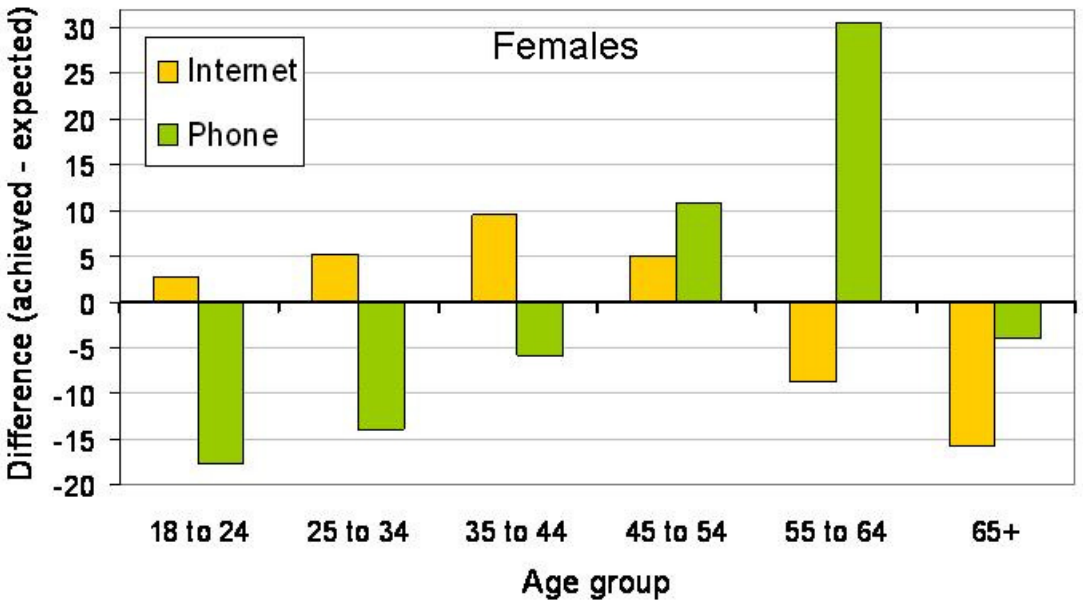
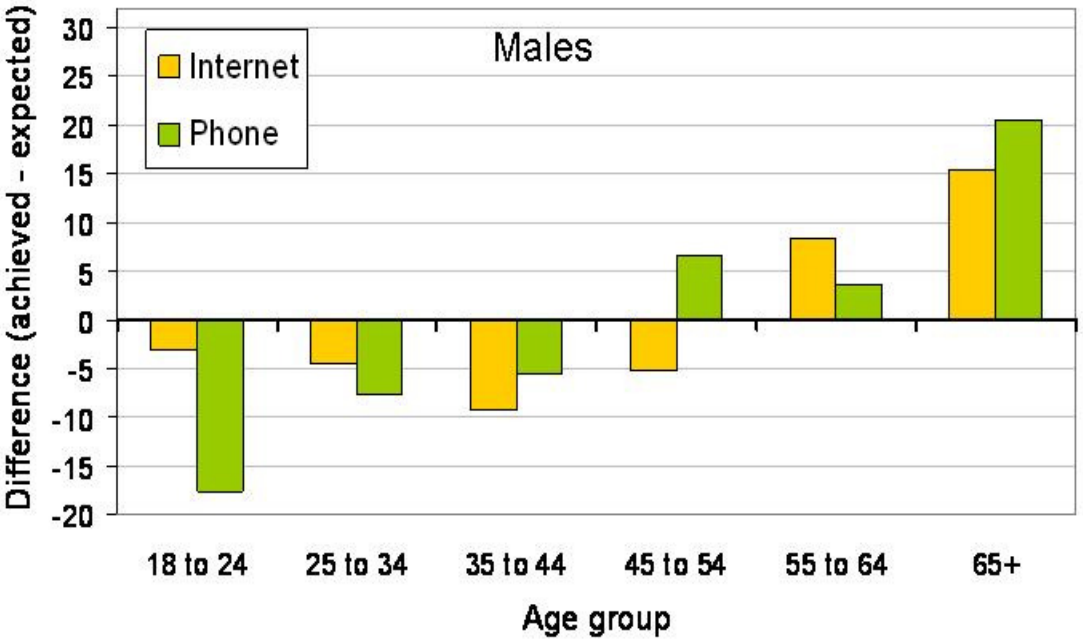


Figure 4: The difference for male participants between the achieved number of participants in each age and sex group and the expected number (based on the total Australian population)

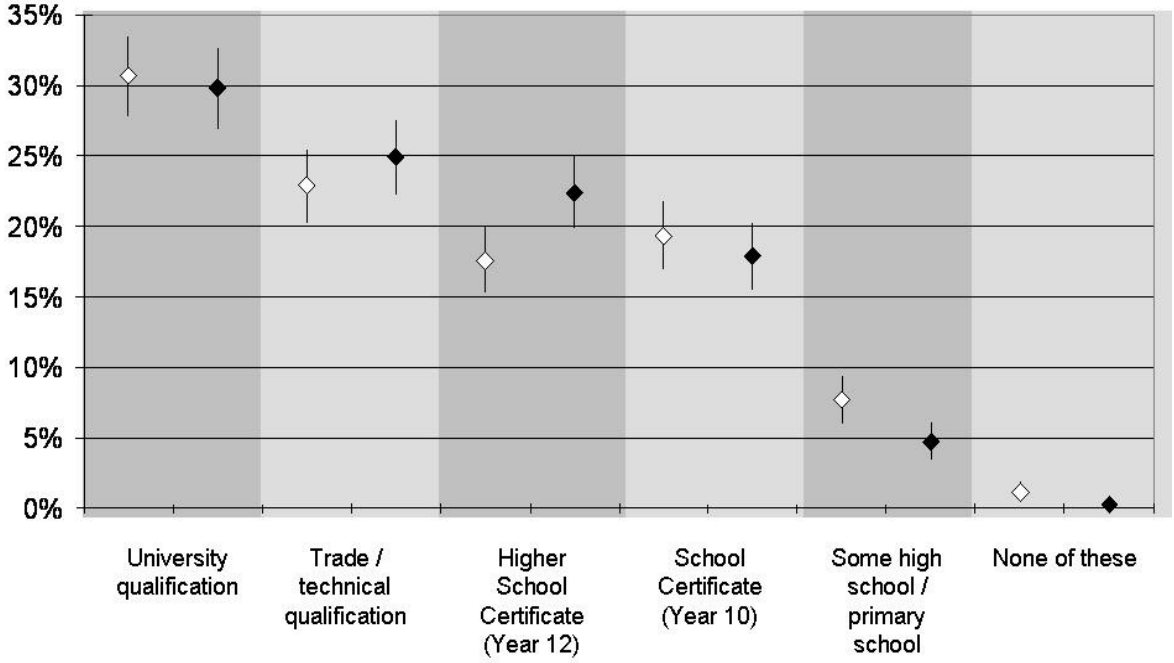


As with age and sex, participants in the two surveys were also very similar according to highest level of education reached, working status, household income, and whether they were from a metropolitan, rural or remote area (figures 5 to 8).

Education level

Participants in both surveys showed a very similar profile according to the highest level of education attained (Figure 5). Slightly fewer internet respondents had completed only some primary or secondary schooling and slightly more had completed Year 12, perhaps reflecting the somewhat younger age structure of the internet sample and the higher numbers of older women surveyed by phone. The younger people might be expected to mirror the related population changes in education attainment over time, while the older were possibly less likely to have completed their schooling.

Figure 5: Highest level of education by survey method showing the percentage of respondents reporting each level of educational attainment and the 95 per cent confidence intervals[12]



Note: The white markers refer to the phone survey participants, the black markers refer to the internet survey participants.

Country of birth

A higher proportion of internet-survey participants were born overseas than those in the phone survey (22 per cent versus 16 per cent), most likely reflecting the younger age profile

[12] Where the confidence intervals overlap between the two survey methods, there was no statistically significant difference.

of the internet sample. Of the total population usually resident in Australia (including children), approximately 29 per cent were born overseas.[13]

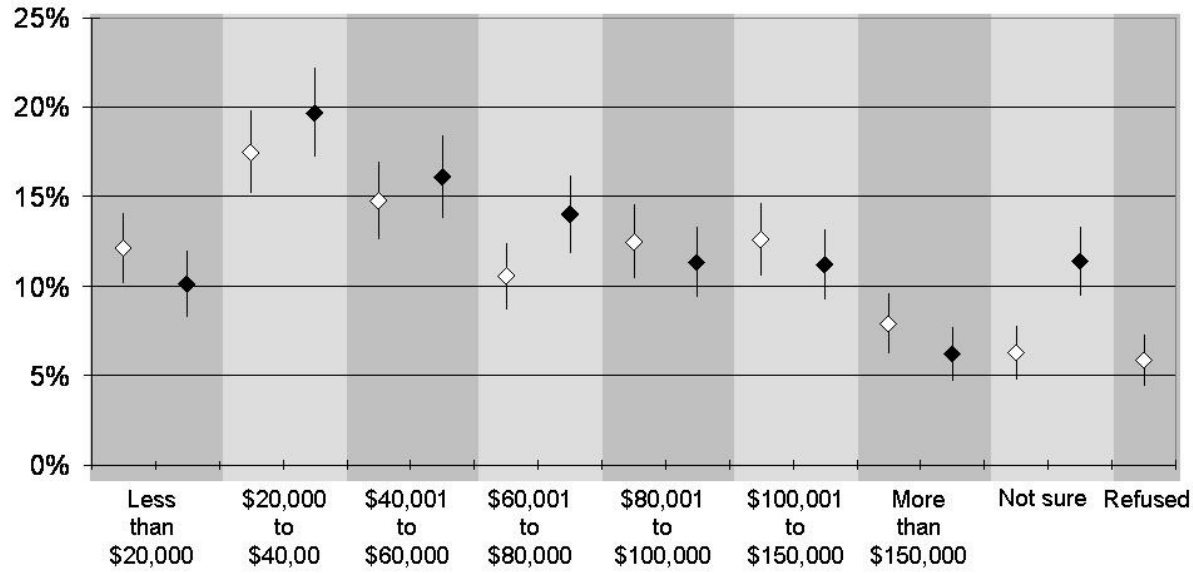
Employment

Similar numbers in each of the two surveys identified themselves as being in paid employment (62 per cent of internet participants, 60 per cent of phone participants). Full-time workers made up 38 per cent of the internet participants and 40 per cent of the phone participants, which compares well with national Census data. Overall, 57 per cent of people aged over 15 years in Australia are in paid employment and 37 per cent are employed full time (59 per cent and 39 per cent respectively among those aged 20 years and over).[14]

Household income

Nominated household income levels were also very similar across the two samples, with no real difference in the proportion of participants in each income category (Figure 6). The phone survey produced a small number of respondents who refused to indicate their income, an option that was not available to the internet survey participants. If the percentage of refusals and 'not sure' responses in the phone survey are combined, the result roughly corresponds to the percentage of 'not sure' responses in the internet survey. This outcome might be expected, since the response in the internet survey is likely to include people who would have refused to answer the income question if a refusal option had been available.

Figure 6: Household income by survey method showing the percentage of respondents in each income category and the 95 per cent confidence intervals[15]



Note: The white markers refer to the phone survey participants, the black markers refer to the internet survey participants. Note that the internet survey did not offer an option of 'refused'.

[13] ABS, *2006 National Census*.

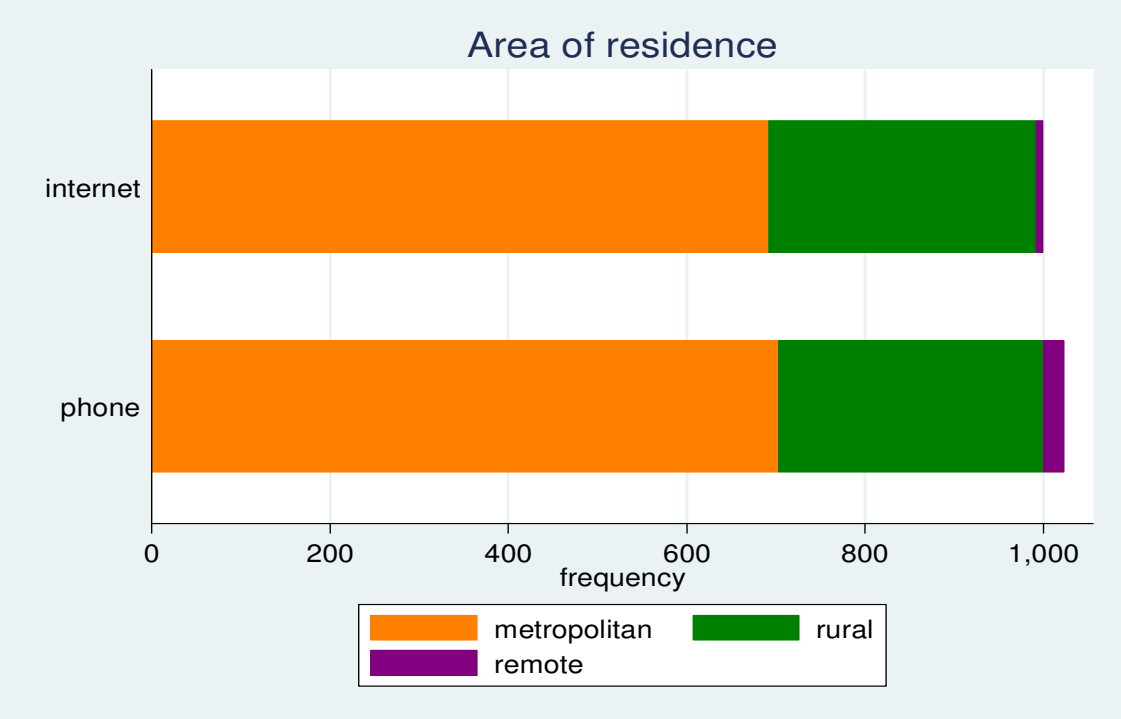
[14] ABS, *2006 National Census*.

[15] The overlapping confidence intervals between the percentages in the two surveys suggest that there was no real difference between them.

Area of residence

As well as being representative by state of residence (Figure 1, above), both surveys produced similar proportions of participants living in each of metropolitan, rural and remote areas (Figure 7). The proportion of survey participants living in metropolitan areas matched well with data for Australia overall (approximately 75 per cent of people live in metropolitan areas).[16]

Figure 7: Place of residence by survey method



The characteristics of the two survey samples reported above indicate that internet-based sampling can generate respondent profiles demonstrating key demographic measures very similar to those generated by RDD. However, there may still be differences between the two types of respondents that cannot be gauged by examining demographic variables. For this reason, it is worth comparing a number of other survey results of an attitudinal or behavioural nature across the two surveys.

'Community-mindedness': blood and organ donation

Similar proportions of participants in each survey reported that they had donated blood in the past (42 per cent of internet participants, 40 per cent of phone participants), and confirmed that they were registered organ donors (51 per cent in both surveys). These results suggest that the 'community-mindedness' of the two survey groups was similar but might reflect a willingness on the part of respondents to participate in surveys, whether by internet or phone.

[16] ABS (Australian Bureau of Statistics), 'Urban and Non-urban population', *Year Book Australia, 2006*, Cat. 1301.0, Canberra, 2007.

4. Discussion

In this paper, we have described the characteristics of two survey samples that were designed to be representative of the adult Australian population: a phone survey sampled through RDD, and a quota-based internet survey sampled through an online panel. The phone survey cost approximately $56 000—nearly 10 times the cost of the online survey, which was completed within a much shorter timeframe.

The two methods produced samples that were close to identical on key sociodemographic variables, including age, sex, state/territory, country of birth, working status, highest level of education, household income and area of residence. The comparison between the two methods demonstrates the power of internet-based sampling to reach a broad and reasonably representative sample on a number of key demographic factors, including factors for which quotas were not set.

Both the age and sex profiles of the two samples were also very similar to official national Census statistics. Quotas on age, sex and state/territory governed the internet-survey sample but these were not linked; the phone survey increased the sampling of men (especially younger men) part-way through to better represent the Australian population. The characteristics of online respondents were very similar to the broader population; in fact, on a number of measures the internet survey proved to be more representative of the Australian population than the random telephone survey. This suggests that imposing just a small number of quotas, even non-linked ones, can help to ensure that other demographic characteristics are quite representative.

The use of quotas to control the respondent profile is a powerful and reliable method of recruiting participants that exhibit a set of desired characteristics, not merely a sample that reflects the broader population. Quota systems can be implemented for both phone and internet surveys, but online quota management allows researchers conducting internet surveys to monitor and manipulate quotas in real time as they fill up. Although very basic demographic criteria were applied in this internet survey, quotas can be extended to precisely define the profile of potential participants by any number of variables. Those variables regarded by researchers as essential to the sampling process can be incorporated into a quota system as desired and invitations to participate sent only to people who possess the characteristics of interest. This level of control over the choice of participants is not available through RDD.

Compared to online sampling, recruiting participants though RDD is expensive and time consuming, and may be no more representative of the general population. Furthermore, any phone survey conducted by multiple interviewers will be subject to some level of interviewer bias (systematic error introduced by each interviewer's influence on response rates and results) that is unavoidable when more than one interviewer is involved, as is standard practice in a modern CATI phone room. While it is difficult to quantify the effects of interviewer bias (since its nature and extent depends on the actual phone room in question), it does present a profound problem for the theory of randomised and controlled survey administration. Conducting surveys online automatically removes any potential systematic errors introduced by multiple interviewers.

4.1 Limitations

Given the characteristics of our survey sample, it would seem that internet methods are an ideal way to carry out population surveys at much lower cost than phone surveys. However, internet surveys have a number of limitations, and there are circumstances in which their use is not advisable.

One of the most important considerations is the subject matter of the survey. If the topic is likely to correlate with the predilection of survey participants to sit on an online panel (or even to use the internet heavily), survey results will not be representative of the wider population, whether it is a population survey or otherwise. For example, if the survey asks questions about use of the internet, or attitudes to internet-based phenomena, care must be taken when selecting the survey method. Conversely, if the subject matter is likely to be independent of online behaviour (including the likelihood of participants to be part of an online panel), there is no reason to avoid the less expensive and more efficient internet-based survey option.

A second consideration is the probability or otherwise that target groups will have online access. Internet access in regional and rural areas (where broadband is unavailable) is poorer than in the cities and thus internet surveys may yield fewer participants from these regions unless they are specifically targeted in the sampling quota. Nevertheless, the internet is used widely in remote regions for education and communication, and our comparison found the number of respondents living in rural and remote Australia to be similar in both surveys. Furthermore, older people tend to use the internet less than younger people, but again age can be added as a criterion for sampling. However, if the involvement of key population groups that tend to have lower levels of internet access is critical to the success of the survey, it is advisable to target these groups by another method. This would likely be the case for Indigenous people, people on welfare or with a disability, those in institutionalised care and other disadvantaged groups. It is probable that these groups would also be harder to reach by RDD and thus would be underrepresented in phone surveys as well.

Nevertheless, our experience indicates that panel-based internet surveys produce results that are no less valid than relatively expensive and time-consuming phone surveys. It may also be that non-response to certain questions is reduced compared with other survey methods, since respondents may feel more comfortable giving feedback in a less personal setting.

5. Conclusions

As landline penetration declines and household internet access rates increase, online survey research is being used more and more by social and (especially) market researchers. Online research is comparatively inexpensive compared with telephone-based surveys, and for population-based surveys quite affordable.

Although telephone surveys are still regarded by some researchers as the ‘gold standard’ in population research, online sampling can now challenge telephone sampling in representing the broader population via the use of good quality online research panels to source respondents. This is especially the case following the introduction of privacy laws governing the matching of electronic lists.

On key demographic characteristics, surveys conducted using internet panels may in fact be more representative of the Australian population than surveys that use RDD or other, more traditional methods. We expect that online survey sampling will become even more representative as access to the internet, particularly in rural and remote areas, improves and these technologies become further incorporated into people’s daily lives.

Acknowledgements

The School of Medicine, University of Western Sydney, provided funding for the phone survey.

References

ABS (Australian Bureau of Statistics), (2006). *2006 National Census*, Canberra.

ABS (Australian Bureau of Statistics), (2007). 'Urban and Non-urban population', *Year Book Australia, 2006*, Cat. 1301.0, Canberra.

ABS (Australian Bureau of Statistics), (2008). *Household Use of Information Technology, Australia, 2007-08: Household access to computers and the internet*, Cat. 8146.0, Canberra.

ACMA (Australian Communications and Media Authority), (2008). *Convergence and Communications Report 1: Australian household consumers' take-up and use of voice communications services*, Australian Government, Canberra.