

# About ~~From the Field~~ Response Rates

*Some unresolved questions*

By Gary Langer

At the ABC News Polling Unit we're occasionally asked about the response rate to our polls. Such inquiries rarely, if ever, specify what is meant by the term. No wonder: it's an issue fraught with complexity.

In fact, there are serious, unresolved questions in analyzing and comparing response rates, and in evaluating the tradeoffs between response rates and other methodological considerations in survey research. As practitioners it's incumbent on us to understand these issues as best we can, and to evaluate their practical significance in our work.

The American Association for Public Opinion Research has issued standard definitions for the sample disposition codes by which response rates are calculated. These are useful, but they're not the full story. Standards have not been promulgated for all the various sample design, sample management and respondent selection procedures that can affect response rates; indeed, AAPOR notes, "Comparisons between polls with different sample designs is not possible." And the AAPOR standards themselves are complex, offering four separate methods of calculating cooperation rates, three for calculating refusal rates, three for contact rates, and six for overall response rates.

*Gary Langer is director of polling for ABC News.*

Broadly, the overall response rate is based on two factors, the contact rate and the cooperation rate. The contact rate in a telephone survey of the general population is the number of households reached divided by the number of working, non-business telephone numbers dialed. The cooperation rate is the number of eligible respondents who complete the survey, divided by the total number of households reached.

The cooperation rate can be addressed by making assiduous efforts to encourage eligible respondents to participate. Higher cooperation, legitimately obtained, should improve the quality of a sample, though it may have no demonstrated practical effect on the data.

At the same time, it's theoretically possible that higher cooperation could degrade a sample in two ways: by bringing in uncooperative respondents who provide nonresponsive answers, and/or by changing random noncooperation to systematic noncooperation. Both these possibilities merit further academic investigation.

Contact rate has considerably more complications than cooperation rate, mainly pertaining to sample design. There are ways to obtain a higher contact rate that have no effect on the quality of the sample. Some even clearly degrade it.

AAPOR's standard definitions for response rate calculation assume a random-digit-dialed

(RDD) sample and systematic within-household selection; they don't specify which RDD sampling and respondent selection procedures pass muster.

Some widely reported polls do not meet even these basic assumptions. Some are drawn from listed samples (e.g., the White Pages), rather than via RDD. Use of a listed sample excludes all non-listed numbers, estimated nationally at about 30% of all households. It should produce a higher contact rate by excluding non-household numbers, but it degrades coverage by also excluding many households that should be part of the sample. (Other polls use voter or surname lists, again potentially improving contact but degrading coverage.)

Likewise, some widely reported polls do no within-household selection of respondents. That can significantly improve the contact rate, but at the expense of in-house randomization. Other polls employ a variety of in-house selection methods: some use at-home samples, some full-household samples; some ask for a specified individual, such as the youngest male. Others use last-birthday or Kish selection methods.

The choice of in-house method also can affect response rates. Kish, for instance, which requires interviewers to get the name, sex and age of all household members, generally is regarded as the most rigorous, but can reduce cooperation as a result of its perceived intrusiveness.

RDD sampling presents pitfalls of its own. Failure to maintain current area codes and exchanges may produce a higher contact rate, but it degrades coverage. Boosting the number of listed numbers required to qualify a working block improves the contact rate, but it also degrades coverage. A panel-back survey would have a lower response rate, given its additional layer of noncontact and noncooperation—even though paneling can be an effective method of measuring real change.

Cleaning the sample by pre-dialing improves the contact rate, but without improving the quality of the sample. Cleaning a sample of business numbers by filtering it through Yellow Pages listings improves the contact rate, but again without improving the sample. More aggressive cleaning, for example by algorithm (i.e., removing all business-type numbers, like 2626 or 4700), would further improve the contact rate—but would degrade coverage.

In AAPOR's definitions, sampled but undialed numbers are counted against response rates. Thus, sample management techniques—using replicates, and especially small replicates, and releasing them sparingly—would improve the response rate, without necessarily improving the sample (depending on whether undialed numbers were randomly or systematically undialed).

It's also important to note that there can be variability in sample dispositions—the raw material upon which response and cooperation rates are calculated. Inadequate policy, poor interviewer training on the application of disposition codes, or lack of supervision can compromise these data. For example, a number that starts as an “initial hang-up” but, when redialed, is not answered, should ultimately be coded as a hang-up (and therefore part of noncooperation) rather than as a no-answer (and therefore part of noncontact).

Other sampling issues beyond response rate calculations arise with other methodologies. RDD-based online surveying via Web-TV, for instance, works only in Web-TV service areas. Other online surveys lack a sampling frame, rendering their response rates moot.

The nature of the study is also relevant. The main way of improving a survey's response rate is to keep it in the field for an extended period of time, re-calling numbers to improve the contact rate and re-contacting eligible respondents in hopes of winning their cooperation. Many surveys, however, are time-sensitive—especially,

faulty questionnaire or weak data analysis. And the importance of response rate in terms of data integrity likely depends to some extent on the subject under study. Some behavioral studies, such as a study of dining-out habits, probably would benefit from a higher response rate; studies focused on opinion appear to be less vulnerable.

Recent published research has shown no substantial effect of lower response rates on measurements of opinion:

- In the Summer 2000 issue of *Public Opinion Quarterly*, Scott Keeter et al. reported on two identical surveys,

**“Editorial content is critical; a high response rate does not compensate for a faulty questionnaire or weak data analysis.”**

but not exclusively, news surveys. Keeping them in the field for weeks would make the results outdated, defeating the purpose of the enterprise.

For example, self-reported stress and depression were quite high immediately after the terrorist attacks of September 11, but declined rapidly in the following two months. A survey that took two months to complete would have missed that trend—it would have had a higher response rate, but far less meaningful data.

On breaking news, an overnight poll gives a sometimes crucial immediate read on public opinion, its lower response rate notwithstanding. And years of accumulated data have shown overnight polls to be coherent with longer-term studies on the same subjects.

Editorial content is critical; a high response rate does not compensate for a

one using a five-day field period and an at-home sample that produced a 36.0% response rate, the other using an eight-week field period and full household sample that produced a 60.6% response rate. “Nonetheless,” Keeter and his colleagues wrote,

the two surveys produced similar results. Across 91 comparisons, no difference exceeded 9 percentage points, and the average difference was about 2 percentage points. Most of the statistically significant differences were among demographic items. Very few significant differences were found on attention to media and engagement in politics, social trust and connectedness, and most social and political attitudes, including even those toward surveys.

- In the Winter 2000 issue of *Public Opinion Quarterly*, Richard Curtin et

al. reported results that “replicate and extend the finding of Keeter et al. that large differences in response rate had only minor effect on cross-sectional analyses.”

- At the AAPOR 2001 conference, Elizabeth Wiebe et al. reported a project to boost cooperation among non-respondents to a Web-TV survey. The effort raised

**“Noncontact and nonresponse should not affect data quality to the extent that they occur randomly.”**

the response rate from 24.7 to 41.9%, but with no significant impact on the data. “The nonresponse follow-up,” they concluded, “was useful in improving response rates to the survey. In terms of improving the representativeness of the sample, there’s little evidence the nonresponse follow-up was necessary.”

- An analysis of ABC News 2000 pre-election tracking data, also presented at the AAPOR 2001 conference, found no significant effect of callbacks on political party identification. The shares of Democrats, Republicans and independents reached on first calls were 37, 31 and 31%, respectively; on second calls, 36-30-33%; on third calls, 38-30-31%; and on fourth-plus calls, 36-30-32%.
- Daniel Merkle and Murray Edelman, in the 2002 book *Survey Nonresponse*, found no relationship between response rates and survey error in a comprehensive study using exit poll data.
- Peter Mariolis of the Centers for Disease Control and Prevention reported to the Statistics Canada Sympo-

sium 2001 on two samples of data on behavioral risk factors collected in Maine in 2000. Despite a 13.3-point difference in response rates, the data were substantively the same; in fact, some data from the lower response-rate survey “seem to be the more accurate.” Mariolis concluded: “Keeter et al. (2000) and Curtin et al. (2002) recently suggested that large differences in response rates in telephone surveys may not have

much impact on most substantive outcomes. The results in this paper support that position.”

- Mariolis further reported on CDC data on behavioral risk factors at the AAPOR 2002 conference, in a presentation entitled “Response Rates and Data Accuracy.” Even though median response rates to ongoing CDC surveys had declined from 71.4% in 1993 to a low of 48.9% in 2000, he found “no consistent or strong relationships between response rates and measures of gender, age or race/ethnicity biases.” Analyzing data on cigarette smoking in Nevada, he found that a 45-point difference in response rates had “a predicted difference in smoking prevalence rates of 1.5 percentage points.”

Such results inevitably must inform cost/benefit considerations. Basic methodological rigor is essential—RDD sampling, systematic selection, and careful interviewer training and supervision. Whatever questions exist about the effect of response rates on data quality, they do not excuse sloppy design and execution. But given the basics, a survey organization can devote resources in four directions:

- It can field more surveys, producing more data on more subjects.
- It can field longer questionnaires, probing opinion more broadly and deeply.
- It can interview larger samples, allowing for more granular and more reliable subgroup analysis.
- Or it can improve response rates by repeatedly re-calling unanswered numbers and re-contacting recalcitrant eligible respondents.

The practical benefits of the first three are most readily apparent.

That is not to say response rate is not a serious concern; it is. There are low-cost ways to increase cooperation that every serious researcher should employ—such as finding the most effective introductory and fallback language, and focusing on interviewer effectiveness.

Unquestionably, the highest possible response rates, legitimately obtained and within the bounds of the project at hand, are desirable in survey research. But poll-to-poll comparisons of response rates are exceedingly difficult, and their value unclear; indeed, given the issues discussed here, a higher response rate is not automatically indicative of better data.

Noncontact and nonresponse should not affect data quality to the extent that they occur randomly. While it is reassuring to see the limited effect of nonresponse in the studies cited above, it doesn’t mean that nonresponse bias does not occur. Instead, the simple fact is that the level at which response rates do begin to affect data quality, which data are and are not affected, and the nature and significance of any effect remain to be established. As practitioners, we eagerly await the results of further study. ●