

“Bullet Poll” Buyers Beware

To our readers: In the June/July issue, we published an interview with Jay Leve, president of Hypotenuse, the research company that conducts the “Bullet Poll.” At the time, we intended to open our pages to other assessments of this innovation in the conduct of surveys. Michael Traugott’s analysis follows here.

We invited Mr. Leve to respond—for publication—to Traugott’s piece. Instead, he and his colleagues had the attorney representing the firm threaten us with legal action—if we did not remove an offending word in the opening paragraph. We have not removed it, because we believe that Traugott’s criticism, while strong, is well within the bounds of vigorous professional exchange on important subjects.

An innovation like the “Bullet Poll” is important. That it’s quick and relatively inexpensive makes it attractive. But if it fails to yield a reliable sampling of opinion—because the field period is too short, or many potential respondents immediately turn off to a voice obviously recorded as part of an entirely computer-generated interview—it may be fool’s gold.

All of us in the field need to pay close attention to possible problems in new techniques. Public Perspective regrets that Hypotenuse would not respond substantively when its technique was challenged. Instead, the organization said only: “Traugott succeeds only in discrediting himself. We stand by our research.”

— Everett Ladd

“CATS” Dissected

By Michael Traugott

The interview with Jay Leve in the last issue of *Public Perspective* had the flavor of a promotional piece for “CATS”—“completely automated telephone surveying” being conducted under the proprietary label of Bullet Polls. With apologies to Norman Bradburn, who coined the phrase “SLOP,” for “self-selected listener opinion polls,” I think we should consider naming this new application “CRAP,” for “computerized response audience polls.”¹ I mean, this new technique is terribly flawed and unlikely to produce reliable results.

Bullet Polls employ a method that may have eventual use under limited conditions, after we understand more about its limitations. There are extreme circumstances under which these techniques can be used to collect vital information quickly—such as information on the locations of electrical generators or particular types of blood (June/July issue, p. 61). However, at present there is no way that consumers of the data can feel comfortable extrapolating the results of such polls to any meaningful population. Problems of low response rates and biased samples are probably

too great at present to draw any inferences with confidence.

These automated polls are being produced because of television stations’ competitive desire to boost their audiences for local news, one of the most profitable segments of their schedules. In this regard, they bear a great deal of similarity to dial-in polls that are used to generate viewer or reader reaction to current topics. A content analysis of the questions asked in any of these polls would provide information about the “seriousness” of the topics addressed. And it would be interesting to know how often, in Bullet Polls, the voice recording device is used to collect open-ended responses. Further, are those responses systematically coded for analysis, or just used for sound bites on the air (p. 61)? Because the technique is best suited to respondents with touch-tone telephones, it is also important to know how many people remain ineligible because they have rotary service (although this number is surely declining over time) or how many clients request that rotary service respondents be included in the survey.

Phantom Opinions

The practice of conducting polls in three to five-hour field periods is an extreme form of commercial pandering to news values and audience reactions, undertaken at the risk of collecting phantom opinions. Television stations use this information, under the guise of serious information gathering, in an attempt to “grow” their market share. For clients, the risk of error from this method appears low because the risk of detection is so small. The topics examined in these surveys are not likely to be measured in other polls, and many of the attitudes found are likely to change over

“
The practice of conducting polls in three to five-hour field periods is an extreme form of commercial pandering to news values and audience reactions, undertaken at the risk of collecting phantom opinions.
”

The Polling Business

the short term as more information becomes available to citizens. Subject matter and fast-paced opinion changes make independent verification unlikely.

It is fundamentally disingenuous to argue that the only thing that distinguishes CATS from other polls is the use of "the recorded voices of a local television anchor" (p. 60) rather than a live voice. There are several steps in the survey process—questionnaire development, pre-testing, and interviewer training, among them—that are eliminated in Bullet Polls because of the speed with which they are fielded.

At the same time, suggesting the need for speed because "Next week's news will be invented next week—we are talking today, tonight, this minute" (p. 62) highlights the fact that this product would not have any commercial value if data could not be produced in a few hours at a low price. Bullet Polls are currently being conducted at the rate of between two and seven per day, on average, depending upon the reference period. The central question is: How can we assess what some of the other "costs" of this methodology might be, in exchange for quick turnaround?

While some field periods extend across days, the "typical" Bullet Poll consists of completed interviews taken between 6:00 and 9:00 pm on a single day (p. 61), sometimes over news that breaks at 5:30 pm. Under these conditions, how much effort can be devoted to questionnaire development, pre-testing, and call-back procedures? It is one thing to offer to make as many call-backs as the client wishes, for example, but how many are possible in a three-hour field period?

Single-Digit Response Rates

There is limited research on the question of response rates and their relationship to representative samples in automated polls. Bullet Polls are obviously being conducted as samples of telephone households without respondent selection: Interviews are being

“ Without compelling evidence to the contrary, it is likely that Bullet Polls produce little more than biased estimates of attitudes, perceptions, or behavior that are unverifiable against known quantities—quantities that have been derived from more rigorous methodologies. ”

conducted with whomever answers the phone. This presents one critical issue of representativeness, for as Paul Lavrakas notes, "surveys that interview the first person who answers the telephone...are not likely to gather data that validly reflect the attitudes, behaviors, experiences, and so on of the population of interest."²

On the response-rate side, Michael Havice has estimated a rate of 7.5% for a computer-digitized voice survey conducted across one week. This rate pales in comparison to those achieved by Havice for a computer-generated random-digit dial (RDD) survey on a similar topic (response rate of 55.2%) and a 74.1% response rate for an equivalent survey based upon a sample generated from phone books, each conducted across three days.³ In a second study using a computer-digitized voice survey, Havice experienced a response rate of 6.8%.⁴ Not surprisingly, he concluded from his analysis of non-contacts that "subjects of a telephone survey find it easier to hang up on a computer than a human."

Other design flaws are that insufficient time is available in a three-hour field period for call-backs, and that trained interviewers cannot engage in securing cooperation or converting refusals. These problems are often encountered with telephone surveys designed with brief field periods.⁵ They're obviously exacerbated when the field period is a matter of hours.

Response rates are inextricably linked to important characteristics of the resulting sample. Jay Leve claims that when his clients "report our results, they

know the Bullet Poll did in fact call as perfect a cross-section of people as was possible given the time constraints" (p. 62). He cites escalating response rates for Bullet Polls but does not provide any statistics. And he further claims that "our results are consistent with our competitors and we are not introducing either sampling or response bias," although he provides no support for this assertion.

A Representative Sample?

Although it is unclear whether he is referring to raw or weighted data, Leve discusses the demographic representativeness of his samples. Havice also suggests that computerized polls can be used to produce demographically representative samples.⁶ But demography is only one consideration in evaluating the representativeness of a sample. While Havice does not present data in his methodological analyses about possible differences in attitudes held by participants and non-participants in his various survey types, there is other, limited research suggesting that either attitudes or strength of attitudes (or both) can vary between respondents and non-respondents to computerized polls.

This research is based upon comparisons between "call-in" polls, conducted in a single day, and other types of telephone surveys on the same topics conducted in local areas across periods varying between three days and three weeks. SLOP and CRAP are not the same thing, but interviews collected in a single day under conditions resulting in low response rates do share some things in common. And it seems reasonable to draw some inferences from survey data

where low response rates are known to occur to others that produce equally low response rates.

Patrick Cotter et al, compared the demographic and attitudinal characteristics of respondents to an Alabama "call-in" poll on the subject of abortion, conducted in one day, to those from a three-day RDD field period.⁷ The demography of the two samples and the views of those who held "strong" and "not so strong" opinions were equivalent in the two surveys, but there was a much greater proportion of respondents with "strong" opinions who dialed in than appeared in the RDD sample, making the overall results in the two surveys different. Benjamin Bates and Mark Harmon concluded from a similar experiment in Texas that "persons with strongly held opinions and a pro-change, activist stance are more likely to respond to phone-in polls."⁸

Another source of potential bias that requires further investigation is the recorded voice used in the CATS method. Since all of the interviews in a single market are "conducted" by the "trusted anchor" of the client's news broadcast, are regular viewers of the competition less likely to respond to such interviews?

Automated Exit Polling

Bullet Polls are also being used to generate a new kind of exit poll. One of the crucial design elements of current exit polls is that interviews with voters leaving the polls in scientifically selected samples of precincts will eliminate vote over-reporting that is common in surveys based upon household samples.⁹ Even after screening problems are overcome with the Bullet Poll methodology, going to the telephone to contact voters on election day reintroduces a significant source of reporting error. The combination of extremely low response rates and reporting bias make Bullet Polls a risky alternative to standard exit polls.

One final issue is the disclosure of methodology under which Bullet Poll

results are produced. What level of information is being provided to audience members about the conditions of collection and resulting quality of these data? Under the American Association for Public Opinion Research Standard for Minimal Disclosure, good professional practice dictates that a description of the sample procedure should be provided, including a "clear indication of the method by which the respondents were selected," "size of samples and, if applicable, completion rates and information on eligibility criteria," and "a discussion of the precision of the findings, including, if appropriate, estimates of sampling error."¹⁰ As is the case with many other media-sponsored telephone polls, it is virtually impossible to produce any of this information for a Bullet Poll, particularly an estimate of sampling error for a survey with a response rate in the 10% range.

Disposable Factoids

In conclusion, this methodology is being used to produce throwaway factoids of dubious value with the suggestion that they reflect real public opinion. Without compelling evidence to the contrary, it is likely that Bullet Polls produce little more than biased estimates of attitudes, perceptions, or behavior that are unverifiable against known quantities—quantities that have been derived from more rigorous methodologies. These data are being used to support the commercial interests of local television stations, under the guise of news gathering.

Bullet Polls may seem like a great buy for a television station, especially when the results are not easily subject to verification. However, you only get what you pay for in the survey business, and unfortunately, Bullet Poll results simply can't be trusted.

Endnotes:

¹ A good review of actual problems that have arisen with SLOP can be found in Richard Morin's, "Numbers from Nowhere: The Hoax of Call-In 'Polls,'" *Washington Post*, February 9, 1992.

² Paul J. Lavrakas, *Telephone Survey Methods*, Applied Social Research Methods Series, Vol. 7 (Newbury Park: Sage Publications, 1987), p. 82.

³ Michael Havice, "How Response Rates Compare for Human and Digitized Phone Surveys," *Journalism Quarterly*, Vol. 66, Spring 1989, pp. 137-42.

⁴ Michael Havice, "Measuring Nonresponse and Refusals to an Electronic Telephone Survey," *Journalism Quarterly*, Vol. 67, Autumn 1990, pp. 521-30.

⁵ For a summary of recent research, see Robert M. Groves and Lars E. Lyberg, "An Overview of Nonresponse Issues in Telephone Surveys," Chapter 12 in Groves et al, *Telephone Survey Methodology* (New York: John Wiley & Sons, 1988).

⁶ See, for example, Michael Havice, "How Response Rates Compare for Human and Digitized Phone Surveys." There may be a suggestion that the sample for the digitized voice survey was younger, supporting the notion that samples from computerized surveys may be younger, more educated, and more likely to use other new communication technologies. See also, Robert LaRose and David Atkin, "Audio Information Services and the Electronic Media Environment," a paper presented to the annual conference of the Midwest Association for Public Opinion Research, Chicago, November 1989.

⁷ Patrick R. Cotter, David K. Perry, and James G. Stovall, "Active and Passive Indicators of Public Opinion: Assessing the Call-In Poll," *Journalism Quarterly*, Vol. 71, Spring 1994, pp. 169-75.

⁸ Benjamin Bates and Mark Harmon, "Do 'Instant Polls' Hit the Spot? Phone-in vs. Random Sampling of Public Opinion," *Journalism Quarterly*, Vol. 70, Summer 1993, pp. 369-80.

⁹ Warren J. Mitofsky, "A Short History of Exit Polls," in Paul J. Lavrakas and Jack K. Holley (eds.), *Polling and Presidential Election Coverage* (Newbury Park: Sage Publications, 1991).

¹⁰ This language comes from Article II, Standard for Minimal Disclosure, American Association for Public Opinion Research, "Code of Professional Ethics and Practices," March 1986.



Michael Traugott is
professor of communication
studies and political science,
University of Michigan