U.S. Department of Transportation

National Highway Traffic Safety Administration

NHTSA

DOT HS 813 242

January 2022

Detecting Change in Community Traffic Safety Attitudes

DISCLAIMER

Suggested APA Format Citation:

Fors Marsh Group (2022, January). *Detecting change in community traffic safety attitudes* (Report No. DOT HS 813 242). National Highway Traffic Safety Administration.

Technical Report Documentation Page

1. Report No. DOT HS 813 242	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Detecting Change in Community Traffic Safety Attitudes		5. Report Date January 2022
		6. Performing Organization Code
7. Author Fors Marsh Group, LLC		8. Performing Organization Report No.
9. Performing Organization Name and Address Fors Marsh Group, LLC 1010 N. Glebe Road, STE 510 Arlington, VA 22201		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. DTNH2217D00034 693JJ918F000187
12. Sponsoring Agency Name and Address National Highway Traffic Safety Administration Office of Behavioral Safety Research 1200 New Jersey Avenue SE Washington, DC 20590		13. Type of Report and Period Covered Final Report
		14. Sponsoring Agency Code NPD-320
15. Supplementary Notes The contracting office representatives on the project were Mary Byrd and Kelly Sheppard.		

16. Abstract

Historically, evaluations of community traffic safety programs have collected data using nonprobability intercept surveys rather than more rigorous probability surveys, or they used secondary data instead of primary data collection of any kind. The technique used to select a sample of respondents from the broader population can affect the interpretability of the results of an evaluation. For example, compared to rigorous probability sampling, nonprobability sampling methods make it more difficult to generalize findings to the broader population. However, it is not always feasible to use probability sampling when collecting data based on operational complexity, statistical complexity, cost, and timing constraints. This document may serve as a tool to build on basic statistics knowledge of community traffic safety program evaluators by highlighting sound practices for using nonprobability sampling methods and secondary data sources. This document presents two nonprobability sampling methodologies (opt-in online panel surveys with quota sampling and intercept surveys with quota sampling) with practical steps for increasing the rigor of community traffic safety program evaluations. This document also identifies and describes existing data sources that may be of use to local traffic safety officials looking to evaluate their road safety programs as well as limitations of these data sources and suggestions regarding their use.

17. Key Words traffic safety, research, attitudinal surveys, program evaluation			18. Distribution Statement This document This document is available to the public from the DOT, BTS, National Transportation Library, Repository & Open Science Access Portal, https://rosap.ntl.bts.gov.
19 Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21 No. of Pages 38	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Table of Contents

Executive Summary **1**
Plan 1: Opt-in Online Panel Survey With Quota Sampling **2**
Overview 2
Target Population and Study Area Size Considerations 3
Implementation of Opt-in Online Survey Panels 4
Sample Design 4
Survey Instrument Design 4
Survey Administration 5
Incentives 5
Survey Estimation 6
Comparison to Probability Sampling Methods 7
Plan 2: Intercept Survey With Quota Sampling **11**
Overview 11
Target Population and Study Area Size Considerations 11
Implementation 11
Sample Design 11
Survey Instrument Design 13
Interviewer Training 14
Survey Administration 15
Incentives 15
Survey Estimation 15
Comparison to Probability Sampling Methods 16
Secondary Data Sources for Program Evaluation **21**
Federal Data 21
Overview 21
Advantages 21
Limitations and Suggestions 21
Highway Safety Plans 22
Overview 22
Advantages 22
Limitations and Suggestions 22

Vision Zero Action Plans 23
Overview 23
Advantages 23
Limitations and Suggestions 23
Social Media 24
Barriers to Social Media Data Use 24
Local Data Sources 24
Conclusion 25
References 28
Appendix: Glossary of Terms 30
Sampling Concepts 30
Quality Evaluation Criteria 31
Survey Error Concepts 32
Weighting Concepts 32

Executive Summary

The mission of the National Highway Traffic Safety Administration is to save lives, prevent injuries, and reduce healthcare and economic costs due to motor vehicle crashes. One way NHTSA furthers this mission is by providing practical information to State and local officials who develop, implement, and evaluate traffic safety programs. Historically, evaluations of community traffic safety programs have collected data from samples of residents using nonprobability intercept surveys in lieu of more rigorous probability surveys, or they used secondary data instead of primary data collection of any kind. The method of sampling or the technique used to select a sample of respondents from the broader population can affect the interpretability of the results of an evaluation. For example, compared to rigorous probability sampling, nonprobability sampling methods make it more difficult to generalize findings from the sample to the broader population. However, it is not always feasible to use probability sampling when collecting data based on operational complexity, statistical complexity, cost, and timing constraints. Thus, many evaluators of community traffic safety programs could benefit from statistically sound practices for using nonprobability sampling methods and secondary data sources.

This document presents two nonprobability sampling methodology plans that provide practical steps and suggestions for increasing the rigor of community traffic safety program evaluation. These plans are opt-in online panel surveys with quota sampling and intercept surveys with quota sampling. This document also identifies and describes existing (secondary) data sources that may be of use to local traffic safety officials looking to evaluate their road safety programs as well as limitations of these data sources and suggestions regarding their use.

Each nonprobability methodology plan provides:

- a thorough description of the sampling method,
- a discussion of factors to consider when assessing the appropriateness of the approach,
- the steps involved in conducting a study using the sampling approach,
- a comparison of the nonprobability sampling method with address-based sampling (ABS), a probability-based approach,
- definitions of key terms, and
- a discussion of the evaluative criteria used to assess the sampling methodology.

In general, nonprobability sampling methods are more biased, less precise, and require more assumptions about the population from which data are collected. These methods should be avoided for use when the goal is to accurately generalize the results from survey respondents to a broader population. Alternatively, nonprobability sampling methods can be used when the use of probability sampling methods is prohibitive and when core objectives of the study are to lower cost, obtain responses faster, or target hard-to-reach populations. Readers of these plans should consider their available resources and contextual circumstances when assessing which of these sampling methods may meet their needs. The contents of this document do not have the force and effect of law and are not meant to bind the public in any way. This document is intended only to provide clarity to the public regarding existing requirements under the law or agency policies

Plan 1: Opt-in Online Panel Survey with Quota Sampling

This methodology plan presents practical steps and suggestions for increasing the rigor of community traffic safety program evaluation through an opt-in online panel survey with quota sampling. We begin with a thorough description of opt-in online panel surveys and quota sampling. Next, we present factors to consider (e.g., target population and study area size) when assessing the appropriateness of this approach for a specific study. We then break down the steps necessary to conduct an opt-in online panel survey with quota sampling (i.e., sample design, questionnaire, survey administration, incentives, survey estimation). Finally, we provide a comparison between opt-in online panel surveys with quota sampling and probability sampling methods. Definitions of key terms used throughout this plan can be found in the glossary of terms at the end of this report.

Overview

An opt-in online panel survey used in combination with quota sampling can be a useful non-probability approach for effectively targeting the general population or a subset of the general population while obtaining survey estimates of acceptable quality. A hypothetical survey assessing attitudes toward nighttime seat belt use among the general population will be referred to throughout this guide as an illustrative example for the implementation of an opt-in online panel survey with quota sampling. Opt-in online panel surveys, which are also referred to as volunteer web panel surveys, use methods such as web advertisements or email invitations to recruit panel members. Through these avenues, people self-select into the panel and participate in surveys later. Personal background information (e.g., demographic, behavioral, attitudinal) is commonly collected. This information can be used later to select sample respondents with specific attributes, which can help reduce recruiting effort and cost (Baker et al., 2010; Couper, 2000). For example, an opt-in online panel may collect information about parental or childcare status, which can be used to draw a sample for a study of car seat safety. If the attributes of interest (e.g., parental status) are not collected upon empanelment, they can be identified through screening questions that are generally asked at the beginning of the survey.

Quota sampling is a sampling technique that, by design, yields a sample of respondents with characteristics like those of the broader population of interest (Baker et al., 2013). The goal of quota sampling is to maintain the proportions of respondent characteristics (e.g., demographic) in the sample as they occur in the population. Mutually exclusive subgroups are first identified in the population for the development of quota cells. These subgroups are identified by certain demographic, behavioral, and/or attitudinal characteristics. For data collection, the maximum number of survey participants with these characteristics is predetermined through the establishment of how many participants are needed for each quota (quota cell size). Each quota cell corresponds with a unique set of respondent characteristics (e.g., college-educated females age 34 to 49). Once the desired number of participants is reached for a particular quota cell, data collection is discontinued for that cell and more effort can be expended on filling other quota cells that do not have a target number of completed surveys (Callegaro et al., 2014).

One approach to specifying quotas entails identifying a source of benchmarks and then specifying the number of interviews/surveys/respondents in each cell in a manner that will result in sample characteristics that resemble the population. Specifically, the population proportions of demographic characteristics (the joint population distribution) within the area is calculated based on external benchmarks from a source such as the American Community Survey (ACS) or the

Current Population Survey (CPS), and the number of survey participants in each quota cell can be determined by multiplying the total sample size by the calculated joint population distribution. For example, as shown in Table 1, suppose the desired total sample size for the example study is 1,000, and the population distribution of age group (16–35, 36–60, and 60+) by gender (male, female) is 15% (males 16–35), 19% (males 36–60), 16% (males 60+), 14% (females 16–35), 20% (females 36–60), and 16% (females 60+). The number of survey participants needed in this case would be 150 (i.e., 1,000 x 15%), 190 (i.e., 1,000 x 19%), 160 (i.e., 1,000 x 16%), 140 (i.e., 1,000 x 14%), 200 (i.e., 1,000 x 20%), and 160 (i.e., 1,000 x 16%) for the six age-by-gender quotas.

Table 1. Example Quota Cell Sizes (N)

		Age 16–35	**Age 36–60**	**Age 60+**	**Total**
Male	Sample size	150	190	160	500
	% of population	(15%)	(19%)	(16%)	(50%)
Female	Sample size	140	200	160	500
	% of population	(14%)	(20%)	(16%)	(50%)
Total	Sample size	290	390	320	1,000
	% of population	(29%)	(39%)	(32%)	(100%)

Note that this approach specifies quotas that are proportional to the population. This can result in accurate estimates overall but may lead to imprecise estimates for some subgroups. The alternative would be to specify larger quotas for subgroups where more precise estimates are needed, such as if you were particularly interested in older (60+) females, and then to correct for this oversampling through a weighting process (described later in the Survey Estimation section of this report).

Behavioral and attitudinal variables can also be used to design the quotas if they are available in the external benchmarks. The demographic, behavioral, or attitudinal characteristics used to design the quotas should be selected based on whether they are related to the survey outcomes. For example, sex and urban/rural dwelling status could be used in the quota design for our example survey if that information were available in both the panel and benchmark data, as both were found to be associated with seat belt use (Li & Pickrell, 2019; Beck et al., 2017).

When accurate benchmarks are not available for the target population, special considerations may be needed for designing quotas, as using quotas for a different population could introduce bias. For example, suppose that the target population is nighttime drivers, but benchmarks are only available for people of driving age, regardless of nighttime driving (as in Table 1). Then, setting quotas for nighttime drivers based on the Table 1 benchmarks could introduce bias, as nighttime drivers may have different characteristics than others of driving age with respect to age and gender.

Target Population and Study Area Size Considerations

Quota sampling can be incorporated into an opt-in online panel survey to make the sample characteristics closely resemble those of the target population and to improve how well the survey estimates values from the population.

Opt-in online panel surveys are appropriate for the general population, common subsets of the general population (e.g., teens, minorities), and some hard-to-reach populations (e.g., caregivers of children). They might be inappropriate or challenging to design for hard-to-reach populations with socially undesirable characteristics (i.e., people who are involved in illegal or stigmatized activities), as these people will tend to not provide such information when they are recruited into the panel. Opt-in online panel surveys are feasible and commercially available for States and likely feasible for large metropolitan areas, but they may not be appropriate when the population size of a local area is especially small. If there are not enough existing panel members for that area, more effort will be required to recruit additional panel members. However, the number of completed surveys might ultimately still be low. Feasibility for a survey in a local area can be further constricted due to additional screening criteria (e.g., surveying only teens or minorities within a small area), which may further limit the number of completed surveys within the target population.

Implementation of Opt-in Online Survey Panels

Sample Design

When the target population is a subset of the general population within a local area, the opt-in online panel should additionally be restricted to the eligible members if the requisite background information (e.g., geographic location) is available in the panel; otherwise, screening questions should be added in the questionnaire to ensure survey participants meet the eligibility criteria. If there are doubts about the availability of background information from the panel, then it may make sense to ask screener questions on key study eligibility criteria in the survey to ensure that the respondents are in the target population. In the case of the example study, age information available in the panel could be used to restrict the sample to either adults or people of driving age. If the population of interest includes additional eligibility criteria (in this example, frequency of nighttime driving) that are unavailable from the panel, then screening questions should be added to assess these criteria. Assuming that benchmarks are available for the target population (e.g., frequent nighttime drivers), then the quota design considerations and procedures described above remain the same.

Survey Instrument Design

There are several questionnaire design aspects to consider as panel surveys are conducted online (Dillman et al., 2014). Questions should display similarly across different devices (including mobile phones), platforms, browsers, and user settings. The surveys should be programmed for the lowest likely screen resolution so that the survey content appears in the intended way on the screen for most users and the need for scrolling is minimized. The questionnaire region should be center-aligned horizontally on the page so that the questions look similar and scrolling is minimized regardless of monitor width.

A hybrid design should be considered to balance the use of the scrolling design (i.e., multiple questions per page) and the page-by-page design (i.e., one question per page). A scrolling design is better for a long questionnaire and for grouping questions by topic. A page-by-page design is more accommodating of skip patterns and can reduce data loss due to overlooked items or unsaved answers. A hybrid design, for which the scrolling design and page-by-page design are used for different types of questions in one questionnaire, can reduce the response burden, improving survey completion rate and response quality.

Additional survey design considerations

- Create informative and appreciative welcome and closing screens that appeal to survey participants.
- Allow survey participants to navigate backwards to previous survey items and to stop the survey and complete it at later time.
- Do not include a progress indicator unless the questionnaire is very short.
- Consider providing "not applicable" and/or "don't know" options; also, consider the effect of their inclusion on respondent behavior. Do not require responses unless necessary for progressing through the survey.
- Ask screening questions up front and background (e.g., demographic) questions at the end. For demographics, consider if the necessary information is already available in the panel or whether there are any data quality concerns about the panel information.
- Avoid using abbreviations that participants may not know.
- Ensure accessibility and compliance with the American Disabilities Act.

Survey Administration

Potential survey participants can be identified based on the background information of the panel and contacted for a specific study screening. Contacts are usually made via emails that include information such as the study name, incentive information, time commitment, and a link to the screening questionnaire. The screening questionnaire should not take too long for a participant to complete (e.g., less than 5 minutes), and should contain both the general screening criteria (e.g., panel members who have previously participated in the same survey should be excluded) as well as study-specific screening criteria (e.g., frequency of nighttime driving if frequent nighttime drivers is the population of interest).

After completing the screening questionnaire, participants who qualify for the survey should complete the consent form before being directed to the main questionnaire for the study. Panel members who are eligible for the study and complete the consent form should be immediately redirected to the full survey. Surveys are self-administered and should be accessible at any time during the designated fielding period. Participants should be able to complete the survey only once. Quotas should be monitored, so that when the maximum number for a quota cell is reached, data collection is discontinued for that cell.

Incentives

Previous research has shown that prepaid cash incentives in address-based surveys are more effective for increasing response rate than promised (i.e., contingent) incentives or gifts (Singer & Ye, 2013). However, for an opt-in online panel survey, sending prepaid cash incentives is not practical. A more cost-effective way is to send the promised incentive after the sample member completes the survey. The promised incentives might include e-rewards currency or some other honorarium, such as prepaid Visa cards or charity donations. Opt-in online panels frequently use these incentives to recruit new panelists and have policies in place regarding the reimbursement of participants for completing surveys. Also note that while the use of incentives to collect data is common in the private sector, it is not something that the Federal Government does as a routine matter. The Federal Government cannot provide these types of rewards without first justifying their use as a necessary expense of collecting the survey data.

Survey Estimation

For those conducting analysis of survey data, survey weighting is a commonly employed mechanism that allows for the generalization of survey results from the survey sample to the population. In practical terms, weights are commonly employed in a survey data set through a weight variable that has a numerical value for each respondent indicating how many population members each respondent is thought to represent. These weights can then be used for estimating common survey statistics, such as means, proportions, and totals. For example, in a survey of the driving-age population, the proportion of people that drives frequently at night could be estimated as the total weights of frequent nighttime drivers divided by the total weights of all respondents.

In a probability-based survey context, survey weights are commonly employed to account for the study design and to correct for biases (Valliant et al., 2018). Although these concepts do not translate neatly to the nonprobability context, survey weights can still be used to correct for selection bias. A straightforward and common way of doing so is to compute calibration weights (e.g., poststratification or raking weights), which are created to ensure that the sample distribution of key characteristics matches that of the population benchmarks. These benchmarks are commonly estimated from a high-quality external data source, such as the ACS. In the context of quota sampling, recall the example used in the sample design section above (Table 1) and suppose that the quotas cannot be followed exactly. In particular, the final numbers of completed survey per group deviate from the expected 150, 190, 160, 140, 200, and 160 for the six age-by-gender quotas. The use of calibration weighting on the age-by-gender categories will ensure the survey-weighted proportions of age-by-gender are still 15%, 19%, 16%, 14%, 20%, and 16%, which will match the proportions in the population.

More importantly, there are many situations in which it will be impractical to specify quotas for all relevant dimensions simultaneously but where calibration weighting on a broader set of variables (or levels of variables) could further correct bias. Uncorrected bias can arise if there are important population characteristics that were not incorporated in the quota design (e.g., quotas correctly reflect the population's age and sex distributions, but highly educated people are overrepresented and have different behaviors of interest than others). The ideal adjustment variables used in calibration weighting are those that are highly correlated with survey variables of interest (Little & Vartivarian, 2005), which may include demographic, behavioral, and attitudinal characteristics. Note that quotas are commonly specified based on the joint distribution of relevant variables (e.g., fully cross-classifying age group by gender in Table 1), but calibration can accommodate a larger number of variables.

Calibration weighting should also account for the screening criteria, which may affect the choice or availability of appropriate external benchmarks used in weighting. If there are differences in survey eligibility that are not reflected in population benchmarks, then weighting an eligible sample (e.g., frequent nighttime drivers in the general population) to population benchmarks that include some ineligible people (e.g., non-drivers in the general population) would introduce bias in estimates. Ideally, population benchmarks used should reflect screening criteria. If such benchmarks are unavailable, then an indirect calibration weighting approach, which involves treating the target population as a domain (i.e., subpopulation) of a larger, well-defined population for which benchmarks are available, could be used. Then, by generating weights for the larger population, weights are indirectly generated for the target population.

Comparison to Probability Sampling Methods

Probability sampling is a general sampling approach whereby all members of the population have a known and non-zero probability of being selected into the survey. This allows results from a survey sample to be generalized to the larger population that the sample members are intended to represent under known mathematical measures of uncertainty. Address-based sampling (ABS), a sampling approach that involves the selection of a random sample of addresses from a list of postal addresses from the U.S. Postal Service, is one of the most commonly used probability sampling methods in practice, and allows for better geographic targeting in a local area compared to other probability sampling methods.

Compared to an opt-in online panel survey with quota sampling, ABS has better coverage of the target population, which reduces the bias generated from any population segments with certain characteristics missing in the sample and ensures that the survey results can be generalized to the target population. However, ABS is often much more complex in terms of operations and more expensive in implementation than an opt-in online panel survey with quota sampling.

The quality of a survey can be assessed in terms of its exchangeability, positivity, composition, and precision. The first three focus on the bias component of survey error and the last one focuses on the variance component of survey error (Mercer et al, 2017).

Exchangeability refers to whether there are any unobserved variables that are correlated with both the inclusion of population members in the sample and the outcomes being measured in the survey. For example, exchangeability could be an issue if people who do not wear seat belts also do not answer surveys for some unidentified or unexplained reason.

Positivity refers to whether any population segments with certain characteristics have no chance of being included in the sample. For example, people without an email address cannot participate in many on-line surveys.

Composition refers to whether the distribution of observed variables in the sample is similar to the distribution of observed variables in the population. Greater composition indicates a more representative sample and thus less potential bias in the estimates.

Precision is the inverse of the degree of uncertainty surrounding an estimate. That is, an estimate that has low sampling variance will have high precision.

Table 2. Quality of Opt-in Online Panel With Quota Sampling Versus Address-Based Sampling

	Opt-in Online Panel Survey With Quota Sampling	**Address-Based Sampling**
Exchangeability	Auxiliary variables available in the panel can be used in designing the quotas to attempt to adjust for bias if the variables correlate with both the inclusion of population members in the sample and the survey outcomes to reduce risk.	Random sampling breaks the association between unobserved variables and probability of inclusion. In addition, auxiliary variables appended to the address list can be used in the sample design to target the specific population of interest and balance the sample allocation among different subpopulations to reduce risk. However, ABS runs the risk of bias related to people who actually read and respond to mail solicitations.
Positivity	Opt-in online panel surveys are subject to bias due to the positivity issue, which cannot be easily adjusted. There are likely substantial parts of the population that are unaware of the panel, or are unable to participate (e.g., due to lack of internet access), and thus have no chance of being selected. The impact of the positivity depends on the extent to which people excluded from the sample may differ systematically from the overall population.	The extent of coverage error depends on the amount of under-coverage of the addresses and the extent of the differences between covered and non-covered addresses. Because the addresses are usually supplied by the USPS, ABS generally has good coverage.
Composition	Quota sampling and calibration weighting can be used to adjust for bias due to the composition issue under certain circumstances.	Calculation of design weights, nonresponse weighting, and calibration weighting can adjust for bias due to composition issues.
Precision	Assuming the estimates are approximately unbiased, the precision of an opt-in online panel survey depends on how large the overall sample size is, the extent to which the sample resembles the population, and how weighting is conducted.	The precision of ABS depends on factors such as how large the overall sample size is, whether a multistage stratified sampling design is implemented, and how weighting is conducted. Highly unequal probabilities of selection in ABS can reduce precision.

If there is concern about exchangeability, one can evaluate whether the auxiliary variables correlate with the survey outcomes. Regression models can be used to test the existence of such relationships in data from similar surveys (if such surveys exist). The relationship between auxiliary variables and inclusion of population members in the survey cannot be tested, as the information for the non-sampled cases is typically unknown for opt-in online panel surveys.

If there is a potential for positivity, the distribution of demographic and other characteristics for panel members and survey participants can be compared to the distribution of these characteristics from the external benchmarks to determine whether any portion of population that is of interest to the researcher is missing.

If composition is a concern, calibration weighting should be conducted if the actual number of completed surveys in each quota cell deviates from the number of completed surveys prescribed in the original quota design or if there are important population characteristics that were not incorporated in the quota design.

If there is concern about precision, one can compare the point estimates to the external benchmark first. If the difference is small, consider the standard error estimates and resulting confidence intervals for the estimates.

Table 3. Feasibility of Opt-in Online Panel With Quota Sampling and Address-Based Sampling

	Opt-in Online Panel Survey With Quota Sampling	Address-Based Sampling
Operational Complexity	Operational complexity is relatively low. Panel construction can be the most complex part as it takes effort to set everything up. Survey fielding and panel maintenance are fairly standard after panel is constructed.	Operational complexity is typically high. Considerable effort is required to create the address list; activities include out-of-scope address removal, address updates, deduplication, standardization, and the appending of auxiliary data. Complexity is further increased if face-to-face interviews are used as the data collection mode, as interviewer training and scheduling are needed.
Statistical Complexity	Statistical complexity varies. Quota sampling and calibration weighting are typically needed, although more effort (e.g., model-based estimations) can be expended to attempt to reduce bias and increase precision.	Statistical complexity is typically high. On the sampling side, the complexity depends on whether a multistage or single stage sampling design is taken. On the weighting side, weighting needs to be conducted to account for design weights, nonresponse weighting, and calibration weighting.
Data Collection Mode	Web is the primary data collection mode.	Face-to-face interviews or mail are the primary data collection modes.
Length of Fielding Period	Fielding is relatively fast. The actual time depends on how quickly the number of surveys needed to fulfill each quota cell can be fielded.	Fielding is typically slow, especially when waiting for responses to return through the mail.
Cost	Cost is generally low (unless a panel is being newly constructed). Factors that influence cost include overall sample size, effort expended to follow up with sample members, field the target number of completed surveys in each quota cell, and the value of incentives used. Furthermore, when the scope or the population size of a local area is too small, additional costs are necessary to recruit more panel members.	Cost varies. Factors that influence cost include overall sample size, how much effort is expended to follow up with non-respondents, and whether incentives are used. If ABS is used with web data collection, costs will typically be higher than an opt-in panel, but not dramatically so. Cost will be higher by orders of magnitude if face-to-face interviews (rather than mail or web) are used as the data collection mode.

Plan 2: Intercept Survey With Quota Sampling

This methodology plan presents practical steps and suggestions for increasing the rigor of community traffic safety program evaluation through an intercept survey with quota sampling. We begin with a thorough description of intercept surveys and quota sampling. Next, we present factors to consider (e.g., target population and study area size) when assessing the appropriateness of this approach for a specific study. We then break down the steps necessary to conduct an intercept survey with quota sampling (i.e., sample design, questionnaire design, interviewer training, survey administration, incentives, survey estimation). Finally, we provide a comparison between intercept surveys with quota sampling and probability sampling methods.

Overview

An intercept survey used in combination with quota sampling can be a useful nonprobability approach for effectively obtaining survey estimates of reasonably acceptable quality under certain circumstances.

Intercept surveys are one of the most commonly used convenience sampling approaches and are conducted by approaching potential participants for an interview as they pass by in a public place (i.e., "intercepting" them; Schonlau & Couper, 2017). Intercept surveys can be conducted in a variety of settings, such as shopping malls or, in the traffic-safety context, Departments of Motor Vehicles (DMV; Baker et al., 2013). For such efforts, systematic sampling is often employed by inviting every nth (e.g., 5th, 10th, 100th) person that passes by a specific spot in a mall or outside DMV office to participate. Alternatively, other random sampling methods could be applied to randomly select people to participate in the study.

Quota sampling can be incorporated into an intercept survey to make the sample characteristics closely resemble those of the target population (improve composition) and to improve the precision of the survey estimates.

Target Population and Study Area Size Considerations

Intercept surveys are appropriate for the general population as well as subsets of the general population (e.g., older adults). They might be inappropriate or challenging to design for hard-to-reach or stigmatized populations (e.g., caregivers of children or drug-impaired drivers, respectively), as the recruiting effort for such populations is high, which drives up the overall cost and length of fielding period. Intercept surveys are appropriate when the population size of a local area is small (e.g., cities or small counties), as they are more likely to cover the population of interest and can be more efficiently used to recruit the population members for the study (compared to a large geographical area or large population).

To illustrate the implementation of an intercept survey with quota sampling, a hypothetical study assessing the attitudes of older drivers about traffic safety attitudes will be used as an example.

Implementation

Sample Design

The first step of sampling for an intercept survey involves the selection of intercept locations within a local area. A wide variety of locations (e.g., shopping malls, gas stations, recreation parks) can be chosen to conduct an intercept survey, and bias can be generated if a small number of locations are selected, as the people who come from one intercept location may have similar characteristics to one another and these may differ from the characteristics of people drawn from other locations. A diverse set of locations should be selected directly and (ideally) randomly to obtain participants with varied characteristics, and a larger number of people can be selected for the locations with larger aggregation of population members. Alternatively, probability-proportional-to-size (PPS) sampling can be applied to select the locations if the population distribution (e.g., the proportion of the overall population represented by different DMV offices) is a primary consideration, after which the same number of people can be selected within each location for easier logistics. For a survey of traffic safety attitudes among older drivers, local DMVs could be chosen as the intercept locations, as the desired demographic can be easily targeted in these locations because in many States they have to take a vision test. Suppose two locations need to be selected among four local DMV offices, and the proportional distribution of older drivers among the four DMV offices is 10% (DMV location 1), 20% (DMV location 2), 30% (DMV location 3), and 40% (DMV location 4). By using a PPS sampling method, the two DMV offices would be selected from among the four using the following selection probabilities: DMV location 1 = 0.1; DMV location 2 = 0.2; DMV location 3 = 0.3, and DMV location 4 = 0.4. This approach ensures that DMV offices with a larger proportion of older drivers will have a higher chance of being selected. Note that this is a simple example of how PPS sampling is conducted, since you would ideally collect information from all four locations. In that situation, and if possible, one could use quota sampling to collect 10% of the desired total from location 1, 20% from location 2, etc. In practice, a fairly large number of locations should be selected from an even larger list of potential sites to reduce the bias generated due to homogeneity within the same location.

As indicated above, quota sampling can be applied to determine the number of survey participants for each local jurisdiction. The distribution of demographic characteristics (e.g., the population proportions for different gender-by-age groups) within a local jurisdiction is calculated based on external benchmarks, such as those provided by the ACS or the CPS, and the number of survey participants in each quota cell is determined by multiplying the sample size by the calculated joint population distribution. Each selected intercept location within a local jurisdiction can follow the same quota design within that local jurisdiction to reduce the study's operational complexity. For example, suppose the sample size for a particular DMV office is 1,000, and the population distribution of age group (65 to 74, 75 or older) by gender (male, female) within a local jurisdiction area served by that DMV is shown in Table 4. The number of survey participants needed in this case would be 240 (i.e., 1,000 x 24%), 260 (i.e., 1,000 x 26%), 230 (i.e., 1,000 x 23%), and 270 (i.e., 1,000 x 27%) for the four age-by-gender quotas.

Table 4. Example Quota Cell Sizes

		Age 65-74	**Age 75 or older**	**Total**
Male	Sample size	240	230	470
	% of population	(24%)	(23%)	(47%)
Female	Sample size	260	270	530
	% of population	(26%)	(27%)	(53%)
Total	Sample size	500	500	1,000
	% of population	(50%)	(50%)	(100%)

This method implicitly assumes that either: a) population distributions are similar across intercept locations; or, b) within the quota cells, the survey outcomes do not depend on the intercept location (e.g., traffic safety attitudes among males aged 75 or older are similar across DMVs). Alternatively, different quotas can be designed for each intercept location within a local jurisdiction to reduce the bias if an external benchmark is available at the intercept location level, although the benchmark is typically only available at the jurisdiction level, and different quotas among locations would add operational complexity.

The timing of survey collection windows should also be carefully considered and selected. In general, data collection periods should occur randomly across the locations and hours of operation (e.g., weekdays and weekends; mornings and afternoons). Also, avoid holidays and festivals that may attract large numbers of people from outside the community and thus complicate survey logistics.

Survey Instrument Design

The ideal questionnaire design for an intercept survey depends on whether the survey is administered by an interviewer or is self-administered by survey participants via paper, tablet, or laptop. Regardless of the mode used, several good practices apply:

- Provide a clear and persuasive introduction to appeal to potential participants and include an appreciation after the survey is conducted.
- Keep the survey short (i.e., 5 to 10 minutes to complete).
- If multiple modes are used, keep the question format and wording as similar as possible across the survey data collection codes.
- Consider providing "not applicable" and/or "don't know" options and the effect of their inclusion on respondent behavior. Do not require responses unless necessary for progressing through the survey.
- Ask screening questions up front and background (e.g., demographic) questions at the end.
- Avoid using abbreviations that participants may not know.

When feasible, administering an intercept survey on a tablet is typical as it is easier to carry compared to a laptop, provides for easier data entry compared to paper surveys, and costs less than surveys verbally administered by interviewers due to a reduced need for training. There are several questionnaire-design best practices and considerations that are specific to surveys conducted on a tablet:

- Questions should display similarly across different platforms and/or browsers that may be used.
- Because the survey should be short, employ a page-by-page design (i.e., one question per page). This is more accommodating of skip patterns and can reduce data loss due to overlooked items or unsaved answers.
- Assuming the survey is short, it may be beneficial to provide a progress indicator. If the survey is long, a progress indicator is not suggested.
- The questionnaire region should be center-aligned horizontally on the page so that the questions look similar and scrolling is minimized regardless of monitor width.
- Allow survey participants to navigate backwards to previous survey items and to stop the survey and complete it at later time.

In some cases, considerations such as the expense of acquiring tablets or participants' unwillingness to touch shared surfaces may make the use of tablets untenable.

There are also questionnaire design considerations unique to surveys that are administered via paper (Fanning, 2005):

- If skip patterns are needed, create visual navigation guides (e.g., arrows) and use them consistently.
- Determine any intuitive question groupings based on question content, type (e.g., matrix questions that use the same response scale), and logical order. Color and space can be used to make question groupings more readily apparent.

Finally, best practices for surveys administered by interviewers include:

- Break complex questions into a series of simpler questions.
- Avoid asking ranking questions, nominal questions with many response options, and questions with a check-all-that-apply format to reduce the response burden.
- Provide short and simple transition statements when reading questions.
    - For example, consider the question "How likely do you think it is that a person would receive a ticket for sending text messages while driving? Would you say very likely, somewhat likely, somewhat unlikely, or very unlikely?"; the phrases "would you say" and "or" are short but important transition statements.

Interviewer Training

The amount of interviewer training necessary depends on whether the intercept survey is administered by interviewers or self-administered by participants. If the survey is self-administered by participants (e.g., via a paper, tablet, or laptop), the training for the staff should include:

- Instructions on how to use the software or system if the survey is conducted on a tablet or laptop.
- Strategies for persuading people to participate in the survey without biasing the survey.
- Instructions on how to implement the sampling procedures.
  Pre-prepared answers for questions that survey participants may ask (e.g., about understanding of specific survey questions and wording) or issues they may encounter (e.g., trouble navigating the software).

If interviewers administer the survey, training should also include:

- Instructions on how to read the questions, as well as what information should or should not be read.
- Techniques for probing when answers that survey participants provide are unclear.
- Mock interviews so that the interviewers can understand how to control their pace and flow and how to manage the pitch and tone of their voice.

Finally, the principal investigators should occasionally monitor interviewers to adjust procedures if needed, address any difficulties, and provide timely feedback to interviewers.

Survey Administration

The interviewers at each intercept location should first identify and establish contact with (i.e., “intercept”) the potential survey participants. Interviewers should convey all information that people need to make an informed decision about whether to participate, including: the study’s name and purpose, incentive information, time commitment, and the voluntary nature of the research. Next, interviewers should ask screening questions to determine the study eligibility of potential participants. These questions may address both general screening criteria (e.g., people who have previously participated in the same survey or who are not from the community should be excluded) as well as study-specific screening criteria (e.g., being a driver 65 or older).

Surveys completed at a data collection site can be self-administered by participants (on paper or a tablet) or administered by an interviewer. If a survey is designed to be self-administered but a participant is uncomfortable or unable to complete the survey themselves on paper or on a tablet, the interviewer can read the questions to the participant and record their answers for them. Once a participant completes the survey, the interviewer should acknowledge the participant’s time and effort, and make note of any issues encountered. Quotas should be continually monitored at the local jurisdiction level so that when the maximum number for a quota cell is reached, data collection is discontinued for that cell. For example, if target sample size for male participants 75 or older is reached, additional 75-or-older male participants should be deemed ineligible to complete the survey.

Incentives

Because intercept surveys are typically administered immediately after the participant is invited to participate, either prepaid or promised (i.e., postpaid) incentives can be used. Previous research has shown that prepaid cash incentives are more effective for increasing response rate than promised incentives or gifts, although these standard findings may not apply to intercept surveys. It is not ideal to distribute prepaid incentives directly to people when they are intercepted. However, incentives can be handed out to people after they agree to participate but before they start to complete the survey. Incentives provided to participants might include a small amount of cash, free products, gift cards, or coupons. As mentioned previously, the use of incentives to collect data is common in the private sector, but it is not something that the Federal Government does as a routine matter. The Federal Government cannot provide these types of rewards without first justifying their use as a necessary expense of collecting the survey data.

Survey Estimation

For an intercept survey, calibration weighting should generally be conducted using the external population benchmarks at the jurisdiction level, although calibrations weights can be created for each intercept location if the population benchmarks are available at the intercept location level. The intercept locations can be treated as clusters when doing survey estimation, so that the correct variance estimates and confidence intervals can be obtained.

As previously mentioned, weighting often increases the variability of survey estimates without guaranteeing the removal of bias. An alternative is to use a model-based approach without applying any types of survey weights. For intercept surveys, statistical model-based approaches include:

- logistic regression if a survey outcome of interest is a proportion estimate,
- linear regression model if the survey outcome of interest is a total estimate, or
- multi-level models, since intercept locations and survey participants can be treated as two analysis levels in the model.

Because the sample drawn for an intercept survey is just a small fraction of the target population and the person-level auxiliary information is typically not available for the non-sampled cases, the population totals of auxiliary variables should be obtained (if available) as alternative auxiliary information to build the statistical models.

Comparison to Probability Sampling Methods

Compared to an intercept survey with quota sampling, ABS can append auxiliary variables that can be used in sample design. This reduces the bias generated from unobserved variables that are correlated with both the inclusion of population members and the outcome being measured (i.e., the exchangeability issue). ABS also has better coverage of the target population, which reduces the bias generated from any population segments with certain characteristics missing in the sample (i.e., the positivity issue), and ensures the survey results can be generalized to the target population. However, as discussed previously, ABS can be more complex in terms of operations and more expensive to implement than an intercept survey under certain circumstances. Table 5 and 6 provide comparisons between intercept surveys with quota sampling and ABS in terms of quality and feasibility. Table 5 summarizes the results of the quality assessment using the four concepts of exchangeability, positivity, composition, and precision.

Table 5. Quality of Intercept Survey With Quota Sampling Versus Address-Based Sampling

	Intercept Survey With Quota Sampling	**Address-Based Sampling**
Exchangeability	Intercept surveys are subject to bias due to the exchangeability issue. Auxiliary variables are typically not available up front and cannot be used in the sample design stage. Questions should be added to the survey instrument so that the relevant variables can be used in survey estimation stages to adjust for the potential bias.	Random sampling breaks the association between unobserved variables and probability of inclusion. In addition, auxiliary variables appended to the address list can be used in the sample design to target the specific population of interest and balance the sample allocation among different subpopulations, so that the risk of exchangeability can be reduced.
Positivity	Intercept surveys are subject to bias due to the positivity issue (which cannot be easily adjusted), as not all locations can be covered in practice. The impact of the positivity depends on the extent to which people excluded from the sample differ systematically from the overall population.	The extent of coverage error depends on the amount of under-coverage of the addresses and the extent of differences between covered and non-covered addresses. Because the addresses are usually supplied by the USPS, ABS generally has good coverage of a target population, especially if the population is not hard to reach.
Composition	People in one intercept location may be different from those in other intercept locations. Quota sampling and calibration weighting can be used to adjust bias due to the composition issue under certain circumstances.	Calculation of design weights, nonresponse weighting, and calibration weighting can adjust for bias due to composition issues.
Precision	It is difficult to determine. Assuming the estimates are approximately unbiased, the precision of an intercept survey depends on how large the overall sample size is, how many locations are selected, how the sample is allocated across locations, the homogeneity within each location, the extent to which the sample resembles the population, and how weighting is conducted.	The precision of ABS depends on factors such as how large the overall sample size is, whether a multistage stratified sampling design is implemented, and how weighting is conducted. Highly unequal probabilities of selection in ABS can reduce precision.

As described previously with opt-in surveys, regression models can be used to test the existence relationships between survey variables and auxiliary variables.

The distribution of demographic and other characteristics for people in intercept locations and survey participants can be compared to the distribution of these characteristics from the external benchmarks to determine whether any portion of the population is missing. Calibration weighting should be conducted if the actual number of completed surveys in each quota cell deviates from the number of completed surveys prescribed in the original quota design or if there are important population characteristics that were not incorporated in the quota design.

If possible, compare the point estimates to the external benchmark first. If the difference is small, look at the standard error estimates/confidence interval.

Table 6. Feasibility of Intercept Survey With Quota Sampling and Address-Based Sampling

	Intercept Survey With Quota Sampling	**Address-Based Sampling**
Operational Complexity	Operational complexity varies depending on how locations are selected, the data collection mode used, and the procedures used to implement the sample design and facilitate participation. Interviewer recruitment and training can be more complex and time-consuming than other operational steps (e.g., survey instrument design and survey fielding).	Operational complexity is typically high. Considerable effort is required to create the address list; activities include out-of-scope address removal, address updates, deduplication, standardization, and the appending of auxiliary data. Complexity is further increased if face-to-face interviews are used as the data collection mode, as interviewer training and scheduling are needed.
Statistical Complexity	Statistical complexity varies. On the sampling side, the complexity depends on what sampling method is taken to select intercept location and how the quota is designed to balance participants' characteristics. On the estimation side, calibration weighting can be conducted, although more effort (e.g., model-based estimation) can be taken to attempt to reduce bias and estimate variance.	Statistical complexity is typically high. On the sampling side, the complexity depends on whether a multistage or single stage sampling design is taken. On the weighting side, weighting needs to be conducted to account for design weights, nonresponse weighting, and calibration weighting.
Data Collection Mode	Surveys can be administered by onsite survey interviewers or can be self-administered by survey participants (with participants recruited by interviewers). Surveys can be completed on paper, on a tablet, or on a laptop.	Face-to-face interviews or mail are the primary data collection modes for collecting survey responses, although telephone and web can be used as supplemental modes for purpose of reminders or data collection if the contact information can be appended to the address list.
Length of Fielding Period	Fielding is typically faster than a mail survey but slower than an online survey. The actual time depends on how many times and locations need to be sampled and how quickly the number of surveys needed to fulfill each quota cell can be fielded.	Fielding is typically slow since face-to-face interviews and mail are the primary modes of survey data collection.

	Intercept Survey With Quota Sampling	Address-Based Sampling
Cost	Cost varies. Factors that influence cost include: number of intercept locations, sample size within each location, number of interviewers recruited, amount of interviewer training offered, amount of interviewer travel required, effort expended to follow up with sample members and field the target number of completed surveys in each quota cell, and value of incentives used.	Cost varies. Factors that influence cost include overall sample size, how much effort is expended to follow up with non-respondents, and whether incentives are used. If ABS is used with web data collection, costs will typically be higher than an opt-in panel but not dramatically so. Cost will be higher by orders of magnitude if face-to-face interviews (rather than mail or web) are used as the data collection mode.

Secondary Data Sources for Program Evaluation

Federal Data

Overview

Multiple sources of Federal data could be used to track change over time in certain traffic safety behaviors. The Fatality Analysis Reporting System (FARS) is a nationwide census providing publicly available data regarding fatal injuries suffered in motor vehicle traffic crashes. The Fatality and Injury Reporting System Tool (FIRST[1])—is an online reporting system that incorporates data from FARS as well as from the Crash Report Sampling System (CRSS) (formerly General Estimates System (GES)) when generating injury estimates. Queries can be run at the State level.

In addition, data regarding commuting behavior are available from the American Community Survey (ACS[2]). Specifically, data are available on means of transportation to work, including driving alone and carpooling; time leaving home to go to work; travel time to work; and number of vehicles in the household. These questions are included in the ACS each year, and data are available by sex and can be analyzed at the county level for localities with populations of 100,000 and above. County estimates are also available for the overall proportion of workers 16 years and older who use a vehicle to commute to work alone.

Another source of data that could be used to track driving-related crashes and injuries is the National Emergency Medicine Services Information System (NEMSIS[3]). NEMSIS is a national database that stores emergency medical services (EMS) data from U.S. States and Territories. NEMSIS provides a framework for standardizing EMS data collection, storage, and sharing nationwide and includes data on vehicular and pedestrian injuries that are submitted by each State.

Advantages

One advantage of using Federal data is that they are publicly and freely available on the NHTSA, U.S. Census Bureau, County Health Rankings, and NEMSIS websites. ACS data are also available at the level of the individual respondent (i.e., microdata) through the U.S. Census Bureau. A variety of weighted analyses—including t-tests and regressions, among others—can be conducted using these person-level data. Federal data are thus easy to obtain and can also be analyzed at the county level where the population is large enough.

Limitations and Suggestions

Federal data do have limitations. ACS person-level data may not always be available at the desired geographic level, such as census tract or even county, when the populations for these specific localities are small. In addition, although the sample size for the ACS is very large at the national level, sampling error can have a meaningful effect on estimates among smaller geographic areas. An important limitation of NEMSIS data is that it only includes crashes where EMS was involved. Another limitation of Federal data, overall, is that they do not directly speak

---

1 https://cdan.dot.gov/query

2 https://www.census.gov/programs-surveys/acs

3 https://nemsis.org

to attitudes or awareness about traffic safety. However, local traffic officials can use such data to supplement surveys they are conducting, as such data provide insights into specific behaviors that various programs and initiatives may be targeting.

Highway Safety Plans

Overview

State Highway Safety Plans (HSPs) are State-level planning documents that describe the allocation of funding and the kinds of activities that are planned for the fiscal year. Specifically, they provide a framework for reducing traffic-related injuries and fatalities. Each State has its own HSP that references and reports on various types of data collected in that State. Types of data that are sometimes included in HSPs include overall rates and percentages of crashes and fatalities, which in some cases are reported by certain groups (e.g., impaired drivers, younger/older drivers, pedestrians, motorcyclists), and other factors, such as the timing of crashes (e.g., time of day, day of the week) and safety measures involved (e.g., seat belt or helmet use). HSPs often also describe surveys and may include measures of attitudes and awareness about traffic safety, perceptions of enforcement strategies, self-reported driving behaviors such as seat belt use, personal assessments of potentially dangerous driving behaviors (e.g., the risk associated with these behaviors or their approval by others), and self-reports of observed traffic behaviors in other drivers. For example, the 2018 New York State HSP included results from a driver behavior survey that included self-reported behaviors such as driving over the speed limit and seat belt use. The 2018 Michigan State HSP included results from a telephone survey that focused on impaired driving, seat belt use, and other driver behaviors. In addition, some States collect data on attitudes as part of broader safety evaluations. These evaluations often use multiple types of data, including survey, administrative, and observational data.

Advantages

Data that are available in HSPs may directly correspond to the specific research questions that local traffic officials have regarding attitudes and awareness about traffic safety, as some States report results from various survey and program evaluation efforts in their HSPs. In these cases, local traffic officials can reference their States' HSPs or contact their State Highway Safety Office to obtain more information if needed. In addition, because HSPs are annual plans, certain measures, including outcomes such as crash rates as well as survey data on attitudes, are reported over time.

Limitations and Suggestions

Limitations of data reported in HSPs include the fact that the methods of data collection as well as the specific kinds of data that are collected vary widely by State. The quality of data and the level at which the data are aggregated (e.g., at the county or city level) vary by State as well. Thus, measures of attitudes and awareness regarding traffic safety may not be readily available in every State's HSP. In addition, the types of analyses that can be conducted with the data that can be obtained will depend on the type and level of data (e.g., person level versus aggregate data) available. However, although survey and evaluation data tend to be reported less consistently, certain measures of outcomes (e.g., rates of crashes or fatalities within a given State) are collected more consistently and over time. Such measures can be useful in augmenting survey efforts that local officials may be planning. For example, outcome measures can be used alongside surveys aimed at evaluating a traffic safety program. Local traffic safety officials should first review their State's HSPs to better understand the tools and data available and contact their State's Highway Safety Office to obtain more information.

Vision Zero Action Plans

Overview

Cities and counties where Vision Zero programs are in place usually publish Vision Zero action plans, which sometimes contain results or descriptions of surveys focused on understanding the public's main safety concerns. For example, as part of Denver Vision Zero (2017), an intercept survey was conducted assessing general traffic safety concerns, including behavior and infrastructure concerns and suggestions regarding what the city can do to improve traffic safety. Similarly, as part of Vision Zero Philadelphia (2017), a short survey was conducted to measure the personal impact that traffic crashes have had on people, people's fear of traffic when walking, and the public's familiarity with speed limits. Similar surveys have been conducted as part of Vision Zero Alexandria (2017), Vision Zero (Washington) DC (2015), and Vision Zero Anchorage (2016) to measure the public's traffic safety concerns.

As part of these Vison Zero action plans, a series of interactive online maps have also been implemented in cities and counties across the United States. Such maps are used to crowdsource data from the public online. These maps allow the public to indicate where crashes, near-misses, unsafe driving conditions, risky traffic behaviors, or unsafe conditions for pedestrians have been observed. Some Vision Zero programs (e.g., Vision Zero Somerville) identify these areas in the online maps themselves to indicate high-volume crash sites and the severity of crashes (e.g., if injuries or fatalities occurred). Such maps have been implemented as part of Vision Zero programs in New York City, Philadelphia, Boston, San Luis Obispo (California), Charlotte (North Carolina), and Montgomery County (Maryland).

Advantages

When available, descriptions of surveys and summaries of results can be easily accessed through a city or county's Vision Zero action plan published online. Because the data are available at the city or county level, they are granular enough to be useful in better understanding a specific locality's traffic safety concerns. As a part of some Vision Zero action plans, interactive maps are publicly available and can be easily accessed online as well. Although the underlying data are not directly available, these data may be requested from specific Vision Zero programs.

Limitations and Suggestions

Generally, information provided about surveys conducted as part of city and county Vision Zero programs tends to be quite limited. For example, relatively little information is provided regarding how data are collected, and the surveys mentioned tend to be relatively short and sample a small number of people. Even when sample sizes are adequate, these surveys can suffer from selection bias if the samples are not drawn from a probability-based frame or adjusted using techniques such as those discussed above. The types of analyses that can be conducted using Vision Zero data may also vary, depending on whether person-level data can be obtained. While not all Vision Zero programs have conducted surveys their action plans may, however, reference outcome measures specific to a given city. These metrics can provide additional insight into behaviors related to a particular locality's traffic safety (e.g., crashes and factors involved in crashes such as drinking alcohol) that may help to supplement survey efforts undertaken by local traffic safety officials.

Social Media

Barriers to Social Media Data Use

Though the use of social media data (e.g., from Twitter or Facebook) in social science research has grown over the years, acquiring and analyzing this type of data presents barriers. One significant barrier is that social media data sometimes must be purchased from third-party vendors, which can be quite expensive. Pricing depends on the type of services the vendor provides, which can range from monitoring social media platforms only (e.g., Twitter, Instagram, Reddit) to monitoring platforms and the web (e.g., news sites, blogs).

It is also difficult to accurately track data by location because users can choose to disable geolocation so that their activities cannot be tracked by location. Although a user may list their hometown in their profile and their location could be identified that way, there is no way to accurately track whether they are tweeting from that location if geolocation is not enabled. Furthermore, when geolocation is enabled, activities are tracked based on the location from which someone posted, which may not be the location the user has listed in their profile. In addition, it may be difficult to get enough data for smaller localities. Given these constraints, using social media data to track attitudes in a specific location can be difficult and prone to error.

Analysis of social media data may also present technological and expertise-related barriers. Experience in data science, natural language processing methods, and structured query language (SQL) is necessary to analyze social media data.

It is also important to note that tweets and posts may not always reflect public opinion. For example, tweeting is a social, self-directed process, and is often driven by specific events. Even when sentiment can be measured from individual tweets, it may be difficult to generalize findings to the population of a given locality, as Twitter users may differ in a variety of ways from non-Twitter users.

Local Data Sources

A variety of local data sources may be readily available to traffic safety officials in specific localities. These include police data (e.g., reports on crashes and fatalities), EMS data, as well as public health and hospital record data (e.g., traffic crash-related injuries and hospitalizations). Some of these data may be available at the person level and could therefore be analyzed using t-tests, regressions, and other statistical techniques. Although these data may not speak to attitudes and traffic safety awareness directly, local officials should consider using these locality-specific data sources, as they provide insights into behaviors and outcomes (e.g., fatalities and crashes) that can be used in conjunction with survey efforts. In addition, because these data sources as locality-specific, they are granular enough to track at the community level.

Conclusion

This document presents two nonprobability sampling methodology plans intended to provide information to local traffic safety officials who are seeking to develop and implement survey studies on traffic safety attitudes and awareness.

Table 7. Comparison of Non-Probability Sampling Methods

	Opt-in Online Panel Survey With Quota Sampling (Plan 1)	Intercept Survey With Quota Sampling (Plan 2)
Typical Target Population	General population Common subset of general population Hard-to-reach population not involved in illegal or stigmatized activities	General population Common subset of general population
Study Area Size	Population size of a local area should be of sufficient size (e.g., large counties, groups of counties).	Population size of a local area is small (e.g., cities or small counties).
Cost	Cost is generally low Factors that influence cost include overall sample size, effort expended to follow up with sample members and field the target number of completed surveys in each quota cell, and the value of incentives used. When the scope or the population size of a local area is too small, additional costs are necessary to recruit more panel members.	Cost can vary. Factors that influence cost include: number of intercept locations, sample size within each location, number of interviewers recruited, amount of interviewer training offered, amount of interviewer travel required, effort expended to follow up with sample members and field the target number of completed surveys in each quota cell, and value of incentives used.
Length of Fielding Period	Fielding is relatively fast. The actual time depends on how quickly the number of surveys needed to fulfill each quota cell can be fielded.	The length of fielding period can vary. Fielding is typically faster than a mail survey but slower than an online survey. The actual time depends on how many times and locations need to be sampled and how quickly the number of surveys needed to fulfill each quota cell can be fielded.

	Opt-in Online Panel Survey With Quota Sampling (Plan 1)	Intercept Survey With Quota Sampling (Plan 2)
Flexibility of Data Collection Mode	Web is the primary data collection mode, although telephone and mail may be used as supplemental modes for the purpose of reminders or data collection if the contact information is available in the panel.	Surveys can be administered by onsite survey interviewers or can be self-administered by survey participants (with participants recruited by interviewers). Surveys can be completed on paper, on a tablet, or on a laptop.

An opt-in online panel survey is a relatively inexpensive and efficient way to collect survey responses, and it is advantageous that information from the panel can be used to assist with sampling design and survey estimation. However, opt-in online panel surveys are typically not appropriate for hard-to-reach populations, such as those whose members are involved in illegal or stigmatized activities, and extra effort and cost will be required to recruit additional panel members to obtain the desired number of survey responses if the population size of a local area is particularly small. While national panels may have over a million members, it is unlikely that they would have enough active panelists from any one medium or small town to assess a change in attitudes.

An intercept survey can be conducted through a variety of modes (e.g., paper, tablet or laptop, face-to-face interviews) and is appropriate for collecting survey responses from the general population when the population size of a local area is small. However, it is not appropriate for use with hard-to-reach populations, and the expenses associated with interviewer training and travel can drive up the survey's cost.

Quota sampling, in conjunction with survey estimation approaches such as calibration weighting and other model-based estimation techniques, can be incorporated into an opt-in online panel survey or an intercept survey to reduce bias and improve the precision of the survey estimates. Note, however, that the bias of an opt-in online panel survey or an intercept survey cannot be eliminated using these techniques.

The nonprobability sampling methods described here should be considered as options when the budget, the timeliness, or the accessibility of a target population prohibits the use of probability sampling. However, these methods are generally subject to greater bias than probability sampling methods and should be used with caution if the study goal is to accurately estimate population values as opposed to detecting change.

When primary data collection for program evaluation is not feasible, there are secondary data sources are available to local traffic safety officials that may be helpful in assessing attitudes and awareness about traffic safety. Data sources such as Vision Zero Action Plans and State HSPs may provide survey data on attitudes and awareness. However, there is considerable variation regarding whether such data are made available as well as its quality. Other data sources, such as Federal data, provide estimates of commuting behavior as well as crash and injury rates. Though such data sources do not directly address local attitudes and awareness about traffic safety, they do offer insights into driving behaviors and can therefore be used to augment survey efforts.

Measuring change in traffic safety attitudes is an important component of program evaluation to determine whether a traffic safety countermeasure achieved its specific objective, and the traffic safety community benefits from solid evaluations that indicate what is and what isn't working. When appropriate, officials may want to partner with researchers and experts at local college or university public health or statistics departments for additional insight.

References

Baker, R., Blumberg, S. J., Brick, J. M., Couper, M. P., Courtright, M., Dennis, J. M., Dillman, D. A., Frankel, M. R., Garland, P., Groves, R., Kennedy, C., Krosnick, J., Lavrakas, P. J., Lee, S., Link, M., Piekarski, L., Rao,. K., Thomas, R. K., & Zahs, D. (2010). Research synthesis: AAPOR report on online panels. *Public Opinion Quarterly, 74(*4*)*, 711–781. doi: 10.1093/poq/nfq048

Baker, R., Brick, J. M., Bates, N. A., Battaglia, M., Couper, M. P., Dever, J. A., Gile, K. J., & Tourangeau, R. (2013, June). *Report of the AAPOR task force on nonprobability sampling*. American Association for Public Opinion Research www.aapor.org/aapor_main/media/mainsitefiles/nps_tf_report_final_7_revised_fnl_6_22_13.pdf

Beck, L., Downs, J., Stevens, M., & Sauber-Schatz, E. (2017). Rural and urban differences in passenger-vehicle-occupant deaths and seat belt use among adults United States, 2014. *MMWR Surveillance Summaries, 66(*17*),* 1–13.

Callegaro, M., Baker, R., Bethlehem, J., Goritz, A. S., Krosnick, J. A., & Lavrakas, P. J. (2014). Online panel research: History, concepts, applications and a look at the future. In M. Callegaro, R. Baker, J. Bethlehem, A. S. Goritz, J. A. Krosnick, & P. J. Lavrakas (Eds.), *Online Panel Research: A Data Quality Perspective*, First Ed. John Wiley & Sons.

Couper, M. P. (2000). Web surveys: A review of issues and approaches. *Public Opinion Quarterly, 64(*4*),* 464–494.

Deville, J. C., & Särndal, C. E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association, 87*(418), 376-382.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: the tailored design method*. John Wiley & Sons

Fanning, E. (2005). Formatting a paper-based survey questionnaire: Best practices. *Practical Assessment Research & Evaluation, 10*(12), 1-14.

Harter, R., Battaglia, M. P., Buskirk, T. D., Dillman, D. A., English, N., Fahimi, M., Frankel, M. R., Kennel, T., McMichael, J. P., McPhee, C. B., Montaquila, J., Yancey, Y., & Zukerberg, A. L. (2016, January 7). *Address-based sampling* (Web page with report, video, and portal links). American Association for Public Opinion Research. www.aapor.org/Education-Resources/Reports/Address-based-Sampling.aspx

Horvitz, D. G., & Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association, 47*(260), 663-685.

Li, R., & Pickrell, T. M. (2019, February). *Occupant restraint use in 2017: Results from the NOPUS controlled intersection study* (Report No. DOT HS 812 594). National Highway Traffic Safety Administration. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812594

Little, R. J. & Vartivarian, S. (2005). Does weighting for nonresponse increase the variance of survey means? *Survey Methodology, 31,* 161–168.

Mercer, A. W., Kreuter, F., Keeter, S., & Stuart, E. A. (2017). Theory and practice in nonprobability surveys: Parallels between causal inference and survey inference. *Public Opinion Quarterly, 81*(1), 250–271.

National Emergency Medicine Services Information System. (n.d.). NEMSIS. https://nemsis.org/

National Highway Traffic Safety Administration. (n.d.). *Fatality Analysis Reporting System (FARS)* (Web page and portal). https://www.nhtsa.gov/research-data/fatality-analysis-reporting-system-fars

National Highway Traffic Safety Administration. (n.d.). *NCSA Tools, Publications, and Data* (Web page and portal). https://cdan.dot.gov

Schonlau, M., & Couper, M. P. (2017). Options for conducting web surveys. *Statistical Science, 32*(2), 279–292.

Singer, E., & Ye, C. (2013). The use and effects of incentives in surveys. *Annals of the American Academy of Political and Social Science, 645(*1*),* 112-141.

Valliant, R., Dever, J. A., & Kreuter, F. (2018). *Practical tools for designing and weighting survey samples*, 2nd ed. Springer.

Appendix: Glossary of Terms

Sampling Concepts

address-based sampling (ABS)

A sampling approach that involves the selection of a random sample of addresses from a list of postal addresses from the U.S. Postal Service.

convenience sampling

The participants are selected to be in the survey based on convenience or ease for the survey researchers.

intercept survey

An approach to conducting surveys that typically involves recruiting participants when they pass by a particular location (e.g., the entrance of a shopping mall). This approach has been extended to the web in recent years by intercepting people who are reaching or browsing a specific website.

multistage stratified sampling

A sampling approach in which population units are partitioned into different subpopulations, called strata, based on geographic and/or demographic characteristics. Next, within each stratum, geographic subsets or other natural clusters are sampled independently, followed by possible further subsampling until a simple random sample or census of the lowest level cluster is obtained.

network sampling

A sampling approach often used with hard-to-reach populations in which employing a probability sampling method is technically infeasible or cost prohibitive. It involves first selecting a small initial sample and then identifying additional participants through their social networks. Also known as snowball sampling, chain-referral, or link-tracing.

nonprobability sampling

A sampling approach in which some members of the population have no chance of being selected into the survey and the selection probability for those who are included is unknown.

opt-in online panel surveys

Surveys in which personal background information for participants (e.g., demographics) is collected upon initial recruitment and panel members can be sampled for a specific online survey based on this background information. A routing approach can also be applied, wherein panelists are invited to complete a survey and assigned dynamically to one of many simultaneously fielding surveys based on the respondent's characteristics and the survey's needs.

probability-proportional-to-size (PPS) sampling

A sampling method in which the size measure for the population unit is available before sampling, and the chance of selecting a population unit is proportional to its size measure.

probability sampling

A sampling approach in which all members of the population have a known and non-zero probability of being selected into the survey. Assuming all sample members complete the survey, probability sampling allows results from a survey sample to be generalized to the larger population that the sample members are intended to represent under known mathematical measures of uncertainty. In practice, factors such as nonresponse or measurement error may affect this inference in ways that may or may not be quantified (Horvitz & Thompson, 1952).

quota sampling

In this special case of sample matching, a sample is drawn so that the distribution of selected characteristics is like that of the larger population with respect to the characteristics used in designing quotas.

sample matching

An approach in which a participant whose data were obtained from a nonprobability survey is paired with a similar person whose data were obtained through a probability survey.

sampling frame

A full list of population members.

simple random sample

A sampling approach in which all population units are sampled independently and have an equal probability of selection.

Quality Evaluation Criteria

composition

Refers to whether the distribution of observed variables in the sample is similar to the distribution of observed variables in the population.

exchangeability

Refers to whether there are any unobserved variables that are correlated with both the inclusion of population members in the sample and the outcomes being measured in the survey.

positivity

Refers to whether any population segments with certain characteristics have no chance of being included in the sample.

precision

The inverse of an estimate's variance. That is, an estimate that has high precision will have low variance.

Survey Error Concepts

expected value

Reflects the long-run average of a random variable. In the context of surveys, the expected value of an estimator typically reflects the average quantity that would be obtained if the survey were repeated an infinite number of times.

bias

The expected difference between the true population value and the expected value obtained by repeating a survey an infinite number of times.

variance

A statistic that characterizes variable (i.e., random) errors in an estimate. Variance is the expected squared difference between an estimate and the long-run average of the estimate. For instance, the sampling variance refers to random variability that arises due to not sampling the entire population for a survey.

confidence interval

A random interval, typically centered on a survey estimate, that indicates where the population value is thought to be located.

Weighting Concepts

weighting

A statistical adjustment method that attempts to generalize the results from the survey respondents to the larger population they are intended to represent. Respondents who are underrepresented or overrepresented are assigned a larger or smaller weight, respectively.

design weights

These are equal to the reciprocal of the known probability of selection in a probability sample. They are used to return to representativeness when the sample design uses unequal probabilities of selection, either intentionally—as when certain minority populations or other less-common elements are oversampled to improve precision of estimates related to them—or when practical issues necessitate unequal selection probabilities.

nonresponse weighting

If there are available covariates that are known for the sampled elements before the conduct of the survey (e.g., geographic location in a cross-sectional study, data from a previous wave in a panel study), then these covariates can be used to predict the probability of a given sample member actually responding, and weights equal to the reciprocal of this estimated probability of response can be computed. If the covariates are categorical and of limited dimension, the response rate within each of the categories can be computed and its reciprocal can be used as the weight. If some of the covariates are continuous, logistic regression can be used to predict the probability of response.

calibration weighting

A weighting approach in which the survey demographic characteristics are made to be equal to the known population totals with respect to the selected characteristics (Deville & Särndal, 1992).

DOT HS 813 242
January 2022

U.S. Department
of Transportation

National Highway
Traffic Safety
Administration

NHTSA

15461-012022-v3b