

**Differences In Response Rates Using Last Versus Final Dispositions In Telephone Surveys**

**Christopher McCarty**

**University of Florida**

## **Abstract**

Most of the literature addressing response rates focuses on how to improve them or how low response rates affect survey data. Surprisingly little has been published on the calculation of response rates and the issues that affect those calculations. In an increasingly competitive survey industry, the comparability of response rates is more important than ever. This article examines one specific factor that has gained attention in recent years -- the effect of using final dispositions versus last dispositions in the calculation of response rates.

Using data from a university-based survey lab, response rates based on last dispositions are found to be, on average, 8 percent higher than those based on final dispositions using a conservative scheme for evaluating call histories. Most of the difference is accounted for by cases where the last disposition is a business or there is no eligible respondent, cases that most survey researchers would accept as final dispositions regardless of previous codings. Five rules are presented for accepting an ineligible last disposition when an eligible disposition occurred in the call history. By accepting these rules, the response rates using last dispositions are on average 2 percent different from those generated using final dispositions.

## **Background**

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 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 subjects 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. Why do we calculate response rates? What are they supposed to tell us? One way of interpreting response rates is as a measure of effort. By flooding a study with sample, that is loading it with untouched numbers without regard for the number that have been loaded, a survey lab can achieve the target number of interviews at a lower cost than a smaller sample that has been worked thoroughly. To some clients low response rates may mean that the survey vendor is trying to cut corners.

Others use response rates as a measure of the extent to which the data represent the responses of the entire population. Some studies have shown that the characteristics of samples based on fewer attempts, and lower response rates, are less representative of the population than are surveys with more attempts (Traugott 1987, Zapka, Chasantaber, Bigelow and Hurley 1994, Lin and Schaeffer 1995). Recent studies, however, suggest that the effect of nonresponse may not be as pronounced as once thought (Curtin, Presser and Singer 2000, Keeter, et al 2000) and

that low response rates may not necessarily indicate bias. Most people who look at response rates are probably more concerned with the issue of representation than that of cutting corners.

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

Frey (1989) 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) 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, in fact, 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 create their own coding schemes for assigning dispositions rather than using a standard. Disposition categories that are too broad may exclude certain cases as ineligible where a more detailed coding scheme would not. Response rates can also differ based on constraints such as the type of sampling frame and the population being surveyed. This article, however, focuses on one particular source of variability in response rates that is not due to effort, the use of last versus final dispositions.

When calculating a response rate, only one of potentially many dispositions is used to represent what happened to a sampled telephone number. The “last disposition” of a telephone call is the disposition coded on the last call attempt. Depending on the call limit for a given survey, that could be anywhere from the first call, if the call limit is set to 1, or the twenty-fifth call if the call limit is set to 25.

The “final disposition” is the disposition considered the most appropriate for the call out of all dispositions coded on all attempts. For example, it is not unusual to call “soft refusals” more than once. It is also not unusual for all attempts subsequent to a soft refusal to be “no answers” or “answering machines.” Most survey researchers would agree that in the evaluation of a call history, a soft refusal is a more appropriate representation of what happened to the case than a no answer, even if the no answer was the disposition on the last attempt. While the coding of this case may seem obvious, there are other call history scenarios that are more complicated.

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 last disposition as the final disposition, although it clearly refers to the last contact. Anecdotal evidence from working with various survey labs and discussing this issue with colleagues suggests that many labs do just that – they use the last disposition as the final disposition.

Why do researchers favor using last disposition rather than a final disposition? Most survey software today are built on relational database engines. The purpose of relational

databases is to store information efficiently, and most importantly not store information redundantly. This means that there will be a database where the unit is a piece of sample, and a separate database where the unit is a telephone call. These databases will be linked by a sample id number or phone number. Frequently the sample management software will include the last disposition of the case as a field in the sample database. But this database 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 last disposition field in the sample database, which can be done quickly. This is the reason it is favored in the industry.

Are there significant differences between response rates using the last disposition and those using final dispositions? If there are no differences, then an argument can be made for using last dispositions in the calculation of response rates, given the widespread use of the practice and the absence of a standard for evaluating call histories to determine final dispositions. While final dispositions are necessary for other measures, such as refusal rates and cooperation rates, response rates could be calculated more easily and monitored on a daily basis if last dispositions could be used.

## **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 last versus final dispositions. This is the approach used in this article.

The University of Florida Survey Research Center conducts a variety of surveys for the state of Florida and other states, and academically funded survey projects. This mix of government sponsored and academically sponsored survey work is typical of a university-based survey lab. Six surveys were selected for this analysis varying in size, field time, population and sample frame. This is intended to reflect the spectrum of call histories that could be generated. For example, RDD surveys will, by and large, have more business and government agency dispositions than 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:

RDD 1 = Random-digit dial survey of Floridians about economic issues conducted for the month of June 2001.

RDD 2 = Random-digit dial survey using sample cleaned by directory-matching and attended dialing of a Midwestern state about health insurance.

RDD 3 = Random-digit dial survey of county residents about health issues.

Listed 1 = Listed sample of commercial HMO members interviewed using CAHPS customer satisfaction survey.

Listed 2 = Listed sample of Medicaid HMO members interviewed using CAHPS customer satisfaction survey.

Listed 3 = Listed sample of elderly Floridians who received services from Department of Elder Affairs and were interviewed about end-of-life issues.

For each of the six surveys in Table 1, several response rates are listed. AAPOR sanctions six response rate calculations that differ in the way partial interviews are treated and whether a proportion of cases of unknown eligibility are assumed to be ineligible. For the purposes of this article I am using RR1, which is the most conservative of the six response rates. The formula for RR1 is as follows:

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

Where

I = Complete interview

P = Partial interview

R = Refusal and break-off

NC = Non-contact

O = Other

UH = Unknown if household/occupied housing unit

UO = Unknown, other

In other words, we divide the number of completed cases by the number of eligible cases.

All ineligible cases (as defined in Table 2) have been excluded from the denominator.<sup>1</sup>

Referring again to Table 1, the row labeled “last disposition” divides the number of completed cases by the number of eligible cases using the last 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. While nobody would use this as a definition of final disposition, this represents the worst response rate that could possibly be calculated, and is thus a useful benchmark<sup>2</sup>.

---

<sup>1</sup> Note that the response rates formulas (RR3 and RR4) that incorporate the “e” factor make it possible to remove large numbers of cases from the denominator simply by assuming that some proportion are not eligible. By setting “e” to .5, half of the dispositions of unknown eligibility are removed and the response rate may rise dramatically.

<sup>2</sup> One reviewer suggested that a worse response rate would be to divide completed interviews by total sample released. It is unlikely that such a response rate would ever be used. In some cases, however, all sample released may be considered eligible, as with panel survey. Thus, sometimes eligibility is not determined by the outcome of a call, but by being on the list.

The difference between response rates based on last dispositions and those based on ever being eligible are potentially large. For the Listed 2 survey, the last dispositions yielded a response rate 18 percent higher than the ever-eligible response rate. On average the last-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 response rate using final dispositions, as calculated above, but also applying each of five rules derived by analyzing cases that were ineligible using last disposition, but eligible using the ever-eligible criterion. These rules are:

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

Assuming that these rules are acceptable, in most cases the difference between a response rate based on last disposition and one based on a final disposition using all these rules 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 and 18 percent to 5 percent for both. The remaining cases that are unaccounted for by the rules have a last disposition that is ineligible, such as a “disconnected number”, and a previous disposition of

“answering machine” or “callback”, indicating the household definitely had a potential respondent at some point. These should be coded as eligible even though the last disposition is ineligible.

The response rates achieved by applying each of the five rules separately demonstrates which rules account for most of the difference. As seen in Figure 1, Rule 1 achieves the most gains in RDD surveys while Rule 2 achieves the most gains 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 RDD surveys where there is often little or no screening criteria that affect respondent eligibility.

Most readers would probably consider these five rules to be reasonable. None of the rules indicate that a potential respondent lives in the household, or alleged household, prior to the last call. Further, two of the five rules (Rules 1 and 2) simply give precedence to the last disposition over any previously coded attempt. Table 1 also shows that Rules 1 and 2 account for a large proportion of the gains of the final disposition over the ever-eligible disposition. The latter is important because Rules 1 and 2 account for a large proportion of the gains, and do not require any evaluation of call history to code. This implies that in these six surveys, using the last disposition as the final disposition results in less error than we might think, so long as Rules 1 and 2 are acceptable.

Figure 1 shows the effect of these rules very clearly. The biggest gains occur in listed surveys by allowing cases where an interviewer talked to a previously eligible respondent who was later found to be ineligible. These gains 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 where business numbers are still likely to be found. Rule 3, regarding fax and data lines, affects both RDD and listed surveys equally.

It should be noted that one reason last dispositions tend to be higher than ever-eligible dispositions is that coding a case as “ineligible” tends to finalize it. For example, in virtually all CATI software a case coded “Disconnected number” will not be reintroduced into the sample pool. If last dispositions are used to calculate response rates, the case coded as “Disconnected number” never has an opportunity to become eligible again. Other cases like this are business and non-working numbers.

## **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 calculation of response rates, and even less using empirical data.

In this paper I have argued that the difference between a response rate calculated using last dispositions and one calculated using final dispositions is, given certain assumptions, not that different. So long as the researcher, and the client, are willing to accept a last disposition of ‘Business’ or ‘No eligible respondent’ as taking precedence over any other historical disposition, the response rate using the last disposition is on average only five percent higher across the six surveys examined in this article. Application of all five rules reduces this to two percent. This suggests that the common practice of providing clients a response rate based on last disposition is not unreasonable, and is certainly justified for interim reports.

These five rules can also be considered 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 recode the case as ineligible according to these five rules. This provides a conservative response rate that is predisposed towards accepting previous eligibility in the call history.

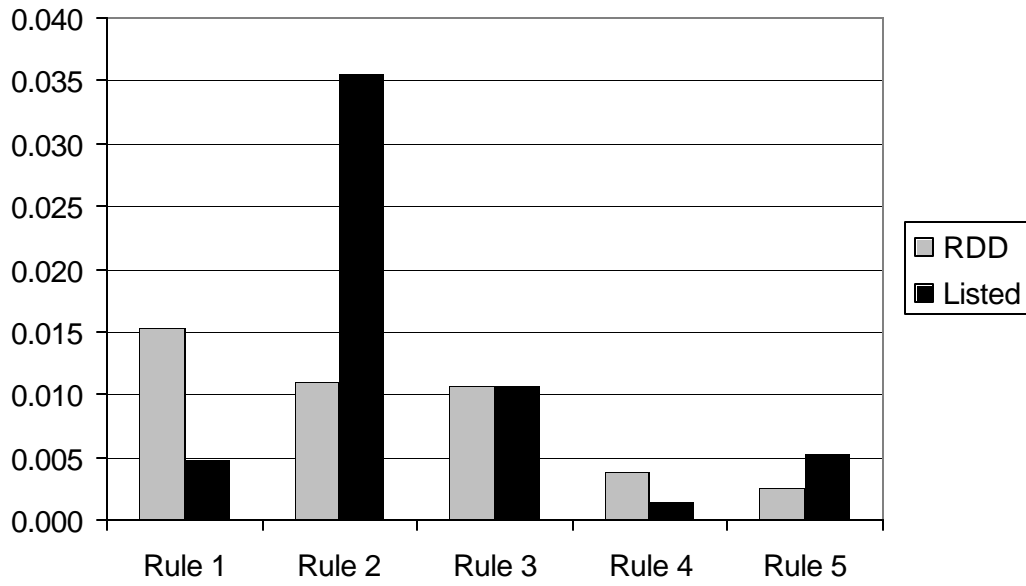
**Table 1. Productivity comparison across surveys.**

	RDD 1	RDD 2	RDD 3	Listed 1	Listed 2	Listed 3
Sample released	3500	29174	4450	30084	16268	1100
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	8010	1002	7630	3823	402
Response rate using last 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

**Table 2. Eligibility of standard AAPOR disposition codes.**

AAPOR Disposition Description	AAPOR Disposition Code	Eligibility (E=eligible, I=ineligible)
Completed interview	1100	E
Partial interview	1200	E
Strong refusal	2110	E
Soft refusal	2120	E
Resp never available	2210	E
Answering machine, message	2221	E
Answering machine, no message	2222	E
Dead	2310	E
Physically, mentally unable	2320	E
Language unable	2330	E
Miscellaneous unable	2340	E
Busy	3120	E
No answer	3130	E
Answering machine, don't know if household	3140	E
Technical phone problem	3150	E
Temporary phone problem	3151	E
Fax/data line	4200	I
Non-working number	4310	I
Disconnected number	4320	I
Number changed	4430	I
Cell phone	4410	I
Call forwarding	4420	I
Business/Government/Other Org	4430	I
Institution	4520	I
Group quarters	4530	I
No eligible respondent	4700	I
Quota filled	4800	I
Callback, Resp not selected	5100	E
Callback, Resp selected	5200	E

**Figure 1. Average gain in response rate over "ever eligible" response rate from applying rule, by type of sample .**



## References

- 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* 6:45-64.
- Brick, John Michael. 1997. An Experiment In Random-Digit Dial Screening. Technical Report National Center For Education Statistics. U.S. Department of Education, Washington D.C.
- Cohen, Steven B. and Barbara Lepidus Carlson. 1995. Characteristics Of Reluctant Respondents In The National Medical Expenditure Survey. *Journal of Economic and Social Measurement* 21:269-297.
- Dillman, Don. 1978. *Mail And Telephone Surveys: The Total Design Method*. New York: Wiley.
- Frey, James. 1989. *Survey Research By Telephone*. Newbury Park: Sage.
- Goyder, John. 1987. *Silent Minority: Nonrespondents On Sample Surveys*. Boulder: Westview Press.
- Groves, Robert, Robert Cialdini and Mick Couper. 1992. Understanding The Decision To Participate In A Survey. *Public Opinion Quarterly* 56: 475-496.
- I-Fen, Lin and Nora Cate Schaeffer. 1995. Using Survey Participants To Estimate The Impact Of Nonparticipation. *Public Opinion Quarterly* 59:236-259.
- Keeter, Scott. 1995. Estimating Telephone Noncoverage Bias With A Telephone Survey. *Public Opinion Quarterly* 59:196-218.
- Kristal, Alan, Emily White, Julie Davis, Gayle Corycell, Trivellore Raghunathan, Susan Kinne, Ting-Kwong Lin. 1993. Effects Of Enhanced Calling Efforts On Response Rates, Estimates Of Health Behavior, And Costs In Telephone Health Survey Using Random-Digit Dialing. *Public Health Reports* 3:372-380.
- Lavrakas, Paul. 1987. *Telephone Survey Methods: Sampling, Selection And Supervision*. Newbury Park: Sage.
- Lessler, Judith T. and William D. Kalsbeek. 1992. *Nonsampling Error In Surveys*. New York: Wiley.
- Piazza, Thomas. 1993. Meeting The Challenge Of Answering Machines. *Public Opinion Quarterly* 57: 219-232.
- Roth, Philip L., Craig A. Be Vier. 1998. Response Rates In HRM/OB Survey Research: Norms And Correlates, 1990-1994. *Journal of Management* 24: 97-118.

The American Association for Public Opinion Research. 2000. *Standard Definitions: Final Dispositions Of Case Codes And Outcome Rates For Surveys*. Ann Arbor, Michigan: AAPOR.

Traugott, Michael W. 1987. The Importance Of Persistence In Respondent Selection For Pre-Election Surveys. *Public Opinion Quarterly* 51:48-57.

Zapka, J.G ., L. Chasantaber, C. Bigelow and T. Hurley. 1994. Methodological Issues For Health-Related Surveys Of Multicultural Older Women *Evaluation & The Health Professions* 17:485-500.