
EDITORIAL POSTLUDE

HERBERT JACK ROTFELD

Mistaking Precision for Reality

Noting that dozens of organizations conducted numerous opinion surveys prior to the elections last November, *The Wall Street Journal's* November 21 online edition posted tabulations and graphs comparing the past two months of polls and outcomes in various Senate races. At the most casual glance, the polls' generally weak predictive power was patently obvious. Even where the correct winner was selected, the vote margins were way off. In November 2006, pundits and poll readers generally predicted Democrats *might* take over the house, though no one predicted a victory by the numbers revealed when people actually voted, and many doubted the possibility of a Democratic majority in the Senate. As polls opened in 2004, these same people predicted that John Kerry would win. In 2000, they foresaw a popular landslide for George Bush.

Despite past failures and ignoring current research errors, public opinion polls continue to generate the primary basis for political news, with all major news organizations spending increasing amounts of money to collect data. The numbers so overwhelm the news that people might perceive the polls as correct and the vote results as the errors. As a guest on Chris Matthews' MSNBC program "Hardball" in October 2006, comedian Robin Williams promoted his new movie in which a talk show host runs for president. With his expertise of having played a politician in a movie, the entertainer accused the U.S. voting system as being either fraud-filled or riddled with bias, with his "proof" of such an assertion that polls failed to predict elections. His statement was unchallenged by the show's journalist host or the roomful of college students of which some (we hope) studied research methods.

Small shifts of public opinion have become the lead of all major news programs, with reporters giving the statistical sampling error as if that and that alone would explain any limitations to the data. These margins of error

“with 95% confidence” are a function of sample size, a statement of statistical precision that nineteen of every possible twenty random samples conducted the same way would get close to the same answer. Left unstated in any report is that every 20th sample would be far off from this margin, and that we do not know if this is one of those samples. More deceptive to their audiences is that they mention only statistical precision as if that was the sole source of potential errors in the data. Accuracy requires going beyond precision, including all other research biases that move the data away from reality.

In their ongoing failure to note all the additional ways that consumer research numbers are not necessarily accurate, reporters allow the poll reports to supplant reality, a situation that goes unchallenged by experts in survey research who should know better (Plissner 1999; Rotfeld 2005).

Admittedly, possessing numbers provide an undeserved image of accuracy in any realm. Anyone wearing a digital watch asserts knowing the exact time, though the highest quality instrument repeats any errors from the source used for setting the time, compounds those errors with limited attention to detail in synchronization with the source, and ignores errors that come over time as batteries grow weak. A digital readout in a speedometer misleads drivers to think they know the exact speed of the car they drive, though different tires or simple tire wear would change the reading.

To the public mind, some wonder how a small sample can represent a large population, a criticism often raised by television program producers who dislike the methods for audience ratings (e.g., see Miller 2004). As a response, radio and television news programs periodically provide an interview with a supposed survey expert, so audiences hear the often repeated comparison of a survey sample with a blood test. They say, “People opposed to doing research with samples of the public should ask their doctor to take all of their blood for testing.” This is a common metaphor, but it is not a valid explanation of the problems with samples in consumer research unless every blood cell is unique, such that blood samples only provide a statistical statement of probable content of the rest of the circulatory system.

For a better metaphor, a sample of people is like taking an x-ray of one part of a body as a basis for concluding about bone structure for the entire skeleton. Bones of different size and shape are not randomly distributed in the body as different types of people are not randomly found in cities, states, or the country. Every sample frame has bias and distortions, while the people selected for the sample from that frame add additional biases by not being available for the interviewers or not responding if they are. To further invalidate a biological metaphor, consumer research is not a straightforward test, like that of DNA matching or blood levels of alcohol. Though telepathy

is not among interviewers' talents, they ask respondents their opinions and beliefs hoping the answers are true.

As a simple test of frame error of all random sample telephone polls, ask your students how many of them own telephone land lines. Probably few do. Specific estimates vary—it is difficult to argue one has an exact measurement of people who avoid being measured—but articles in various business magazines generally agree that less than a third of people under the age of twenty-five have an available directory-listed nonmobile phone by which they can be contacted for a telephone survey. Nonresponse rates for all survey research studies are now extremely high and getting worse, especially as the increasingly popular caller ID and answering machines allow people to screen calls to avoid pollsters, telemarketers, and academic researchers.

Admittedly, the tendency to present precision as the sole basis for assessing research information has driven the teaching of social science or business research for many years. Textbooks on research methods devote limited attention to qualitative research biases, with a majority of chapters and, one assumes, the resulting class time, on statistical data analysis. For many academic journals, as in the news reports, statistical precision statements are the primary focus of attention as if that alone states how close the data represents reality.

In journals other than *JCA*, the nonstatistical biases may be a quick list of "limitations" at the end of the paper that readers could easily ignore. Yet these qualitative biases impact how the study can be interpreted and what can be validly concluded from findings. As part of the statement of how the study might not represent reality, they should be integrated into any discussion of the implications or conclusions, or sometimes part of the explanation of the research method, but they should never be relegated to a special section or listed at the end.

All data needs to be approached with skepticism. All findings need to be interpreted.

Anyone with long experience in marketing knows that a majority of new product launches will probably fail even if extensive research precedes the launch, sometimes because the test itself did not model the product as it would actually be consumed. For the most famous modern example, testing a new formula for their product asking people to try a small sip of the liquid, makers of Coca-Cola mistakenly believed that the new formula would be preferred by people who drank a whole can of the redesigned product. It was not till New Coke was on the market that they saw the error from seeing the data from an inappropriate research method as a statement of reality (Gladwell 2005).

In my own studies of media vehicles' standards for acceptable advertising, we would ask the managers what was their most common reason for not publishing, printing, or broadcasting certain advertising. Consistently, for every type of vehicle in every country for over two decades, the most common answer would be in terms of whether the advertising "fit" with their editorial content or whether it might offend their general audience. The sole exception was a mail survey of daily newspapers, with a majority of publishers reporting that their most common reason for rejecting advertising was they thought it might be deceptive. For many years, I repeated the speculation that this difference might be because daily newspapers were the one medium whose origins were tied to providing news, while all others grew out of a focus on entertainment (Rotfeld, Lacher, and LaTour 1996). Yet an undergraduate student in my class this fall, Elizabeth Stone, speculated that because they are news vehicles, the daily newspapers' publishers might have been more prone to falsely make a socially desirable response in the anonymous survey, crediting themselves with policing the advertising for deceptive claims. Maybe she is right, since a socially desirable response bias might explain an avoidance of concerns for advertising deception encountered from newspaper managers at Federal Trade Commission hearings (Rotfeld 2003), but there is no way we can know for sure.

In the past year, I have had many meetings with faculty groups who asked what types of manuscripts we would prefer to receive, if a particular study is "appropriate," or if *JCA* publishes qualitative research, questions easily answered by reading the journal. As shown by many published articles, qualitative data often can provide greater insight into the consumers' minds (e.g., Wolburg 2006). In the past few months, I have penned many rejection letters in which the reviewers indicated that the paper seemed to have data in search of a theory, or statistics that do not seem to answer the conceptual question raised, problems from authors thinking that numbers alone possess intrinsic value. In the past few weeks, several authors were directed by reviewers to revise and resubmit papers with a clearer explanation of how the data are appropriate, saying that the authors need to discuss issues beyond the simple precision of the statistics.

With studies of the consumers' interests, it is important to recognize the need for seeing people beyond the numbers. For research, it is important to not confuse statistical precision alone as a statement of reality.

REFERENCES

- Gladwell, Malcolm. 2005. *Blink: The Power of Thinking Without Thinking*. New York: Little Brown & Company.

- Miller, Chris. 2004. *Blood on My Briefcase: 30 Years in the Advertising Wars*. Xlibris Corporation.
- Plissner, Martin. 1999. *The Control Room: How Television Calls the Shots in Presidential Elections*. New York: The Free Press.
- Rotfeld, Herbert Jack. 2003. Desires Versus the Reality of Self-Regulation. *Journal of Consumer Affairs*, 37 (Winter): 424–427.
- Rotfeld, Herbert Jack. 2005. A Snapshot or a Painting: Metaphors, Myths, Misuses and Misunderstandings of Marketing Research Information by Journalists and Other People Who Should Know Better. *Journal of Consumer Marketing*, 22 (#1): 4–5.
- Rotfeld, Herbert J., Kathleen T. Lacher, and Michael S. LaTour. 1996. Newspaper Standards for Acceptable Advertising. *Journal of Advertising Research*, 36 (September/October): 37–48.
- Wolburg, Joyce M. 2006. College Students' Responses to Antismoking Messages: Denial, Defiance and Other Boomerang Effects. *Journal of Consumer Affairs*, 40 (Winter): 294–323