Comment on David Moore's "What Single Polls and Their Marginals Tell Us (Or Not!)"

Howard Schuman

Tags: survey practice

Survey Practice

Vol. 5, Issue 1, 2012

Comment on David Moore's "What Single Polls and Their Marginals Tell Us (Or Not!)"*

David Moore writes that someone with a name that sounds much like mine "appears to say that as far as marginal results go, it doesn't matter whether question wording is tendentious, or whether non-opinion and intensity are measured—that such results don't really tell us much 'no matter the wording.'" None of this is true for me.

The present tempest in a teapot arose out of a short article I wrote for the June 2011 issue of *Survey Practice* about the importance of measuring strength of opinion. I presented evidence that some strength measures were better than others in the case of the issue of requiring permits for guns. The article reported data I gathered that deliberately included a measure of intensity, though it also indicated that a different and apparently more useful measure of attitude strength was obtained as well. How could David conclude from that article that I don't care whether "intensity" is measured?

After David's first article in *Survey Practice*, in which he accused me of "nihilism"—

"the [philosophical](#) doctrine suggesting the [negation](#) of one or more putatively [meaningful](#) aspects of [life](#)," according to *Wikipedia*—I suggested that he look at my most recent and fullest attempt to reflect on five decades of experience studying the question-answer process in surveys: *Method and Meaning in Polls & Surveys* (2008), especially Chapter 1 entitled "Ordinary Questions, Survey Questions, and Policy Questions," though chapter 2 is relevant as well because the marginals for open and closed questions are compared. He would find there a much more comprehensive and nuanced approach than was possible in a two page article in *Science* magazine written more than 25 years ago primarily to alert natural science readers to research on the question-answer process in surveys.

To avoid prolonging this controversy further, let me acknowledge here and now that if no other good evidence were available about public attitudes toward bike lanes in New York City, I would pay attention to poll results based on a single question that asked a reasonable sample of the relevant population

their views of bike lanes. I would certainly not wish the question to be tendentious, whatever that means in this case, nor to ignore non-opinion and intensity. If two versions of the question were available, all the better, since each tells us something different about public attitudes.

What I would object to, however, is reporting the response percentages as though they reflect in a simple sense "public opinion"—an important concept but one very difficult to specify—and leaving it at that. For example, if non-opinion is tapped by the proportion of people who say "don't know," then I would immediately call to mind the evidence in my 1981 book (with Stanley Presser) of split-sample experiments that the DK% is not an absolute but depends on how much such a response is encouraged, discouraged, are even disallowed altogether.¹ Moreover, if my interest was in predicting how New Yorkers would respond to an actual referendum that asked about their support or opposition to bike lanes, I would try to word a question to fit the wording of the referendum—which might mean either including or excluding a DK option, depending on how the referendum was worded and what is known about the likely turnout and its composition.

Equally important, I would ask other questions as well, for example, whether respondents knew what bike lanes were for and where they are located, whether they themselves had ridden a bicycle in a bike lane, or had ever been bothered by a bicycle in a bike lane, and still other points. Finally, I would do my best to integrate these results in my own mind because for me the crucial point about working with univariate results—whether for a single question or a host of related questions or a single question with a single measure of attitude strength—is that JUDGMENT is critical. One should avoid—as too often happens with popular presentations of poll results—a mechanical report of response percentages as though they should be accepted at face value as reflecting "public opinion," rather than assessed carefully and critically. This is the nub of the issue for me: I don't believe that univariate percentage results should be taken as though they are themselves unofficial referenda. One always must use judgment and be aware of the many challenges when operationalizing ideas in the form of survey questions.

In my 2008 book, referenced above, I distinguished two views of survey results:

survey fundamentalism: "the naïve acceptance of the numbers in a survey report as a literal picture of public opinion..."

survey cynicism: "the equally naïve belief...that poll results are worthless because investigators can readily produce whatever they wish by means of clever question wording or statistical mumbo jumbo."

---

¹ H. Schuman and S. Presser, 1981. Questions & Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context (1981; 1996), pp. 122–125.

I've tried in my books and articles (including a number that have appeared in POQ) to steer between those two extremes, though probably not with success in every case. As far as the present exchange is concerned, it is avoidance of *survey fundamentalism* that has been paramount. But I certainly don't embrace survey cynicism, since almost everything I've written over the past 50 years has been based on taking survey data seriously.

Moreover, Chapter I in *Method and Meaning in Polls and Surveys* regards univariate results so seriously that it includes a discussion of the need to resolve a difficult ethical issue that arose when The Michigan Society for Medical Research asked us to do a survey measuring support for its use of pound animals in research. The results would clearly be treated as a quasi-referendum to be made public and to influence legislation. As Director of the University of Michigan's Survey Research Center at that time, I agreed to do the survey but only if an opponent of such use of pound animals played a part in questionnaire construction. The Medical Society agreed and I then acted as an intermediary to facilitate their joint development of a questionnaire, revising questions as they worked out the issues and wording to be included. For the most part the two hostile parties were able to resolve disagreements over wording, but in two instances where this proved impossible, I created split-sample experiments to determine whether the preferred wording of each made a difference in results. (In one case, the wording variation had little effect on the marginals; in the other case there was a clearly reliable effect and the two parties had to decide how to interpret and report the data.) Not only did this approach to taking marginals seriously resolve disagreements, but a side benefit was that both parties discovered that writing survey questions is not as simple as they had thought.

In still another part of the same chapter, I review a report by Lipset and Ladd that uses marginals to draw essentially referendum-like conclusions. It seemed evident that some (though not all) of their questions were biased in wording and that their argument using the results was misleading. Thus I took their univariate results seriously, but not uncritically.

Finally, at the end of this same chapter I stated that one of the contributions of univariate results has been to help us see the views of others unlike ourselves and those we know personally. It is all too easy to forget that our own friends and neighbors do not tell us how distant parts of the country (and world!) think about issues. Univariate results based on good samples of a general population help to reduce the egocentric tendency to overgeneralize from too restricted a set of contacts.

To quote from the end of the chapter: "Thus polls and surveys that are well done in terms of sampling a general population contribute to creating a more cosmopolitan citizenry, and reports of single variable results can therefore serve a positive function. Yet at the same time, there is lack of public sophistication regarding the extent to which survey results are shaped by how questions

are framed and worded, how much or little they are answered carefully and knowledgeably, and how greatly they may be restricted to a particular point in time." (p. 28) It is the responsibility of the survey researcher or pollster to take account of both sides of this dilemma.

I have done hundreds of split-sample experiments on the form, wording, and context effects of survey questions—many reported in the book *Questions and Answers in Attitude Surveys*, but also many in recent years. These results have given me a healthy respect for unintended problems with question wording, and the sense that judgment is essential when reporting results.² This is, especially true for univariate results because they are ordinarily presented as specific percentages, indicating that X % of the population believes or feels such and such, suggesting an exactness that is seldom warranted and not just because of sampling error. Analytic results reporting change over time and differences by gender, education, etc. are more often presented as qualitative differences (men believe such and such more than women; response X has changed positively over recent years), with less emphasis on exact percentages, thus making them somewhat less likely than marginals to be reified.

I don't wish to argue further with David. I tried in my 2008 book to write succinctly about what I've learned over the years from work with survey questions and survey data. Whether others find the book useful is not for me to determine. But I hope that conclusions about my views will not be based on a few sentences written a long time ago.

ACKNOWLEDGMENTS

I am grateful to Eleanor Singer for providing very constructive comments on a previous draft of this Comment.

---

² For a recent example of an unintended question wording, or better, question framing effect, see Schuman, Corning, and Schwartz "Framing Variations and Collective Memory: 'Honest Abe' vs. 'The Great Emancipator'" (Social Science History, 2012 forthcoming).