Kypros Kypri, Amanda Wilson, John Attia, Paschal Sheeran, Peter Miller, and Jim McCambridge; Social desirability bias in the reporting of alcohol consumption: a randomized trial, Journal of Studies on Alcohol and Drugs Vol. 77, Issue 3, p. 526-531 (2016)

Available from: http://dx.doi.org/10.15288/jsad.2016.77.526

Social desirability bias in the reporting of alcohol consumption: A randomised trial

Kypros Kypri PhD*[1]

Amanda Wilson PhD[1]

John Attia PhD[1]

Paschal Sheeran PhD[2]

Peter Miller PhD[3]

Jim McCambridge PhD[4]

[1]Centre for Clinical Epidemiology & Biostatistics, School of Medicine & Public Health, University of Newcastle, Australia

[2]Department of Psychology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

[3]School of Psychology, Deakin University, Geelong, Australia

[4]Department of Health Sciences, University of York, Great Britain

* School of Medicine and Public Health
Room 4104 HMRI Building
University of Newcastle, New Lambton Heights, NSW 2305 Australia
kypros.kypri@newcastle.edu.au +61 2 4042 0536
(Fax +61 2 4042 0044)

Words: 2128 (Abstract 270)

Conflicts of interest

None.

Funding

The study was funded with an Australian Research Council Discovery Grant (DP1093809). KK's contribution to the research was funded by a National Health & Medical Research Council Senior Research Fellowship (APP1041867) and a University of Newcastle Senior Brawn Research Fellowship.

Authors' contributions

JM and KK conceived of and designed the trial. AW conducted the pilot study and developed trial materials and procedures under the direction of KK. AW and PM implemented the trial. KK led the drafting of the manuscript with input from all authors.

Trial Registration

Australian New Zealand Clinical Trials Registry, ACTRN12611000254998 (http://www.anzctr.org.au/)

Key words: social desirability bias; reactivity; behaviour; randomised controlled trial; methodology; AUDIT; context

Abstract

Objective: In order to investigate reporting of alcohol consumption, we manipulated the contexts of questions in ways designed to induce social desirability bias.

Method: We undertook a two-arm, parallel group, individually randomised trial at an Australian public university. Students were recruited by e-mail to a web-based "Research Project on Student Health Behaviour". Respondents answered nine questions about their physical activity, diet and smoking and were unknowingly randomised to a group presented with either (A) three questions about their alcohol consumption, or (B) seven alcohol dependence and problems questions under a prominent header "Alcohol Use Disorders Identification Test", followed by the same three alcohol consumption questions.

Results: 3594 students (mean age 27, SD 10) responded and were randomised: 1778 to group A; and 1816 to group B. Outcome measures were the number of days they drank alcohol, the typical number of drinks they consumed per drinking day, and number of days they consumed ≥6 drinks. The primary analysis included participants with any alcohol consumption in the preceding four weeks (1304 in group A; 1340 in group B) using between group, two-tailed t-tests.

Results: In groups A and B, respectively, means (and SDs) of the number of days drinking were: 5.89 (5.92) versus 6.06 (6.12), p=0.49; typical number of drinks per drinking day: 4.02 (3.87) versus 3.82 (3.76), p=0.17; and number of days consuming ≥6 drinks: 1.69 (2.94) versus 1.67 (3.25), p=0.56.

Discussion: We could not reject the null hypothesis because earlier questions about alcohol dependence and problems showed no sign of biasing respondents' subsequent reports of alcohol consumption. These data support the validity of university students' reporting of alcohol consumption in web-based studies.

Introduction

Behavioural science increasingly relies on study participants to report on their own behaviour rather than on direct observation by researchers or other objective measurement (Baumeister et al., 2007). It is, however, widely accepted that self-report of sensitive information, such as one's use of psychoactive substances, including alcohol, is subject to bias (Dillman, 2007).

The inclusion or exclusion of particular questions, and the order in which they are presented can profoundly influence survey responses [see, e.g., (Sudman et al., 1996), for a review]. For instance, Lasorsa (2003) observed that self-reported attention to news reports about an upcoming election was 21% lower when participants answered questions about their political knowledge and then indicated their attention to the news, as compared to reporting their news attention first and then their political knowledge (Lasorsa, 2003). Participants used their answers to the first (political knowledge) question to infer how much attention they paid to the news ("I don't seem to be very knowledgeable, so I probably don't give the news much attention!"). Such *context effects* in surveys are problematic for researchers wishing to obtain valid reports of attitudes and behaviour.

Context effects encompass or at least overlap conceptually with the *socially desirable response set*, in which study participants report behaviour in ways they think will be viewed more favourably, or less unfavourably, by researchers (Crowne & Marlowe, 1960). Many studies in the alcohol field have used scales measuring socially desirable response set to assess and adjust for such tendencies. However, Babor, Brown and Del Boca (1990) cite evidence that heavy drinkers tend to score highly on measures of sociopathy and, conversely, that non-heavey drinkers tend to be more conventional. They argue that social desirability scales merely reflect these stable individual characteristics related to drinking levels, and

therefore that statistical adjustment of participant reports risks introducing rather than reducing measurement error. Experimental studies are needed to help overcome the limitations of observational study designs, to determine whether misreporting can be induced.

The Alcohol Use Disorders Identification Test (AUDIT) is perhaps the most widely used self-report instrument for alcohol use and contains three sub-scales pertaining to alcohol consumption, dependence symptoms, and other harmful consequences (Saunders et al., 1993). Given the widespread use of the AUDIT for screening in research, it is important to determine whether asking questions about dependence symptoms and harmful consequences could bias subsequent data. Here we present an experiment investigating whether study participants under-report their alcohol use if they first answer questions about dependence symptoms and harmful consequences.

Methods

Design and setting

We conducted a parallel groups, individually randomised trial (Figure 1) in a large public university in Australia.

Procedure

Approximately 24,000 students, without restriction by age or enrolment type, were invited to complete a web survey in a message sent to their student e-mail address by the university administration, with the subject line “Study of Student Lifestyles”. The message contained a hyperlink to a webpage with a Participant Information Sheet attached and a button to start the survey. Participants were then presented with five pages including 13 questions about their

demographic characteristics and four behavioural domains, taken from a national university student survey (Kypri et al., 2010).

Page 1 ("Your Demographic Details") asked participants their age, gender, what type of accommodation they lived in, and which university they attended. Page 2 ("Physical Activity") asked four questions about physical activity; page 3 ("Diet") asked three questions about diet, page 4 ("Smoking") asked a question about smoking, and page 5 ("Recent Drinking") asked "Have you had a drink containing alcohol in the past 4 weeks?". Participants were then randomised by the web server without their knowledge to group A or B.

Blinding

We adopted this procedure to ensure that participants were blind to the true nature of the study which was presented as a survey in order to minimise the potential for research participation effects (McCambridge et al., 2014). Procedures were computerised and automated ensuring that the researchers were blind to randomization, allocation, and assessment of outcome. The use of deception (the study was not, in fact, about student lifestyles) and blinding were considered ethical (McCambridge et al., 2013) given the low risk to participants and likely benefits of quantifying a phenomenon that may bias many studies in the behavioural and health sciences (McCambridge et al., 2014). The University of Newcastle's Human Research Ethics Committee approved the study.

Interventions

Group A (the control group) were shown a page without a title containing the three questions used to assess the study outcomes:

1. On how many days in the last 4 weeks did you have a drink containing alcohol? (Response options: 1, 2, 3…28)
2. How many Standard Drinks containing alcohol did you have on a typical day when you were drinking in the last 4 weeks (Please refer to the Standard Drinks guide on the right)? (Response options: 1, 2, 3…24, 25-29, 30-34, 35-39, 40-49, 50 or more)
3. On how many days in the last 4 weeks did you have 6 or more Standard Drinks on one occasion?

Group B (the experimental group) were asked the dependence symptoms and harmful consequences questions from the AUDIT, as shown in Figure 2, with that title also fully presented in red capital letters. They were then presented with the same page of three questions used to assess the study outcomes, shown above.

Pilot research

We pilot tested the web instrument and procedures with 17 volunteer university students (9 men and 8 women) who completed the questionnaire and then provided feedback in a face-to-face interview, using an iterative procedure for developing web survey and intervention material (Hallett et al., 2009). The purpose was to check whether the experimental manipulation had been effected, namely, that participants presented with AUDIT items 4-10 (Figure 2) would feel some discomfort in answering, but not consider it peculiar to be asked such questions in the context of such a study. Three questions directly concerned the experimental manipulation:

> Do you think the option 'prefer not to answer' is necessary for any of the questions?
>
> What did you think of the introduction to Section C [containing the AUDIT]?
>
> Do you think the language was judgemental?

Participants were generally positive about the survey but some felt discomfort answering questions relating to the alcohol problems items (the experimental manipulation). Eight

participants noted that they thought other students might find the questions confronting but they themselves did not. Four said they felt the language was not judgemental, one said it was “a little judgemental” and that “I was aware it was about my drinking”, while three said it might be confronting for other students.

Several participants commented on the particular question about needing a drink first thing in the morning after a night drinking (Figure 2): e.g., “I didn't know there were people like that” and “I wouldn't want people to know if I did”. The AUDIT questions were widely seen as judgementally framed, with the word 'failed', in particular, perceived as subjective and loaded.

We concluded that the material achieved the aims of: (1) having face validity; (2) not being so long as to evoke suspicion about the true motives of the research; and (3) being likely to make respondents feel mildly uncomfortable and on guard about the possibility that they might be negatively evaluated on the basis of their answers.

Outcome measurement

The three pre-determined co-primary outcomes were the number of days drinking in the preceding four weeks (possible range 1-28), the typical number of drinks per drinking day (a count variable starting at 1), and the number of days they consumed ≥6 drinks (possible range 0-28).

Sample size estimation

The study was powered to detect a difference of 0.10 standard deviation units (i.e., a Cohen’s *d* of 0.10), i.e., a very small effect on mean weekly alcohol consumption. On the conservative

assumption of a 15% response fraction, we expected 3632 participants, of whom approximately 2904 (80%) would have consumed alcohol in the preceding four weeks (1452 individuals per group). As the outcome measures were taken within the survey (and not at a subsequent session) we expected close to 100% completion. 1400 individuals per group provides power of 0.80, with an alpha of 0.05, and a two-sided test, to detect the hypothesised difference.

Analysis

The primary outcomes were analyzed using t-tests which are robust to skew in the outcome when the sample size is large. For the primary analysis, and according to an analysis plan developed before the study, we included in the t-tests only participants who reported drinking in the four weeks preceding baseline. As secondary analyses we conducted t-tests for comparison of all randomized participants, and chi-squared tests for the proportion in each group who indicated any drinking in the preceding four weeks.

Results

Of 3594 individuals who completed the survey and were randomised, 2644 reported consuming alcohol in the preceding 4 weeks and were included in the primary analysis. Two thirds (66%) of participants were women and the mean age was 27 years (SD 10). The proportions in groups A (73.3%) and B (73.8%) who reported any drinking in the preceding four weeks did not differ significantly (p=0.76).

Primary analyses

Drinkers in groups A (n=1304) and B (n=1340) reported that they had a mean (SD) of 5.89 (5.92) versus 6.06 (6.12) drinking days in the preceding four weeks (t=-0.94, p=0.49); 4.02

(3.87) versus 3.82 (3.76) drinks per typical occasion (t=1.38, p=0.17), and 1.69 (2.94) versus 1.67 (3.25) days in which they had ≥6 drinks, respectively.

Secondary analyses

Participants (including non-drinkers) in groups A (n=1778) and B (n=1816) reported that they had a mean (SD) of 4.32 (5.70) versus 4.47 (5.89) drinking days in the preceding four weeks (t=-0.76, p=0.49); 2.68 (3.99) versus 2.55 (3.86) drinks per typical occasion (t=0.98, p=0.33), and 0.97 (2.79) versus 0.97 (3.03) days in which they had ≥6 drinks, respectively.

Discussion

This study aimed to determine whether answering questions about alcohol dependence symptoms and other problems makes respondents under-report their drinking. Participants who answered such questions (group B) reported the same of level of alcohol consumption on each of the three measures in both the primary analyses (drinkers only) and secondary analyses (all participants) compared to participants who were not presented with these questions (group A).

Given the randomised design in a large sample, the only factors that could have influenced responses to consumption questions pertained to the inclusion of the full AUDIT title and questions about dependence symptoms and harmful consequences presented prior to those about consumption. The fact that this experimental manipulation had no influence on participants' reports of their consumption demonstrates that participants were not knowingly or unwittingly influenced by the prior questions.

Participants were under the impression they were answering general questions about aspects of their health and well-being. The questions regarding diet, smoking and exercise were selected to support this impression and were not in themselves of interest to the researchers. Upon completion of the study, participants were sent an e-mail link to a debriefing page describing the study aims and procedures and explaining the need for concealment of the true nature of the study (McCambridge et al., 2012).

The study generates the hypothesis that use of the AUDIT for screening prior to assessing alcohol-related issues in more depth does not bias self-reports of alcohol consumption. This hypothesis needs testing in population groups other than university students and in samples with higher initial response fractions.

We hypothesised under-reporting arising from the context of a health-focused study in which participants might have given responses they felt were more in line with low-risk drinking. However, the weak evidence base in this area makes a competing hypothesis plausible too: that having been presented with questions about alcohol use disorders, respondents in the experimental group were more likely to accurately report their drinking because they did not want to misrepresent themselves in such a serious context. The findings of this experimental comparison show unequivocally that there was no main effect of the context we created and that it is possible to study such phenomena in ways that permit strong inferences concerning effect size and causation.

Although alcohol consumption was the target behaviour for the purpose of this study, the type of experimental manipulation used here has wider relevance. Developing and evaluating interventions with small effects at the individual level–that produce a benefit when

aggregated across a large population–is a key challenge for the health sciences. If self-reports are subject to even small biases, estimates of small intervention effects, e.g., of web-based alcohol interventions (Kypri et al., 2014), may be compromised. The findings of this study suggest that students' reports of their drinking may be sufficiently valid to constrain context effects. The possibility that such biases may operate in studies of psychoactive substance use and other health behaviours deserves further investigation using experimental designs.

Acknowledgements

The study was funded with an Australian Research Council Discovery Grant (DP1093809). KK's contribution to the research was funded by a National Health & Medical Research Council Senior Research Fellowship (APP1041867) and a University of Newcastle Senior Brawn Research Fellowship. We are grateful to Dean Pung for programming the web pages and study database.

References

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the Science of Self-Reports and Finger Movements. *Perspectives on Psychological Science, 2*(4), 396-403.

Babor, T.F., Brown, J., & Del Boca, F. Validity of self-reports in applied research on addictive behaviors: Fact or Fiction? *Behavioral Assessment* 12:5-31, 1990.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354.

Dillman, D. A. (2007). *Mail and internet surveys : the tailored design method* (2nd ed.). Hoboken, N.J.: Wiley.

Hallett, J., Maycock, B., Kypri, K., Howat, P., & McManus, A. (2009). Development of a Web-based alcohol intervention for university students: processes and challenges. *Drug Alcohol Rev, 28*(1), 31-39. doi: DAR008 [pii]
10.1111/j.1465-3362.2008.00008.x

Kypri, K., Paschall, M. J., Langley, J. D., Baxter, J., & Bourdeau, B. (2010). The role of drinking locations in university student drinking: findings from a national web-based survey. *Drug Alcohol Depend, 111*(1-2), 38-43. doi: S0376-8716(10)00156-0 [pii]
10.1016/j.drugalcdep.2010.03.018

Kypri, K., Vater, T., Bowe, S. J., Saunders, J. B., Cunningham, J. A., Horton, N. J., & McCambridge, J. (2014). Web-based alcohol screening and brief intervention for university students: a randomized trial. *JAMA, 311*(12), 1218-1224. doi: 10.1001/jama.2014.2138

Lasorsa, D. L. (2003). Question-order effects in surveys: The case of political interest, news attention, and knowledge. *Journalism & Mass Communication Quarterly, 80*(3), 499-512.

McCambridge, J., Kypri, K., Bendtsen, P., & Porter, J. (2013). The use of deception in public health behavioral intervention trials: a case study of three online alcohol trials. *Am J Bioeth, 13*(11), 39-47. doi: 10.1080/15265161.2013.839751

McCambridge, J., Kypri, K., & Elbourne, D. (2014). Research participation effects: a skeleton in the methodological cupboard. *J Clin Epidemiol, 67*(8), 845-849. doi: 10.1016/j.jclinepi.2014.03.002

McCambridge, J., Kypri, K., & Wilson, A. (2012). How should debriefing be undertaken in web-based studies? Findings from a randomized controlled trial. *J Med Internet Res, 14*(6), e157. doi: 10.2196/jmir.2186
v14i6e157 [pii]

Midanik, L. (1982). The validity of self-reported alcohol comsumption and alcohol problems: A literature review. *British Journal of Addiction, 77*, 357-382.

Saunders, J. B., Aasland, O. G., Babor, T. F., de la Fuente, J. R., & 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*(6), 791-804.

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology* San Francisco: Jossey-Bass Publishers.

Wilson, P. (1981). Improving the methodology of drinking surveys. *Journal of th Royal Statistical Society, 30*(3), 159-167.

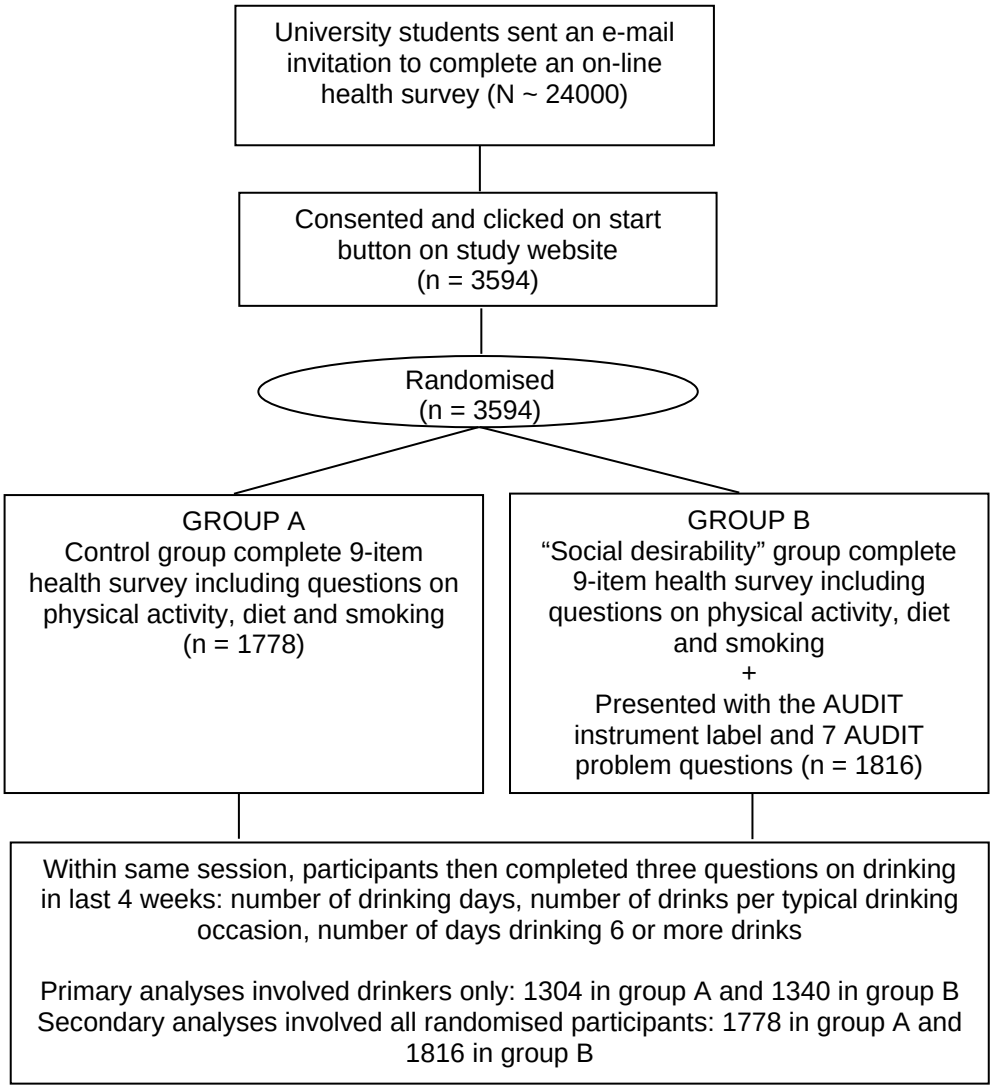

Figure 1. Trial flow diagram

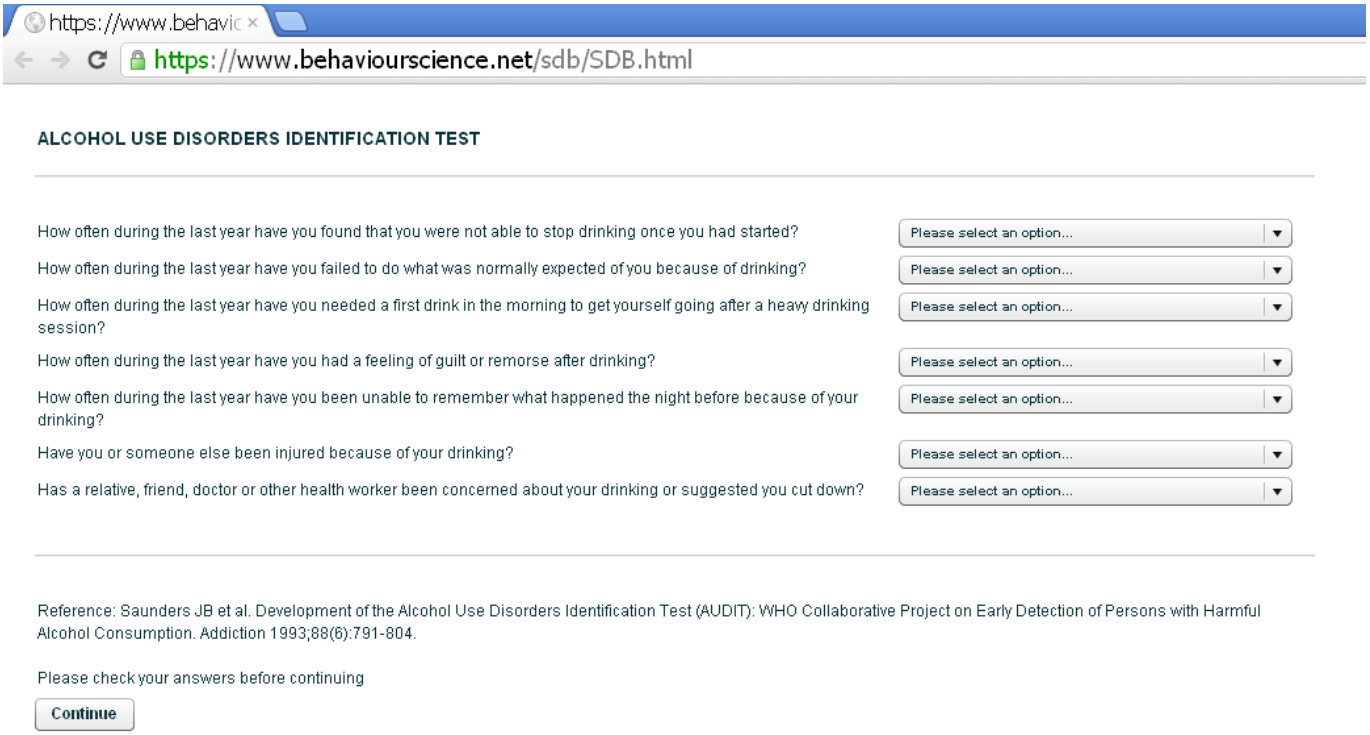

Figure 2. Web-page presented to group B to induce social desirability bias