

DIFFERENCES IN RESPONSE RATES USING MOST RECENT VERSUS FINAL DISPOSITIONS IN TELEPHONE SURVEYS

CHRISTOPHER MCCARTY

University of Florida

Introduction

Although response rates have always been important in survey research, they have become central in discussions about survey quality during recent years. It is not unusual for clients, particularly those from state and federal agencies, to have expectations about the response rate that they will get on a survey. At the same time, the number and types of surveys have ballooned, yielding a combination of modes, populations, and topics that have increased the complexity of both data collection and analysis. This has caused some confusion as clients, who have little concrete data for comparison and sometimes little experience in survey research, develop expectations for response rates that are out of line with current trends for the type of survey they commission.

Some of the confusion about response rates has to do with what they mean. Response rates were originally intended as a measure of the extent to which the data represent the responses of the entire population, that is, as an indicator of nonresponse bias (Hansen and Hurwitz 1946). Some studies have shown that the characteristics of samples based on fewer attempts, and subsequent lower response rates, are less representative of the population than are surveys with more attempts (Lin and Schaeffer 1995; Traugott 1987; Zapka et al. 1994). Recent studies, however, suggest that the effect of nonresponse may not be as pronounced as was once thought (Curtin, Presser, and Singer 2000; Keeter et al. 2000) and that low response rates may not necessarily indicate bias.

An interpretation of response rates by some who sponsor surveys, such as government administrators, is that they are a measure of effort. To some clients low response rates may mean that the survey vendor is trying to cut corners. For example, by flooding a study with sample—loading it with untouched numbers without regard for the number that have been previously loaded—a survey lab can achieve the target number of interviews at a lower cost than

by using a smaller sample that has been worked thoroughly. But the response rate, including the uncontacted “fresh numbers” in the denominator, will naturally be lower than in the thoroughly worked sample. Hence, the belief on the part of some survey sponsors that the response rate is a proxy for effort.

During the past 30 years much of the literature on response rates has focused on how to improve them (Brehm 1994; Dillman 1978; Kristal et al. 1993; Piazza 1993; Roth and BeVier 1998) and on the effect of low response rates on data quality (Cohen and Carlson 1995; Goyder 1987; Groves, Cialdini, and Couper 1992; Groves and Couper 1998; Lin and Schaeffer 1995). Very few books or articles talk about how response rates are calculated or the effects on response rates of different schemes for disposition coding (see AAPOR 2000; and Lynn et al. 2001).

Frey (1989) is an exception. Frey discusses how some rate calculations create the appearance of “better research.” Among the techniques used to artificially inflate rates are sample substitution and ad hoc decisions about which cases count as eligible. Lavrakas (1987), another exception, presents a table using disposition data from the 1979 Chicago Metropolitan Survey to show the effects of different calculations on the rate. The rate varies from a low of 33.7, achieved by dividing all completes by all numbers, to a high of 71.9, achieved by dividing “all potential completions” by “all eligibles.”

While the response rate is sometimes used as a proxy for level of effort, many parts of the survey process affect response rates yet have nothing to do with the amount of work or resources a survey organization invests. For example, survey labs may create their own coding schemes for assigning dispositions rather than use a standard method (Lynn et al. 2001). Broader disposition categories may classify certain cases as ineligible when a more detailed coding scheme would not.¹ Response rates can also differ based on constraints such as the type of sampling frame and the population being surveyed.²

This research note focuses on one particular source of variability in response rates that is not due to effort—the use of “most recent” versus “final” disposition coding. When calculating a response rate, only one of potentially

1. For example, the Behavioral Risk Factor Surveillance Survey (BRFSS) uses 11 disposition codes while the American Association of Public Opinion Research sanctions 29 codes. The Centers for Disease Control (CDC) BRFSS protocol calls for replacing code 03 (a nonworking number) after one call and finalizing the case as ineligible. The AAPOR codes break nonworking numbers into at least four categories, two of which are eligible (3150, technical phone problem; 3151, temporary phone problem) and two of which are ineligible (4310, nonworking number; 4320, disconnected number). The CDC coding is predisposed to a higher response rate given the broader coding scheme.

2. One way to discourage interpretation of the response rate as a measure of effort would be to develop an independent index of effort based on individual call data. This index would reflect (among other things) the number of calls per case, the timing of the calls, and the length of time telephone numbers were fielded. By reporting an effort index along with the response rate, it is more likely that the response rate would be interpreted as it was originally intended, that is, as a measure of the potential for nonresponse bias.

many dispositions is used to conclusively represent what happened to a sampled telephone number. The “most recent disposition” of a sampled telephone number is the disposition coded on the last call attempt. Depending on the call limit for a given survey, the most recent disposition could result, for example, from the first call, if the call limit is set to one, or from the twenty-fifth call, if the call limit is set to 25.

The “final disposition” is the disposition considered the most appropriate for the case out of all dispositions coded on all attempts. For example, it is not unusual to recontact “soft refusals.” It is also not unusual for all attempts subsequent to a soft refusal to be “no answers” or “answering machines.” A soft refusal may be viewed as a more appropriate representation of what happened to the case than a no answer, even if the no answer was the disposition on the most recent attempt. While the coding of this case may seem obvious, there are other call history scenarios that are more complicated. Such cases point to the fact that the most recent disposition coding may not be the best final disposition coding.

Currently, the American Association for Public Opinion Research (AAPOR) does not sanction a particular set of rules for assigning final dispositions. The *Standard Definitions* manual does acknowledge the work of Lessler and Kalsbeek (1992), one of the few books to address the issue of calculating response rates, suggesting one give “predominant consideration to circumstances of the last contact” (AAPOR 2000). This might suggest to some that they use the most recent disposition as the final disposition. Anecdotal evidence from working with various survey labs and discussing this issue with colleagues suggests that many labs do just that—they use the most recent disposition as the final disposition. For example, the CDC User’s Guide for the BRFSS stipulates the assignment of a final disposition for each of its 11 codes after a certain number of calls. However, it does not address what disposition is to be assigned when the call history contains mixed dispositions. The implication of the guide is that the most recent disposition should be used after a certain number of calls are made.

Why do researchers most favor using the recent disposition rather than a final disposition? One possible reason is the structure of many computer-assisted telephone interview (CATI) programs. Most survey software programs today are built on relational data-base engines. The purpose of relational data bases is to store information efficiently and, most important, not to store information redundantly. This means that there will be a data base in which the unit is a piece of sample and a separate data base in which the unit is a telephone call. These data bases will be linked by a sample ID number or phone number. Frequently the sample management software will include the most recent disposition of the case as a field in the sample data base. But this data base has room for only one disposition of all the dispositions realized for a case. The last one is easily collected and substituted on each call. The

software also usually includes a report that summarizes the most recent disposition field in the sample data base, which can be done quickly.

Are there large differences between response rates using the most recent disposition and those using final dispositions? If there are no differences, then an argument can be made for using the most recent dispositions in the calculation of the response rates, given the widespread use of the practice and the absence of a standard for evaluating call histories to determine final dispositions.³ The purpose of this research note is to test (across different surveys using real data) the differences between response rates calculated using the most recent dispositions and response rates modifying the most recent disposition with a conservative algorithm for determining the final disposition.

Method

While it is possible to create an algorithm to evaluate call histories that examines all possible combinations of disposition histories, it is not practical. In practice only a fraction of all possible combinations actually occur. Another approach is to compare response rates from various types of surveys and to determine which patterns of call histories are responsible for the difference between response rates using the most recent versus the final dispositions. This is the approach used in this research note.

The University of Florida Survey Research Center (UFSRC) conducts a variety of surveys for the state of Florida and for other states and academically funded survey projects. This mix of government-sponsored and academically sponsored survey work, while typical of many university-based survey labs, may differ from the types of surveys in many other survey organizations, such as market research firms, political polling firms, or large consulting and university-based survey organizations that conduct mostly federally funded research. Six surveys were selected for this analysis, varying in size, duration of field period, population, and sample frame. Three of the sample frames were from random digit dialing (RDD), and three were from lists of people enrolled in a health plan or a social service program. These six surveys reflect the spectrum of call histories that would typically be generated in UFSRC. For example, RDD surveys will, by and large, have more business and government agency dispositions than will listed household samples. Surveys of Medicaid HMO clients will have more disconnected numbers than surveys of commercial HMO clients. The following survey descriptions refer to the data in table 1:

3. One reviewer notes that, in studies with high settings for call attempts, response rates using the most recent dispositions can be misleading early in the survey before the sample is worked sufficiently. In general, all response rates will be misleading until the survey sample has been worked and the survey nears completion. Conversely, as the sample is worked (particularly on surveys with high call limits), the most recent dispositions and the final dispositions will tend to converge.

Table 1. Productivity Comparison across Surveys

	RDD 1	RDD 2	RDD 3	Listed 1	Listed 2	Listed 3
Sample released	3,500	2,9174	4,450	3,0084	16,268	1,100
Call limit	15	12	12	10	10	10
Average attempts per case	5.6	5.1	4.1	5.1	4.1	2.2
Completed interviews	501	8,010	1,002	7,630	3,823	402
Response rate using most recent disposition	.28	.41	.32	.43	.49	.40
Response rate using ever eligible	.21	.33	.28	.31	.31	.39
Response rate using all five final disposition rules	.26	.40	.32	.38	.44	.40
Response rate using only Rule 1	.23	.34	.29	.32	.32	.39
Response rate using only Rule 2	.22	.36	.28	.35	.38	.39
Response rate using only Rule 3	.22	.34	.28	.32	.33	.39
Response rate using only Rule 4	.22	.33	.28	.31	.31	.39
Response rate using only Rule 5	.22	.33	.28	.32	.32	.39
Response rate using only Rule 1 and Rule 2	.24	.37	.30	.36	.38	.39

NOTE.—The following are with regard to the most recent call. Rule 1: Business always coded as business. Rule 2: No eligible respondent always coded as no eligible respondent. Rule 3: Disconnected number previously coded as no answer or temporary phone problem always coded as disconnected number. Rule 4: Fax/data line previously coded as no answer or temporary phone problem always coded as Fax/data line. Rule 5: Nonworking number previously coded as a temporary phone problem always coded as a nonworking number.

- RDD 1. Random-digit-dial survey of Floridians about economic issues conducted in the month of June 2001. This sample did not include directory matching or attended dialing (a service provided by some sampling firms in which a call center dials the telephone number and a human listens to the call outcome, eliminating many business and nonworking numbers).
- RDD 2. Random-digit-dial survey using sample of a Midwestern state about health insurance in which the sampling frame was cleaned by directory matching and attended dialing to purge businesses and nonworking numbers.
- RDD 3. Random-digit-dial survey of Hernando County, Florida, residents about health issues. This sample did not include directory matching or attended dialing.
- Listed 1. Listed sample of commercial health maintenance organization (HMO) members interviewed using the Consumer Assessment of Health Plans (CAHPS) customer satisfaction survey.
- Listed 2. Listed sample of Medicaid HMO members interviewed using the CAHPS customer satisfaction survey.
- Listed 3. Listed sample of elderly Floridians who received services from the Department of Elder Affairs and were interviewed about end-of-life issues.

For each of the six surveys in table 1, several response rate calculations, based on different disposition coding rules, are listed. The American Association of Public Opinion Research sanctions six response rate definitions that differ in the way partial interviews are treated and whether a proportion of cases of unknown eligibility are assumed to be ineligible. Here we employ RR1, which is the most conservative of the six response rate definitions. The formula for RR1 is

$$RR1 = \frac{I}{(I + P) + (R + NC + O) + (UH + UO)}$$

where I = complete interview, P = partial interview, R = refusal and break-off, NC = noncontact, O = other, UH = unknown if household/occupied housing unit, and UO = unknown, other.

All cases falling into the categories NC, UH, and UO would be considered eligible using this formula. All ineligible cases (the 4000 series codes in table 2) have been excluded from the denominator. Other formulas (such as RR3 and RR4) incorporate an *e* (multiplier) term that makes it possible to remove cases from the denominator by assuming that some proportion are not eligible. For example, by setting *e* to .5, half of the dispositions of unknown eligibility would be removed and the response rate would rise.

Referring again to table 1, the row labeled “most recent disposition” divides

Table 2. Eligibility of Standard AAPOR Disposition Codes

AAPOR Disposition Description	AAPOR Dis- position Code	Eligibility (E = Eligible, I = Ineligible)
Completed interview	1,100	E
Partial interview	1,200	E
Strong refusal	2,110	E
Soft refusal	2,120	E
Respondent never available	2,210	E
Answering machine, message	2,221	E
Answering machine, no message	2,222	E
Dead	2,310	E
Physically, mentally unable	2,320	E
Language unable	2,330	E
Miscellaneous unable	2,340	E
Busy	3,120	E
No answer	3,130	E
Answering machine, don't know if household	3,140	E
Technical phone problem	3,150	E
Temporary phone problem	3,151	E
Fax/data line	4,200	I
Nonworking number	4,310	I
Disconnected number	4,320	I
Number changed	4,430	I
Cell phone	4,410	I
Call forwarding	4,420	I
Business/government/other organization	4,430	I
Institution	4,520	I
Group quarters	4,530	I
No eligible respondent	4,700	I
Quota filled	4,800	I
Callback, respondent not selected	5,100	E
Callback, respondent selected	5,200	E

NOTE.—American Association for Public Opinion Research, *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys, 2000*.

the number of completed cases by the number of eligible cases using the most recent disposition only and the classification in table 2. The row labeled “ever eligible” divides the number of completes by the number of eligible cases, categorizing a case as eligible if it ever received an eligible disposition. For example, a call history that contained any one of the AAPOR codes labeled with an E in table 2 would be considered eligible. The difference between response rates based on the most recent dispositions and those based on the case ever having been eligible is potentially large. For the Listed 2 survey,

the most recent dispositions yielded a response rate 18 percent higher than the ever-eligible response rate. On average, the most-recent disposition rate is 8 percent higher than the ever-eligible rate across the six surveys. Listed samples tend to show bigger differences than RDD samples.

The next set of rows in table 1 show the final disposition response rates calculated using the RR1 formula but also applying each of five disposition coding rules derived by analyzing cases that were ineligible using the most recent disposition but eligible using the ever-eligible criterion. These rules are:

- Rule 1. Business on most recent call always coded as business, regardless of previous coding
- Rule 2. No eligible respondent on most recent call always coded as no eligible respondent, regardless of previous coding
- Rule 3. Disconnected number on most recent call previously coded as no answer or temporary phone problem always coded as disconnected number
- Rule 4. Fax/data line on most recent call previously coded as no answer or temporary phone problem always coded as Fax/data line
- Rule 5. Nonworking number on most recent call previously coded as a temporary phone problem always coded as a nonworking number.

In most cases the difference between a response rate based on the most recent disposition and one based on a final disposition using all these rules combined is very small, on average about 2 percent and at most 5 percent. For example, in the Listed 1 and Listed 2 surveys, the application of these five rules reduced the difference from 12 percent and 18 percent, respectively, to 5 percent for both. The remaining cases that are unaccounted for by the rules have a most recent disposition that is ineligible, such as “disconnected number,” and a previous disposition of “answering machine” or “callback,” indicating that the household definitely had a potential respondent at some point. These should be coded as eligible even though the most recent disposition is ineligible.

The response rates achieved by applying each of the five rules separately demonstrate which rules account for most of the difference. As seen in figure 1, Rule 1 achieves the most response rate increase in RDD surveys, while Rule 2 achieves the most response rate increases in listed-sample surveys. This is because Rule 1, dealing with attempts coded as “business,” is unlikely to apply with listed samples but very likely to apply with RDD. Conversely, Rule 2, which addresses respondent eligibility, is an important consideration in many listed surveys, such as current customers of HMOs, but less so in an RDD survey in which there are different screening criteria.

Two of the five rules (Rules 1 and 2) simply give precedence to the most recent disposition over any previously coded attempt. Table 1 shows that Rules 1 and 2 account for a large proportion of the increase of response rates

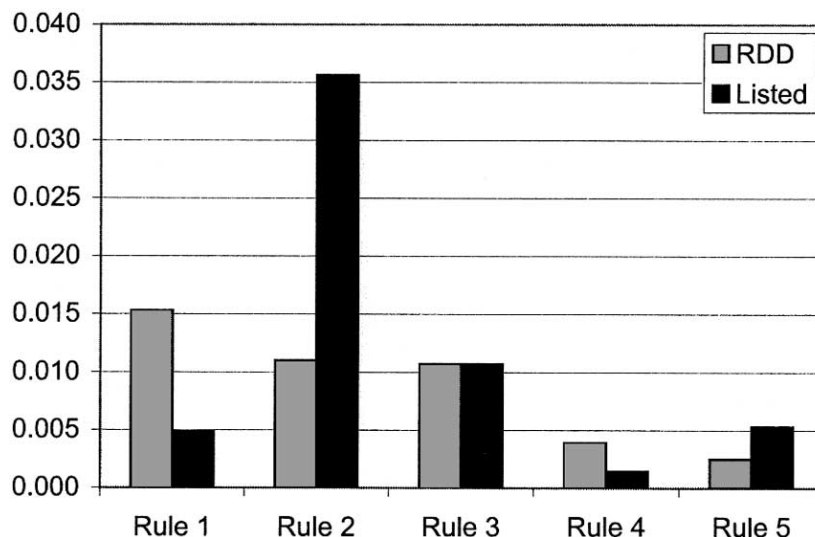


Figure 1. Average gain in response rate over “ever eligible” response rate from applying disposition coding rule, by type of sample.

associated with the final disposition over the ever-eligible disposition. This observation is important because Rules 1 and 2 do not require any evaluation of call history to code. Therefore, in these six surveys, using the most recent disposition as the final disposition results in less error than might have been expected so long as Rules 1 and 2 are deemed reasonable. These two rules do place credence in interviewer efforts to vet the information provided by contacted parties to determine business status or respondent ineligibility.⁴

As seen in figure 1, the biggest increase in response rates occurred in listed surveys in which cases initially treated as potentially eligible were found on the most recent call to be ineligible. These increases are twice those of any other rule for RDD or listed surveys. Rule 1, which allows potentially eligible cases that were later found to be businesses to be counted as ineligible, is clearly more important for RDD surveys in which business numbers are still commonly found. Rule 3, regarding fax and data lines, affects both RDD and listed surveys equally.

4. Automated dialing systems may produce different results than interviewer-initiated dialing in ascertaining eligibility. Automated dialers may finalize cases at a faster rate than interviewer-initiated dialings because they more rapidly identify disconnected numbers, businesses, and nonworking numbers. These numbers are never reintroduced into the sample pool; hence, the response rate outcome based on “most recent disposition” for automated dialers may be higher than the outcome for interviewer dialing.

Conclusion

There are many things that affect the response rate for a telephone survey. Most of the literature on the topic of response rates is focused on procedural changes that will yield higher response rates. Very little has been written on the classification of sample outcomes for the calculation of response rates, and even less such research has used empirical data. While AAPOR (2000) has made great strides in the standardization of response rate calculations, there is no standard algorithm for the evaluation of call histories to determine final dispositions.

In this note, I have shown that the difference between a response rate calculated using most recent dispositions and one calculated using final dispositions is, given certain assumptions, fairly small—at least in the accumulated experience of one survey lab, involving studies of different kinds. So long as the researcher and the client are willing to accept a most recent disposition of “business” or “no eligible respondent” as taking precedence over any other historical disposition, the response rate using the most recent disposition is on average only 5 percent higher than a “final” disposition created using these codes across the six surveys examined in this article. Application of all five coding rules reduced the difference between most recent and final dispositions to 2 percent on average.

The five disposition coding rules discussed here can also be considered as a starting point for constructing an algorithm for evaluating call histories to assign a final disposition. One possible algorithm is to code cases as eligible if there was ever any eligible disposition in the call history and to recode the case as ineligible according to these five rules.

References

- American Association for Public Opinion Research (AAPOR). 2000. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. Ann Arbor, MI: AAPOR.
- Brehm, John. 1994. “Stubbing Our Toes for a Foot in the Door? Prior Contact, Incentives, and Survey Response Improving Response to Surveys.” *International Public Opinion Research* 6(1):45–64.
- Cohen, Steven B., and Barbara Lepidus Carlson. 1995. “Characteristics of Reluctant Respondents in the National Medical Expenditure Survey.” *Journal of Economic and Social Measurement* 21(3/4):269–97.
- Curtin, Richard, Stanley Presser, and Eleanor Singer. 2000. “The Effects of Response Rate Changes on the Index of Consumer Sentiment.” *Public Opinion Quarterly* 64(4):413–28.
- Dillman, Don. 1978. *Mail and Telephone Surveys: The Total Design Method*. New York: Wiley.
- Frey, James. 1989. *Survey Research by Telephone*. Newbury Park, CA: Sage.
- Goyder, John. 1987. *Silent Minority: Nonrespondents on Sample Surveys*. Boulder, CO: Westview.
- Groves, Robert, Robert Cialdini, and Mick Couper. 1992. “Understanding the Decision to Participate in a Survey.” *Public Opinion Quarterly* 56(4):475–96.
- Groves, Robert, and Mick P. Couper. 1998. *Nonresponse in Household Interview Surveys*. New York: Wiley.
- Hansen, M. H., and W. N. Hurwitz. 1946. “On the Theory of Nonresponse in Sample Surveys.” *Journal of the American Statistical Association* 41(236):517–29.
- Keeter, Scott, Carolyn Miller, Andrew Kohut, Robert M. Groves, and Stanley Presser. 2000.

- "Consequences of Reducing Nonresponse in a National Telephone Survey." *Public Opinion Quarterly* 64(2):125–48.
- Kristal, Alan, Emily White, Julie Davis, Gayle Corycell, Trivellore Raghunathan, Susan Kinne, and Ting-Kwong Lin. 1993. "Effects of Enhanced Calling Efforts on Response Rates, Estimates of Health Behavior, and Costs in a Telephone Health Survey Using Random-Digit Dialing." *Public Health Reports* 108(3):372–80.
- Lavrakas, Paul. 1987. *Telephone Survey Methods: Sampling, Selection, and Supervision*. Newbury Park, CA: Sage.
- Lessler, Judith T., and William D. Kalsbeek. 1992. *Nonsampling Error in Surveys*. New York: Wiley.
- Lin, I-Fen, and Nora Cate Schaeffer. 1995. "Using Survey Participants to Estimate the Impact of Nonparticipation." *Public Opinion Quarterly* 59(2):236–58.
- Lynn, Peter, Roeland Beerten, Johanna Laiho, Johanna, and Jean Martin. 2001. "Recommended Standard Final Outcome Categories and Standard Definitions of Response Rate for Social Surveys." Working Paper no. 2001–23. Colchester: University of Essex, Institute for Social and Economic Research.
- Piazza, Thomas. 1993. "Meeting the Challenge of Answering Machines." *Public Opinion Quarterly* 57(2):219–32.
- Roth, Philip L., and Craig A. BeVier. 1998. "Response Rates in HRM/OB Survey Research: Norms and Correlates, 1990–1994." *Journal of Management* 24(1):97–118.
- Traugott, Michael W. 1987. "The Importance of Persistence in Respondent Selection for Pre-election Surveys." *Public Opinion Quarterly* 51(1):48–57.
- Zapka, J. G., L. Chasantaber, C. Bigelow, and T. Hurley. 1994. "Methodological Issues for Health-Related Surveys of Multicultural Older Women." *Evaluation and the Health Professions* 17(4): 485–500.

Copyright of Public Opinion Quarterly is the property of American Association for Public Opinion Research and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.