RANDS COVID-19 SURVEY 2 NATIONAL CENTER FOR HEALTH STATISTICS (NCHS)

AMERISPEAK PROJECT REPORT

August 28, 2020

Client Contact:
Paul Scanlon | wyv6@cdc.gov
NORC Account Manager:
 J. Michael Dennis | Dennis-Michael@norc.org
NORC Project Manager:
John Dombrowski | Dombrowski-John@norc.org

STUDY INTRODUCTION

NORC conducted the RANDS COVID-19 Survey 2 on behalf of NCHS using NORC's AmeriSpeak® Panel for the sample source. The AmeriSpeak sample was supplemented by nonprobability online opt-in sample (Dynata). This project was an extension of ongoing questionnaire evaluation research for the Collaborating Center for Questionnaire Design and Evaluation Research (CCQDER) within the National Center for Health Statistics' (NCHS) Division of Research and Methodology (DRM). This survey, the second of two, additionally focused on the general population's experience with COVID-19, including experiences with health care during the pandemic, impacts on health (mental and physical) and financial security, and quarantine and isolation attitudinal and behavioral measures.

In addition to supporting a better understanding of the general population's experience with COVID-19, this project also placed an emphasis, as in previous iterations of the RANDS surveys, on new measurement items and web and phone probes. Five experiments were included in the survey in order to study these more (the first two were also included in RANDS COVID-19 Survey 1):

1) **NHIS vs. Alternate COVID-19 Test**: Comparing two ways of asking about receiving diagnostic tests for COVID-19, with 1) being a standalone question not specifying what kind of COVID-19 test the respondent received and 2) being a series of two questions asking specifically about a test for *having the COVID-19 virus at the time of the test* vs. an *antibody test to determine if one's had it in the past*.

2) **Quarantine Probe:** Comparing differences in responses to a close-ended (multiple-punch) vs. open-ended probe asking respondents to describe their interpretation of quarantine/isolation behavior. This was expanded in Survey 2 to also include a close-ended (multi-punch) vs. open-ended probe asking respondents about the medical provider's use of telemedicine.

3) **Immunocompromised Test:** Comparing 1) a single close-ended question about having a condition that weakens one's immune system vs. 2) a pair of questions - one close-ended question about having a condition that weakens the immune system and another about taking medicine or have any medical procedure that compromised one's immune system.

4) **Symptoms Probe:** Comparing two different open-ended probes of 1) "what" respondents were thinking about vs. 2) "why" respondents answered a certain way when describing their coronavirus symptoms at their worst point.

5) **Time-of-Reference Test:** Comparing the difference in time reference between 1) 4-month period vs. 2)12-month period. This difference was asked across a series of questions asking about mental and emotional well-being (i.e., medication use and counseling and therapist visits).

This study was offered in English only, and it was administered via online web survey and phone interview.

This AmeriSpeak Project Report supplements the information provided in the *NORC Card*, which provides an in-depth profile of sample quality metrics for the study, the data collection field period, interview sample size, response rate statistics, the design effect, and sampling margins of error, among other statistics. Please refer to the NORC Card for information useful for compliance with the AAPOR Transparency Initiative, in addition to information provided in this AmeriSpeak Project Report.

For more detailed information on the AmeriSpeak panel recruitment and management methodology, please see the Appendix ("Technical Notes on AmeriSpeak Methodology") attached to this AmeriSpeak Project Report.

STUDY-SPECIFIC DETAILS

Sampling

A general population sample of U.S. adults age 18 and older was selected from NORC's AmeriSpeak Panel for this study. For additional details on the selection of the sample, please see the memo on Sampling and

Weighting prepared for and previously shared with NCHS, "NORC Memo Covid-19 RANDS Sampling and Weighting 2020.05.18 Final.docx". During the RANDS COVID-19 Survey 2 sample preparation, n=12 cases from the original COVID-19 Survey 1 sample were removed, as they were no longer active panelists (i.e., withdrawn from the panel).

The sample for this specific study was selected from the AmeriSpeak Panel using sampling strata based on age, race/Hispanic ethnicity, education, gender, and household income (96 sampling strata in total). The size of the selected sample per sampling stratum is determined by the population distribution for each stratum. In addition, sample selection takes into account expected differential survey completion rates by demographic groups so that the set of panel members with a completed interview for a study is a representative sample of the target population. If panel household has one more than one active adult panel member, only one adult in the household is eligible for selection (random within-household sampling). Panelists selected for an AmeriSpeak study earlier in the business week are not eligible for sample selection until the following business week.

For technical information about the AmeriSpeak Panel, including recruitment process and panel management policies, please see the Appendix.

The AmeriSpeak panel sample was supplemented with respondents from the Dynata nonprobability online opt-in panel.

Field

A small sample of English-speaking AmeriSpeak web-mode panelists were invited on Friday, July 24th for a pretest. In total, NORC collected 39 pretest interviews. The initial data from the pretest was reviewed by NORC and NCHS approved NORC's data review.

The following changes were made before fielding the Main survey to collect the 10,000 interviews.
> Question/Response Text Updates: PROBE_AUTOIM
> New Questions added: PROBE_SUSPECT
> Logic Update: SUSPECT

In total NORC collected 11,483 interviews in the Main survey, 11,061 by web mode (AmeriSpeak & Dynata) and 422 by phone mode (all AmeriSpeak).

Please see NORC Card for field period, sample sizes, and the AAPOR response rate documentation for the AmeriSpeak sample.

Sample Performance Summary

The sample performance summary is below.

Distribution of Completed Interviews in Delivered Analysis File
by Sample Source (Unweighted)

Dynata Sample Source		NORC AmeriSpeak Source		Total	
N Interviews	%	N Interviews	%	N Interviews	%
5,502	47.9%	5,981	52.1%	11,483	100%

Response Rate Reporting for AmeriSpeak sample

AmeriSpeak Survey 2 Sample

Weighted AAPOR RR3 Recruitment rate: 35.1%

Weighted Household retention rate: 83.6%

Survey completion rate: 69.1%

Weighted AAPOR RR3 cumulative response rate: 20.3%

AmeriSpeak Longitudinal Sample

Weighted AAPOR RR3 Recruitment rate: 35.1%

Weighted Household retention rate: 83.6%

Survey completion rate: 80.2%

Weighted AAPOR RR3 cumulative response rate: 23.5%

Gaining Cooperation of AmeriSpeak Panelists for the Study

To encourage study cooperation, NORC sent email reminders and text message reminders to sampled web-mode panelists on the following dates:

- Monday, August 3rd (Initial Soft-launch invitation)
- Wednesday, August 5th (Full/Remaining sample invitation)
- Friday, August 7th (Reminder to soft-launch sample)
- Saturday, August 8th (Reminder to remaining-launch sample)
- Tuesday, August 11th
- Monday, August 17th

Furthermore, NORC sent out an invitation/reminder notification letter via USPS to n=3,164 (2 of the 12 inactive cases were a part of the mail efforts) low-completion rate panelists on August 5th, in order to boost completion rate and sample representativeness. For more details on the mail efforts (rationale, selection criteria, etc.), please see the memo NORC prepared and previously shared, "NORC Memo on Field Exp to NCHS 2020.05.27.docx".

To administer the phone-survey, NORC dialed the sampled phone-mode panelists throughout the field period.

Panelists were offered the cash equivalent of $5 for completing the survey.

Data processing

NORC prepared a fully labeled data file of respondent survey data and demographic data for NCHS.

NORC applied standard cleaning rules to the survey data for quality control. In total, 193 cases (117 AmeriSpeak; 76 Dynata) were removed from the final data set of completed interviews based on two cleaning rules. Descriptions of the cleaning criteria and the counts from each are below (counts are overlapping).
- Removing Speeders (i.e., those that completed the survey in less than 1/3 the median duration)
    - n=183 (107 AmeriSpeak; 76 Dynata) removed for speeding.
- Removing Respondents with High Refusal Rates (i.e., those that skip or refused more than 50% of the eligible questions)
    - n=35 (33 AmeriSpeak; 2 Dynata) removed for high refusal rates.

Of those 193 cases removed:
- n=168 (94 AmeriSpeak; 24 Dynata) were marked with one of the two flags above

- n=25 (23 AmeriSpeak; 2 Dynata) were marked with both of the flags above

Because the behavior of these respondents indicated they did not meet the criteria to be eligible for the main study interview, the survey responses from these respondents were removed from the final data set prior to post-stratification weighting. These respondents were not counted toward the total number of interviews delivered

AmeriSpeak Probability-based Sample Weights

Standard weighting procedures are used to develop weights for the AmeriSpeak probability-based sample. This involves computing a design weight to account for the selection probabilities under the sample design, and weighting adjustments for unknown eligibility, nonresponse, and frame coverage (panel base sampling weights). The design-weighted data are then calibrated to the known distribution of the target population for a set of demographic variables through raking ratio adjustments to create final study weights.

AmeriSpeak Panel Weights: Since the sampling frame for this study is the AmeriSpeak Panel, which itself is a sample, the starting point of the weighting process for the study is the AmeriSpeak panel weight[1]. To develop the panel weight, NORC first computed the panel base weight as the inverse of the probability of selection from the NORC National Frame (the sampling frame that is used to sample housing units for AmeriSpeak) or other address-based sample frames (supplemental panel samples were selected from frames developed from the USPS Delivery Sequence Files). The sample design and recruitment protocol for the AmeriSpeak Panel involves unequal sampling rates across the sampling strata and additional subsampling of initial nonresponding housing units for in-person nonresponse follow-up (NRFU). The panel base weights reflect all the variations in panel sample selection probabilities. The panel base weights are then adjusted to account for unknown eligibility and nonresponse among eligible housing units. These adjustments were conducted using weighting classes defined by some household characteristics provided by commercial data vendors, including partisan score, political party identification, the presence of young adult(s), and minority status. To produce the final household panel weights, the household-level nonresponse adjusted weights are post-stratified to match the number of households per census division obtained from the most recent Current Population Survey (CPS). Final household weights are assigned to each eligible adult in the recruited household. These person-level weights are then adjusted to compensate for nonresponding adults within a recruited household. Finally, the nonresponse adjusted person-level panel weights are raked to population totals associated with the following variables:

Panel Weighting Variables & the Variable Categories
Age: 18-24, 25-29, 20-39, 40-49, 50-59, 60-64, and 65+
Gender: Male and Female
Census Division: New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific
Race/Ethnicity: Non-Hispanic White, Non-Hispanic Black, Hispanic, and Non-Hispanic Other
Education: Less than High School, High School/GED, Some College, and BA and Above
Housing Tenure: Home Owner and Other
Household phone status: Cell Phone-only, Dual User, and Landline-only/Phoneless

The external population totals are obtained from the Current Population Survey, except for Household Phone Status, which is determined by the National Institutes of Health bi-annual survey on wireless substitutions[2]. The weights adjusted to the external population totals are the *final panel weights*.

---

[1] The AmeriSpeak panel weight existed prior to this study; the weighting procedures are described here for clarity and completeness.

[2] Blumberg SJ, Luke JV. Wireless substitution: Early release of estimates from the National Health Interview Survey, January–June 2019. National Center for Health Statistics. May 2020. Available from: https://www.cdc.gov/nchs/nhis.htm

Study-specific base weights are computed as the product of the final panel weights and the inverse of the probabilities of selection under the study sample design. These base weights reflect the cumulative probabilities of selection into the study sample. Since not all sampled panel members respond to the interview, an adjustment is needed to compensate for non-respondents. This nonresponse adjustment decreases potential nonresponse bias associated with sampled panel members who did not complete the survey.

The *nonresponse adjusted weights* for the study are adjusted via a raking ratio adjustment method to general population known population totals associated with the following socio-demographic characteristics:

> **Study-Specific Post-Stratification Weighting Variables & the Variable Categories**
> **Age:** 18-24, 25-29, 20-39, 40-49, 50-59, 60-64, and 65+
> **Gender:** Male and Female
> **Census Division:** New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific
> **Race/Ethnicity:** Non-Hispanic White, Non-Hispanic Black, Hispanic, and Non-Hispanic Other
> **Education:** Less than High School, High School/GED, Some College, and BA and Above

These socio-demographic characteristics were weighted to the Current Population Survey.

At the final stage of weighting, any extreme weights were trimmed based on a criterion of minimizing the mean squared error associated with key survey estimates. Weights after trimming are re-raked to the same population totals to produce the *final study weights* for our probability-based sample.

AmeriSpeak Longitudinal Sample Weights

Longitudinal weights for the AmeriSpeak sample are calculated for those that completed both RANDS COVID-19 Survey 1 and Survey 2. The 6,800 completes from Survey 1 comprised the starting sample for Survey 2's longitudinal portion. The post-stratification weights described in the Survey 1 Project Report were used as the starting *study-specific base weights* for the longitudinal sample.

To minimize potential nonresponse bias, nonresponse adjustments for the Survey 2 completes (5,452 completes) are computed across respondent demographics. The *nonresponse adjusted weights* are adjusted via a raking ratio adjustment method to general population known population totals associated with the following socio-demographic characteristics:

> **Study-Specific Post-Stratification Weighting Variables & the Variable Categories**
> **Age:** 18-24, 25-29, 20-39, 40-49, 50-59, 60-64, and 65+
> **Gender:** Male and Female
> **Census Division:** New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific
> **Race/Ethnicity:** Non-Hispanic White, Non-Hispanic Black, Hispanic, and Non-Hispanic Other
> **Education:** Less than High School, High School/GED, Some College, and BA and Above

These socio-demographic characteristics were weighted to the Current Population Survey.

At the final stage of weighting, any extreme weights were trimmed based on a criterion of minimizing the mean squared error associated with key survey estimates. Weights after trimming are re-raked to the same population totals to produce the *final study weights* for our probability-based longitudinal sample.

Probability and Nonprobability Joint Sample TrueNorth™ Calibration Weights

In addition to the 5,981 probability-based AmeriSpeak sample, which was always designed to be a stand-alone nationally representative sample, we included 5,502 non-probability sample from Dynata to enhance the overall sample size. As there is no known "design" to nonprobability samples, units in the nonprobability

sample are simply given a design weight of one. The nonprobability sample is then calibrated to the same known distributions of the population as described above for the probability sample. Therefore, the nonprobability sample weights are calibration weights.

To create the final calibration weights for the combined probability and non-probability sample, first we calculate weights for the AmeriSpeak probability-based sample as described above. We then calculate statistical corrections made to the non-probability sample via NORC's TrueNorth™ calibration weighting service, an innovative hybrid calibration approach developed at NORC based on small area estimation methods in order to explicitly account for potential bias associated with the nonprobability sample[3,4]. The purpose of TrueNorth calibration is to adjust the weights for the nonprobability sample so as to bring weighted distributions of the nonprobability sample in line with the population distribution for characteristics correlated with the survey variables. Such calibration adjustments help to reduce potential bias, yielding more accurate population estimates.

At the core of the TrueNorth method, *small area modeling* is conducted in the following steps:

- First, we identify a set of four key response variables from the survey using a machine learning approach called gradient boosted tree modelling. Ideally, the key response variables are associated with the largest bias in the nonprobability sample and also are highly correlated with other response variables.
- Second, we define a set of 36 domains in the data, where each domain is a specific, relevant subgroup for data analysis and reporting.
- Third, we fit domain-level small area models for each of the response variables identified earlier using weighted probability sample and nonprobability sample domain-level estimates as input. The model included covariates, domain-level random effects, and sampling errors. The covariates were external data available from the Current Population Survey (CPS), American Community Survey (ACS), General Social Survey (GSS), and National Health Interview Survey (NHIS).
- Fourth, the fitted small area models generate predicted values for each domain and for each response variable.

After the steps, the weighted AmeriSpeak sample and the calibrated nonprobability sample are used to develop a small area model to support domain-level estimates, where the domains were defined by race/ethnicity, age, education, and gender. The dependent variables for the models were key survey variables.

Finally, the combined probability and nonprobability sample weights were derived such that for the combined sample, the weighted estimate reproduced the usual demographic benchmarks as well as small domain estimates (derived using the small area model) for key survey variables.

Design Effect and Sampling Margin of Error Calculations

Probability-based sample:
Study design effect: 1.96
Study margin of error: +/- 1.77%

Probability-based Longitudinal sample:

[3] Ganesh, N., Pineau, V., Chakraborty, A., Dennis, J.M., (2017). "Combining Probability and Non-Probability Samples Using Small Area Estimation." *Joint Statistical Meetings 2017 Proceedings.*
[4] Yang, Y. Michael, Nada Ganesh, Ed Mulrow, and Vicki Pineau. (2018). "Estimation Methods for Nonprobability Samples with a Companion Probability Sample," Proceedings of the Joint Statistical Meetings, 2018.

Study design effect: 1.82
Study margin of error: +/- 1.79%

TrueNorth probability-nonprobability sample:
Study design effect: 1.79
Study margin of error: +/- 1.32%

Under TrueNorth, the margins of error were estimated from the root mean squared error associated with the small area model, along with other statistical adjustments. A TrueNorth estimate of margin of error is a measure of uncertainty that accounts for the variability associated with the probability sample as well as the potential bias associated with the nonprobability sample.

Deliverables

The following files were created for NCHS as part of the study deliverables:

- Survey interview data file in both SPSS and SAS formats
- Codebook in Excel format
- Final programming questionnaire in Word document
- Project report documenting study procedures
- NORC Card (for AmeriSpeak Survey 2 and Longitudinal samples)

HOW TO DESCRIBE AMERISPEAK AND NORC @ THE UNIVERSITY OF CHICAGO

For purposes of publication, when describing AmeriSpeak and its panel methodology, we recommend using the following language:

> Funded and operated by NORC at the University of Chicago, **AmeriSpeak®** is a probability-based panel designed to be representative of the US household population. Randomly selected US households are sampled using area probability and address-based sampling, with a known, non-zero probability of selection from the NORC National Sample Frame. These sampled households are then contacted by US mail, telephone, and field interviewers (face to face). The panel provides sample coverage of approximately 97% of the U.S. household population. Those excluded from the sample include people with P.O. Box only addresses, some addresses not listed in the USPS Delivery Sequence File, and some newly constructed dwellings. While most AmeriSpeak households participate in surveys by web, non-internet households can participate in AmeriSpeak surveys by telephone. Households without conventional internet access but having web access via smartphones are allowed to participate in AmeriSpeak surveys by web. AmeriSpeak panelists participate in NORC studies or studies conducted by NORC on behalf of governmental agencies, academic researchers, and media and commercial organizations.
>
> For more information, email AmeriSpeak-BD@norc.org or visit AmeriSpeak.norc.org.

If editors or reviewers are requesting anything more specific or any other detail, please reach out to us to make certain you are using accurate language.

NORC at the University of Chicago is best described as follows:

NORC at the University of Chicago is an independent research institution that delivers reliable data and rigorous analysis to guide critical programmatic, business, and policy decisions. Since 1941, NORC has

conducted groundbreaking studies, created and applied innovative methods and tools, and advanced principles of scientific integrity and collaboration. Today, government, corporate, and nonprofit clients around the world partner with NORC to transform increasingly complex information into useful knowledge. Please visit www.norc.org for more information.

TECHNICAL OVERVIEW OF THE AMERISPEAK® PANEL
NORC'S PROBABILITY-BASED HOUSEHOLD PANEL

Updated June 16, 2020

This technical overview provides the basic information about AmeriSpeak®, a large probability-based panel funded and operated by NORC at the University of Chicago. AmeriSpeak is designed to be representative of the U.S. household population, including all 50 states and the District of Columbia. U.S. households are randomly selected with a known, non-zero probability from the NORC National Frame as well as address-based sample (ABS) frames, and then recruited by mail, telephone, and by field interviewers face to face. AmeriSpeak panelists participate in NORC studies or studies conducted by NORC on behalf of governmental agencies, academic institutions, the media, and commercial organizations.

The construction of the AmeriSpeak panel started in 2014 with pilot samples. In 2015, about 7,000 households were recruited from a sample of around 60,000 addresses. In 2016, about 128,000 addresses were sampled to expand the panel to around 20,000 recruited households. About 51,000 addresses were selected for the 2017 recruitment, which led to the expansion of the regular AmeriSpeak panel to 23,000 recruited households. The AmeriSpeak Panel expanded to approximately 30,000 households in 2018 and 35,000 households in 2019 through further recruitment efforts. The 2020 recruitment is expected to expand the panel by another 5,000 households. The current panel size is 48,900 panel members age 13 and over residing in over 40,000 households.

In addition to the regular panel for general population studies, AmeriSpeak also contains sub-panels to support studies of special populations, including AmeriSpeak Latino, AmeriSpeak Teen, and AmeriSpeak Young Adult 18-34 (which features an oversample of African Americans, Hispanics, and Asians). AmeriSpeak is also the probability sample source for TrueNorth®, the NORC calibration solution for combining probability and non-probability samples for estimation through small area modeling that leverages data from AmeriSpeak, the American Community Survey, Current Population Survey, and other data sources for improved statistical efficiency.[5]

Panel Sample Frame

The primary sampling frame for AmeriSpeak is the 2010 NORC National Frame, a multistage probability sample that fully represents the U.S. household population. We provide a brief description of how the National Frame was constructed after the 2010 Census.

The primary sampling units (PSUs) in the first stage sample selection are 1,917 National Frame Areas (NFAs), each of which is an entire metropolitan area (made up of one or more counties), a county, or a group of counties with a minimum population of 10,000. A total of 126 NFAs are selected in the first stage, including 38 certainty NAFs, 60 urban NFAs, and 28 non-urban NFAs. The largest 38 NFAs, those with a population of at least 1,543,728 (0.5 percent of the 2010 Census U.S. population), were selected into the National Frame with certainty. The certainty NFAs consist of areas with high population density and are dominated by census tracts with city-style mailing addresses. These areas contain 56 percent of the U.S. household population while only accounting for about 8 percent of the geographic area of the United States. The remaining 1,879 NFAs are stratified into urban areas where city-style addresses predominate, and rural areas that are less likely to have city-style addresses. The latter stratum comprises 81 percent of the geographic area, but only 14 percent of the population. A sample of 60 and 28 NFAs is selected systematically

---

[5] For more information about TrueNorth, see http://amerispeak.norc.org/our-capabilities/Pages/TrueNorth.aspx.

from the urban and rural stratum, respectively, with selection probability proportional to size (PPS) where the measure of size is the number of housing units per NFA.

Within the 126 selected NFAs, the secondary sampling units (SSUs) are segments defined from Census tracts or block groups, where each segment contains at least 300 housing units according to the 2010 Census. Within the certainty NFAs, a sample of 896 segments was selected using systematic PPS sampling, where the size of a segment is the number of housing units. Implicit stratification was achieved by sorting the segments by location (NFA, state, and county), by principal city indicator, by ethnic and income indicators. From each urban and rural NFA, a sample of 8 and 5 segments was selected, respectively, using systematic PPS sampling where the measure of size is the number of housing units per segment. A total of 618 segments are selected from the non-certainty NFAs[6]. Overall, a stratified probability sample of 1,514 segments was selected into the National Frame in the second stage sampling.

Within the selected segments, all housing units are listed using the U.S. Postal Service Delivery Sequence File (DSF). In the 123 segments where the DSF coverage is deemed inadequate, the DSF address list is enhanced with in-person field listing to improve coverage. The final National Frame, consisting of all listed households in the sample segments, is estimated to provide over 97 percent coverage of the U.S. household population. It contains almost 3 million households, including over 80,000 rural households that are added through in-person listing. In addition to NORC's National Frame, the DSF is used as a supplemental sample frame in four states. Although nationally representative, the National Frame does not include households from Alaska, Iowa, North Dakota, and Wyoming. Since 2016, the annual panel recruitment sample has included a small address-based sample (ABS) from these four states to assure AmeriSpeak presence in in all U.S. States and Washington, D.C.

In 2017, an enhanced DSF frame was also used to develop a new Latino Panel with adequate representation of Spanish-language-dominant Hispanics. Census tracts with high incidence (at least 30%) of Spanish-dominant Hispanics were targeted for this recruitment. Furthermore, within these Census tracts, households that were flagged as Hispanic based on consumer vendor data (that are typically used for direct-mail marketing) were oversampled. This new AmeriSpeak Latino Panel contains approximately 5,400 Hispanic panelists with 24% of those panelists being Spanish-language dominant. As of February 2020, 63.4% of the recruited adults in the AmeriSpeak Panel were sourced from the National Frame, 27.1% from targeted ABS frames, and 9.5% from voter registration files[7]. Proper weights allow the full use of the combined sample.

Panel Sample Selection

For panel sample selection between 2014 and 2018, National Frame segments are stratified into six sampling strata based on the race/ethnicity and age composition of each segment, as below:

- Hispanic, high youth segments
- Hispanic, not high youth segments
- Non-Hispanic Black, high youth segments
- Non-Hispanic Black, not high youth segments
- Other, high youth segments
- Other, not high youth segments

Hispanic segments are those where Hispanics make up at least a third of the population and the Hispanic share in the population is greater than that of non-Hispanic Black. Similarly, non-Hispanic Black segments are

---

[6] A sample of 5 segments was selected from each of the 28 non-urban NFAs. However, 2 sample segments were later subsampled out in Montana due to cost.

[7] Voter registration files were used as supplemental sample source for the AmeriSpeak Young Adult Panel.

those where non-Hispanic Black make up at least a third of the population and the non-Hispanic Black share in the population is greater than that of Hispanic. Finally, High Youth refers to segments in which 18-24 year old adults are at least 12% of the total adult population. The above stratification is used to oversample housing units in areas with higher concentration of young adults, Hispanics, and non-Hispanic African-Americans. The resulting household sample is referred to as the initial AmeriSpeak sample or sample for initial panel recruitment.

To support the second stage of panel recruitment, initially sampled but nonresponding housing units are subsampled for a nonresponse follow-up (NRFU)[8]. At this stage, consumer vendor data are matched to the pending housing units, and housing units that are flagged as having a young adult[9] (18-34 years of age) or minority (Hispanic[10], non-Hispanic Black[11]) are oversampled for the NRFU sample. Overall, approximately one in five initially nonresponding housing units are subsampled for NRFU using the same six sampling strata defined above. Due to NRFU, these initially nonresponding housing units have a much higher selection probability compared to the housing units that were recruited during the first stage of panel recruitment.

A 2-phase state-based ABS sample design was used for the 2019 AmeriSpeak recruitment. NORC's National Frame is designed to represent the U.S. household population nationally. At the state level, however, the panel may have more significant clustering effects from the use of the National Frame, especially for states with a small population. The primary objective of the 2019 design is to improve state-level representation by selecting the recruitment sample mostly from areas that are outside the National Frame. A stratified systematic sample was selected in the first phase, where each state constitutes a sampling stratum and the sample was allocated to the strata proportional to the square root of the state population. In the second phase, young adults, Hispanic, non-Hispanic Black, and conservatives are oversampled based on commercial data sources to improve their representation in the panel. Because the 2019 design did not use NRFU face-to-face recruitment, the 2019 design did not involve geographic clustering.

The overall AmeriSpeak Panel sample design reflects the cumulative design features of the NORC National Frame and the annual recruitment samples thus far. These design features are captured in the final panel weight for each recruited household and each panelist. There are mainly two reasons why the sampling design for AmeriSpeak Panel recruitment deviates from EPSEM (Equal Probability of Selection Method) sampling: (a) oversampling of housing units in segments with a higher concentration of young adults and minorities results in the sample selection probabilities being higher for housing units in these segments; and (b) NRFU subsampling results in initially nonresponding housing units having a much higher selection probability. Furthermore, within the NRFU samples, selection probabilities vary for housing units depending on the appended commercial data flags to target specific demographics for improved efficiency. The initial and NRFU sampling procedures are examined and possibly modified each year to more efficiently recruit subpopulations who are less likely to respond to the recruitment survey.

Panel Recruitment Procedures

The highly efficient recruitment procedures set AmeriSpeak apart from other national panels. AmeriSpeak Panel recruitment is a two-stage process: (i) an initial recruitment using USPS mailings, telephone contact, and

---

[8] A small fraction of initially nonresponding housing units are not eligible for NRFU, including "hard refusals" and those with an appointment for a call back from NORC.

[9] A young adult flagged household refers to a household where MSG or TargetSmart indicated there was an 18-24 year old adult in the household. In 2016 and 2017, a slightly different definition was used, and a young adult flagged household was defined as having an 18-34 year old adult in the household by MSG or 18-30 year old adult by TargetSmart.

[10] A Hispanic flagged household refers to a household where MSG or TargetSmart indicated the presence of a Hispanic adult in the household.

[11] A non-Hispanic Black flagged household refers to a household where MSG or TargetSmart indicated the presence of a non-Hispanic Black adult in the household.

modest incentives, and (ii) a more elaborate NRFU recruitment using FedEx mailings, enhanced incentives, and in-person visits by NORC field interviewers.

For the initial recruitment, sample households are invited to join AmeriSpeak online by visiting the panel website AmeriSpeak.org or by calling a toll-free telephone line (in-bound/outbound supported). Both English and Spanish languages are supported for online and telephone recruitment. The initial recruitment data collection protocol features the following: 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 contact by NORC's telephone research center for sample units with a matched telephone number.

For the second-stage NRFU recruitment, a stratified random sample is selected from the nonrespondents of the initial recruitment. Units sampled for NRFU are sent a new recruitment package by Federal Express with an enhanced incentive offer. Meanwhile, NORC field interviewers make personal, face-to-face visits to the pending cases to encourage participation. Once the households are located, the field interviewers administer the recruitment survey in-person using CAPI or else encourage the respondents to register online or by telephone.

Panel Recruitment Response Rate and Other Panel Statistics

A sample household is considered recruited or responded if at least one adult in the household joins the panel. The weighted household response rate (AAPOR RR3) is about 6% for initial recruitment and 28% for NRFU recruitment. We report two recruitment response rates: one for all the panel recruitment years (2014-2019) and one for the recruitment years with NRFU (2014-2018). For all recruitment years, the cumulative weighted household response rate is 24.1%; for recruitment years with NRFU, and the cumulative weighted household response rate is 34.2%. For client studies requiring a panel recruitment response rate exceeding 30%, the sampling frame may be restricted to the panelists recruited in the NRFU years. The panel recruitment response rate calculation methodology is consistent with AAPOR guidelines and fully documented.[12] The annual panel retention rate is about 85%.

For individual client surveys based on the AmeriSpeak Panel, the AAPOR RR3 response rate is between 10% to 20% depending on specific study parameters such as target population, survey length, time in the field, salience of subject, and the like. This response rate takes into account panel recruitment rate, panel retention rate, and survey participation rate.[13]

Other important panel statistics with respect to the 2014-2019 recruited households are as follows: 62% are recruited in the initial stage and 38% are recruited via NRFU; 85% of the active panelists prefer to do web or online surveys, while 15% prefer to participate in telephone surveys; 21% of the recruited households are

---

[12] See http://amerispeak.norc.org/research/Pages/WhitePaper_ResponseRateCalculation_AmeriSpeak_2016.pdf
[13] A properly calculated cumulative AAPOR response rate for panel-based research takes into account all sources of non-response at each stage of the panel recruitment, management, and survey administration process (see https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf, page 48-9). A common misapplication of the term "response rate" in online panel surveys is to represent the survey-specific cooperation rate as the "cumulative survey response rate." *See* "Response Rate Calculation Methodology for Recruitment of a Two-Phase Probability-Based Panel: The Case of AmeriSpeak" authored by Robert Montgomery, J. Michael Dennis, N. Ganesh. The paper is available at https://amerispeak.norc.org/research/.

non-Internet[14]; 82% are cell-phone only or cell-phone mostly; 19% are African-American and 20% Hispanic; and 33% have household income below $30,000 (compared to CPS benchmark of 26%).[15]

Impact of Non-Response Follow-Up

NRFU is instrumental in producing the industry-leading response rate for AmeriSpeak Panel recruitment. Moreover, due to the more intensive effort, NRFU recruitments better represent hard to reach groups and are therefore more representative of the target population. For example, initial recruitment tends to under-represent young adults 18-34 years of age. NRFU recruitment corrects for this bias by bringing the age distribution of the panel closer to ACS benchmarks.

Overall, NRFU recruitment significantly improves the representation of the panel with respect to demographic segments that are under-represented among the respondents to the initial recruitment, including young adults (persons 18 to 34 years of age), African Americans, Hispanics, lower income households, renters, cell-phone only households, and persons with lower educational attainment (e.g., no college degree). To the extent that these demographic characteristics are correlated with substantive survey variables, NRFU helps to reduce potential non-response bias in the sample estimates. NORC's research indicates that NRFU respondents are indeed somewhat different from initial respondents for many common survey variables. For example, compared to the panelists recruited during the initial stage, NRFU panelists tend to be more conservative politically, more likely to attend church, less interested in current events or topics in the news report, less knowledgeable about science, less likely to be in favor of gun control policies, less likely to read a print newspaper (more likely to read the news online and use social media), more likely to eat at fast food restaurants and so on[16]. These observations illustrate that NRFU recruitment is critical for achieving a more balanced panel and for making the substantive estimates in AmeriSpeak studies more accurate. Even though NRFU panelists are more reluctant to complete surveys, the addition of NRFU panelists reduced total absolute bias on average 5 to21 percentage points when compared to the initial stage recruits (among examined surveys).[17]

Mixed-Mode Data Collection

The AmeriSpeak Panel supports mixed-mode data collection to improve response rate and the representativeness of the complete surveys. During the recruitment survey, AmeriSpeak panelists are offered an opportunity to choose their preferred mode—web or phone—for future participation in AmeriSpeak surveys. A recruited household can consist of both web- and phone-mode panelists residing in the same household. Panelists predominantly prefer web over phone mode. As of February 2020, 85% of the active panelists prefer to do web or online surveys, while 15% prefer to participate in telephone surveys. The telephone mode encompasses panelists without internet access, panelists whose only internet access is via a smartphone, and panelists with internet access but are unwilling to share an email address.

---

[14] The non-internet households (HHs) are those that do not select "High-speed, broadband internet at home (such as cable or DSL)" or "Dial-up internet at home" response options when they are asked "What kind of internet access do you have? Please select all that apply" item in the recruitment survey. The non-internet HHs include those that only use internet on a cell connection or mobile phone.

[15] For transparency purposes, unweighted percentages are presented in this section. Hence, these results do not take into account selection probabilities. The base weighted distributions that take into account selection probabilities can be provided upon request.

[16] *See* "The Undercounted: Measuring the Impact of 'Nonresponse Follow-up' on Research Data and Outcome Measures" authored by Ipek Bilgen, J. Michael Dennis, N. Ganesh. The paper will be soon available at https://amerispeak.norc.org/research/.

[17] *See* "Nonresponse Follow-up Impact on AmeriSpeak Panel Sample Composition and Representativeness" authored by Ipek Bilgen, J. Michael Dennis, N. Ganesh. The paper is available at https://amerispeak.norc.org/research/.

To the extent that non-internet households or "net averse" persons are different from the rest of the population, mixed-mode surveys have better population coverage and produce more accurate population estimates. NORC's telephone interviewers administer the telephone surveys using a data collection system supporting both the phone and web modes, providing an integrated sample management and data collection platform. For panelists using smartphones for web-mode surveys, the NORC survey system renders an optimized presentation of the survey questions for these mobile users.

Panel Management and Maintenance

Panel management and maintenance are crucial for panel health and efficiency. NORC maintains strict panel management rules to limit respondent burden, reduce panel attrition, and minimize the risk of panel fatigue. On average, AmeriSpeak panelists are invited to participate in client studies two to three times a month. AmeriSpeak works with NORC clients to create surveys that provide an appropriate user experience for AmeriSpeak panelists. AmeriSpeak will not field surveys that in our professional judgment will result in a poor user experience for our panelists. AmeriSpeak also has a designated website and a telephone number for panelist communications.

Panel maintenance is a dynamic process because the AmeriSpeak Panel is supplemented and refreshed regularly over time to grow the panel, compensate for panel attrition, and improve panel representation for specific subpopulations. For example, the Latino Panel and Teen Panel are created to support studies of Hispanics and teenagers, respectively; the 2019 recruitment is primarily designed to improve sample representation at the state level. As panelists are added or/and removed from the panel, the panel refreshment process takes place to ensure that the refreshed panel fully represents the corresponding target population.

ABOUT NORC AT THE UNIVERSITY OF CHICAGO

As one of the world's foremost independent research institutions, NORC at the University of Chicago delivers objective data and meaningful analysis to help decision-makers and leading organizations make informed choices and identify new opportunities. Since 1941, NORC has applied sophisticated methods and tools, innovative and cost-effective solutions, and the highest standards of scientific integrity and quality to conduct and advance research on critical issues. Today, NORC expands on this tradition by partnering with government, business, and nonprofit clients to create deep insight across a broad range of topics and to disseminate useful knowledge throughout society.

Headquartered in downtown Chicago, NORC works in over 40 countries around the world, with additional offices on the University of Chicago campus, the DC metro area, Atlanta, Boston, and San Francisco.

ADDITIONAL RESOURCES

To learn more about AmeriSpeak or to share an RFP, please contact AmeriSpeak at AmeriSpeak-BD@norc.org. Information about AmeriSpeak capabilities and research papers are available online at AmeriSpeak.NORC.org.