SURVEYS · QUESTIONNAIRE DESIGN

Asking Sensitive Questions with Care and Precision

Questions about illegal behaviour, health, money, sex or socially disapproved views produce refusals and misreporting. This article explains what makes a question sensitive, how respondents misreport and the design and mode choices, indirect techniques and ethical safeguards that improve both accuracy and respect for respondents.

ModQu research article · 2,619 words · 24 references (APA 7th edition)

Some questions are easy to ask and easy to answer. Others touch on matters that people would rather keep private or that might make them look bad: drug use, sexual behaviour, income, illness, voting, prejudice, cheating, unpaid taxes. When surveys ask such questions, some respondents refuse, some say they do not know and some give answers that are not true. The results can be badly biased. Studies comparing survey answers with records have found, for example, that many people who did not vote report that they did (Silver et al., 1986; Ansolabehere & Hersh, 2012).

Asking sensitive questions well is therefore one of the central skills of questionnaire design. It requires understanding why questions are sensitive, how respondents respond to sensitivity and which design choices reduce the problem. It also requires care for respondents, who are entitled to privacy and respect. This article covers both.

What makes a question sensitive

Tourangeau and Yan (2007) distinguished three kinds of sensitivity.

- **Intrusiveness.** Some questions are seen as an invasion of privacy, whatever the answer. Questions about income or religion may feel intrusive to some respondents simply because they touch on private matters.
- **Threat of disclosure.** Some questions carry a risk that the answer could become known to others, such as employers, family members or authorities, with harmful consequences. Questions about illegal behaviour are an example.
- **Social undesirability.** Some questions have answers that are socially approved or disapproved. Respondents may want to present themselves favourably, by over-reporting desirable behaviour, such as voting or exercising, and under-reporting undesirable behaviour, such as drinking or prejudice.

Sensitivity is not a fixed property of a question. It depends on the respondent, the setting and the culture. A question about alcohol may be harmless in one country and highly sensitive in another; a question about income may be routine for one respondent and embarrassing for another. Reviews of research on social desirability bias stress that it arises from the interaction between respondents' needs for approval, the perceived privacy of the situation and the content of the question (Krumpal, 2013).

How respondents respond to sensitivity

Respondents can deal with a sensitive question in several ways.

- **Refusing** to answer the whole survey, so that people with sensitive characteristics are under-represented.
- **Skipping** the question or answering "don't know", producing item nonresponse.
- **Misreporting**, by giving an answer that is more socially acceptable than the truth.

Misreporting is the most damaging, because it is invisible in the data. Evidence suggests that much misreporting on sensitive topics is deliberate rather than the result of poor memory: respondents edit their answers before reporting them (Tourangeau & Yan, 2007). Psychologists distinguish between unconscious **self-deception**, in which people genuinely believe a favourable view of themselves, and conscious **impression management**, in which they adjust their answers for others (Paulhus, 1984). Survey design can do more about the second than the first.

Choosing the mode

The single most powerful influence on reporting of sensitive information is the mode of data collection. People report more sensitive behaviour when they answer privately than when they answer to an interviewer.

Self-administration

A large body of experiments has found that self-administered modes produce more reports of sensitive behaviour than interviewer-administered ones. In an experiment comparing different ways of asking about sexual behaviour and drug use, self-administration on a computer produced higher reporting than interviewer administration (Tourangeau & Smith, 1996). A study of adolescents found that reports of sensitive behaviour, such as male-male sex and drug use, were several times higher with audio computer-assisted self-interviewing, in which respondents hear questions through headphones and enter answers themselves, than with self-administered paper questionnaires (Turner et al., 1998).

Web and automated modes

A comparison of telephone interviews, interactive voice response and web surveys found that the web produced the most reports of sensitive information and the most accurate reports when checked against records (Kreuter et al., 2008). Research on interviews conducted by text message on smartphones found more disclosure of sensitive information in text than in voice interviews (Schober et al., 2015).

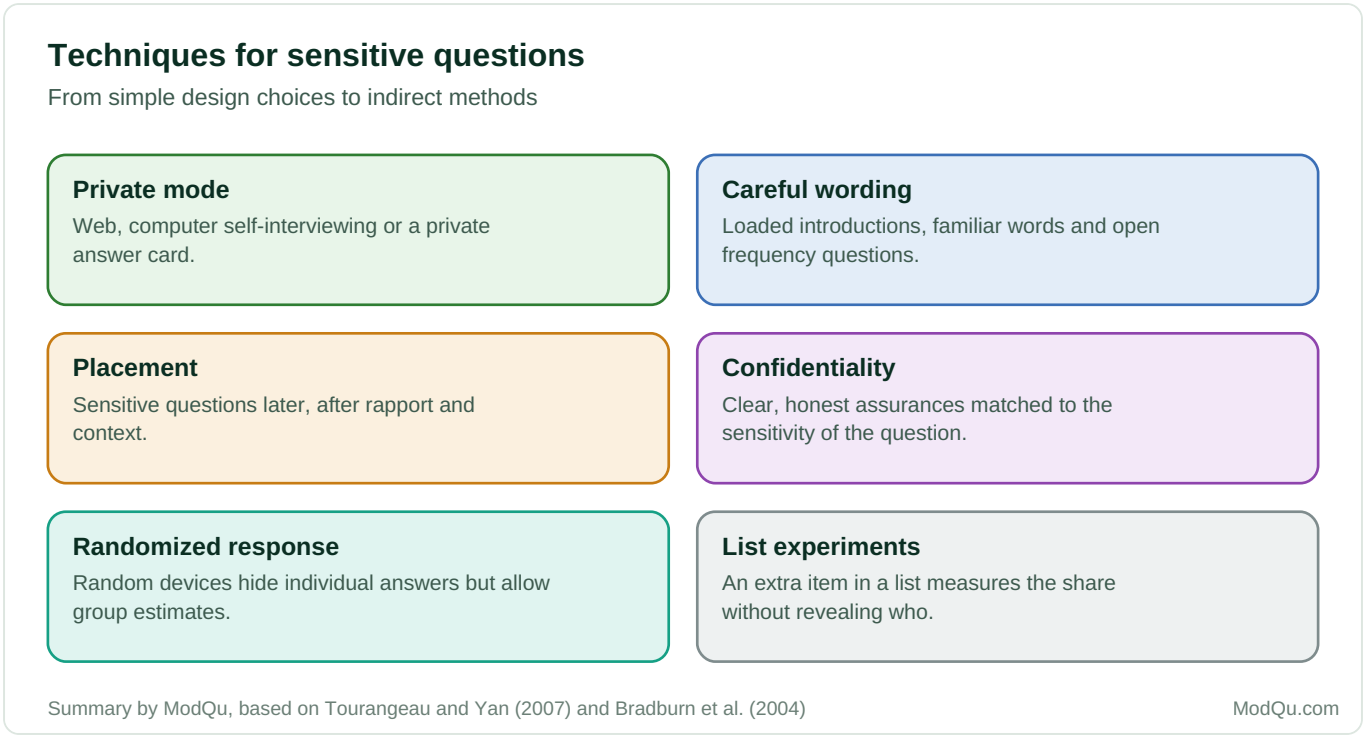


Figure 1. Main techniques for improving answers to sensitive questions. Summary by ModQu, based on Tourangeau and Yan (2007) and Bradburn et al. (2004).

When an interviewer is needed

Many surveys need interviewers for other reasons, such as reaching people without internet access or keeping response rates high. In these surveys, sensitive sections can be self-administered: the interviewer hands over a laptop or tablet, or a paper form in a sealed envelope, so that the respondent answers privately. Simple techniques such as showing a card with lettered answer options and asking the respondent to give only the letter also reduce the sense of disclosure.

Wording sensitive questions

Loading the question

Question wording can make undesirable answers easier to give. Classic advice includes several techniques, sometimes called **loading** (Bradburn et al., 2004; Bradburn & Sudman, 1979):

- **Everybody does it:** "Even the calmest parents get angry at their children sometimes. Did your children do anything in the past seven days to make you angry?"
- **Assume the behaviour:** "How many cigarettes do you smoke each day?" rather than "Do you smoke?", with an option for none.
- **Authorities recommend it:** "Many doctors now think that drinking wine reduces heart attacks. How many glasses of wine did you drink last week?"

- **Reasons for not doing it:** "In talking to people about elections, we often find that a lot of people were not able to vote because they weren't registered, they were sick, or they just didn't have time. Which of the following statements best describes you?"

These techniques can increase reporting of undesirable behaviour, but they must be used with care, because they can also encourage over-reporting or offend respondents. Face-saving answer options for voting questions, which give respondents acceptable reasons for not having voted, have reduced over-reporting in several studies (Belli et al., 1999).

Familiar words and open frequencies

Using the words respondents themselves use for sensitive behaviour, rather than formal or technical terms, can make questions easier to answer. For frequency questions, open answer formats or answer ranges that extend well beyond typical behaviour avoid signalling that a high answer is unusual. If the top category for alcohol consumption is "more than 7 drinks a week", respondents who drink 20 may feel that they are extreme; a scale that goes higher can make them more comfortable reporting accurately (Bradburn et al., 2004).

Long questions and context

For questions about threatening behaviour, longer questions with more introduction have been found to increase reporting, perhaps because they give respondents more time to recall and signal that the topic is a normal one to discuss (Bradburn & Sudman, 1979). Placing a sensitive question after related, less sensitive ones can also ease the transition.

Asking about time periods

For sensitive behaviour, the reference period also matters. Respondents may be more willing to admit that they have ever done something than that they did it recently, because past behaviour seems less connected to who they are now. A common approach asks first about lifetime behaviour and then about recent behaviour, which eases respondents into the topic. The order also produces useful data on both lifetime and current prevalence (Tourangeau & Yan, 2007).

Not knowing versus not telling

Some answers to sensitive questions are inaccurate not because respondents are hiding the truth but because they do not know it or cannot remember it. Questions about income, debt or medical history may be sensitive and difficult at the same time. Separating the two problems helps: offering respondents time to consult records, or ranges instead of exact amounts, addresses difficulty, while privacy and confidentiality address sensitivity.

Placement and context

Sensitive questions are usually placed later in the questionnaire, after respondents have become engaged and built some trust in the survey. Placing them too early can cause break-off before any data are collected. Placing them in a section with a clear explanation of why the information is needed helps respondents understand that the question is relevant, not idle curiosity.

Confidentiality and anonymity

Respondents are more willing to answer sensitive questions when they believe their answers will be protected. A review of experimental studies found that assurances of confidentiality increased response and reduced item nonresponse for sensitive topics, but had little effect, or even a negative one, when topics were not sensitive (Singer et al., 1995). Elaborate assurances for harmless questions can make respondents suspicious that something sensitive is coming.

Assurances must be honest. If data will be linked to other records or shared with others, respondents must be told. Where possible, surveys can collect sensitive data without identifiers, or store identifiers separately from answers, so that the promise of confidentiality is backed by practice.

Indirect techniques

When even private modes and careful wording are not enough, indirect techniques can be used. These methods are designed so that no one, including the researcher, can know an individual respondent's answer, while still allowing estimates for the group.

Randomized response

The **randomized response technique**, introduced by Warner (1965), uses a random device, such as a coin or a die, that only the respondent sees. In one common version, the respondent answers the sensitive question truthfully if the device shows one outcome and answers "yes" regardless if it shows another. Because the researcher knows the probabilities of the random device, the share of "yes" answers due to the sensitive behaviour can be estimated for the group, but no individual answer reveals anything certain about the respondent.

A meta-analysis of validation studies found that randomized response produced more valid estimates of sensitive behaviour than direct questioning, although the improvement was not universal and the method was less valid when the topic was very sensitive (Lensvelt-Mulders et al., 2005). The method has costs: it requires much larger samples, because the randomisation adds noise, and some respondents find it confusing or do not follow the instructions. Recent statistical work has developed better ways to design and analyse randomized response studies (Blair et al., 2015).

List experiments

The **list experiment**, or **item count technique**, randomly divides respondents into two groups. The control group sees a list of several non-sensitive items and reports how many apply to them, without saying which. The treatment group sees the same list plus the sensitive item and also reports only the number. The difference between the average counts in the two groups estimates the share of respondents to whom the sensitive item applies.

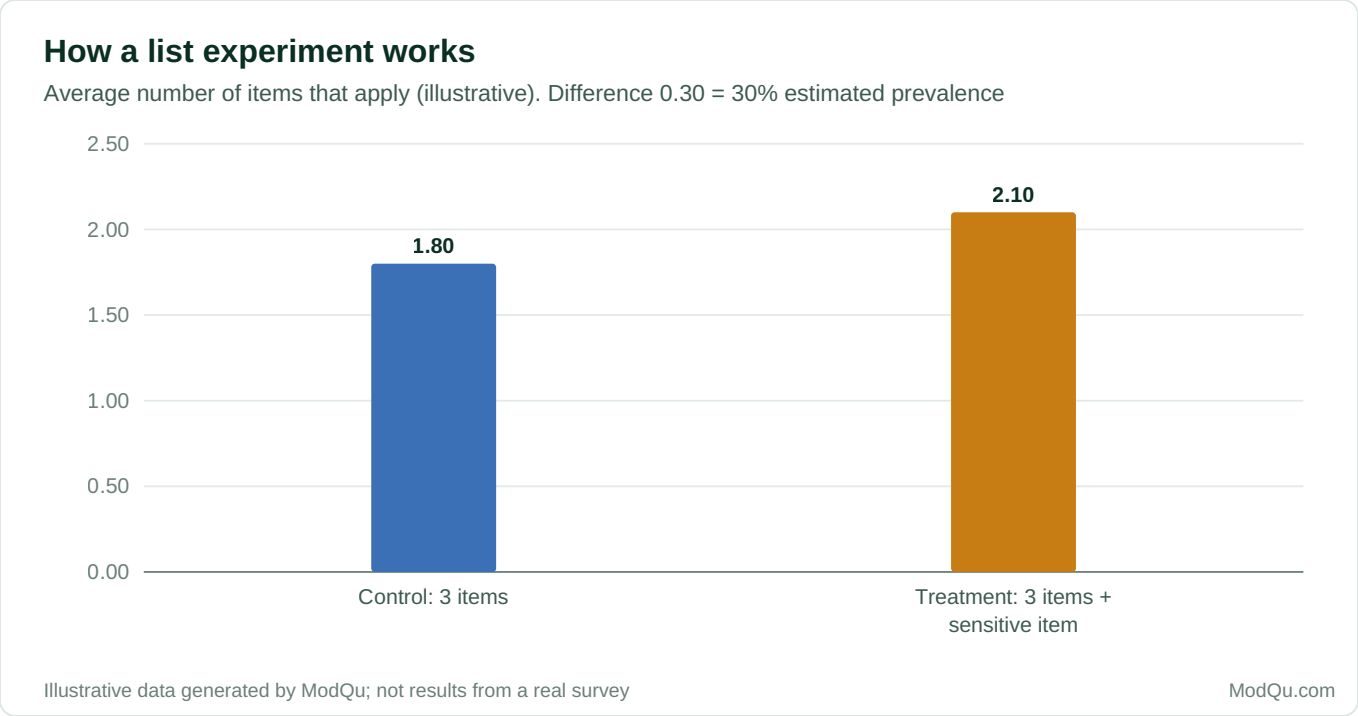


Figure 2. How a list experiment estimates a sensitive behaviour. The difference in average counts between the two groups (2.10 − 1.80 = 0.30) estimates that 30 per cent have the characteristic. Illustrative data generated by ModQu.

List experiments have been used to study racial attitudes, which respondents are reluctant to express directly (Kuklinski et al., 1997), and voter turnout, where they reduced over-reporting compared with direct questions (Holbrook & Krosnick, 2010). Their weaknesses include large sampling error, because the estimate depends on a difference between two averages, and problems when respondents would answer yes to all or none of the control items, which can reveal their answer to the sensitive item. Guidance on designing the control list to avoid these problems and on statistical analysis has improved the method considerably (Blair & Imai, 2012; Glynn, 2013).

How well do indirect techniques work?

A study that compared direct questions, randomized response and list experiments against known true values from official records of a referendum vote found that the indirect methods reduced bias compared with direct questioning, with randomized response performing best in that study, but that all methods had substantial variance (Rosenfeld et al., 2016). Indirect techniques are best used when the topic is highly sensitive, the sample is large and the goal is to estimate the prevalence of a behaviour rather than to analyse individual answers.

The bogus pipeline

A different approach, mainly used in laboratory research, is the **bogus pipeline**, in which respondents are led to believe that a device, such as a lie detector, can verify their answers. A meta-analysis found that this increased reporting of socially undesirable attitudes and behaviour (Roese & Jamieson, 1993; Jones & Sigall, 1971). Because it relies on deception, it is rarely appropriate in general surveys.

Measuring social desirability

Some surveys include scales that measure each respondent's tendency to give socially desirable answers, such as the Marlowe-Crowne scale, which asks about behaviour that is desirable but rare, or undesirable but common (Crowne & Marlowe, 1960). Scores can be used to identify respondents prone to such answers or as control variables in analysis. However, these scales mix genuine virtue with impression management and add length, so their use in general surveys is limited (Paulhus, 1984).

Validating answers

Where possible, the accuracy of answers to sensitive questions should be checked against an external source. Records of voting, hospital admissions, convictions or benefit receipt, linked with consent, show how far survey answers can be trusted. Such validation studies have been the main source of evidence about misreporting and about which methods reduce it (Ansolabehere & Hersh, 2012; Tourangeau & Yan, 2007).

Interviewers and sensitive questions

When interviewers are used, their characteristics and behaviour can affect answers to sensitive questions. Respondents may answer differently depending on the interviewer's sex, age or ethnicity, particularly when the question concerns attitudes related to those characteristics. Training interviewers to read sensitive questions in a neutral, matter-of-fact tone and to accept any answer without reaction helps to reduce these effects. The presence of other people during the interview, such as a spouse or parent, can also reduce reporting, so interviewers should try to conduct sensitive sections in private (Tourangeau & Yan, 2007).

Ethics and care for respondents

Sensitive questions raise ethical as well as methodological issues. Respondents should know before they agree to take part that the survey includes sensitive topics, and they should be able to skip any question without penalty. Questions about traumatic experiences, such as violence or abuse, can cause distress, and surveys that ask them should provide information about sources of support. Sensitive data must be stored securely and reported only in ways that cannot identify individuals. A sensitive question should be asked only when the answer is genuinely needed for the purpose of the research.

In business surveys

Business surveys ask sensitive questions more often than is sometimes recognised. Employee surveys ask about harassment, discrimination and views of managers; customer surveys ask about income, health conditions or debt; market research asks about alcohol, gambling or personal care products. In employee surveys especially, respondents may fear that their answers will be traced back to them. Reporting only for groups large enough to protect anonymity, using an independent organisation to collect the data and explaining these protections clearly are all important for honest answers (Krumpal, 2013).

A checklist for sensitive questions

1. Ask only what is needed, and explain why it is needed.
2. Use a self-administered mode, or a private section within an interview.
3. Place sensitive questions after rapport has been built.
4. Use familiar words, neutral wording and answer ranges that do not make true answers look extreme.
5. Consider face-saving options and loaded introductions where over- or under-reporting is likely.
6. Give honest assurances of confidentiality that match the sensitivity of the question.
7. Always allow respondents to decline.
8. Consider indirect techniques for highly sensitive topics with large samples.
9. Validate answers against records where possible.
10. Report results with an honest account of possible misreporting.

Conclusion

Sensitive questions are among the hardest in survey research. Respondents may refuse, skip or misreport, and misreporting can bias results without leaving any trace in the data. The most effective remedies are private modes of data collection, careful wording, thoughtful placement and credible confidentiality. Indirect techniques such as randomized response and list experiments can go further, at the cost of precision. Above all, sensitive questions should be asked with respect for respondents, whose willingness to share private information is the foundation on which this kind of knowledge rests.

References

Ansolabehere, S., & Hersh, E. (2012). Validation: What big data reveal about survey misreporting and the real electorate. *Political Analysis, 20*(4), 437–459. https://doi.org/10.1093/pan/mps023

Belli, R. F., Traugott, M. W., Young, M., & McGonagle, K. A. (1999). Reducing vote overreporting in surveys: Social desirability, memory failure, and source monitoring. *Public Opinion Quarterly, 63*(1), 90–108. https://doi.org/10.1086/297704

Blair, G., & Imai, K. (2012). Statistical analysis of list experiments. *Political Analysis, 20*(1), 47–77. https://doi.org/10.1093/pan/mpr048

Blair, G., Imai, K., & Zhou, Y.-Y. (2015). Design and analysis of the randomized response technique. *Journal of the American Statistical Association, 110*(511), 1304–1319. https://doi.org/10.1080/01621459.2015.1050028

Bradburn, N. M., & Sudman, S. (1979). *Improving interview method and questionnaire design: Response effects to threatening questions in survey research*. Jossey-Bass.

Bradburn, N. M., Sudman, S., & Wansink, B. (2004). *Asking questions: The definitive guide to questionnaire design for market research, political polls, and social and health questionnaires* (Rev. ed.). Jossey-Bass.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*(4), 349–354. https://doi.org/10.1037/h0047358

Glynn, A. N. (2013). What can we learn with statistical truth serum? Design and analysis of the list experiment. *Public Opinion Quarterly, 77*(S1), 159–172. https://doi.org/10.1093/poq/nfs070

Holbrook, A. L., & Krosnick, J. A. (2010). Social desirability bias in voter turnout reports: Tests using the item count technique. *Public Opinion Quarterly, 74*(1), 37–67. https://doi.org/10.1093/poq/nfp065

Jones, E. E., & Sigall, H. (1971). The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin, 76*(5), 349–364. https://doi.org/10.1037/h0031617

Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and web surveys: The effects of mode and question sensitivity. *Public Opinion Quarterly, 72*(5), 847–865. https://doi.org/10.1093/poq/nfn063

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity, 47*(4), 2025–2047. https://doi.org/10.1007/s11135-011-9640-9

Kuklinski, J. H., Cobb, M. D., & Gilens, M. (1997). Racial attitudes and the "New South". *Journal of Politics, 59*(2), 323–349. https://doi.org/10.2307/2998167

Lensvelt-Mulders, G. J. L. M., Hox, J. J., van der Heijden, P. G. M., & Maas, C. J. M. (2005). Meta-analysis of randomized response research: Thirty-five years of validation. *Sociological Methods & Research, 33*(3), 319–348. https://doi.org/10.1177/0049124104268664

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*(3), 598–609. https://doi.org/10.1037/0022-3514.46.3.598

Roese, N. J., & Jamieson, D. W. (1993). Twenty years of bogus pipeline research: A critical review and meta-analysis. *Psychological Bulletin, 114*(2), 363–375. https://doi.org/10.1037/0033-2909.114.2.363

Rosenfeld, B., Imai, K., & Shapiro, J. N. (2016). An empirical validation study of popular survey methodologies for sensitive questions. *American Journal of Political Science, 60*(3), 783–802. https://doi.org/10.1111/ajps.12205

Schober, M. F., Conrad, F. G., Antoun, C., Ehlen, P., Fail, S., Hupp, A. L., Johnston, M., Vickers, L., Yan, H. Y., & Zhang, C. (2015). Precision and disclosure in text and voice interviews on smartphones. *PLOS ONE, 10*(6), e0128337. https://doi.org/10.1371/journal.pone.0128337

Silver, B. D., Anderson, B. A., & Abramson, P. R. (1986). Who overreports voting? *American Political Science Review, 80*(2), 613–624. https://doi.org/10.2307/1958277

Singer, E., Von Thurn, D. R., & Miller, E. R. (1995). Confidentiality assurances and response: A quantitative review of the experimental literature. *Public Opinion Quarterly, 59*(1), 66–77. https://doi.org/10.1086/269458

Tourangeau, R., & Smith, T. W. (1996). Asking sensitive questions: The impact of data collection mode, question format, and question context. *Public Opinion Quarterly, 60*(2), 275–304. https://doi.org/10.1086/297751

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin, 133*(5), 859–883. https://doi.org/10.1037/0033-2909.133.5.859

Turner, C. F., Ku, L., Rogers, S. M., Lindberg, L. D., Pleck, J. H., & Sonenstein, F. L. (1998). Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. *Science, 280*(5365), 867–873. https://doi.org/10.1126/science.280.5365.867

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association, 60*(309), 63–69. https://doi.org/10.1080/01621459.1965.10480775