

Estimating a Socially Undesirable Behaviour

Philip Gendall, Janet Hoek, and Matt Blakeley

This article reports the findings of an experiment designed to evaluate the effect of four questionnaire formats on the willingness of drivers convicted of drunken driving to admit to this socially undesirable behaviour. Four samples of convicted drunk drivers each received a different version of a mail questionnaire about driving behaviour and attitudes. Two versions of the questionnaire included a battery of attitude statements before the 'key' question on drunken driving convictions. One set of attitude statements downplayed the perceived social undesirability of drunken driving, while the other set emphasised it. The other two versions of the questionnaire contained no attitude statements. One simply asked the 'key' question directly, while in the other version drunken driving was embedded in a list of eight traffic offences. The highest admission rate was achieved for the questionnaire which used attitude statements to reduce the social undesirability of drunken driving before asking the behavioural question. However, the inclusion of negatively-framed preceding attitude statements and embedding the behaviour concerned in a list of other related behaviours also yielded a higher admission rate than simply asking a direct question.

Keywords: questionnaire design, response rates, social desirability bias, attitude statements, survey questions

Introduction

Most surveys rely on respondents' reports of their own behaviour, rather than on direct observation or external verification of this behaviour. But there are many reasons why respondents may give inaccurate answers to survey questions. They may misunderstand questions, lack sufficient interest in the topic to give accurate answers, fail to remember the correct answers, or deliberately give incorrect answers (see Locander & Burton 1976; Dillman 1978; Kalton & Schuman 1982; Sudman & Bradburn 1982; Belson 1986; Converse & Presser 1986). One explanation for the latter is that some respondents feel more comfortable giving socially acceptable answers, a phenomenon known as social desirability bias.

Social desirability bias occurs when a respondent replies to a question "... in a way that conforms to dominant belief patterns among groups to which the respondent feels some allegiance or identification..." (Dillman 1978, p62). In other words, some respondents answer questions about moral or ethical issues in ways which they perceive are socially desirable, presumably in order to create a favourable impression with the interviewer or survey designer (Phillips & Clancy 1971/72; Locander & Burton 1976).

Tests using variables that can be compared with actual statistics, such as criminal convictions or alcohol consumption, suggest that social desirability bias can have a major effect on surveys about sensitive issues (Kalton & Schuman 1982). Behaviours such as donations to charity, for example, are typically over-reported, while less desirable behaviours such as smoking and drinking are likely to be under-reported.

Because survey estimates of socially undesirable behaviours are lower than the population parameters, researchers generally assume that the question form or questionnaire design which yields the highest estimate is the most accurate (Kalton & Schuman 1982). Thus pre-

testing or split ballot testing of alternative questionnaires seeks the highest estimate of socially undesirable attitudes or behaviours. The reverse applies for behaviours or attitudes likely to be perceived as socially desirable.

However, even if these assumptions are valid, researchers still do not know how accurate the 'best' alternative is, unless they can compare their results with known facts. Peterson (1984) illustrated this by asking respondents their age; a question typically open to inaccurate responses. In Peterson's research, four alternative methods of asking respondents their age were used on sample members whose ages were already known. The most accurate alternative was found to be the simple "How old are you?", while the least accurate method involved asking people to classify themselves into five-year categories. However, the latter had the lowest refusal rate, while the most accurate method had the highest refusal rate. In other words, there was a trade-off between accuracy and refusal rate.

Sudman and Bradburn (1982) devote an entire chapter of their book to what they describe as "threatening questions" (which, by implication, are likely to be subject to social desirability bias). They describe a wide variety of methods for asking such questions, most of which are based on the notion of reducing the perceived seriousness of the behaviour or issue under investigation. These methods include making the question very casual ("Have you happened to have....?"); making the behaviour concerned appear common place ("A lot of people... Have you ever...?"); embedding the behaviour in a list of behaviours with varying levels of seriousness; and including attitudinal questions about the behaviour before the behavioural questions.

However, while these suggestions seem logical enough, there is little empirical evidence to support them, and virtually no evidence as to their relative effectiveness. This article presents the findings of an experiment designed to evaluate the relative accuracy of four methods of estimating the occurrence of a socially undesirable behaviour.

Method

The objective of this study was to determine which of four alternative questionnaires would yield the highest level of self reporting of a conviction for drunken driving, from a sample of respondents who had all committed this offence in the previous twelve months. The sample was selected from published court records and Ministry of Transport records and was divided into four subsamples, each of which received a different version of a mail questionnaire.

Two versions of the questionnaire included a battery of attitude statements before the 'key' question about drunken driving convictions. One 'positive' set of attitude statements downplayed the perceived social undesirability of drunken driving, while the other 'negative' set emphasised it. Both sets of attitude statements are shown in Figure 1.

These two versions of the questionnaire then asked the 'key' behavioural question in the following direct form:

"In the past twelve months have you been convicted of being drunk in charge of a motor vehicle?"

The other two versions of the questionnaire contained no attitude statements. One asked the 'key' question directly, while the other embedded drinking and driving in a list of eight traffic

Figure 1. Preceding attitude statements

Version 1	Version 2
Positive statements.	Negative statements.
Traffic officers often misuse the powers they are given.	There need to be more traffic officers on the road to make driving safer.
First-offence drunk drivers should be allowed to keep their licences.	Courts should have the power to remove vehicles of repeat traffic offenders.
Penalties for drunk driving are too harsh.	Traffic officers should have more power to deal with drunk drivers.
Alcohol doesn't have much effect if you are a good driver.	Drinking and driving is the most serious traffic offence.
Drinking and driving is not as dangerous as it's made out to be.	Convicted drunk drivers should be made to attend a defensive driving course.

offences ranging in seriousness from a parking ticket to causing death through the misuse of a motor vehicle:

"Now I'd like to ask you about any traffic offences you've committed.

Could you please indicate if in the past twelve months you have been fined for, or convicted of, any of the following traffic offences? PLEASE TICK ALL THAT APPLY

- A parking offence
- Running a red light/stop sign
- Failing to indicate correctly
- Misuse of a vehicle causing death
- Drinking and driving
- Not wearing a safety belt
- Speeding
- Not having a warrant of fitness
- None of these

Procedure

In September 1989, a short, self-completion questionnaire, covering letter and reply-paid envelope were mailed to a total of 303 motorists convicted in the previous 12 months of driving under the influence of alcohol. Respondents were told that this was a survey on driving habits and attitudes among a random sample of the population, conducted by the Marketing Department of Massey University. Two weeks later a similar package was sent to sample members who had not responded to the initial mailout, and the process was repeated two weeks after this for all remaining non-respondents.

Of the 303 drivers initially selected, two had died and 49 had 'Gone - no address'. This reduced the eligible sample size to 252. From this eligible sample 159 valid replies were received, representing an overall response rate of 63%. The response rates for the different versions of the questionnaire varied between 61% and 67% (see Table 1).

Table 1. Analysis of responses

Response	Version 1		Version 2		Version 3		Version 4		Total	
	n	%	n	%	n	%	n	%	n	%
Valid Response	41	53	37	49	39	53	42	55	159	52
Ineligible ¹	13	17	14	19	10	14	14	19	51	17
Refusals	-	-	-	-	1	1	4	5	5	2
No Response	23	30	24	32	24	32	17	22	88	29
Total	77	100	75	100	74	100	77	100	303	100
Response rate ²		64%		61%		61%		67%		63%

Note.

1. Includes 'Gone-no address', overseas or dead.

2. Response rate = [Valid responses/(Total sample-Ineligibles)]x100.

Results

The average admission rate for the four versions of the questionnaire was 75%, but ranged from 62% to 85% (see Table 2). These admission rates compare favourably with the 65% admission rate achieved by Locander and Burton (1976) using a randomised response technique. However, this difference may be the result of cultural differences between New Zealand and the USA or changes in social values over time, rather than the result of question effects.

Version 1 (the version containing attitude statements which downplayed the seriousness of drunken driving) yielded the highest admission rate (85%), and Version 4 (which asked the

'key' question directly) the lowest admission rate (62%). The difference between these two admission rates is significant at the 5% level.

Version 2 (which contained attitude statements emphasising the seriousness of drunken driving) and Version 3 (which embedded the question in a list of traffic offences) achieved similar admission rates; 76% and 77%, respectively. However, the differences in admission rates between these versions and either Version 1 or Version 4 are not significant.

The above results are concerned solely with valid replies to the survey, but accuracy of valid responses is only one measure of data quality. A more comprehensive view of data quality requires consideration of the number of sample members who either elect not to respond to the survey or who do not respond to particular questions.

Table 2. Analysis of admission rates for drink driving convictions

Response	<u>Version 1</u>		<u>Version 2</u>		<u>Version 3</u>		<u>Version 4</u>		<u>Total</u>	
	n	%	n	%	n	%	n	%	n	%
Admitted offence	35	85	28	76	30	77	26	62	119	75
Denied offence	3	7	5	14	5	13	16	38	29	18
No answer	3	7	4	11	4	10	-	-	11	7
Total	41	100	37	100	39	100	42	100	159	100

Note.

Version 1: contained five attitude statements downplaying the importance of drunken driving, followed by the 'key' question asked directly.

Version 2: contained five attitude statements emphasising the importance of the offence, followed by the 'key' question, asked directly.

Version 3: contained no attitude statements; embedded the 'key' question in a list of traffic offences.

Version 4: contained no attitude statements; asked the 'key' question directly.

Version 4, the most direct form of the questionnaire, had the highest response rate (67% after allowing for refusals and complete non-response) but the lowest admission rate among those who did respond (see Tables 1 and 2). This was the only questionnaire not affected by item non-response, but was the most affected by complete refusals (although the numbers involved are too small to draw any conclusions from). However, there was no evidence of a consistent inverse relationship between response rate and admission rate.

Discussion

In our survey, we assumed that drunken driving had a social stigma attached to it and, as a result, that respondents would be reluctant to admit to this behaviour. Previous research conducted in the USA suggested that, depending on the question design employed, between

35% and 48% of respondents would not admit to a conviction for drunken driving (Kalton & Schuman 1982). However, only 25% of our respondents failed to report a drunken driving conviction. Thus, although our results support the belief that the survey was subject to social desirability bias, the evidence was less conclusive than expected.

The questionnaire which produced the highest admission rate contained positively-worded attitude statements (Version 1). This suggests that the attitude statements successfully downplayed the importance of drunken driving and hence reduced the perceived cost to respondents of admitting the behaviour. Many researchers claim that every attempt must be made to reduce the perceived social undesirability of a behaviour if research is to yield high, and presumably accurate, admission rates to that behaviour (Kalton & Schuman 1982; Sudman & Bradburn 1982). Our research corroborates this claim.

The version of the questionnaire which contained negatively-worded attitude statements (Version 2) also produced a higher admission rate than the version which simply asked the direct question (Version 4). This was contrary to our expectations, since this treatment attempted to emphasise the social undesirability of drunken driving. It appears that the presence of attitude statements had an effect on respondents' willingness to admit to a drunken driving conviction, regardless of their content. There are several possible reasons for this.

Sudman and Bradburn (1982) speculated that including attitudinal questions about a behaviour would lead respondents to believe the researcher was more interested in their responses to these attitudinal questions than their response to the behavioural question. Respondents might then perceive the behavioural question to be for classification purposes only, and as a result feel less threatened by it. Another possibility is that, having answered a larger number of questions about a topic, respondents may feel more comfortable about the topic and so lose their inhibitions about admitting to the behaviour.

Nevertheless, while the mere presence of attitude statements appeared to reduce the social undesirability of a conviction for drunken driving, the version of the questionnaire containing 'positive' attitude statements produced a higher admission rate than the version containing 'negative' statements. Given that the only difference between the questionnaires was in the wording of their respective attitude statements, it is reasonable to assume that the wording of these statements was responsible for this effect (although the difference in admission rates is not significant).

The logic for embedding a socially undesirable behaviour in a list of related behaviours also focuses on reducing the perceived importance of the behaviour to respondents, thereby reducing the 'costs' associated with admitting to the behaviour (Dillman 1978). The embedded question may achieve this by reducing the perceived undesirability of the behaviour in relation to other, more undesirable behaviours, by making respondents more accustomed to admitting behaviour, or by reducing the time respondents spend considering each behaviour.

Whatever the reason, our results suggest that embedding a socially undesirable behaviour in a list does increase the admission rate for such behaviours above those achieved by a direct question, though the difference may be small. However, given that it has not been established why this particular format produces increased admission rates, researchers should exercise caution when using lists in preference to direct questions. For example, if lists are more

successful because they reduce the time respondents spend considering each behaviour, then their use may increase the risk of errors due to inaccuracies in respondents' recall.

The treatment of non-respondents also requires consideration. In the absence of any knowledge about the actual rate of drink driving offences, the best estimate of this parameter would normally be the proportion of valid responses which contained an admission to the behaviour. However, this decision assumes no non-response bias, and, for a survey investigating a socially undesirable behaviour, it is easy to imagine why non-respondents may differ from respondents.

Consequently, if the objective is to estimate the incidence of a socially undesirable behaviour, it is clearly not justifiable to consider only valid responses to a survey, since the admission rate to the behaviour will almost certainly depend, at least in part, on the response rate achieved.

Unfortunately this study showed no evidence of a consistent relationship between response rate and admission rate, despite the fact that the questionnaire with the highest response rate achieved the lowest admission rate. However, because all sample members had committed the behaviour concerned, it was possible to learn something about the accuracy of the methods used as well as their relative performance.

On the basis of our results, estimates of a socially undesirable behaviour such as drunken driving based on self reporting of respondents to mail surveys should be increased by between 15% and 25%. This assumes that some attempt has been made to increase the likelihood of respondents admitting to the behaviour. For other types of socially undesirable behaviour, the size of the scaling factor required will depend on the perceived seriousness of the behaviour compared to drunken driving.

Conclusions

The average admission rate to a conviction for drunken driving was 75%, confirming that under-reporting can lead to inaccurate estimates of socially undesirable behaviours or attitudes. To reduce this bias researchers can attempt to reduce the perceived social undesirability of the behaviour concerned. One effective way of achieving this in a self-completion survey is to include questions, such as attitude statements, which downplay the importance of the behaviour, before asking the behavioural question.

In some cases, there may be a trade-off between overall response rate, or individual item response rate, and admission rate, but our study did not examine this issue in detail. On the basis of our results, the best advice for researchers is that normal estimates of socially undesirable behaviour such as drunken driving should be inflated by at least 15%, even when attempts have been made to reduce the cost to respondents of admitting the behaviour.

Finally, the fact that none of our treatments yielded a 100% admission rate means that further research will be required to identify the 'optimum' questionnaire format for estimating socially undesirable behaviour. This research could investigate the effect of combining techniques thought to reduce social desirability bias. It would be interesting, for example, to study the effect of combining attitude statements with a behavioural question embedded in a list.

References

- Belson WA (1986). *Validity in Survey Research*. London: Gower Publishing Co.
- Converse JM & Presser S (1986). *Survey Questions: Handcrafting the Standardised Questionnaire*. New Delhi: Sage Publications.
- Dillman DA (1978). *Mail and Telephone Surveys: The Total Design Method*. New York: John Wiley and Sons.
- Kalton G & Schuman H (1982). The effect of the question on survey responses: A review. *Journal of the Royal Statistical Society*, 145 (1), 42-57.
- Locander WB & Burton JP (1976). The effect of question form on gathering income data by telephone. *Journal of Marketing Research*, 13, 189-192.
- Peterson RA (1984). Asking the age question: A research note. *Public Opinion Quarterly*, 48, 379-383.
- Phillips DL & Clancy KJ (1971/72). Some effects of 'social desirability' in survey studies. *American Journal of Sociology*, 77, 921-939.
- Sudman S & Bradburn N (1982). *Asking Questions*. San Francisco: Jossey-Bass.
- Philip Gendall is Professor and Head of Department, Janet Hoek is a Lecturer and Matt Blakeley was a BBS(Hons) student, in the Department of Marketing, Massey University.**