How to Overcome Social Desirability Bias for Very Sensitive Topic Research

Online is an effective methodology to gain answers to sensitive questions because the perceived anonymity from the point of view of the participant is high, and there is less risk of social desirability bias caused by the participant wanting to “please the interviewer”.

Usage of a technique called Randomized Response can be effective in eliciting candid answers for topics which are especially sensitive.

There are a variety of ways to use this technique, but the easiest one to understand is called “the unrelated question with the known distribution” design (Greenberg et al 1969).

The researcher presents two possible questions to the respondent and only one answer list. One of the questions is the sensitive one, and the other is some question to which the researcher already knows how the answers should distribute (for example, month of birth, or the last digit of the respondent’s telephone number.) The answer list is Yes/No or True/False.

The respondent uses some randomization technique (for example flipping a coin) to decide which question they should answer and they then check the appropriate box. The researcher has the answer to the question... but no idea which question was being answered! The respondent knows the researcher doesn’t know, so they can be 100% honest if the sensitive question is the one they have to answer.

The technique works because you know the two essential probabilities: the probability of the respondent answering the sensitive question and the probability of the respondent answering “yes” to the non-sensitive question.

An example:
1. Interview a sample of 100 people
2. Ask them to toss a coin to decide whether to answer a sensitive question or a non-sensitive one.
3. The non-sensitive question has a known probability of a “yes” answer of 60%.
4. The study results show 70% answering yes
5. We know that in 50% of cases the non-sensitive question was answered, and that the non-sensitive question has a 60% probability of being a Yes. Therefore 30 out of the 100 answers are yes answers to the non-sensitive question (60% of 50%)
6. There were 70 yes answers in total, 30 were from the non-sensitive question, so the remaining 40 are from the sensitive one

7. We know that the sensitive one was answered 50 times, 40 out of 50 is 80% -- which is the answer to the sensitive question

The technique is effective, but has some limitations:
• We don’t know which of the 100 respondents answered the sensitive question so are unable to do any sub-analyses on them or use this answer as an analysis variable.
• The technique has been shown to increase non-response since people don’t trust either how it would work or how it might protect their anonymity.
• People could still be lying to themselves to protect their own self-esteem
• The effective sample size is reduced by 50% and, because of the number of random elements being used, the variance is high – meaning we can have less confidence in the answers. (The throwing of heads/tails may not be 50/50 in this case, and this particular sample may not match the known probability of the non-sensitive question)