REPORTING ALCOHOL CONSUMPTION: THE PROBLEM OF RESPONSE VALIDITY

Jennifer J. Waterton
Social & Community Planning Research, London

SUMMARY

Sample survey estimates of alcohol consumption
are subject to severe response biases. This
paper presents the results of a small-scale
methodological survey using two matched samples
of adult males conducted in November 1982.
One group reported their alcohol consumption over
the previous seven days in the conventional way.
The other group were interrogated by micro-
computer and entered their answers directly from
the keyboard. The results suggest that computer
interrogation may be a fruitful avenue for fur-
ther research for surveys of sensitive issues.

ALCOHOL SURVEYS AND THE PROBLEM OF RESPONSE
VALIDITY

Survey measurement of alcohol consumption in a
specified population is of interest to those con-
cerned with the economic, medical and social
implications of patterns of drinking behaviour.
Estimates of consumption inferred from self-
report survey data may be compared with estimates
of sales data to produce a 'coverage rate'. The
coverage rate is defined as:

$$\frac{\text{Estimate of per capita consumption (pcc) from survey}}{\text{Estimate of pcc from sales data}} \times 100\%$$

The independent estimate from sales figures is
not ideal. In particular, assumptions must be
made about a) the amounts of homemade alcohol
consumed, b) the amount of alcohol bought in one
area but consumed in another and c) the time
elapsing between purchase and consumption. Des-
pite these limitations, the coverage rate can
provide a rough guide to the accuracy of the
self-report survey data. Over the last 20 years,
alcohol surveys have achieved coverage rates
between 20% and 70%, with most concentrated in
the 40%-60% range (Mäkelä, 1971, Room 1971,
Pernanen, 1974, Wilson, 1981). This substantial
underreporting occurs irrespective of whether
respondents are asked questions relating to a
'typical week' in their drinking history, say,
or whether a timeframe immediately prior to the
interview is specified during which time the res-
pondent is asked to recount all drinking occa-
sions and beverage amounts drunk. The former
approach, namely the Quantity Frequency approach,
is extremely complicated for respondents, as they
are required to generalise about their 'usual
frequency of drinking occasion', 'usual amounts
of each type of alcoholic beverage consumed', etc.
The latter approach yields unbiased estimates and
is in general easier for the respondent, thus it
is to be preferred. Current practice is to ask
about the past 7 days.
Several factors contributing to the shortfall
in the coverage rates may be identified:

i)   an inadequate sampling frame which fails to
     include areas and institutions where the
     heavy consumers are concentrated.
ii) the heavy consumers are harder to locate,
     (even if they are included in the sampling
     frame). They may comprise a large propor-
     tion of the 'non-contact' category.
iii) the 'refusers' on an alcohol survey may be
     more likely to be heavy consumers than the
     non-refusers.
iv) respondents may genuinely forget some of
     their drinking occasions of the past 7 days.
     In addition, there may be difficulties
     associated with recalling the amounts con-
     sumed, especially if the drinking occasion
     took place in a private house, when standard
     measures would not be poured.
v)   respondents may deliberately underreport the
     frequency with which they drink and the
     amounts consumed on each occasion, because
     of the social stigma attached to heavy con-
     sumption.

The general question of validity of self-
reported consumption is reviewed by Midanik
(1982). Bias due to response error may be cap-
able of modification by implementing alternative
methodologies. Techniques such as randomised
response (Warner 1965) may be applied to self-
report consumption data, and this acts to pre-
serve confidentiality of response. Alternative-
ly, prospective studies with respondents filling
in a diary each day may reduce the response bias
due to forgetting (Poikolainen & Kärkkäinen,
1983). A third possibility in alcohol surveys
is to take blood samples to validate self-
reported consumption (e.g. Chick et al., 1981)
and it may be conjectured that prior knowledge
that blood samples would be taken might increase
respondent truthfulness. A fourth possibility
is to use a self-completion questionnaire rather
than ask respondents to report their drinking
behaviour to an interviewer. A fifth possibil-
ity, and one which has only recently become
feasible with the advent of portable micro-
computers, is to replace interviewers with com-
puters, thus preserving anonymity of response in
a way which is not possible even with self-
completion methods.

Computer Interrogation
The evidence on the question of whether accur-
acy may be improved by using computers instead of
interviewers as interrogators is limited. The
author is not familiar with any field studies of
computer interviewing techniques in the area of
alcohol research. However, studies carried out
in specialised institutions have been reported,
(Card et al., 1974, Slack et al., 1968). In one
such recent study by Lucas et al. (1977), 36
volunteer male patients in a Glasgow hospital
were questioned about their drinking behaviour
and any alcohol related problems they experienced.
Identical information was elicited from the res-
pondents by computer and by trained medical
staff. Most of the results reflected good
agreement for the two data collection methods.
On the question of consumption, however, the
average amount elicited by computer was 30%
higher than that by direct interview. The
sample was a self-selected group of volunteers,
so it would be unwise to attach general

substantive significance to this result.

A Computer Interview Experiment on Adult Males in Edinburgh

For the purpose of the methodological experiment, a split sample was employed. A clustered stratified design was used: one polling district was selected with probability proportional to size from each of ten strata constructed from the City of Edinburgh Parliamentary Constituencies. The stratification of wards (and hence of polling districts) was effected on the basis of the percentage of non-manual workers they contained. The simple random sampling procedure within each of the ten selected districts ensured that, after rejecting the women whose names were initially chosen from the electoral register listings, fifty names remained within each district. Of these, twenty-five were assigned at random to the computer interview, and the remaining twenty-five to a direct interview. Thus it was a 3-stage sample, with wards as a nominal stage only.

If the information obtained from the electoral register proved to be incomplete or out-of-date at the fieldwork stage, an auxiliary procedure (Blyth and Marchant, 1973) was used to select respondents. Interviewing was carried out between October and December, 1982.

Interviewers were recruited by newspaper advertising and trained by the author. No interviewer conducted both computer interviews and direct interviews. This restriction arose because the interviewing periods for the two types of interview were to coincide, while simultaneously making maximum usage of the hired computers: the logistics of rotating the four computers efficiently among a larger pool of interviewers seemed prohibitively complicated. The two disjoint groups of interviewers were similar regarding characteristics such as their age, social class and experience, which might be expected to affect the response they obtained. It was decided that only male interviewers were to be used on the project: this was partly because only male respondents were included in the study and partly because of the weight and size of the computers.

The micro-computers used in the study were Hewlett-Packard HP-85s, which are roughly the size of a typewriter and weigh about 20lbs. The machines have a normal keyboard and separate numeric pad, four "soft" (function programmable) keys, a small thermal printer, a cassette drive and a 5-inch square display screen with high resolution. They are readily transportable in leather carry cases.

The survey was introduced to both groups of respondents as a study which had been designed to investigate whether the method of data collection affected response. Both groups were told that their responses would not be identified with their name at any stage.

The interview comprised three sections. The first and last parts, concerned principally with demographic details, were asked by the interviewers, and were identical for both groups. It was in the middle section, where the 7-day retrospective drinking history data was collected, that the alternatives were implemented.

The format of the middle section of the com-

puter interview was as follows: the interviewer first ascertained whether the respondent had consumed any alcoholic beverage in the seven days prior to interview. If so, the use of the computer was explained to the respondent (if not, the interviewer proceeded to the last part of the interview). When the interviewer was satisfied that the respondent was capable of proceeding on his own he withdrew to a place in the room where the display screen (and hence the respondent's answers) was not visible to him. For each of the last seven days in reverse temporal order, questions about the respondent's consumption of various kinds of alcoholic beverage appeared on the screen. There were only three possible answers to every question: Yes, No or a number. If a "Yes" or "No" answer was appropriate, the respondent pressed the corresponding "soft" key. If a number was required, he typed in the number and then pressed the "Return" key. The respondent's typed answers appeared on the display screen only, not on the line printer. The respondent had been told beforehand that information was required on a daily basis. This meant that if he had been involved in several drinking occasions on the same day, some mental addition would be necessary. After all the data for a particular day had been input from the keyboard, and before moving on to ask about the previous day's consumption, (the seven days being asked about in reverse order), the computer reproduced the respondent's replies on the screen and asked if he wished to change any of them. This facility gave the respondent the opportunity to correct any typing mistakes or any values which, on reflection, he wished to alter.

The consumption data was stored with the respondent's reference number on a cassette. The matching of the data to the demographic records was carried out at the coding stage. Interviewers were not able to access the consumption data themselves. This aspect of anonymity of response was emphasised to respondents.

The interviewers on the direct survey were instructed to reproduce exactly the wording of the questions appearing on the computer screen, to insist that daily totals (not occasion totals) were reported to them and to refuse to give any memory prompts, so as to maintain comparability with the computer interview. All else being equal, any difference in the levels of reporting which emerges can then be ascribed to the changed methodology, with the caveat that the strategy of allocating interviewers (as described above) means that any methodological effect is confounded with an interviewer effect. In the presentation of the results it is assumed that there were no systematic biases between the two groups of interviewers.

RESULTS

Although the computer is not used in the initial section of the interview, it was thought that the sight of an interviewer carrying a machine might make respondents hesitant about participating. The response rates are presented in Table 1. The base for the response and refusal rates is calculated as (Total – Out of Range). The difference in refusal rates was examined further by calculating the corrected χ^2

statistic. The observed frequencies, and the value of the statistic are given in Table 2. This value of the χ^2 statistic is significant at the 8% level. Thus there is limited evidence of a higher refusal rate on the computer interview sample. The impression from talking to interviewers was that, not surprisingly, older respondents were more wary of the computer. However, the difference in the composition of the two samples is not severe enough to invalidate comparisons between the two sets of results.

There were five respondents not included in the table who, although they were assigned to a computer interview, co-operated only on condition that they received a direct interview. These five individuals are omitted from subsequent analyses. Two of these five had below average consumption and the remaining three were above average. Their inclusion into the direct group would not substantially alter any of the conclusions reached.

The differences in the socio-economic composition of the ten wards selected as shown in Table 3 suggest that an analysis at polling district level may be of interest. However, given the small cell sizes, any conclusions drawn about the relative merits of the two techniques within polling districts must remain tentative.

The ordering in Table 3 corresponds to the ordering on the variable 'percentage non-manual'. There is a considerable amount of variation between polling districts within wards. In particular the district selected in SW Corstorphine was not representative of the ward as a whole.

There were seventy-four respondents who reported no alcohol consumption during the week prior to the interview: thirty-two out of one-hundred-and-forty-five respondents interviewed by computer and forty-two out of one-hundred-and-seventy-five interviewed directly. The corrected χ^2 statistic for detecting a difference in the observed proportions is 0.08, which is clearly not significant. To conform with the conventions on other alcohol surveys, these seventy-four respondents were removed from the analysis of amounts consumed. The results presented are therefore based on the remaining 246 interviews, of which 133 were direct.

Table 4 contains the mean number of units of alcohol drunk by respondents, broken down by type of interview and beverage type. One unit of alcohol is defined to be equivalent to 9 grammes of absolute alcohol. Standard measures of alcoholic beverages are readily equated to an equivalent number of units of alcohol. The figures in parentheses are the estimated standard errors of the means.

The mean total for those interviewed directly who had consumed some alcoholic beverage in the week prior to interview may be compared with results from recent surveys of similar populations. Knight and Wilson (1981) found a weekly average of 19.5 units for Scotland's adult males. A study in two waves – the first in 1978, the second in 1981 following up the population of Lothian males aged 17 or over in 1978 reported averages of 21.1 and 17.9 units respectively (Kendall et al., 1982).

Table 4 shows that the overall amounts reported to the computer are 33% higher than those obtained by direct questioning. On closer examination of the data it is clear that there is a general tendency throughout the population to report higher values to the computer.

In Table 5 the subgroup means for other variables believed to affect consumption are presented with their corresponding significance probabilities. For this analysis, the dependent variable was taken to be the square root of total consumption, in order to render the data distributions more symmetrical and to stabilise the variances in population subgroups.

In addition, a linear regression of age on the square root of consumption gives a regression coefficient of -0.029, a value significant at the 0.001 level. A multi-way analysis of variance was performed to assess whether the differences noted in Table 4 could be explained on the basis of demographic factors. As shown in Table 6, for the purpose of predicting consumption, the method of interview is second in level of significance only to age of respondent, and is clearly an important explanatory variable, even after controlling for the demographic differences between the two groups.

Those interviewed by computer were asked three questions by the interviewer about their concentration, their accuracy and the ease with which they were able to operate the computer. The results, which indicate a high level of acceptability among respondents, appear in Table 7.

The direct interview took on average 15 minutes to complete. The computer interview took 28 minutes on average but this may be accounted for by the time taken to explain the use of the computer, and the additional questions posed by the interviewer.

Practical Issues

There are practical difficulties associated with a survey of this kind. A restriction to male interviewers with access to a car was imposed due to the weight and size of the computers. Finding a firm who will hire computers for short periods of time may prove problematic. In particular, if it is possible to hire only a few machines, the fieldwork may take longer than one would wish.

The increased cost of computer interviewing because of equipment costs and extra interviewing time should not be seen as a major obstacle. The rapid progress in hardware makes it likely that smaller computers with extensive capabilities will soon be available at a fairly low cost. It may then prove economical to purchase computers outright and recover the capital expenditure over a number of surveys.

Several difficulties that had been anticipated did not arise: no-one reported the lack of a suitable flat surface on which to place the computer; interference by other household members did not prove a cause for concern; no respondent appeared to have any difficulty reading the screen; no respondent objected to running the computer from their own domestic power supply.

DISCUSSION AND CONCLUSIONS

The tentative conclusion of this methodological experiment is that computer interviewing is a feasible alternative and may be appropriate for

surveys of threatening or sensitive issues. The sample sizes were small, however, making subgroup comparisons unsatisfactory, and the exclusion of women respondents prohibits any generalisation to the adult population. If the finding of a somewhat higher consumption for the computer respondents is to be interpreted as a feature of the changed methodology, it must be assumed that the interviewers were successful in carrying out the instructions to reproduce the wording appearing on the display screen exactly and in refraining from prompting respondents in any way. In fact, the repetitive nature of the questions (identical sequences of questions were asked for each of the seven days) meant that the group undergiong the face-to-face interview often did not listen to the whole question before answering. This phenomenon has an analogy in the computer interview situation: as the interview progresses, it seems likely that the respondent would not actually read the words appearing on the screen, having learned by experience what they would be.

Despite all caveats about over-interpretation, unless computer interviews produce deliberate over-reporting (which seems unlikely), it does seem that computer interviewing may result in improved accuracy and may be appropriate for surveys of other threatening or sensitive issues. It remains to elucidate the reasons for this improvement. Several credible explanations are listed below. It is likely that they all contribute in reducing the bias of forgetting and selective response known to exist in alcohol surveys. Some of these could be set up as testable hypotheses in future surveys to establish their relative importance.

1. The presence of the computer encouraged the respondents to take the interview more seriously. The computer represents an obvious investment of time and money on the part of the survey researchers. This attitude is reflected in the answer to the question about the effect of the computer on concentration may be expected to reduce the bias due to inadvertent forgetting.

2. The computer does not react in any way to the responses typed in. It is believed that people may modify their responses in such a way as to avoid the displeasure of the interviewer. With computer interviewing this potential source of bias is removed.

3. The computer places no expectations on the respondent about the length of time they should take to answer. The respondent may read and re-read the question without feeling a pressure to answer quickly.

4. People may be attracted by the novelty of the approach, particularly if they are unfamiliar with computers. If this is the only factor at work, then the advantages of computer interviewing would diminish over time as their use became more widespread.

5. The computer preserves anonymity of response and thus reduces the threat of reporting socially undesirable behaviour.

References

Blyth, W and Marchant, L (1973) A Self-weighting random sampling technique. Journal of the Market Research Society, 15, 157-162.

Card, W I, Nicholson, M, Crean, G P, Watkinson, G E, Evans, C R and Russell, D (1974) A comparison of doctor and computer interrogation of patients. International Journal of Biomedical Computing, 5, 175.

Chick, J, Kreitman, N and Plant, M (1981) Mean cell volume and gamme glutamyltranspeptidase as markets of drinking in working men. Lancet, 1249-1251.

Kendell, R E, de Roumanie, M, Ritson, E B and Kendrick, S (1982) Community Response to Alcohol Related Problems. Geneva: WHO.

Knight, I and Wilson, P (1980) Scottish Licensing Laws. HSMO.

Lucas, R W, Mullin, P J, Luna, C B and McInray, D C (1977) Psychiatrists and a computer as interrogators of patients with alcohol-related illnesses: a comparison. British Journal of Psychiatry, 131, 160-167.

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

Poikolainen, K and Kärkkäinen, P (1983) Diary gives more accurate information about alcohol consumption than questionnaire. Alcohol epidemiology Section of the International Council on Alcohol and Addictions.

Room, R (1971) Survey vs sales data for the US. The Drinking & Drug Practices Surveyor, 3, 15-16.

Slack, W V and Van Cura, L J (1968) Patient reaction to computer based medical interviewing. Computers and Biomedical Research, 1, 527.

Warner, S L (1965) Randomised response: a survey technique for eliminating evasive answer bias. Journal of the American Statistical Association, 60, 63-69.

Table 1   An analysis of response rates by interview type

	Computer	Direct	Overall	
Successful Interviews	145	175	320	* This includes people
Refusals	32	21	53	who were ineligible,
Non-contacts	20	18	38	people in hospital,
Out of Range*	52	43	95	derelict addresses etc.
Total	249	257	506	
Refusal rate	16.2%	9.8%	12.6%	
Response rate	73.6%	79.9%	77.1%	

Table 2   An analysis of refusal rates by interview type

	Computer	Direct	Total
Refusals	32	21	53
Successful + Non contacts	165	193	358
Total - Out of Range	197	214	411

$\chi^2_{(1)} = 3.23$

Table 3   Socio-economic composition of polling districts

Ward	Non-Manual*	Manual*	Students	Unclassified	Total
Granton	7	28	0	0	35
Lochend	6	31	1	0	38
Maymarket	6	15	8	1	30
Inch	7	18	0	1	26
Stenhouse	9	22	0	1	32
Firrhill	16	20	1	0	37
Calton	11	8	8	1	28
SW Corstorphine	14	24	0	0	38
Stockbridge	22	5	0	2	29
Morningside	26	0	5	1	32
Total	124	171	23	7	325

*   As defined in "Classification of Occupations 1980", OPCS.

Table 4   Mean consumption in units of alcohol by beverage type

(Figures in parentheses are estimated standard errors)

	All beverages types		Beer		Wine		Spirits	
Combined sample	22.6	(1.55)	16.8	(1.21)	1.4	(0.22)	4.3	(0.64)
Computer sample	26.1	(2.58)	19.0	(1.82)	1.7	(0.34)	5.4	(1.27)
Direct sample	19.6	(1.82)	15.0	(1.59)	1.2	(0.29)	3.4	(0.51)
(Computer ÷ Direct) x 100	133.4		126.4		150.0		158.5	
Computer - Direct	6.5		4.0		0.5		2.0	
Level of significance of difference	0.016							

Table 5   Mean consumption by demographic subgroups

Variable	Subgroup names	Subgroup means	Significance probability p
Employment Status	Part time/full time employed	24.1	0.004
	Part time/full time education	23.5	
	Unemployed	30.6	
	Retired	11.3	
	Out of work due to disability	15.0	
Social Class	I & II	21.6	0.72
	III NM	19.0	
	III M	24.4	
	IV & V	24.2	
	Student	21.1	
	Unclassified	23.1	
Marital Status	Married	20.0	0.016
	Single	29.9	
	Widowed	23.0	
	Other	18.1	
Ward	Granton	23.2	0.14
	Lochend	27.9	
	Haymarket	26.8	
	Inch	19.0	
	Stenhouse	26.4	
	Firrhill	19.6	
	Calton	14.6	
	SW Corstorphine	21.3	
	Stockbridge	36.0	
	Morningside	15.5	

Table 6   Multi-way analyses of variance taking the square root of total consumption as the dependent variable

Factors and covariates in the Analysis of Variance*	Significance level of main effects
Type of interview	0.016
Type of interview	0.028
Employment status	0.199
Marital status	0.242
Social class	0.140
Polling district	0.273
Age	0.001

*   The dependent variable is taken as the square root of total consumption.   The classic Analysis of Variance approach in which the effect of a variable is assessed after adjusting for all other factors and covariates has been adopted.

Table 7   Respondent's attitudes to computer interviewing

Question	Response	Relative frequency of response
Do you think you were more or less accurate than you would have been if the questions had been asked by the interviewer?	More accurate	.30
	Less accurate	.02
	About the same	.68
Do you think you concentrated more or less than you would have done if asked the questions by the interviewer?	Concentrated harder	.65
	Concentrated less	.03
	About the same	.32
How did you find the method of interviewing?	Very easy	.53
	Easy	.44
	Neither easy nor difficult	.03
	Difficult	.00
	Very difficult	.00