Response Rate Calculation Methodology for Recruitment of a Two-Phase Probability-Based Panel: The Case of AmeriSpeak

Authors: Robert Montgomery, J. Michael Dennis, Nada Ganesh

Abstract

Response rate is, at best, a partial measure of the quality of a statistical sample of a population. Nonetheless, it is arguably the most widely cited metric of the quality of a statistical survey. Transparency in response rate reporting is increasingly a priority for AAPOR and for the survey research community during a time when panel-based sampling designs are increasingly complex. Because of the complexity of the probability-based panels, reviewers of panel-based surveys are challenged to understand without ambiguity how a panel-based response rate is actually calculated and how it compares to the response rates of other panels. Moreover, documenting the response rate calculation methodology is instrumental in complying with the guidelines from the AAPOR Transparency Initiative, to build trust with the users and consumers of the panel-based survey data, and to advance the micro-field of response rate methodology. In this brief paper, we attempt to provide transparency and document an application of the AAPOR Standard Definitions in calculating the response rate for the panel recruitment stage for NORC's AmeriSpeak Panel, which involves a two-phase probability-based sample design. By documenting and making public our interpretation of the AAPOR Standard Definitions in their application to the AmeriSpeak Panel recruitment, our goal is to invite feedback from the survey research community as well as facilitate the development of response rate calculation methodology for probability-based household panels.

Introduction

Online panels have become a key source of data collection for many surveys across a wide range of topic areas, but due to the wide variation in design and methodology between panels, it is often difficult to compare or even understand fully the quality of specific panels. The increasing complexity of various panel-based sampling frames and data collection modes adds to this problem. In fact, one of the most important and distinctive design features of an online panel is the use of probability-based recruitment or volunteer opt-in recruitment for building the sample panel. Calculation for virtually all standard AAPOR response metrics apply for probability-based online panels, but they all do not apply to volunteer opt-in online panels, leading to inappropriate apples to oranges comparisons of response metrics among online panels. Transparent documentation and reporting of online panel design and response metrics, although only a partial solution, is a critical component in the effort to bring transparency to the reporting and use of social science and public opinion data collected using online panels.

In this article we describe the calculation of the household recruitment rate for NORC's AmeriSpeak® Panel. AmeriSpeak is a probability-based mixed-mode panel designed to be representative of the US household population. Randomly selected US households are sampled with a known, non-zero probability of selection from the NORC National Frame, and then contacted by US mail, telephone, and field interviewers (face to face). (NORC at the University of Chicago 2015) In order to provide a clear understanding of the AmeriSpeak methodology and avoid the complexity associated with unique design and fielding of individual surveys, we focus only on the household recruitment stage. AmeriSpeak employs a two-phase panel recruitment method which provides improved demographic representation of the panel sample and a higher response rate over a one-phase recruitment approach.

Literature Review

Callegaro and Disogra (2008) lay out a framework for the calculation and reporting of response rates for online panels which includes three stages: recruitment, connection, and survey completion. Each of these stages has an associated success rate that can be computed for probability-based panels, such as the AmeriSpeak Panel. This framework was adapted for the seventh and subsequent versions of the AAPOR Standard Definitions document which describes the calculation of response rates for Web panels. (Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 7th edition. AAPOR 2011) (Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 9th edition. AAPOR 2016) Further expansion of this method to include multi-stage and multi-mode designs is described specifically for probability-based online panels. (Disogra and Callegaro 2016)

In recent years, several probability-based online panels have provided descriptions of their recruitment methodologies and response rates. For example, The American Trends Panel (ATP) of the Pew Research Center was recruited in early 2014 from respondents to a national cellphone and landline RDD telephone survey, the 2014 Political Polarization and Typology Survey. (Pew Research Center 2015) Pew reports that the cumulative panel response rate accounting for the response rate to the recruitment telephone survey, agreement to join the panel, and the average response rate of a panel introduction survey (such as a recruitment profile survey) is on the order of 3.5% for each wave of recruitment. The Gallup Organization recruits members for the Gallup Panel using an RDD phone survey, attempting to recruit all household members ages 13 and older. Households that agree to participate are mailed a "membership packet", and those households that return a completed packet become part of the panel. (Tortora 2009) The reported panel response rate for the portion of the Gallup panel recruited as part of the February 2006 World Affairs survey is 8%. The panel response rate was calculated as the product of initial recruitment, completion of the welcome packet, and the world affairs survey response rate. (Arens and Steiger 2006)

The design of the AmeriSpeak Panel incorporates a second phase of recruitment, Phase 2, in which a sub-sample of households that have not responded to initial Phase 1 recruitment efforts is selected for an intensive non-response follow-up recruitment effort, including an in-person component. The recruitment rate for the AmeriSpeak Panel is therefore a weighted version of the calculation recommended by AAPOR that accounts for the both Phase 1 recruitment and Phase 2 non-response follow-up recruitment, described in the following section.

Recruitment Methodology

The AmeriSpeak Panel is recruited using a multi-stage address-based sample (ABS) design. The sample frame is the NORC National Frame, an area probability sample frame constructed by NORC based on the US Postal Service Delivery Sequence File (USPS DSF), and providing sample coverage of approximately 97 percent of US households. The National Frame contains a probability sample of almost 3 million households, including over 80,000 rural households added through the in-person listing of households that were not recorded on the USPS DSF. For the 2014-2015 AmeriSpeak recruitment, a stratified clustered random sampling approach was used to select sample units from the National Frame including an oversample of households in Census tracts or block groups with high proportions of young adults and/or Hispanics and/or African-Americans, and attempts were made to recruit all English-speaking members age 18 and older in the sampled households. (Beginning with 2016 AmeriSpeak recruitment, attempts were made to recruit all English- and Spanish-speaking members age 18 and older in sampled households.)

As mentioned prior, recruitment is a two-phase process. In the initial mail recruitment to sampled addresses, households are invited to join AmeriSpeak by visiting the panel website AmeriSpeak.org or by calling a toll-free telephone number. Study invitations are communicated via an over-sized pre-notification postcard, a USPS recruitment package in a 9"x12" envelope (containing a cover letter, a summary of the privacy policy, FAQs, and a study brochure), two follow-up post cards, and also email contacts and telephone follow-up for sample units matched to vendor address lists containing that contact information. Phase 1 recruitment occurs over a period of 9 weeks. The Phase 2 non-response campaign targets a stratified random sub-sample of the non-responders from the initial recruitment. Nonresponse strata are defined based on consumer vendor data identifying whether a household is likely to include a household member who is a young adult, of Hispanic origin, or not of Hispanic origin and Black. Units sampled for the nonresponse follow-up campaign are sent by express mail a new recruitment package with an enhanced incentive offer. NORC field interviewers then make personal, face-to-face visits to the respondents' homes to encourage participation in AmeriSpeak by taking the action of joining AmeriSpeak via the AmeriSpeak.org web portal, by completing the enrollment with the staff via the CAPI Instrument in the staff's tablet PC, or by calling the AmeriSpeak toll-free telephone number. Phase 2 recruitment occurs over a minimum of 8 weeks.

All cooperating households first register for AmeriSpeak by providing name and contact information for validating the sample unit, followed by documenting informed consent and agreement to the study's Privacy Policy and Terms and Conditions. Participants are also asked their preferred mode of responding to AmeriSpeak surveys, web or CATI. Afterwards, AmeriSpeak panelists complete an introduction survey of about 15 minutes by web or by CATI phone to collect information useful for sampling, weighting, and analysis of sample representativeness. Figure 1 illustrates the recruitment process.

Figure 1: AmeriSpeak Panel Recruitment Methodology
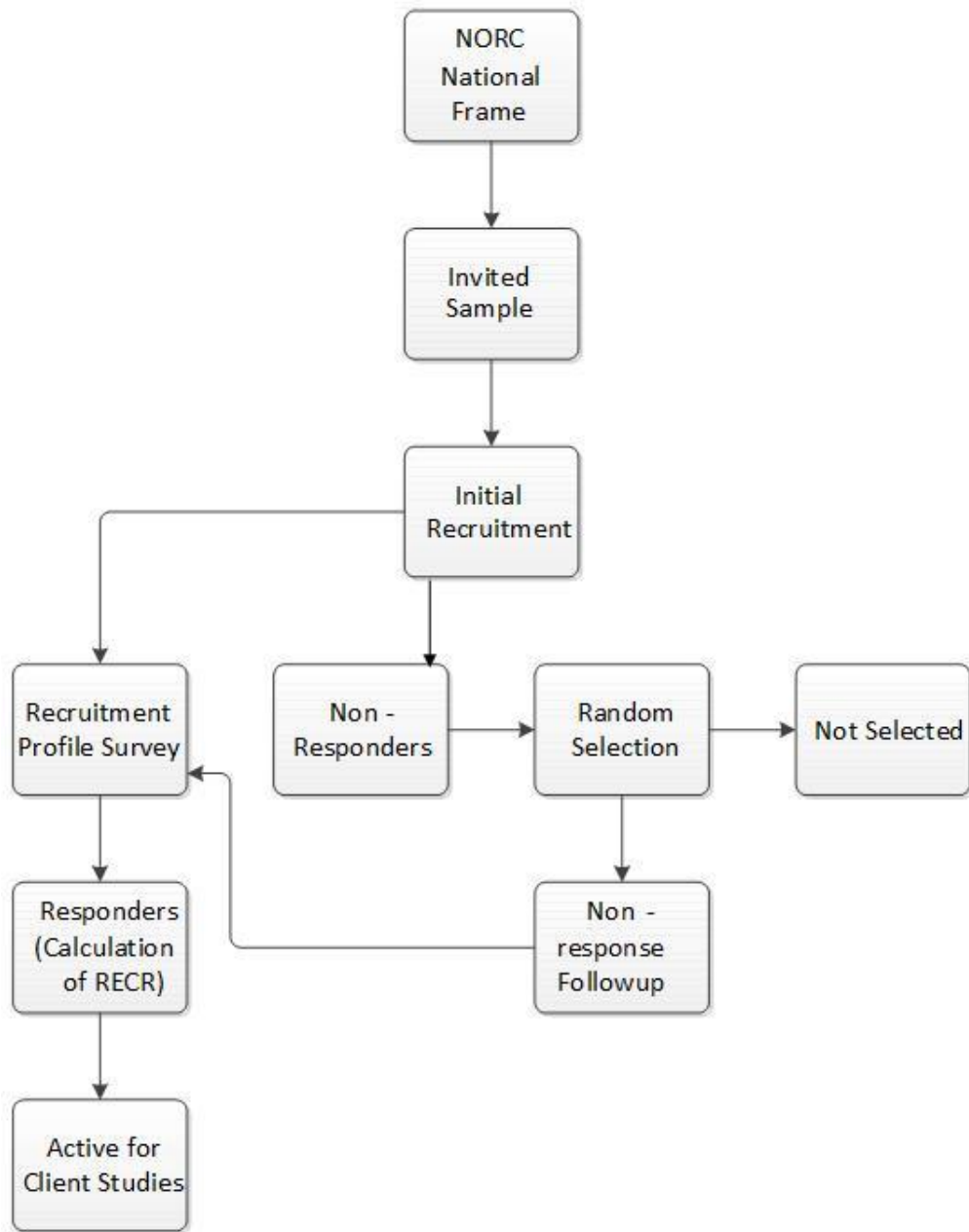
Response Rate Calculation

The AmeriSpeak panel **household recruitment rate**, *RECR*, is defined as the weighted AAPOR 3 response rate:

$$RECR = \frac{IC_{BW}}{IC_{BW} + (R_{BW} + NC_{BW} + UO_{BW}) + e_{BW} * UH_{BW}}$$

where

- IC_{BW} = base weighted total number of recruited households

- R_{BW} = base weighted total number of age eligible but non-recruited households (at least one household member is identified as 18+ years of age)

- NC_{BW} = base weighted total number of non-contacted households

- UO_{BW} = base weighted total number of contacted and known households but with unknown age eligibility

- UH_{BW} = base weighted total number of housing units with unknown eligibility

- e_{BW} is the observed eligibility rate, which is calculated as

$$e_{BW} = \frac{(IC_{BW} + R_{BW} + UO_{BW} + NC_{BW})}{(IC_{BW} + R_{BW} + UO_{BW} + NC_{BW} + J_{BW} + OOS_{BW})}$$

- J_{BW} = base weighted total number of age ineligible households (all households members are less than 18 years of age)

- OOS_{BW} = base weighted total number of out of scope housing units (i.e., vacant, vacation home, etc.)

Note that households with disposition codes UO and NC are assumed to have at least one 18+ year old person in the household. This is a conservative assumption which marginally decreases the reported recruitment rate.

As of November 2015, the weighted household recruitment rate (*RECR*) for the AmeriSpeak panel was 36.9% using this calculation. This compares favorably with Pew's ATP cumulative panel response rate of 3.5% and the Gallup Panel response rate at 8%, and is attributable to AmeriSpeak's two-phase panel recruitment strategy.

Application to Client Surveys

The RECR for the AmeriSpeak panel as reported above is for the full recruited panel sample, representative of the entire US. Depending on the geographic and demographic sample requirements of a specific client survey, the RECR may differ. For example, the RECR may be higher or lower in specific states or regions of the country. Additionally, in computing final response rates for client surveys, the

client survey response rate (calculated as completed interviews divided by panel members invited to take the survey) is multiplied by the *RECR* and the weighted retention rate.

Conclusion

In an environment of declining response rates, a rapidly changing landscape, and the constant introduction of novel survey methodologies it is vital that survey practitioners strengthen their commitment to transparency and openness. The AAPOR Transparency Initiative is an important step in this direction, requiring documentation of survey methodologies, including the calculation of response metrics. Our goal in documenting the response rate calculation for the AmeriSpeak panel recruitment is to not only fulfill our obligation as members, but to help build trust in users of the panel, as well as provide a basis of comparison between different online panels.

References

2011. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 7th edition. AAPOR.* The American Association for Public Opinion Research.

2016. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 9th edition. AAPOR.* The American Association for Public Opinion Research. .

Arens, Zachary, and Darby Miller Steiger. 2006. "Time in Sample: Searching for Conditioning in a Consumer Panel." *Public Opinion Quarterly.*

Callego, Mario, and Charles Disogra. 2008. "Computing Response Metrics for Online Panels." *Public Opinion Quarterly* 72 (5): 1008-1032.

Disogra, Charles, and Mario Callegaro. 2016. "Metrics and Design Tool for Building and Evaluating Probability-Based Online Panels." *Social Science Computer Review* 34 (1): 26-40.

NORC at the University of Chicago. 2015. *AmeriSpeak Advantage.* January. http://norc.org/PDFs/AmeriSpeak%20Advantage%20updated%20%202015%2001%2026-FINAL%20WEB%20FILE.PDF.

Pew Research Center. 2015. "Building Pew Research Center's American Trends Panel." Washington, D.C.

2011. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 7th edition. AAPOR.* The American Association for Public Opinion Research.

2016. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 9th edition. AAPOR.* The American Association for Public Opinion Research. .

Tortora, Robert. 2009. "Attrition in Consumer Panels Methodology of Longitudinal Surveys, ed. Peter Lynn, pp. 235–49. Hoboken, NJ: Wiley." In *Methodology of Longitudinal Surveys*, edited by Peter Lynn, 235–49. Hoboken, NJ: Wiley.