

The New York Times

August 27, 2006

The Public Editor

Precisely False vs. Approximately Right: A Reader's Guide to Polls

By JACK ROSENTHAL

LAST March, the American Medical Association reported an alarming rate of binge drinking and unprotected sex among college women during spring break. The report was based on a survey of “a random sample” of 644 women and supplied a scientific-sounding “margin of error of +/- 4.00 percent.” Television, columnists and comedians embraced the racy report. The New York Times did not publish the story, but did include some of the data in a chart.

The sample, it turned out, was not random. It included only women who volunteered to answer questions — and only a quarter of them had actually ever taken a spring break trip. They hardly constituted a reliable cross section, and there is no way to calculate a margin of sampling error for such a “sample.”

The Times published a correction explaining the misrepresentation, and the news media that used the story would probably agree with what Cliff Zukin, a Rutgers authority on polls, told Mystery Pollster, a polling blog: how unfair it is to publish a story “suggesting that college students on spring break are largely drunken sluts.”

The story also threatened larger harm. Its general point was indisputable; vacationing collegians often behave recklessly. But there was a larger recklessness in the misrepresentation of the survey. Now that everyone has a phone and calls are cheap, polling organizations have blossomed, and each such example of bad polls risks undermining public confidence in good ones.

Another example surfaced last week in The Wall Street Journal. It examined a “landmark survey,” conducted for liquor retailers, claiming to show that “millions of kids” buy alcohol online. A random sample? The pollster paid the teenage respondents and included only Internet users.

Such misrepresentations help explain why The Times recently issued a seven-page paper on polling standards for editors and reporters. “Keeping poorly done survey research out of the paper is just as important as getting good survey research into the paper,” the document said.

These standards, coming just as the fall campaign heats up, provide a timely reminder of responsible journalism. But the best of intentions are not always met in practice, at The Times or in other media. The standards do not, for instance, discuss how even a punctilious poll story can be given inflated prominence. There is no reason, in any case, to limit such cautions to journalists. Readers, too, need to know something about polls — at least enough to sniff out good polls from bad. Here's a brief guide.

False Precision

Beware of decimal places. When a polling story presents data down to tenths of a percentage point, what the pollster almost always demonstrates is not precision but pretension. A recent Zogby Interactive poll, for instance, showed that the candidates for the Senate in Missouri

were separated by 3.8 percentage points. Yet the stated margin of sampling error meant the difference between the candidates could be seven points. The survey would have to interview unimaginably many thousands for that zero point eight to be useful.

Experienced researchers offer a rule of thumb: rather than trust improbably precise numbers, round them off. Even better, look for whole fractions.

Sampling Error

The Times and other media accompany poll reports with a box explaining how the random sample was selected and stating the sampling error. Error is actually a misnomer. What this figure actually describes is a range of approximation.

There's also a formula for calculating the error in comparing one survey with another. For instance, last May, a Times/CBS News survey found that 31 percent of the public approved of President Bush's performance; in the survey published last Wednesday, the number was 36 percent. Is that a real change? Yes. After adjustment for comparative error, the approval rating has gained by at least one point.

For a typical election sample of 1,000, the error rate is plus or minus three percentage points for each candidate, meaning that a 50-50 race could actually differ by 53 to 47. But the three-point figure applies only to the entire sample. How many of those are likely voters? In the recent Connecticut primary, 40 percent of eligible Democrats voted. Even if a poll identified the likely voters perfectly, there still would be just 400 of them, and the error rate for that number would be plus or minus five points. So to win confidence, a finding would have to exceed 55 to 45.

This caution applies forcefully to conclusions about other subgroups. What could a typical survey tell about, say, college-age women? Out of a random sample of 1,000, a little more than half would be women and only about 70 would be of college age. That's too small a subsample to support any but the most general findings.

Questions

How questions are phrased can mean wide shifts, even with wholly neutral words. Men respond poorly, for instance, to questions asking if they are "worried" about something, so careful pollsters will ask if they are "concerned."

The classic "double negative" example came in July 1992, when a Roper poll asked, "Does it seem possible or does it seem impossible to you that the Nazi extermination of the Jews never happened?" The finding: one of every five Americans seemed to doubt that there was a Holocaust. How much did that startling finding result from the confusing question? In a follow-up survey, Roper asked a clearer question, and the number of doubters plunged from the original 22 percent to 1 percent.

Extreme questions are fine if the poll asks questions at both extremes, says Frank Newport, editor in chief of the Gallup Poll and author of "Polling Matters," an authoritative 2004 book on this subject. The difference between the answers "can give us good insights into evolving social norms," he says. "All data are interesting."

In any case, Warren Mitofsky, head of a leading international polling company, observes that "for political surveys, most of the questions have been asked for many years, have been tested and are not the source of error."

The order of questions is another source of potential error. That's illustrated by questions asked by the Pew Research Center. Andrew Kohut, its president, says: "If you first ask people what they think about gay marriage, they are opposed. They vent. And if you then ask what they think about civil unions, a majority support that."

Answers

People never wish to look uninformed and will often answer questions despite ignorance of the subject. Some 40 years into the cold war, many respondents were still saying yes, Russia is a member of NATO. That's why, says Rob Daves, head of the American Association of Public Opinion Researchers, skillful pollsters will first ask, for new or sophisticated subjects, a scaling question like, How much do you know about this issue: a great deal, some, not at all?

Respondents also want to appear to be good citizens. When the Times/CBS News Poll asks voters if they voted in the 2004 presidential election, 73 percent say yes. Shortly after the election, however, the Census Bureau reported that only 64 percent of the eligible voters actually voted.

Jon Krosnick, an authority on polling and politics at Stanford, uses the term "satisficing" to describe behavior when a pollster calls. If people find the subject compelling, they become engaged. If not, they answer impatiently. Either way, says Kathy Frankovich, director of surveys for CBS News, "people grab the first thing that comes to mind."

Intensity

How strongly people feel about an issue may be the most important source of poll misunderstanding. In survey after survey, half the respondents favor stronger gun controls — but don't care nearly as much as the 10 percent who want them relaxed.

Intensity can be measured by asking a scaled question: Is the issue of abortion so important that you will cast your vote because of a candidate's position? One of several important issues? Not important? Each added question increases the interview length, testing the respondent's patience and the pollster's budget. Nevertheless, on divisive issues, responsible pollsters will ask four, five, even a dozen questions, probing for true feelings.

Public opinion is not precise, and in any case it is constantly churning. Measuring it cannot hope to be precise. What readers can hope for, whether in an individual poll, a consensus from several polls or from the polling profession generally, is the truth — approximately right.

Jack Rosenthal, president of The New York Times Company Foundation, was a senior editor of The Times for 26 years. [Byron Calame](#), the public editor, is on vacation.