Under the Hood

The following is based on a speech evaluating polling practices that was delivered by Gary Langer, director of polling at ABC News, to colleagues attending the 2000 American Association for Public Opinion Research conference.

The simple message is this: Not all polls are created equal. Like all the alleged news that crosses our desks, every single poll needs to be evaluated carefully and judged on its merits.

Of course, our own polling comes first. We spend a lot of time under the hood, sweating the methodological details. And then we work hard to provide complete and thoughtful analysis. Indeed, it's our view that a poll without analysis is just numbers. The hallmark of serious polling is the acuity of the analysis brought to bear.

We know the importance of what we ask. But the time we have to research and write a questionnaire can be as short as literally minutes. Nonetheless, we're constantly reassured by how robust good sampling is when we see how different, honest approaches—even if not perfectly crafted—produce very similar and coherent results.

We're also occasionally reminded of the differences that can occur. Two polls last spring asked whether Congress should investigate the Elian Gonzalez case. One first informed respondents that Congress intended to do just that. Lo and behold, support for hearings was 11 points higher in the poll that told people hearings were already planned.

But 11 points is not the end of the world. In fact, we take unjustified umbrage when data diverge. Polls too often are imbued with an undeserved sense of finality. It's the notion of mathematical certainty, something that I suspect is Biblical in origin. The writing on the wall said "*mene, mene, tekel, uparsin*"—numbered, numbered counted and weighed. The Lord had numbered Belshazzar's kingdom, counted it and come to its end, weighed it and found it wanting. Since then, numbers have been freighted with the power of final judgment.

With polls it's not so. When polls differ it's not contradiction; it's additional information. When we evaluate differences, we learn more about our craft, and often about public views of the issue at hand. Sometimes we just learn what not to do.

Something as simple as question order can matter. One week last March, a Pew poll had Al Gore up by six points, and a virtually contemporaneous CBS poll had George W. Bush up by seven. In comparing them we found that before they got to the horse race, Pew asked nine other questions, including Clinton approval, Congress approval, and open-ended questions on Bush and Gore characteristics. While CBS's opening set of questions included Clinton approval and favorability measures for each primary candidate, mostly they asked questions different than Pew's.

Did that make the difference? Who knows? But it's a dilemma. I've had the following internal debate: You're going into the field. You want to ask Clinton approval. You always ask it first; that's trend. But you don't want to ask the horse race right after it. So you need to put in some kind of buffer, hopefully innocuous. But will that bias the horse race?

Our solution is to resist asking Clinton approval in our election polls. We ask likelihood of voting. We ask attention to the campaign. And then we ask the horse race.

We all know about the spotty results in New Hampshire this year. I won't say much about the final pre-election polls, but I do want to mention one worrisome trend. In 1996, ABC News conducted eleven nights of tracking in New Hampshire. It's a powerful way to work up to a final estimate; with all those data rolling forward you can really get your hands around the dynamics of the race, discern its trajectory, do some modeling and end up with the ability to make some well-informed judgments. Instead, what we saw this year were more one-shot final weekend polls with small samples and no chance to track the trend. To my conservative nature, that's dangerous—a step in the direction of throwing darts.

That's a judgment call. But some polls are simply inferior. Especially in an election year, we try to vet the methodology of virtually every poll we see—and some would curl your hair: autodialers masquerading as polls, listed rather than RDD samples, no respondent selection protocol, unweighted samples, nutty turnout scenarios, and more. One name-brand poll last winter implied a turnout of 43% of the voting-age population in the South Carolina Republican primary, double any reasonable expectation. Needless to say, it's advisable to look closely at purported likely voter scenarios.

Undecideds can be a significant source of differences, especially in election polls. A poll out of Harvard last winter contrived to show that 74% of Americans had no choice for president. Their question broadly invited people not to answer. Gallup recently did a more neutral test in which it got 26% undecided in the horse race by asking it open-ended. That's fine for a test, but remember, the horse race

doesn't seek to measure who's going to win; it asks whom you'd support if the election were today. If the election were today and you were voting, undecided would not be an option. So there's no reason there should be a high undecided on this question. We're confident in this because there are plenty of other questions—about foreign trade, for example—in which ABC finds high undecideds.

Subgroups are also worth close attention. You'll see a fair amount of over-analysis of really small groups. Last spring one heavily flacked poll made fairly prominent reference to its numbers among Hispanic voters, saying they seemed to be "moving back to Gore." Looking closely, though, this poll had a sample size for Hispanics of 70 people. Let's face it: if you think a subgroup is really important, and you want to analyze it responsibly, you need to take the time, trouble—and expense—to sample it adequately.

Subgroups in general are a big deal because of our fascination with swing voters. I call them the group *du jour*. It's human nature to try to reduce this monumental contest to simpler proportions, to the single group, the single archetype, the single person on whom the whole thing rests, like some little old gay Hispanic lady in Ohio.

But in pressing on we tend to forget that in 1996, for example, soccer moms voted essentially like every other woman.

We've compiled a list of everything that's been described in the papers lately as a swing group. It includes elderly women (that's the *New York Times*), parents (so says the Associated Press), mothers (*New York Times* again), seniors (the AP again) Catholics (says Ralph Reed), Cuban-Americans, Hispanics, suburbanites, something called "techno-Republicans" and, most recently, in *The Washington Post*, "Midwestern non-union non-college-educated middle- and low-income white women."

Now, in our last poll of 1,000 people there were 30 who fit that bill. They supported Bush by a four-point margin, while all women supported Bush by a two-point margin.

We checked the 1996 exit poll and compared all women to "Midwestern non-union non-college-educated middle- and low-income white women." The latter were three points better for Clinton. This with a sample size of 82 out of 4,196 respondents.

In a close election everyone counts. But a swing group has to fit two criteria—its majority vote swings from party to party across elections, and it's big enough to make a difference. Using these criteria leaves us exactly two reliable swing voter groups: independents and white Catholics.

Now in this election we might add women and moderates, since they both swung from Bush last fall to Gore this winter and back to a dead heat in the spring. We'll watch them throughout the campaign. We'll watch the issues, the candidate qualities, and the horse race, too. And above all, we'll try to understand what the people we interview are telling us. Because, fundamentally, a pollster's prime responsibility is not simply to gather numbers—but to make sense of them.

—Gary Langer

Have an opinion? Perhaps a reply to something appearing in Public Perspective? Direct submissions to the editor at pubper@opinion.isi.uconn.edu. Submissions should be no more than 750 words. Authors will be contacted prior to publication.

Comprehensive Coverage. Reasonable Rates.

The oldest and largest archive of polling data in the world.

Roper Center for Public Opinion Research

University of Connecticut

860.486.4440

www.ropercenter.uconn.edu