

Methodological (Revised) Report for RICS-IVR 2020 U.S National Police-Race Survey

Paul J. Lavrakas, PhD.

July 3, 2020



Scott Richards, CEO
Scott@ReconnectResearch.com
(310) 613-8686

**APPENDIX A
METHODOLOGICAL DETAILS**

Sponsors and Data Collection Organizations: The sponsor for the survey was Reconnect Research (of Los Angeles, CA) and the Institute for Policy & Opinion Research (IPOR), Roanoke College (of Salem, VA). Survox by Enghouse Interactive (of Phoenix, AZ), conducted the IVR data collection using sample provided by Reconnect Research.

Research Partnership: Reconnect Research developed the questionnaire, with IPOR’s review. Reconnect Research provided the RICS sample for this survey and supplied IPOR with a weighted dataset. IPOR and Reconnect Research analyzed some of the findings independently of each other.

Questionnaire Wording: The IVR questions and response options and the order of the questions are shown below.

The introductory IVR script is shown here:

Please answer our urgent Police and Race Survey, conducted by Reconnect Research and Roanoke College. Para español , presse número sinco. Your call could not be completed and was sent to this survey.
If you’re in a risky or life-threatening situation press 9, otherwise don’t miss this opportunity to have your voice heard! Let's begin!
If you’re 18 or older, press 1. If 17 or younger, press 5. To learn more, press any other key.
Thank you. This five-minute survey is voluntary, and your answers are confidential. Please listen to the entire question and answers before responding. For any question, if you would prefer not to answer, press the star key.
Reconnect Research and the Roanoke College survey center, respected survey research organizations, are conducting an important national Police and Race survey and we would greatly appreciate your feedback.
Reconnect Research and Roanoke College rely on good people like you to participate in this urgent confidential Police and Race Survey.

The wording of the Police-related and Race-related questions was as follows – i.e., this is the order in which the questions were asked. Some questions, as noted below, had the order of the response options reversed, so that each order was randomly assigned to any given respondent.

Do you approve or disapprove of the job Donald Trump is doing as president? If approve, press 1. If disapprove, press 2 (Also asked with response options reversed)
Would you say things in the country are going in the right direction or the wrong direction? If the right direction, press 1. If the wrong direction, press 2 (Also asked with response options reversed)
Are you satisfied or dissatisfied with the job your local police department does? If satisfied, press 1. If neither satisfied or dissatisfied, press 2. If dissatisfied, press 3. (Also asked with response options reversed)
If that’s somewhat satisfied/dissatisfied, press 1. If that’s very satisfied/dissatisfied, press 2
Have you or an immediate family member ever been, or felt that you were being, harassed by the police? If yes, press 5. If no, press 6.

Have you or an immediate family member ever had an experience where a police officer helped keep you safe in a potentially dangerous situation? If yes, press 5. If no, press 6.
Some people think there are not enough police officers in their neighborhood, while others think there are too many. Thinking about your neighborhood: If you think there are not enough police officers, press 1. If too many, press 2. If the right amount, press 3. If you are uncertain, press 4. (Also asked with response options reversed)
Do you think that racial and ethnic discrimination in the United States is a problem or is it not a problem? If a problem, press 1. If not a problem, press 2 (Also asked with response options reversed)
Is it a big problem or a small problem? If a big problem, press 1. If a small problem, press 2. (Also asked with response options reversed)
Since Donald Trump became president, have race relations in the United States become better, worse, or has there been no change? If better, press 1. If worse, press 2. If there has been no change, press 3. (Also asked with response options reversed)
In a difficult or dangerous situation, are police officers more likely to use excessive force if the person is black, or are they just as likely to use excessive force against black and white people? If police officers are more likely to use excessive force if the person is black, press 1. If police officers are just as likely to use excessive force if the person is black or white, press 2. (Also asked with response options reversed)
Have you heard about the protests across the country in reaction to a recent incident where a black man named George Floyd in Minneapolis, died when a police officer knelt on his neck? If yes, press 5. If no, press 6. [Skip to demographics if "No"]
Do you think the anger that led to the protests across the country was fully justified, partially justified, or not at all justified? If fully justified, press 1. If partially justified, press 2. If not at all justified, press 3. (Also asked with response options reversed)
Do you think that recent the looting of businesses was fully justified, partially justified, or not at all justified? If fully justified, press 1. If partially justified, press 2. If not at all justified, press 3. (Also asked with response options reversed)
Since the protests started, have the words and actions of President Trump made the situation, worse, better or had no effect? If worse, press 1. If better, press 2. If no effect, press 3. (Also asked with response options reversed)

The wording of Demographic/Background questions was as follows, and this is the order in which these questions appeared:

Thanks. For statistical purposes only. If you are Male, press 1. if Female press 2.
Did you go to college, tech school or vocational school? For Yes press 5. For No press 6.
Did you graduate from high school? For Yes press 5. For No press 6.
Did you receive a bachelor's degree or higher? Prompt: For Yes press 5. For No press 6.
How old are you? If you are, 18 to 24 Press 1. 25 to 34 press 2. 35 to 49 press 3. 50 to 64 press 4. 65 or older press 5.
Are you of Hispanic, Latino or Spanish origin? For Yes press 5. For No press 6.
Which of the following best represents your race? For White press 1. Black press 2. Asian press 3. Another race or more than one race press 4.
In politics today, do you consider yourself a Republican, Democrat, or Independent? If Republican, press 1. If Democrat, press 2. If Independent, press 3. (Also asked with Democrat and Republican reversed)

To reduce the likelihood of nonignorable primacy effects in the final data, and as noted above, most substantive questions had their response options spoken in a reverse order a random half of the time.

Also, an unfolding approach to measuring educational attainment, the satisfaction with local police question, and the racial/ethnic discrimination question was used. This was done because past R&D by Reconnect Research has shown that this unfolding approach gathers much more accurate data when using IVR than using a single item for each construct.

Target Population: National sample of the U.S., including Alaska and Hawaii. English-speaking adults and Spanish-speaking adults aged 18 years or older, who are using the telephone to place a call (see Levine et al., 2019, in Footnote 1, for more about how Redirected Inbound Call Sampling – aka RICS -- works)¹. Data were gathered throughout the 24-hour day, thus during all daytime and nighttime hours.

Date Collection Field Period Dates: June 18-23, 2020.

Sampling Frame: The sampling frame for this survey is a set of phone calls (and the phone numbers from which the calls were placed and the persons who made them) in the U.S., including Alaska and Hawaii, that were placed during the days and nights of the field period, in which the call did not reach its intended party and was *by happenstance* redirected by the telephone service company handling the call to Redirect Research. This generates a nonprobability sample of the U.S. (see Levine et al., 2019). People not covered by this frame are those who did not place a call during the field period and those who placed a call, but their call reached the party they intended. People in the U.S. who placed a call during this time period that was handled by a telecom company other than the telecom company vendors contracted by Reconnect Research also are not covered by this frame. Thus, the vast majority of telephone calls placed in the U.S. during the field period were not covered by this frame.

Sample Supplier: As described in Levine et al. (2019), when someone dials a nonworking number, rather than playing an “error” message and ending the call, the caller in RICS is redirected by the telecom company that is handling the call to a survey recruitment and data collection system. (Reconnect Research contracts with various telecom companies for any given survey that it conducts to get its sampled calls.) There are two types of phone numbers that come to Reconnect Research: direct inward dialing (DID), also known as toll calls, and toll-free (TFN) calls. The method for Reconnect Research to obtain calls to redirect is different for each type of telephone number. Redirecting DID calls requires forming partnerships with local telephone companies. In contrast, calls to nonworking TFNs are provided by telecommunication companies that are the customer of record for large banks of toll-free numbers. Redirecting TFN calls to nonworking numbers to a data collection system requires access to calls from large banks of nonworking TFNs.

Recruitment of Respondents: The introductory IVR script that was used to screen, and recruit respondents is shown earlier in this appendix under the Questionnaire Wording section.

¹ Levine, B., Krotki, K., and Lavrakas, P.J. (2019). Redirected Inbound Telephone Call Sampling (RICS): Telephone Surveying via a New Sampling Paradigm. *Public Opinion Quarterly*, nfz024, <https://doi.org/10.1093/poq/nfz024>

45,853 redirected calls were used for this survey. Of these, 4,391 started the age-screening question or 9.6%. 389 callers indicated that they were not at least 18 years of age and contact with them was terminated. Of the 45,853 calls, 2,886 started the questionnaire or 6.3%. Of the 45,853 calls, 1,959 answered the final substantive question of the survey or 4.3% (which is the very conservative AAPOR RR1 response rate for this survey).

Sampling Design: As described under the Sample Supplier section above, calls that could not reach the party the caller intended are sent throughout the 24-hour day and night to Reconnect Research by its telecom company vendors. For the current surveying, Reconnect Research connected these redirected calls to Survox's IVR system to start the recruitment and data collection stages. There was no other formal sampling design other than whatever numbers came into Reconnect Research by chance from its telecom vendors.

RICS produces a nonprobability sample, which is remarkably heterogenous in how well it represents the U.S. national population. As Levine et al. (2019) note, RICS unweighted final samples often have a considerably larger portion of low-educational attainment and minority respondents than DFRDD in the unweighted final samples. On other demographic characteristics, RICS unweighted final samples match population parameters, including geographic location, at least fairly well.

Data Collection Mode: Data were gathered via a self-administered Interactive Voice Response (IVR) methodology in English and in Spanish. Two human males (one for English and one for Spanish) recorded the scripts used for the survey.

Sample sizes. A total of 2,886 respondents started the questionnaire. As the questionnaire proceeded, there were breakoffs who did not complete the questionnaire, which is a common occurrence with IVR data collection, compared to data collection performed by interviewers. There was a total of 1,959 completed Wave 1 questionnaires through the final substantive question. 149 of the 1,959 completions were with respondents in Spanish. This unweighted proportion (7.6%) is very close to the portion of U.S. adults who are estimated by Nielsen to be Spanish-dominant.

Of note, a decision was made by IPOR to limit analyses (and thereby to report findings) on only a subset of all those who started the questionnaire. This subset was comprised of respondents who answered at least half of the substantive questions without reporting a "Don't Know" as their answer or without refusing to answer. This totaled 1,917 respondents.

Given that these are nonprobability samples, margins of sampling error (MOSE) are not reported or used in the analyses. However, for those who are curious, the MOSE for a simple random probability sample of 1,917 adults, for a binary variable distributed 50/50, and with a 95% level of confidence, is ± 2.2 pp.

Weighting. Weighting was carried out by using the demographic/background variables that were gathered in the questionnaire and comparing them to their population parameters for adults in the U.S. It also used a geographic area indicator associated with the caller's phone number. With two exceptions, the source of the population parameters used in this weighting was the American Community Survey carried out by the U.S. Bureau of the Census.

The variables that were used in weighting were geographic location, gender, education, age, race, Hispanicity, language of the respondent, and political party affiliation. Party affiliation was used because it was reasoned that many of the questions would be correlated with basic political attitudes and that this nonprobability sample would benefit from such a weighting correction. The population parameter for party affiliation was based on a combination of national surveys estimates of this parameter, in which the Gallup political party affiliation tracking question (<https://news.gallup.com/poll/15370/party-affiliation.aspx>) and the Pew political party affiliation tracking question (<https://www.pewresearch.org/wp-content/uploads/2020/03/conflicts-topline.pdf>) figured prominently.

For the parameter of the portion of U.S. adults who are Spanish-dominant speakers, a value (8%) was used. This value comes from unpublished studies that Nielsen carried out in the prior decade. Those very large national surveys (n > 10K) used area-probability samples, along with multimode recruitment and data collection (mail, phone, and in-person) and were conducted by Nielsen in partnership with RTI. They achieved an AAPOR RR1 of greater than 85%.

An iterative process was used to bring the distribution of the variables used in weighting within +/- 3 pp of their population parameters.

The design effect (*deff*) associated with the weighting was 1.24.

A Total Survey Error Perspective on the Findings

All surveys suffer from error (be that bias and/or variance). Whether that error is ignorable or nonignorable is subject to differing opinion among the consumers of a given survey. Thus, this survey likely has coverage error, nonresponse error, and measurement error associated with the questionnaire and with respondents. This survey also has sampling error, as all sample surveys do.

Coverage Error. It is possible that the callers who were sampled for this survey were somewhat more likely to be people who had a greater concern about the topics being surveyed. This possibility is linked to the inherent nature of RICS to reach more minorities and lower educated adults than what actually makes up the general population of the USA. If this is true, then the findings, even after the weighting was applied, may somewhat over-estimate the concerns of the general population towards issues of police violence and race-related inequities.

Unit Nonresponse Error. Those sampled callers who were more likely to be concerned about police violence and race-related inequities may have been more likely to decide to start and to complete the questionnaire (without breaking off mid-stream). If this is true, then the findings, even after the weighting was applied, may somewhat over-estimate the concerns of the general population towards issues of police violence and race-related inequities.

Respondent-related Measurement Error. Respondents who provide data via IVR are known on occasion to exhibit primacy effects, whereby they are more likely to choose answers which they hear at the start of a list of response options than those answers they hear at the end of the list. That is why

questions that had more than two response options in our survey, and even some with only two response options, used response options that were asked in one of two random orders. Analyses show that there were some statistically significant primacy effects for among these questions. Thus, the value of administering the order of response options in both ways (forward and reverse) and then reporting data that include both sets of response orders. Whether this approach eliminated the primacy effects is unknown, but it is very likely that it appreciably reduced their sizes and thereby raised the accuracy level of these data.

Furthermore, as the questionnaire progressed, there was an increased proportion of item-nonresponse. No imputation was used to replace these missing data. Thus, if there was differential item-nonresponse, whereby certain types of people were more likely to break-off than others, then that could affect the reliability (i.e., representativeness) of the data that were produced later in the questionnaire by those who did not break-off. Of note, an investigation into this phenomenon suggests no meaningful pattern of differential attrition by demographic factors.

Questionnaire-related Measurement Error. Our questions were worded in a way to try to conform as closely as possible with how such questions have been worded in the past by prominent surveys. That does not guarantee that the wording did not contribute bias or variance to the data that were gathered.

Mode-related Measurement Error. IVR has been consistently shown in the research literature to provide data that are less susceptible to social desirability bias.² Given that most of the substantive questions in this survey are on “sensitive” topics, it is very likely that the accuracy (i.e., less bias) of the data exceeds what would have been gathered via an interviewer-administered data collection mode.

Sampling Error. Nonprobability samples do not allow one to use the traditional means of calculating sampling error that is afforded when one has a probability sample. Nevertheless, experience shows that many consumers of survey reports apply their own “sampling error” logic in their mind to thinking about any sample, as though it were a probability sample. In the case of this survey, the size of such a mental “estimate” of sampling error would be approximately ± 2.5 pp (taking into account the *deff*).

Furthermore, the reader is reminded that any statistical tests that are done in bivariate and multivariate analyses assume that probability samples were used to gather the data, if the findings are to be generalized to a larger population of interest.

TSE Summary Perspective. Taking into account the arguably likely nature of coverage error, nonresponse error, and respondent-related measurement error, it seems more likely that this survey has somewhat over-represented the attitudes of those concerned in a negative way about the issues of police violence and related racial/ethnic inequities, than it under-represents such concerns.

² Kreuter, F., Presser, S. and Tourangeau, R. (2009). Social Desirability Bias in CATI, IVR, and Web Surveys: The Effects of Mode and Question Sensitivity. *Public Opinion Quarterly*, 72(5), Pages 847–865.