Charles Darwin University

Effects of question type and order when measuring peak consumption of risky drinking events

Yeung, Jun Ting; Livingston, Michael; Callinan, Sarah; Wright, Cassandra; Kuntsche, Emmanuel; Room, Robin; Dietze, Paul

OXFORD

Article

Effects of Question Type and Order When Measuring Peak Consumption of Risky Drinking Events

Jun-Ting Yeung[iD][1], **Michael Livingston**[iD][2], **Sarah Callinan**[iD][2], **Cassandra Wright**[iD][1,2], **Emmanuel Kuntsche**[iD][2], **Robin Room**[iD][2,3] and **Paul Dietze**[iD][1,*]

[1]Burnet Institute, Melbourne, VIC, Australia , [2]Centre for Alcohol Policy Research, La Trobe University, Bundoora, VIC, Australia , and [3]Centre for Social Research on Alcohol and Drugs, Department of Public Health Sciences, Stockholm University, Stockholm, Sweden

*Corresponding author: Burnet Institute, 85 Commercial Road, Melbourne, VIC 3004, Australia. Tel. +61 3 9282 2134; Fax: +61 3 9282 2100; E-mail: paul.dietze@burnet.edu.au

Abstract

Aims: There is new interest in measuring alcohol consumption during risky drinking events, but there is little guidance on how to best ask such questions. In this study, we contrast two different types of questions on peak consumption over a single heavy drinking occasion. We used a general question that ask respondents to recall the total amount consumed (total consumption question), and location-specific questions that ask respondents to recall consumption in each drinking location (location-specific peak consumption, LSPC).

Methods: Heavy drinkers (≥11 Australian Standard Drinks (ASD) per occasion for males, ≥8 for females) from the second wave of a prospective cohort study were recruited via landline random digit dial from Melbourne in 2012. Respondents were randomly assigned to surveys of different question order, and either first received total consumption ($n = 127$) or LSPC questions ($n = 147$). T-tests compared peak consumption between categories stratified by sex and consumption tercile.

Results: Mean peak consumption was 12.5 ASD. Irrespective of question order, consumption amounts for total consumption and LSPC questions were not significantly different for both sexes. However, drinkers in the highest tercile asked LSPC questions first provided significantly higher consumption estimates in response to the total consumption question than in response to the LSPC questions.

Conclusion: At a population level, LSPC and total consumption questions produce similar estimates of peak consumption for risky drinking events. Except for heavy drinkers, general consumption questions may be sufficient when asking about these drinking events in consumption surveys, without the greater response burden of longer LSPC questions.

INTRODUCTION

Accurate assessment tools are vital in understanding patterns of alcohol consumption and related harms in a given population (Gmel and Rehm, 2004; WHO, 2014). Since self-report data are a fundamental tool in alcohol epidemiology, it is important to be aware of the issues that impact self-reported estimates of alcohol consumption. Underreporting of alcohol consumption is a well-known phenomenon; population survey measurements of alcohol consumption are often lower than the levels reported from alcohol sales data. For example, self-report amounts have been estimated to account for 40–60%

(Midanik, 1982) or 50–77% (Stockwell *et al.*, 2004) of total alcohol sales.

Many factors are likely to underpin under-reporting of alcohol consumption in surveys, including recall bias resulting from poor memory recall (Stockwell *et al.*, 2004), social desirability bias (Davis *et al.*, 2010), compromised brain function (Elofson *et al.*, 2013) and lack of knowledge of standard drink sizes (White *et al.*, 2005). Wilson's (1981) analysis provides a comparative framework in understanding this problem, raising issues such as the discrepancies that arise from the under-representation of heavy drinkers from sampling error and seasonal variations in alcohol consumption. Specific context-oriented questions such as those developed by Clark (1981), for example, aim to improve recall by providing context cues of past drinking occasions. Consistent with theories of memory more broadly, it is argued that the more specific the question, the more helpful it is with recall (Strube, 1987). Surveys using location- and beverage-specific context reinstatement cues show elevated levels of alcohol consumption compared to general questions (Mooney and Gramling, 1991; Wyllie et al., 1994; Casswell *et al.*, 2002; Livingston and Callinan, 2015). It has also been shown to reduce the risk of under-reporting of alcohol consumption (Krosnick, 1999). This anchoring effect of asking about drinking occurring in a variety of contexts is thought to provide a more accurate self-report response as respondents are prompted about features of their drinking occasions that they would have otherwise missed (Single and Wortley, 1994).

Another approach to improving the measurement of alcohol consumption is to focus on consumption amounts during heavy drinking occasions. Drinking patterns of young people typically involve particular occasions of heavy drinking often resulting in experiencing of various forms of harm (Gmel *et al.*, 2011; Kuntsche and Gmel, 2013). By including the amounts consumed over atypical and special occasions, the amount of self-reported consumption increased among light drinkers by over 200% (Bellis *et al.*, 2015). Increased amounts of alcohol consumed on single occasions have been shown to increase lifetime risk of alcohol-related disease, injury and health harms (Rehm *et al.*, 2008; Gmel *et al.*, 2011). Other methods for improving recall accuracy in different contexts have also been noted: providing respondents with information about standard drink sizes (White *et al.*, 2005), recent recall using graduated-frequency and quantity-frequency methods (Stockwell *et al.*, 2004), timeline followback diary measures (Werch, 2009) and ecological momentary assessment (Wright *et al.*, 2016).

One particular measure of interest is peak consumption, which we define as the highest level of alcohol consumption over one occasion in the most recent 12 months. Peak consumption is an example of an atypical consumption event that is important in studying acute harms from alcohol (NIAAA, 2003) and has shown to have good predictive ability in relation to alcohol use disorders (Greenfield *et al.*, 2006; Gmel *et al.*, 2011).

Moreover, question order effects, widely studied in other fields such as psychology (Strack, 1992) and market research (Bradburn and Mason, 1964), have been neglected in alcohol consumption surveys. Placement of questions within questionnaires has been shown to affect responses. For example, the reported levels of citizen satisfaction in a public service survey were significantly different depending on question order in the questionnaire (Walle and Ryzin, 2011). However, in the field of alcohol epidemiology, few studies have examined the effect of question order on self-report estimates.

In one study, Harford (1994) surveyed the frequency of monthly heavy drinking using a general summary question and quantity-specific questions and found that the reported prevalence of heavy drinking was significantly lower when the general question was asked after the quantity-specific questions.

These results indicate how seemingly minor methodological changes, such as the structure of the questionnaire, can impact survey measurement of alcohol consumption. It is important to assess how question type impacts self-reports of alcohol consumption (e.g. context-specific questions compared to general consumption questions). Similarly, it is important to determine any other effects such as question positioning in surveys or question order on self-reporting estimates.

We are unaware of any other studies of question type or of question order relevant to peak consumption. In this study, we compared two self-report estimates of peak consumption, measured in Australian Standard Drinks (1 ASD = 10 g ethanol), derived from two question types: a general approach (a simple question on total consumption) and a specific approach (location-specific peak consumption, LSPC). We also manipulated the structure of the questionnaire in two ways: (a) questions varied between questionnaires (i.e. different respondents were presented different versions of the questionnaire, Version [A] and Version [B]), and (b) questions varied within subject (i.e. within the same questionnaire, different types of questions were asked to the individual respondent to derive peak consumption). This manipulation enabled us to investigate the effect of question order by comparing self-report estimates between questionnaire, and the effect of question type by comparing self-report estimates given by the same respondent (termed within-subject). We were also able to examine whether the two types of questions gave varying estimates of peak consumption across sex and consumption levels.

On the basis of previous research on context-oriented questions (Mooney and Gramling, 1991; Casswell *et al.*, 2002; Livingston and Callinan, 2015), we expected that the context-specific cues provided by the LSPC questions would yield greater self-report consumption estimates than those obtained from the single question on total amount consumed (irrespective of any question order effect).

We also expected a question order effect; we expected that the total consumption question would yield greater self-report estimates when asked after the LSPC questions as respondents will have been primed to report greater amounts, because of context-specific cues, than when the total consumption question precedes the LSPC where they will not have been primed prior to the question being asked. Finally, we examined patterns of effects according to different amounts consumed on the drinking occasions, to examine whether effects were most pronounced in relation to the heaviest drinkers, by categorizing drinkers on the basis of the amounts consumed by participants on their most recent heavy drinking occasion.

METHODS

Study design

Our study draws on data collected as part of the Young Adults and Alcohol Study (YAAS), a longitudinal prospective cohort study examining risky single occasion drinking in young Australians (Dietze *et al.*, 2017). Participants were recruited in 2012 with follow-ups in 2013, 2015 and 2016. We analysed data collected from the first follow-up in 2013 where participants were randomly assigned to two groups: [A] Total-first and [B] LSPC-first survey conditions.

Sample and recruitment

In 2012, a representative sample of 802 young Melburnians was recruited by a contracted survey fieldwork provider from a sampling frame of landline telephone numbers (Dietze *et al.*, 2014). Contacted households, from local government areas in Melbourne, were screened for anyone aged between 18 and 25 years, with the next-birthday method used if there was more than one individual within this age group. Potential participants were further screened for whether they engaged in a heavy drinking occasion (defined as ≥11 ASDs for males, ≥8 ASDs for females), at least once in the past year.

In 2013, 531 (66% of 802) baseline participants responded to follow-up interviews, of which 82 (15% of 531) were excluded because of missing data, 'do not know' or refused responses. The number of participants with valid responses was 449. Further details on the methods have been described elsewhere (Dietze *et al.*, 2014, 2017).

Survey questions

Respondents were administered a structured questionnaire that included sections on drinking variables and harms, familial and other drinking contexts, drinking consequences, satisfaction with life and demographics. Respondents were provided with information about standard drink sizes to facilitate data collection. In 2013, two questions were used to derive self-report estimates of peak consumption. These approaches varied in their degrees of cue specificity.

The first method, here termed the total consumption question, involves a general overall approach in measuring peak consumption. Respondents were asked to self-report the total amount of ASD consumed in their heaviest recent occasion of drinking. In the 2013 YAAS questionnaire, recent was defined as the 'past 12 months'.

The second method, here termed the LSPC question, involved the use of context-specific cues to measure alcohol consumption. Respondents were asked to recall information about their heaviest recent drinking occasion, noting the location where they first consumed alcohol and the quantity of ASD consumed. If respondents indicated that they continued drinking at further locations, the questions were repeated for up to 10 different settings. A list of nine options was provided: (a) private home (own, friends, acquaintances, dealers); (b) pub or bar; (c) nightclub; (d) rave or dance party; (e) music festival or concert; (f) public place (street, park, public toilet); (g) car; (h) work or educational institution; and (i) others. Amounts reported at each location were summed to generate a cumulative estimate for their peak consumption for the drinking event.

Survey questionnaire

Two versions of the questionnaire were created with question order rotated. Version [A], termed Total-first ($n = 222/449$, 49%), asked the total consumption question before the LSPC questions. Version [B], termed LSPC-first ($n = 227/449$, 51%), asked the LSPC questions before the total consumption question.

Figure 1 shows a flowchart of the [A] Total-first and [B] LSPC-first questionnaires respectively. Participants were randomly assigned to one of two questionnaire versions. Both questionnaires were identical in structure, length and method of administration. Respondents were unaware of the structure or difference in questionnaire versions.

Responses were coded using a computer-assisted telephone interview system with data transferred to Stata (version 14.2) for coding,
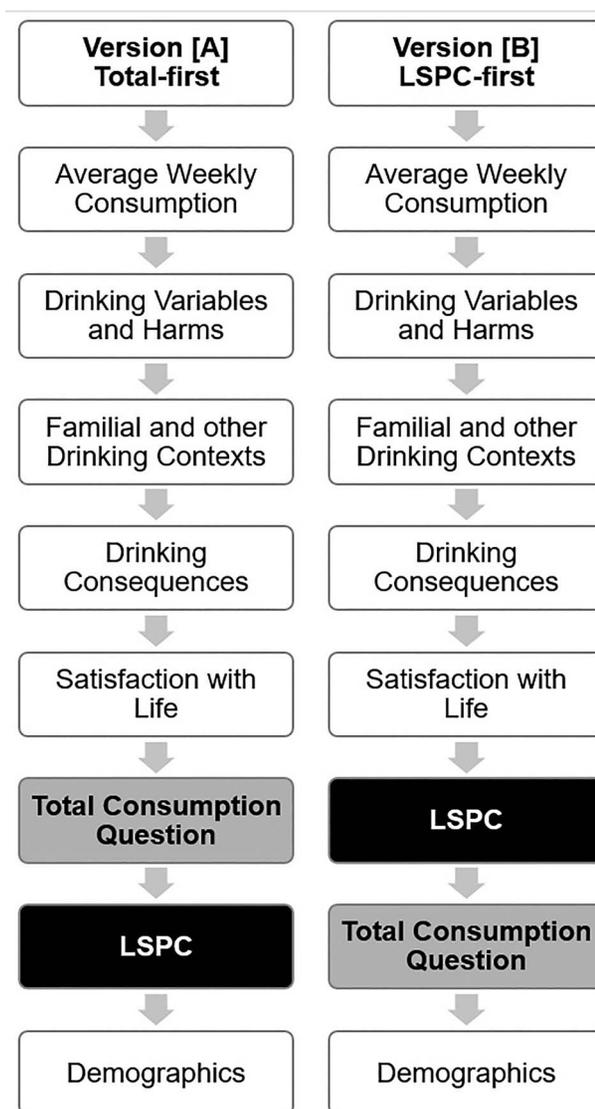

Fig. 1. Structure of the two survey questionnaire versions administered.

cleaning and analysis. To categorize the respondents into consumption terciles (see below), a random number between 0 and 0.01 exclusive was added to each data point using Stata's *runiform* command. This process converts each measurement into a unique value without impacting the overall dataset, and also allows for the clear subdivision of data into terciles.

Statistical analysis

At a conceptual level, LSPC and total consumption questions for those who only drank at one location are for all intents and purposes identical. Therefore, statistical analysis was carried out only on drinkers who reported consuming at two or more locations ($n = 169/449$, 38% single location drinkers; $n = 280/449$, 62% multi-location drinkers).

Bivariate analyses using chi-square were conducted to test for any differences between the two groups on a range of demographic variables such as gender, age, employment, country of birth, average weekly income, sexual orientation, education and ancestry.

Descriptive statistics were generated on self-reported amounts of peak consumption. Since the initial sample already consisted of high-risk drinkers, it was inappropriate to analyse respondents according to generally accepted drinking levels (e.g. <4, 5–6, 7–10, 11–19, 20+ in AIHW, 2017). Subsequently, respondents were retrospectively categorized into consumption terciles ('low', 'medium', 'high') based on their total consumption question amounts reported in wave 2 (2013): male categories were ≤12 ASD, 12–15 ASD, ≥15 ASD; female categories were ≤8 ASD, 8–10 ASD, ≥10 ASD (Demographic characteristics by category provided in Supplementary Table 2).

Two-tailed t-tests were used to analyse the self-report estimates using the different survey questions. Unpaired t-tests compared between-questionnaire estimates of peak consumption to assess effects irrespective of question order. Paired t-tests compared within-questionnaire estimates to assess within-subject question order effects. The difference between each survey question (total consumption minus corresponding LSPC) was calculated as Total-first difference and LSPC-first difference. The patterns observed for those reporting drinking different amounts on an occasion were examined within low, medium and high tercile consumption categories. Based on these values (Total-first difference and LSPC-first difference), observations were excluded as outliers if the difference was more than three standard deviations from the mean difference ($n = 6/280$, 2%). A summary of the exclusion criteria is shown in Fig. 2. The final analytical sample in 2013 ($n = 274$) comprises 34% of the total recruited sample in 2012 ($n = 802$). Demographic and drinking characteristics (average weekly consumption, peak total consumption) for the analysis sample ($n = 274$) were similar to the remainder of the cohort at baseline ($n = 528$, Supplementary Table 1), suggesting that any biases due to loss from follow-up would be small. Given the different screening criteria for males ($n = 143/274$, 52%) and females ($n = 131/274$, 48%), we analysed both sexes separately. The level of statistical significance was determined at $P < 0.05$.

Ethics

Ethics approval for the YAAS cohort study has been obtained from the Alfred Health Research Ethics Committee (Alfred Hospital Ethics approval number 35/12).

RESULTS

Sample demographics

The final sample of respondents who reported drinking in multiple locations was $n = 274$ (Version [A] Total-first $n = 127$; Version [B] LSPC-first $n = 147$). The majority of respondents were aged between 18 and 21 years (67%) and born in Australia (93%) and 52% were male. Table 1 shows other sample demographics.

Bivariate analyses demonstrate that the respondents to each questionnaire version were comparable in terms of gender, age, employment, religion, sexual orientation, education and ancestry (Table 1, $P > 0.05$). The two groups were significantly different in terms of average weekly income and country of birth.

Self-report estimates of peak consumption

Mean peak consumption estimates obtained from the different questionnaire prompts are displayed in Fig. 3. In general, the second question of each survey version elicited slightly higher consumption estimates. However, irrespective of question order (i.e. when estimates were compared between questionnaires), consumption amounts for
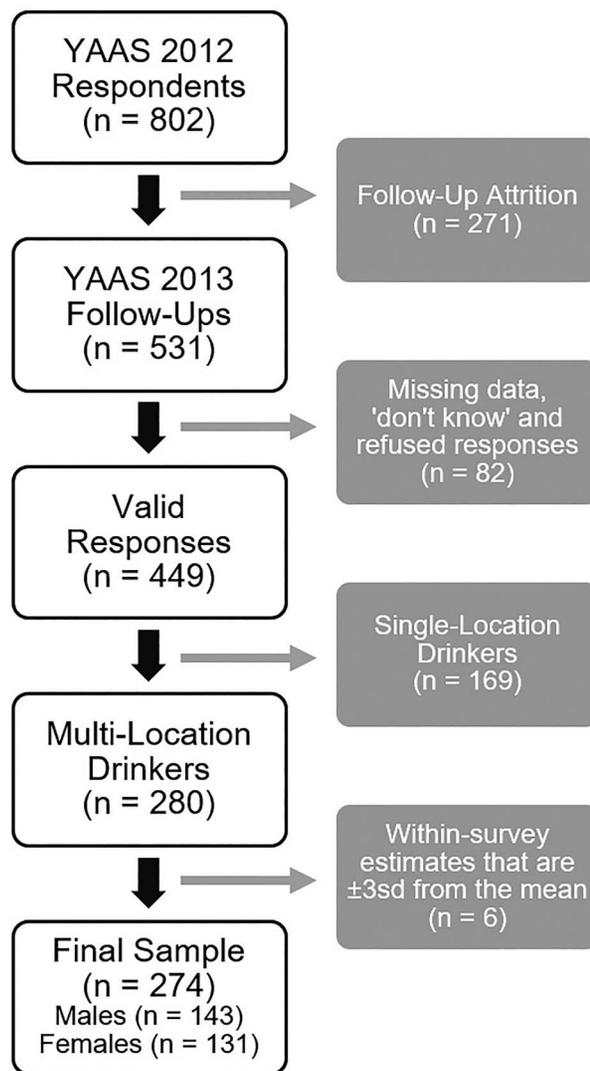

Fig. 2. Summary of exclusion criteria applied.

total consumption question and the LSPC questions were not significantly different for both males and females (Table 2).

Question order effect

To test our second hypothesis, we compared self-report estimates to the same question across different questionnaire versions to investigate any question order effect. When we compared total consumption in the [A] Total-first condition to the [B] LSPC-first conditions, we did not find any significant difference between the estimates. Similarly, there was no significant difference between the estimates obtained from the LSPC questions in the [A] Total-first and [B] LSPC-first versions (Table 2).

We also examined whether there were any within-questionnaire differences in the estimates generated by the two different question types (Table 3). In both [A] Total-first and [B] LSPC-first versions, total consumption and LSPC did not produce statistically significant different estimates. The LSPC estimates of [B] LSPC-first females were on average 0.40 ASD smaller than their total consumption estimates ($P = 0.04$). This difference was not observed in [B] LSPC-first males.

635

Table 1. Demographic characteristics of the study population

		Version [A] Total-first (%) 127	Version [B] LSPC-first (%) 147	Chi-squared	df	*P*-value
n						
Gender	Male	54	51	0.17	1	0.68
	Female	46	49			
Age category	18–21	72	63	2.53	1	0.11
	22–25	28	37			
Employment category	Part-time	54	48	6.50	4	0.17
	Full-time	30	38			
	Unemployed	9	6			
	Full-time student	7	5			
	Other	0	3			
Country of birth	Australia	96	90	3.95	1	**0.05**
	Other	4	10			
Average weekly income (AUD)	< 250	30	15	8.40	2	**0.02**
	250–600	40	48			
	>600	30	37			
Religion	Yes	12	8	1.02	1	0.31
	No	88	92			
Sexual orientation	Heterosexual	93	94	0.09	1	0.76
	Other	7	6			
Education category	<Year 12	5	6	6.06	4	0.20
	Year 12	54	41			
	Tertiary	28	31			
	Diploma	7	13			
	Trade	6	9			
Ancestry category	UK	47	38	4.14	5	0.53
	Ireland	8	7			
	Other Western Europe	13	18			
	Eastern Europe	6	5			
	Asia	10	10			
	Other	16	22			

However, when the analysis was stratified by consumption terciles, [B] LSPC-first high tercile respondents of both sexes had self-report estimates that were significantly different (Table 4). [B] LSPC-first high tercile males (≥15 ASD) had estimates derived from their total consumption question that were on average 2.14 ASD greater than their corresponding estimate derived from the LSPC questions (*P* < 0.01). [B] LSPC-first high tercile females (≥10 ASD) had estimates that were on average 1.09 ASD greater (*P* = 0.04). There were no other statistically significant differences in the low or medium consumption terciles, nor were any order effects apparent.

DISCUSSION

Similarity between self-report estimates of different question types

Contrary to what we expected in our first hypothesis, estimates of peak consumption from total consumption and location-specific consumption were not statistically significantly different. Our analyses show that males and females self-report similar amounts of peak consumption with both question types (Table 2).

The similarities between the estimates may be explained by a question order effect. According to Grice's principle of conversation, the Maxim of Relevance explains how respondents make their responses as relevant to the conversation as possible (Grice, 1975). According to this theory, respondents could have estimated their

subsequent consumption amounts by referring to their preceding estimate. Version [A] Total-first respondents could have used their preceding estimate as a reference to partition their following location-specific consumption response. Similarly, Version [B] Location-first respondents could have used their preceding estimates to provide a summative response for their following total consumption. It seems likely that there is a consistency effect of question order. With the exception of Version [B] Location-first question administered among heavy drinkers, respondents generally tend to match their subsequent response to their preceding response, as in the current case when the different question types relate to the same issue of peak consumption.

As discussed earlier, previous research suggests that the greater the contextual specificity of a question, the greater the estimates of alcohol consumption (Mooney and Gramling, 1991; Wyllie *et al.*, 1994; Casswell *et al.*, 2002). However, our current findings suggest otherwise. The comparison between [A] Total Consumption Question and [B] LSPC (Table 2) found that there were no significant differences between estimates when either question type was first asked in the questionnaire (and therefore unaffected by any question order effect). Additionally, the similarity in self-report estimates in the other comparisons (Table 2) could be due to a consistency effect, where respondents try to provide a similar response to their previous one when they are given two questions relating to the same information in a single survey. However, our study design does not allow to test for such possible explanations of this null effect. To our knowledge, there has been no other research on the interplay
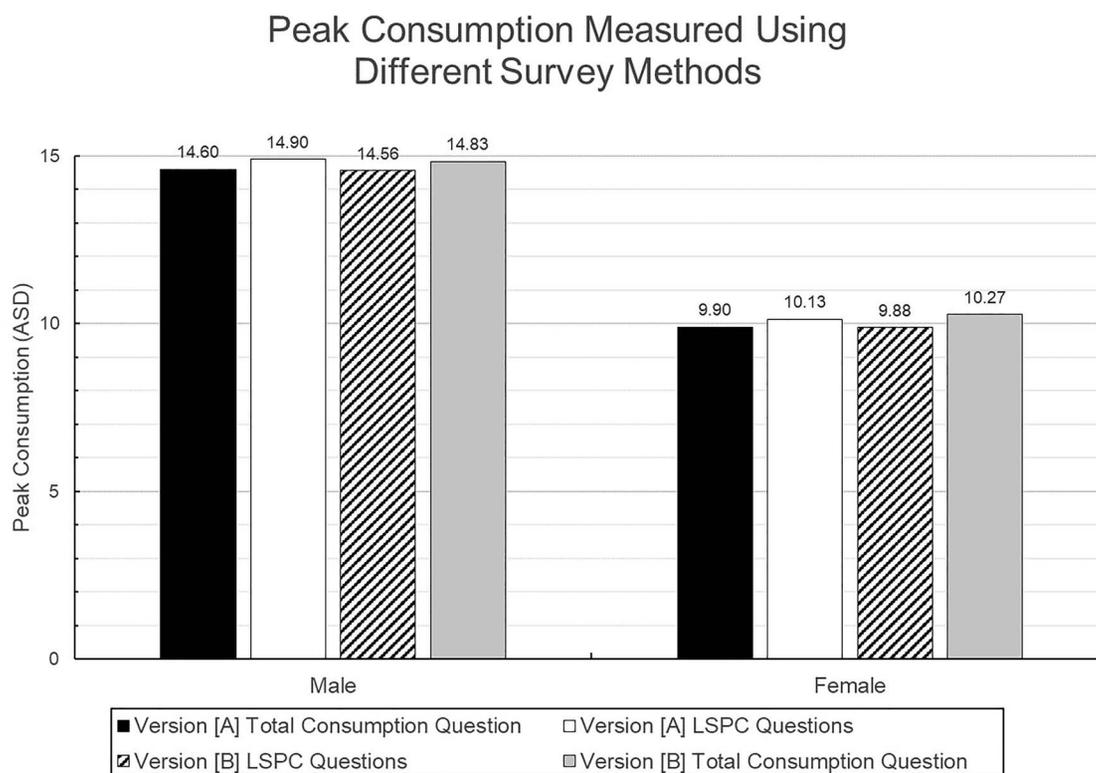
Peak Consumption Measured Using Different Survey Methods

Fig. 3. Peak consumption measured using different survey questions.

between question order and question type. Future research will be needed to determine how these effects interact and how they impact recall in alcohol surveys. Similarly, future studies should investigate the effect of question order compared to different recall biases on self-reporting.

It is important to point out that the aforementioned studies looked at measures of average weekly consumption. In contrast, our study examined estimates for peak consumption in the past 12 months and observed no significant differences between question prompts. Further research is needed to differentiate the effect of context specificity in peak consumption questions compared to average weekly consumption.

Similarity between self-report estimates of different question order

Contrary to our second hypothesis, self-report estimates for both [A] Total-first and [B] Location-first respondents were similar and minimal question order effect was observed (Table 3).

It is unclear what the reasons are behind this unexpected finding. It is possible that [A] Total-first respondents self-report similar estimates because they use their preceding total consumption as a reference when subsequently partitioning it into separate locations. It could also be the case that [B] Location-first respondents try to match their consumption amounts from separate locations amounts with a cumulative estimate for their subsequent total consumption amount.

Future research should consider using a 'think-aloud' protocol to understand the thought process of respondents when they are completing the questionnaire. By using this method, Greenfield found that respondents had varied interpretations with the definition of

'usually' in his survey (Greenfield, 2000), which helped to explain some of the observed variations. A similar analysis could elucidate how respondents derive their estimates of total consumption. It is possible that respondents calculate their peak consumption by going through each drinking location, or by considering the types of alcohol consumed. Although our results provide quantitative data about peak consumption estimates, we cannot offer any qualitative analysis on how respondents derived their estimates.

The length of the recall period is another factor that might impact recall bias. Given its nature, it may be beneficial to include shorter periods of recall to reduce the effects of recall bias. In the initial 2012 survey, however, 60% of the recent heavy drinking occasions surveyed occurred within 1 month of the survey date (Dietze *et al.*, 2014). If the recall period was shorter, certain demographics, like infrequent heavy drinkers, would not be surveyed. Future research should find a suitable balance between sample representation and the effects of recall bias.

Nonetheless, a question order effect was observed among [B] Location-first high tercile drinkers, who were observed to have notable difference between their self-report estimates (Table 4). This may simply reflect a greater number of locations visited—it is possible that our predicted effects would only begin to appear when the number of drinking locations is large and there is an increase in confusion about drinking across locations.

It is worth noting that recalling heavier drinking occasions would presumably be more cognitively taxing than recalling a lighter one, i.e. it would be easier to tally up the drinks from an 8-drink night than a 20-drink night.

Another reason why high tercile drinkers may show inconsistency may be because they could have difficulties with memory recall. High levels of alcohol consumption have been found to impair long-term

Table 2. Unpaired *t*-tests comparing between-questionnaire survey question estimates

	Mean (95% CI)		Mean difference (95% CI) ([A]−[B])	*t*	df	*P*-value
	[A] Total consumption question	**[B] Total consumption question**				
All	12.42 (11.56, 13.28)	12.60 (11.80, 13.39)	−0.18 (−1.35, 0.99)	−0.30	272	0.76
Male	14.60 (13.47, 15.74)	14.83 (13.72, 15.93)	−0.22 (−1.80, 1.35)	−0.28	141	0.78
Female	9.90 (8.90, 10.89)	10.27 (9.37, 11.17)	−0.37 (−1.70, 0.96)	−0.55	129	0.58
	[A] LSPC	**[B] LSPC**				
All	12.68 (11.76, 13.60)	12.27 (11.52, 13.01)	0.42 (−0.75, 1.58)	0.70	272	0.48
Male	14.90 (13.63, 16.17)	14.56 (13.56, 15.56)	0.34 (−1.26, 1.93)	0.42	141	0.68
Female	10.13 (9.12, 11.14)	9.88 (9.05, 10.70)	0.25 (−1.02, 1.53)	0.39	129	0.70
	[A] Total consumption question	**[B] LSPC**				
All	12.42 (11.56, 13.28)	12.27 (11.52, 13.01)	0.15 (−0.98, 1.28)	0.26	272	0.79
Male	14.60 (13.47, 15.74)	14.56 (13.56, 15.56)	0.04 (−1.45, 1.54)	0.06	141	0.95
Female	9.90 (8.90, 10.89)	9.88 (9.05, 10.70)	0.02 (−1.24, 1.29)	0.04	129	0.97
	[A] LSPC	**[B] Total consumption question**				
All	12.68 (11.76, 13.60)	12.60 (11.80, 13.39)	0.09 (−1.12, 1.29)	0.14	272	0.89
Male	14.90 (13.63, 16.17)	14.83 (13.72, 15.93)	0.07 (−1.59, 1.73)	0.08	141	0.93
Female	10.13 (9.12, 11.14)	10.27 (9.37, 11.17)	−0.14 (−1.48, 1.19)	−0.21	129	0.83

memory (Nelson *et al.*, 1986) and can result in blackouts (White, 2003). Although we canvassed a range of negative outcomes, we did not collect data on blackouts linked to the specific drinking event. Future research should consider investigating the relationship between experiencing amnesic episodes and reported alcohol consumption.

The effect of question order between average weekly consumption and peak consumption may also be of interest in future research. These are common survey questions that attempt to capture individuals with different consumption patterns (e.g. frequent weekly drinking, and infrequent one-off drinking). It would be useful to know if these questions are subject to any question order effects. Similarly, given that there are different methods to survey alcohol consumption, researchers should consider which is most appropriate in answering their research question. For example, the use of beverage-specific peak consumption may be useful in understanding the consumption behaviour of those who report consuming alcohol in only one location (who were excluded from our analyses), as the two questions we included were functionally identical for these participants. This means that our study was necessarily limited to multi-location drinkers in our sample and, as a consequence, cannot

be generalized to those who drink in only one location (which was 38% of our overall sample).

Our sample was recruited using landlines at a time when households were shifting to mobile-only, meaning our sample did not capture these mobile-only households. However, in July 2012, the percentage of Australians aged 18 and over with landline access was 78%, which still represented a large proportion of young adults (Commonwealth of Australia (Australian Communications and Media Authority) 2013). Moreover, participants were randomly assigned prospectively into questionnaire versions, making sample representativeness less of an issue for the main question of interest to this study (Rothman *et al.*, 2013).

Implications for future surveys

Our findings have implications for the type of questions used in alcohol surveys. The total consumption question we used is a time-efficient method of deriving a general estimate for peak consumption. In contrast, location-specific consumption can collect specific information regarding consumption in different contexts, but at the expense of a longer survey. At a summative level, our results

Table 3. Paired *t*-tests comparing within-questionnaire survey question estimates to assess within-subject question order effects

Version [A] Total-first	[a] Total consumption question	[A] LSPC	Two-tailed paired *t*-test statistics			
	Mean (95% CI)		Mean difference (95% CI)	*t*	df	*P*-value
All	12.42 (11.56, 13.28)	12.68 (11.76, 13.60)	−0.26 (−0.53, 0.01)	−1.93	126	0.06
Male	14.60 (13.47, 15.74)	14.90 (13.63, 16.17)	−0.29 (−0.73, 0.15)	−1.32	67	0.19
Female	9.90 (8.90, 10.89)	10.13 (9.12, 11.14)	−0.23 (−0.52, 0.06)	−1.56	58	0.12
Version [B] LSPC-first	**[B] LSPC**	**[B] Total consumption question**	**Two-tailed paired *t*-test statistics**			
	Mean (95% CI)		Mean difference (95% CI)	*t*	df	*P*-value
All	12.27 (11.52, 13.01)	12.60 (11.80, 13.39)	−0.33 (−0.69, 0.03)	−1.79	146	0.08
Male	14.56 (13.56, 15.56)	14.83 (13.72, 15.93)	−0.27 (−0.89, 0.36)	−0.85	74	0.40
Female	9.88 (9.05, 10.70)	10.27 (9.37, 11.17)	−0.40 (−0.78, −0.01)	−2.06	71	**0.04**

Table 4. Paired *t*-tests comparing within-questionnaire survey question estimates to assess within-subject question order effects, categorized by total consumption question consumption tercile

	Consumption tercile	Mean difference	*t*	df	*P*-value
Version [A] Total-first difference, Total consumption question minus LSPC (95% CI)					
Male	<12	0.10 (−0.48, 0.68)	0.35	29	0.73
	12–15	−0.38 (−1.37, 0.61)	−0.81	16	0.43
	>15	−0.79 (−1.71, 0.14)	−1.77	20	0.09
Female	<8	−0.23 (−0.46, 0.01)	−2.02	21	0.06
	8–10	−0.06 (−0.50, 0.38)	−0.29	23	0.77
	> 10	−0.54 (−1.65, 0.57)	−1.05	12	0.31
Version [B] LSPC-first difference, Total consumption question minus LSPC (95% CI)					
Male	<12	−0.89 (−2.05, 0.27)	−1.58	26	0.13
	12–15	−0.04 (−0.78, 0.70)	−0.10	26	0.92
	>15	2.14 (1.05, 3.23)	4.10	20	**<0.01**
Female	<8	0.10 (−0.36, 0.57)	0.46	23	0.65
	8–10	0.04 (−0.31, 0.39)	0.24	24	0.81
	>10	1.09 (0.04, 2.13)	2.16	22	**0.04**

are strikingly similar across different question types and different question orders. As most respondents provide similar responses of peak consumption, it may seem redundant to include both questions. Instead, researchers should weigh the advantages and disadvantages of each survey method in deciding which question to include. The location-specific consumption questions provide a mechanism of measuring drinking in different types of locations, which may be suited to some research questions (e.g. how much predrinking occurs in private locations, see Dietze *et al.*, 2014). In general, however, the much shorter total consumption question appears to do a decent job without having the higher response burden of the location-specific consumption questions.

Measuring drinking behaviour is important both in a clinical and public health context. This paper provides relevant findings on one of the ways that this measurement can be improved. Moreover, given the increased administration time of the context-specific questions, it is important to determine whether this extra time spent is worthwhile in terms of producing improvements in recall.

In conclusion, at a population level, there appears to be no major difference between the estimates of peak consumption derived from our total consumption question and our LSPC. Nevertheless, inconsistency between self-report estimates was higher among those who reported drinking more, suggesting that further work needs to be undertaken to determine the best method of assessing drinking events for these heavier drinkers.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Alcohol and Alcoholism* online.

ACKNOWLEDGEMENTS

FUNDING

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no other conflicts of interest.

REFERENCES

Communications report 2012–2013. Canberra: Australian Communications and Media Authority.

AIHW (2017) In Welfare AIH (ed). *National Drug Strategy Household Survey 2016: detailed findings*. Vol. Drug Statistics series no. 31. Canberra: AIHW, 168.

Bellis MA, Hughes K, Jones L *et al.* (2015) Holidays, celebrations, and commiserations: measuring drinking during feasting and fasting to improve national and individual estimates of alcohol consumption. *BMC Med* 13:113.

Bradburn NM, Mason WM (1964) The effect of question order on responses. *J Mark Res* 1:57–61.

Casswell S, Huckle T, Pledger M (2002) Survey data need not underestimate alcohol consumption. *Alcohol Clin Exp Res* 26:1561–7.

Clark W (1981) Public drinking contexts: bars and taverns. In Thomas C, Harford, Gaines LSS (eds). *Social Drinking Contexts: Proceedings Of A Workshop*. Washington: NIAAA, 244.

Davis CG, Thake J, Vilhena N (2010) Social desirability biases in self-reported alcohol consumption and harms. *Addict Behav* 35:302–11.

Dietze P, Agius PA, Livingston M *et al.* (2017) Correlates of alcohol consumption on heavy drinking occasions of young risky drinkers: event versus personal characteristics. *Addiction (Abingdon, England)* 112: 1369–77.

Dietze PM, Livingston M, Callinan S *et al.* (2014) The big night out: what happens on the most recent heavy drinking occasion among young Victorian risky drinkers? *Drug and alcohol review* 33:346–53.

Elofson J, Gongvatana W, Carey KB (2013) Alcohol use and cerebral white matter compromise in adolescence. *Addict Behav* 38: 2295–305.

Gmel G, Kuntsche E, Rehm J (2011) Risky single-occasion drinking: bingeing is not bingeing. *Addiction (Abingdon, England)* 106:1037–45.

Gmel G, Rehm J. (2004) *Measuring alcohol consumption. Contemporary Drug Problems: An Interdisciplinary Quarterly 31*.

Greenfield TK (2000) Ways of measuring drinking patterns and the difference they make: experience with graduated frequencies. *J Subst Abuse* 12:33–49.

Greenfield TK, Nayak MB, Bond J *et al.* (2006) Maximum quantity consumed and alcohol-related problems: Assessing the most alcohol drunk with two measures. *Alcohol Clin Exp Res* 30:1576–82.

Grice HP (1975) Logic and conversation. In Cole P, Morgan JL (eds). *Syntax and Semantics 3: Speech acts*. New York: Academic Press, 41–58.

Harford TC (1994) The effects of order of questions on reported alcohol consumption. *Addiction (Abingdon, England)* 89:421–4.

Krosnick JA (1999) Survey research. *Annu Rev Psychol* 50:537–67.

Kuntsche E, Gmel G (2013) Alcohol consumption in late adolescence and early adulthood - where is the problem? *Swiss Med Wkly* 143:w13826.

Livingston M, Callinan S (2015) Underreporting in alcohol surveys: whose drinking is underestimated? *J Stud Alcohol Drugs* 76:158–64.

Midanik LT (1982) The validity of self-reported alcohol consumption and alcohol problems: a literature review. *Addiction (Abingdon, England)* 77:357–82.

Mooney LA, Gramling R (1991) Asking threatening questions and situational framing: the effects of decomposing survey items. *Sociol Q* 32:289–300.

Nelson, TO, McSpadden M, Fromme K *et al.* (1986) Effects of alcohol intoxication on metamemory and on retrieval from long-term memory. *J Exp Psychol Gen* 115:247–54.

NIAAA (2003) *Recommended Sets of Alcohol Consumption Questions, National Institute on Alcohol Abuse and Alcoholism Task Fore on Recommended Questions of the National Council on Alcohol Abuse and Alcoholism*. Behesda, MD, USA: National Institute on Alcohol Abuse and Alcoholism.

Rehm J, Room R, Taylor B (2008) Method for moderation: measuring lifetime risk of alcohol-attributable mortality as a basis for drinking guidelines. *Int J Methods Psychiatr Res* 17:141–51.

Rothman KJ, Gallacher JE, Hatch EE (2013) Why representativeness should be avoided. *Int J Epidemiol* 42:1012–4.

Single E, Wortley S (1994) A comparison of alternative measures of alcohol consumption in the Canadian National Survey of alcohol and drug use. *Addiction (Abingdon, England)* 89:395–9.

Stockwell T, Donath S, Cooper-Stanbury M *et al.* (2004) Under-reporting of alcohol consumption in household surveys: a comparison of quantity-frequency, graduated-frequency and recent recall. *Addiction (Abingdon, England)* 99:1024–33.

Strack F (1992) "Order effects" in survey research: activation and information functions of preceding questions. In Schwarz N, Sudman SS (eds). *Context Effects in Social and Psychological Research*. New York, NY: Springer New York, 23–34.

Strube G (1987) Answering survey questions: the role of memory. In Hippler H-J, Schwarz N, Sudman S (eds). *Social Information Processing and Survey Methodology*. New York, NY: Springer New York, 86–101.

Walle SVD, Ryzin GGV (2011) The order of questions in a survey on citizen satisfaction with public services: lessons from a split-ballot experiment. *Public Adm* 89:1436–50.

Werch CE (2009) Quantity-frequency and diary measures of alcohol consumption for elderly drinkers. *Int J Addict* 24:859–65.

White AM (2003) What happened? Alcohol, memory blackouts, and the brain. *Alcohol Res Health* 27:186–96.

White AM, Kraus CL, Flom JD *et al.* (2005) College students lack knowledge of standard drink volumes: implications for definitions

of risky drinking based on survey data. *Alcohol: Clin Exp Res* 29: 631–8.

WHO (2014) *Global status report on alcohol and health - 2014*. Geneva: World Health Organization.

Wilson P (1981) Improving the methodology of drinking surveys. *J R Stat Soc Series D* 30:159–67.

Wright CJ, Dietze PM, Crockett B *et al.* (2016) Participatory development of MIDY (mobile intervention for drinking in young people). *BMC Public Health* 16:184.

Wyllie A, Zhang J-F, Casswell S (1994) Comparison of six alcohol consumption measures from survey data. *Addiction (Abingdon, England)* 89:425–30.