Assessment of non-response bias in an Internet survey of alcohol use

Kypros Kypri, PhD[1]*
Shaun Stephenson, BSc (Hons)[2]
John Langley, PhD[2]

[1]School of Medical Practice and Population Health
University of Newcastle
Newcastle, Australia

[2]Injury Prevention Research Unit
Department of Preventive and Social Medicine
University of Otago
Dunedin, New Zealand

* Author to whom correspondence should be addressed
c/o Hunter Population Health
Locked Bag 10 Wallsend NSW 2287
Australia

E-mail: kypros.kypri@hunter.health.nsw.gov.au
Tel: +61 2 4924 6205
Fax: +61 2 4924 6215

Abstract

Background: Falling survey response rates are a growing concern in epidemiological research, principally because prevalence estimates may be biased by selective non-response. Internet-based methods have the potential to yield higher quality data with lower non-response, and at lower cost than traditional methods. Little research exists on non-response bias in Internet surveys of alcohol use. The present investigation draws on a study of the implementation of an Internet-based alcohol survey involving a random sample of 1,910 university students with a response rate of 82% (N=1,564).

Aims: To identify non-response bias and to quantify its impact on estimates of alcohol consumption, the incidence of alcohol-related problems, and the prevalence of hazardous drinking.

Methods: Survey non-response has been characterized in terms of a *continuum of resistance* model, which assumes that individuals who respond only after considerable time and effort, i.e., late responders, will resemble non-respondents in the behaviors of interest. Two methods were used to test this model: (1) comparison of the demographic characteristics of the target sample with those of the respondents, and (2) comparison of alcohol variables for those who responded late, with those who responded early.

Results: The results attained using method 1 showed that bias varied as a function of gender, age, ethnicity, and living arrangement. The results attained using method 2 show that the incidence of alcohol-related problems and hazardous drinking prevalence varied as a function of response latency. If only the early and intermediate respondents had participated, the incidence of alcohol-related problems and the prevalence of hazardous drinking would each have been under-estimated by 3%.

Conclusions: The findings reported here are consistent with the continuum of resistance model, but show that the bias resulting from non-response is arguably too small to be of concern with respect to estimating consumption levels, the incidence of alcohol-related problems, and the prevalence of hazardous drinking.

Key words: alcohol, survey, Internet, student, non-response

Running head: Non-response bias in an Internet survey of alcohol use

A trend of falling response rates in population surveys is a growing concern in epidemiological research, principally because prevalence estimates may be biased by selective non-response (Caetano, 2001). In alcohol research, survey-based under-estimates of aggregate consumption have frequently been assumed to be a result of selective non-participation or under-reporting by heavy drinkers. A review of the literature reveals contradictory findings: in some studies, non-respondents have been found to be, on average, heavier drinkers than respondents (Van Loon et al., 2003; Wild et al., 2001), while in others, abstainers are over-represented among non-respondents (Lahaut et al., 2002; Lemmens et al., 1988). Either result biases estimates of total consumption and the prevalence of hazardous drinking.

New information technologies, in particular, the World-Wide Web, have created the potential for advances in the methods used to conduct survey research. Although the uptake of Internet access has grown exponentially in recent years (Couper, 2000), present levels of coverage are still insufficient to warrant reliance on Internet-based methods for general household surveys. For segments of the population with high levels of Internet access, e.g., college students, Internet-based methods have the potential to yield higher quality data with lower non-response, and at lower cost than traditional methods (Dillman, 2000; Schmidt, 1997).

Survey non-response has been characterized in terms of a *continuum of resistance* model (Lin and Schaeffer, 1995), in which the propensity of sampled individuals to respond is inferred from the level of effort required to elicit their survey response. The model assumes that those who respond only after considerable time and effort (*late respondents*), resemble non-respondents in the behaviors of interest.

The continuum of resistance model gives rise to two main methods of assessing non-response bias. The first of these involves comparison of the

demographic characteristics of the population of interest (or at least of the total target sample), with those of the respondents. Where differences are found, parameter estimates can be adjusted for the distributions of the demographic variables in the population. The utility of this method is low where demographic variables are poorly related to the behavior of interest. In the case of alcohol, however, a variety of demographic variables, most notably gender, age, ethnicity, and residential setting, have been shown to be related to consumption and the experience of problems (Kypri et al., 2002a; Wechsler et al., 1995).

A less commonly used method of assessing non-response bias is to compare those who respond late to a survey, with those who respond early, in terms of the behaviors of interest. Late response is variably defined, but is generally operationalised as either a response after considerable time has elapsed from the point of first contact, or after multiple efforts on the part of the researcher to elicit a response. Parameter estimates for all of the respondents can be subtracted from those for the early respondents alone to produce an estimate of non-response bias.

The assumption underlying comparison of early versus late survey respondents, is that late respondents are near to non-respondents on the continuum of resistance. Non-response bias would therefore tend to be under-estimated using this approach, and the extent of under-estimation will be positively associated with the proportion of non-response.

The present investigation follows two studies of the implementation of an Internet-based alcohol survey involving random samples of university students (Kypri and Gallagher, 2003; Kypri et al, under review) resulting in response rates of 85% (N=150) and 82% (N=1,564) respectively. The aims of this paper are to identify non-response bias in the larger of those studies (Kypri et al., under review), and to quantify the impact of bias on estimates of

alcohol consumption, the incidence of alcohol-related problems, and the prevalence of hazardous drinking.

Method

Sample

The study sample consisted of 1910 students aged 16-29 years, randomly selected from the enrolment database of the University of Otago, Dunedin, New Zealand. New Zealand is a country of 4 million people, with an average annual alcohol consumption of 9.8 liters per capita, compared with 9.1 liters per capita in the USA (World Health Organization, 2003). Like the USA, New Zealand inherited its dominant drinking culture from northern Europe, and accordingly there exists a strong preference for beer over other forms of beverage alcohol (Smart, 1998). Wine and spirit-based drinks have increased in popularity in recent years, particularly among young people, but beer remains the most popular beverage in both countries.

The University of Otago is a public institution, with around 17,000 students, making it the third largest of eight universities in New Zealand. Further information about the university can be obtained at www.otago.ac.nz. The south island city of Dunedin, where the main campus is situated, has a population of 120,000 people. Further information about the city can be obtained at www.dunedin.co.nz.

Students of Maori ethnicity (the indigenous population of New Zealand), who make up 7.3% of this population, were over-sampled by a factor of two, in order to achieve greater explanatory power for Maori, previously identified as at elevated risk of hazardous drinking (Kypri et al., 2002a; Kypri et al., 2002b). The target sample thus included 276 students of Maori descent and 1,634 students of non-Maori descent.

Recruitment

Participants were recruited in a three-phase procedure:

Phase 1: Invitation. A personally addressed and signed letter on university letterhead was mailed to each sampled student (day 1), inviting them to participate in a confidential Alcohol Use Survey (AUS) via the web. The letter notified the recipients that in three days time an e-mail message would be sent to their student e-mail address, and that a hypertext link contained within the message, when clicked, would open their computer's web browser at the site hosting the survey. Attached to the brief letter was a token incentive (a ball-point pen) and an information sheet approved by the University ethics committee, with details of the study. The e-mail invitation was generated using mail merge software so that each message was personally addressed and sent individually rather than as part of a bulk transmission. This message was sent two days after the letter was expected to arrive at students' in-term residential addresses (day 3), so that they had had the opportunity to read the letter and to see the token incentive. Attached to the e-mail was an electronic copy of the mailed letter and the information sheet describing the research. A hypertext link e-mail address of the researcher, and their postal address and phone number were prominently displayed to allow recipients to make inquiries about the study.

Embedded in the e-mailed hypertext link was a unique identifier (equivalent to a password), which was linked to the sampled individual's identifying information (provided by the university) in a separate database. The hypertext link could be used only once to complete the web questionnaire. This prevented the student or any other person over-writing the data on a subsequent occasion.

Phase 2: Reminders. One week after the first e-mail was sent (day 10), the survey database was checked to determine if the student had responded. A polite personalized reminder e-mail was sent to students who had not yet responded. This also contained a second hypertext link to the web questionnaire in case the previous e-mail had been deleted. This was

supplemented with a letter posted to the student's residential address (day 11).

Phase 3: Intensive follow-up. Ten days after the reminder e-mail (day 20), the survey database was checked again. Students who had not yet responded (and had not indicated refusal to participate) were telephoned to check that they had received the e-mail and asked if they were willing to participate. Up to five follow-up telephone calls were allowed for in an attempt to make contact. Those who wished to participate but preferred not to use a computer were offered a pen-and-paper alternative. Data collection was stopped 56 days after posting of the invitation letter.

The three phases of recruitment directly reflect the different levels of effort required to elicit a response, i.e., they are non-arbitrary categories. In the analysis, students who responded in phase 1 (days 1-9) , 2 (days 10-19), or 3 (days 20-56), are referred to as *early respondents, intermediate respondents,* and *late respondents* respectively.

Survey instrument and measures

The survey instrument consisted of a series of web-pages linked to a relational database. Participants were asked to point and click responses to a range of measures, a full list of which is provided elsewhere (Kypri and Gallagher, 2003). Measures used in the analyses presented in this study include the Alcohol Use Disorders Identification Test (Saunders et al., 1993); a 7-day retrospective diary, in which the number of standard drinks (defined as 10 g ethanol) consumed on each day and the duration of the drinking session was recorded; and the Alcohol Problems Scale (Kypri, 2003) a checklist of 14 alcohol-related consequences experienced in the preceding 3 months. Examples of items in this scale were: vomiting, unprotected sex, theft, and arrests for drunken behavior. The entire questionnaire can be viewed at http://ipru.otago.ac.nz/ausdemo. Survey completion time was computed by

subtracting the start time from the finish time as recorded on the server hosting the questionnaire. The median completion time for the web questionnaire was 16.7 minutes (inter-quartile range: 13.6 to 20.5 minutes).

Response rates

Response rates were computed for four groups of sampled individuals: (1) those who completed the entire web questionnaire or the pen-and-paper form, (2) those who fulfilled the core data requirement, i.e., by completing all items up to and including the AUDIT but not the entire questionnaire, (3) those who provided some data but did not meet the core data requirement, and (4) those who did not provide any data. For the purpose of computing an overall response rate, groups 1 and 2 were classified as respondents and groups 3 and 4 were classified as non-respondents.

Analysis of non-response bias

Two methods were used to quantify non-response bias.

Method 1 – Comparison of respondents versus non-respondents. Non-response was examined as a function of four demographic variables: gender, age (16-17 years, 18-19 years, 20-21 years, 22-24 years, 25-29 years), ethnicity (European, Maori, Asian, Pacific Islands, Other), and residential setting (house sharing, hall or college, parents' home, and other). Where the distribution of a demographic variable for respondents differed from the distribution in the sample (indicated by a significant chi square test), the parameter estimate was weighted to reflect the distribution of the demographic variable in the sample.

Method 2 – Comparison of early, intermediate and late respondents.

Mean AUDIT scores for early, intermediate, and late respondents were computed. Estimates of alcohol consumption (in g ethanol), mean APS score, and hazardous drinking prevalence (% scoring 8+ on the AUDIT) with and without the late respondents, were compared.

The results obtained from both methods were used to produce weighted estimates of the extent of any bias with respect to three key indicators: population alcohol consumption (from the drinking diary), the incidence of alcohol-related problems (from scores on the APS), and the prevalence of hazardous drinking.

Results

Survey response

Of 1,910 randomly sampled students 331 (17.3%) did not respond at all. In eight cases (0.4%), the sampled individual started the survey but did not proceed far enough to meet the core data requirement. Seven cases (0.4%) with evidence of a response set or other significant inconsistencies were considered invalid and were reclassified as non-respondents. The total non-response was thus 18.1%. Complete responses were received from 1,520 students (79.6%). In a further 44 cases (2.3%), although the entire survey was not completed, the core data requirement was met. Thus, 1,564 (1520 + 44) students responded, producing an 81.9% (95% CI: 80.1, 83.6) overall response rate.

Temporal pattern of response

Of the 1,564 responses received, 828 were received between days 1 and 9 of the survey (phase 1: early), 436 were received between days 10 and 19 (phase 2: intermediate), and 300 were received between days 20 and 56 (phase 3: late).

Assessment of response bias

Method 1: Comparison of respondents versus non-respondents. Table 1 presents the gender, age, ethnicity, and living arrangement distributions for the respondents (N=1,564) and the sample (N=1910). There were small but statistically significant differences on all four demographic variables.

<Table 1>

Method 2 – Comparison of early, intermediate, and late respondents. Mean AUDIT scores (and 95% CIs) for early, intermediate, and late respondents, were 10.4 (9.9, 10.8), 10.4 (9.8, 11.1), and 11.5 (10.6, 12.3), respectively. Given that mean scores for early and intermediate respondents were effectively the same, they were grouped into a single category for further analysis. Table 2 presents summary statistics for alcohol consumption, alcohol-related problems, and the prevalence of hazardous drinking, for all the respondents, and separately for early/intermediate respondents, and late respondents.

<Table 2>

Extent of the bias

The results attained using method 1 showed evidence of small biases in the distributions of response to the survey as a function of the demographic variables. Accordingly, the distribution of each demographic variable was used to produce weighted estimates of alcohol consumption, the incidence of alcohol-related problems (APS scores), and the prevalence of hazardous drinking (percentage scoring 8+ on the AUDIT), derived from the respondents. From table 3 it can be seen that the effect of non-response bias was less than 2% in either direction.

<Table 3>

The results attained using method 2 (see Table 2) show that alcohol consumption, APS scores, and hazardous drinking prevalence varied as a function of response latency. If only the early and intermediate respondents had participated in the survey, weekly alcohol consumption would have been over-estimated by 3% [1- (105g / 102 g)]; while the incidence of alcohol-

related problems and the prevalence of hazardous drinking would have been under-estimated by 3% [1- (2.78 / 2.88)] and 3% [1- (61.1% / 62.7%)] respectively.

Discussion

Two methods of analysis were used to study non-response bias. The first showed that men in particular were less inclined than were women, to participate in the Alcohol Use Survey. Notwithstanding some small differences between the sampled and respondent distributions of students by age, ethnicity, and residential setting, one could reasonably conclude that the respondents were otherwise representative of the Otago University student population aged 16-29 years. Based on observed associations of the alcohol measures and each demographic variable among the survey respondents, consumption was under-estimated by up to 2%, while the prevalence of hazardous drinking was estimated to within $\pm$0.5%.

The second approach showed that late respondents, i.e., those who participated only after an intensive recruitment effort, were, on average, heavier drinkers than those who responded early or in the intermediate period. The validity of this approach rests on the assumption of the continuum of resistance model: that late respondents would not have participated without the additional recruitment effort, and therefore that their drinking is a good indicator of non-respondents' drinking. Using this method, in the present study, both the incidence of alcohol-related problems and the prevalence of hazardous drinking appear to have been under-estimated, albeit by small amounts.

One apparently contradictory finding was that alcohol consumption measured with a 7-day retrospective diary, was lower (although not significantly so) among late respondents than among early/intermediate respondents. In contrast, AUDIT scores were highest (although not

significantly so), for the late respondents. This discrepancy is probably an artifact of measurement, the result of cyclical variation in drinking levels across the academic year. Late respondents' diary reports were based on 7-day periods in May and early June, closer to the commencement of first semester examinations, than those of the early/intermediate respondents. Although there is no published evidence from New Zealand on the cyclical variation in university student alcohol consumption, it is widely believed within the university community that alcohol consumption decreases in the period immediately prior to examinations. For items 1-3 of the AUDIT, no reference period is stated, instead estimates of the individual's typical consumption are requested. For items 4-10, the reference period is 1 year. These features make the AUDIT superior to the 7-day diary as an indicator of and individual's typical drinking habits. In light of this, it is unlikely that the observed lower alcohol consumption among late respondents reflects the true difference in consumption for those who responded late versus early.

Assessment of non-response bias has inherent limitations. These are highlighted by Gliksman and colleagues (2000), with respect to their large survey of Canadian college students:

> "There can be no definitive guidelines regarding acceptable response rates because the potential bias is based on both the extent of non-response and the difference between respondents and non-respondents, the latter of which we cannot know. Thus, even surveys with response rates as high as 70% and 80% can have sizeable bias, while those with low response rates can have low bias if the difference between respondents and non-respondents is minimal" (2000, p.9).

Gliksman et al. assessed for non-response bias by comparing early and late respondents (method 2), finding that "…although some alcohol measures differed by response time, none of the alcohol problem measures, drug-use measures nor mental health measures differed significantly" (2000, p.8). Their analytical methods are, however, not presented, and the level of bias is not

specified. Furthermore, it is not clear whether any analyses comparing respondents and non-respondents (method 1) were conducted.

A limitation of using method 2 for Gliksman et al.'s data is that the total response to their survey was only 51%, such that the assumption implicit in this method – that late respondents resemble non-respondents – is less likely to be correct, than it would be for a survey with a higher total response. Furthermore, although non-response may be under-estimated in the present study, its potential effect on consumption and prevalence estimates is considerably lower than for surveys with lower overall response.

The findings are consistent with the continuum of resistance model, however, their generalizability is debatable. This was an Internet-based survey of students at a university which draws its students from throughout New Zealand and some other countries. These individuals represent considerable demographic and cultural diversity but are unlikely to completely reflect the survey response behavior of the general population. Arguably, they share more in common with North American or European university and college students. The study is presented as an exemplar of a survey method in which the impact of non-response bias is too small to be of significant concern.

A practical implication of the findings is that the most intensive phase of follow-up, that involving telephone reminders of non-respondents (phase 3), is probably not warranted in future surveys. Limiting the recruitment effort to phases 1 and 2, while probably under-estimating incidence and prevalence rates to a small extent, would free up resources allowing an increase in the target sample size, and a large reduction in the data collection period (from 56 to 20 days). These features make the Internet-based approach increasingly attractive as a means of conducting alcohol surveys.

Acknowledgements

This research was funded by the Health Research Council of New Zealand and the Alcohol Advisory Council of New Zealand. We are grateful to Dr. Dorothy Begg for helpful comments on a draft of this paper.

References

Caetano, R. (2001) Non-response in alcohol and drug surveys: a research topic in need of further attention. *Addiction* **96**, 1541-5.

Couper, M. (2000) Web surveys: a review of issues and approaches. *Public Opin Q* **64**, 464-94.

Dillman, D. A. (2000) *Mail and internet surveys : the tailored design method*. J. Wiley, New York.

Gliksman, L., Demers, A., Adlaf, E. M., Newton-Taylor, B. and Schmidt, K. (2000) *Canadian Campus Survey 1998*. Centre for Addiction and Mental Health, Toronto, Canada.

Kypri, K. (2003) Tertiary student hazardous drinking: Epidemiology and development of a brief intervention trial. PhD thesis, University of Otago, Dunedin.

Kypri, K. and Gallagher, S. J. (2003) Incentives to increase participation in an internet survey of alcohol use: a controlled experiment. *Alcohol and Alcoholism* **38**.

Kypri, K., Gallagher, S. J. and Cashell-Smith, M. (under review) An Internet survey method for studying college student drinking.

Kypri, K., Langley, J. D., McGee, R., Saunders, J. B. and Williams, S. (2002a) High prevalence, persistent hazardous drinking in New Zealand tertiary students. *Alcohol and Alcoholism* **37**, 457-64.

Kypri, K., McGee, R., Saunders, J. B., Langley, J. D. and Dean, J. I. (2002b) Interpretation of items on the AUDIT questionnaire. *Alcohol and Alcoholism* **37**, 465-7.

Lahaut, V. M., Jansen, H. A., van de Mheen, D. and Garretsen, H. F. (2002) Non-response bias in a sample survey on alcohol consumption. *Alcohol Alcohol* **37**, 256-60.

Lemmens, P. H., Tan, E. S. and Knibbe, R. A. (1988) Bias due to non-response in a Dutch survey on alcohol consumption. *Br J Addict* **83**, 1069-77.

Lin, I.-F. and Schaeffer, N. C. (1995) Using survey participants to estimate the impact of nonparticipation. *Public Opinion Quarterly* **59**, 236-258.

Saunders, J. B., Aasland, O. G., Babor, T. F., de la Fuente, J. R. and Grant, M. (1993) Development of the Alcohol Use Disorders Identification Test (AUDIT): World Health Organisation Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption--II. *Addiction* **88**, 791-804.

Schmidt, W. C. (1997) World-Wide Web survey research: Benefits, potential problems, and solutions. *Behavior Research Methods, Instruments, & Computers* **29**, 274-279.

Van Loon, A. J., Tijhuis, M., Picavet, H. S., Surtees, P. G. and Ormel, J. (2003) Survey Non-response in the Netherlands. Effects on Prevalence Estimates and Associations. *Ann Epidemiol* **13**, 105-10.

Wechsler, H., Dowdall, G. W., Davenport, A. and Castillo, S. (1995) Correlates of college student binge drinking. *American Journal of Public Health* **85**, 921-926.

Wild, T. C., Cunningham, J. and Adlaf, E. (2001) Nonresponse in a follow-up to a representative telephone survey of adult drinkers. *J Stud Alcohol* **62**, 257-61.

Table 1. Demographic distributions of the sample and respondents

	% of sample* (N=1,910)	% of respondents* (N=1,564)	(95% CI)	$\chi 2$	p-value
Gender					
Female	55	58	(55, 60)	23.7	<0.01
Male	45	42	(40, 45)		
Age					
16-17	2	2	(2, 3)	10.3	0.04
18-19	38	40	(37, 42)		
20-21	36	35	(32, 37)		
22-24	15	15	(13, 17)		
25-29	9	9	(7, 10)		
Ethnicity					
Asian	12	12	(10, 14)	12.4	0.02
European	68	70	(68, 72)		
Maori	14	14	(12, 16)		
Other	3	3	(2, 4)		
Pacific Islands	2	2	(1, 2)		
Residential setting					
Sharing	65	64	(62, 67)	14.1	0.03
Hall of residence	20	20	(18, 22)		
Parents' home	8	9	(8, 10)		
Other	7	6	(5, 8)		

* Percentages do not always sum to 100 because of rounding

Table 2. Alcohol consumption, alcohol-related problems, and hazardous drinking prevalence, by response latency

	Grams ethanol in preceding week*		**APS score***		**Hazardous drinking prevalence**	
	(Mean)	(95% CI)	(Mean)	(95% CI)	(%)	(95% CI)
All	102	(96, 109)	2.88	(2.75, 3.00)	62.7	(60.2, 65.1)
Early / intermediate	105	(98, 112)	2.78	(2.64, 2.92)	61.1	(58.3, 63.8)
Late	90	(75, 105)	3.32	(3.01, 3.63)	69.3	(63.8, 74.5)

* Fourteen respondents who did not complete all required fields in the drinking diary, and 38 who did not complete all items in the APS were not included in the computation of estimates for these measures.

Table 3. Unweighted and weighted estimates of alcohol consumption, incidence of alcohol-related problems, and hazardous drinking prevalence

	Grams ethanol in preceding week* (mean)	**APS score** (mean)	**Hazardous Drinking Prevalence** (%)
Unweighted	102	2.88	62.7
Weighted for:			
Gender	104	2.90	63.0
Age	102	2.87	62.5
Ethnicity	101	2.84	62.2
Residential setting	103	2.88	62.8