

**Matching Countermeasures to Driver Types and Speeding Behavior
Supporting Statement for Information Collection Request**

SUPPORTING STATEMENT

Part B.

COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS

Table of Contents

B. Collections of Information Employing Statistical Methods.....2

B.1. Describe the potential respondent universe and any sampling or other respondent selection to be used.....2

B.2. Describe the procedures for the collection of information.....8

 B.2.1. Procedure.....8

 B.2.2. Sample Size.....14

B.3. Describe methods to maximize response rates.....15

B.4. Describe any tests of procedure or methods to be undertaken.....16

B.5. Provide the names and telephone numbers of individuals consulted on statistical aspects of the design.....17

References.....18

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

This National Highway Traffic Safety Administration (NHTSA) proposed information collection will employ statistical methods to analyze the data collected from respondents. The following sections describe the procedures for respondent sampling and data tabulation.

The objective of this study is to collect information from a stratified random sample of licensed drivers in the State of Idaho using an address-based multi-mode survey (mail and web) methodology. The sample will be stratified by age, gender, and number of speeding convictions in the previous three years.

The questionnaire will include questions on driver beliefs, attitudes and driving behaviors, with a specific focus on speeding-related behavior and speeding countermeasures, as well as basic demographic information. Survey responses will be combined with information from state driver records. This will provide additional information on driving history which can be used to validate certain self-report survey responses. The driver records data will also be used to conduct a non-response bias analysis for the survey. The data collected in this study will help us better understand the types of speeding that occur and assist in the development of more targeted countermeasures that address the various types of speeding.

B.1. DESCRIBE THE POTENTIAL RESPONDENT UNIVERSE AND ANY SAMPLING OR OTHER RESPONDENT SELECTION TO BE USED.

Participants in this study will be a stratified random sample of licensed drivers in the State of Idaho. The potential respondent universe is all licensed drivers in Idaho age 18 and over. According to a recent analysis of the data, Idaho has 1,065,547 drivers aged 18 and over.

The primary criterion for site selection for this project involved finding a State with a DMV/MVA that was willing to work with the research team to provide driver contact information for the address-based survey or to conduct all mailings required for the study. In addition, the state needed to be willing to provide the necessary corresponding driver records information to be merged with the survey responses. A secondary criterion was finding a State with a serious speeding problem as determined by the number of speeding related crashes and/or fatalities. Initial State contact was conducted through NHTSA's Regional Offices and distribution of a study flyer by NHTSA at the *38th International Forum on Traffic Records & Highway Information Systems*. Based on these activities, three States indicated potential interest in participating in the study (Idaho, New York, and North Dakota). Representatives from each state were contacted and informed about the study. Based on initial contacts, follow-up conversations were held with Idaho and New York to discuss more detailed issues regarding their Division of Motor Vehicle (DMV) database capabilities and driver information protection. While both Idaho and New York met our primary criteria for site selection, on a per capita basis, Idaho ranked near the top among states for speed-related fatal crashes, in contrast to New York, which falls in the bottom half. Therefore, Idaho was selected due to their expressed willingness

to work with the research team, their excellent records system, and their high per capita speeding-related fatal crashes.

Using State driver records, the random sample of licensed drivers age 18 and older will be stratified by age, gender, and number of speeding convictions in the previous three years. Stratification by speeding conviction history is of special interest in this study on speeding, driver speeding types, and speeding countermeasures. The remainder of this section provides (1) an overview of the sampling methodology implemented in the Idaho speeding survey and (2) minimum detectable difference (MDD) calculations for selected contrasts. These calculations were done using Stata analytical software.

NHTSA targeted a sample size of 1,974 respondents, and we anticipate a response rate near 50%, so the initial sample mailing will consist of approximately 3,948 drivers. We also include estimates here for 1,692 respondents at a 43% return rate, which is adequate to provide detectable differences comparable to the earlier National Survey on Speeding Attitudes and Behaviors (a source of some of the questions included in this survey questionnaire). The sampling frame will be stratified by age (three categories), gender (male/female), and speeding convictions (0, 1, and 2+ in the last three years). Taken together, those variables form 18 strata ($3 \times 2 \times 3 = 18$). Each record in the sampling frame will be assigned to one and only one stratum. Within each stratum, Battelle will conduct a simple random sample to select drivers for the sample. The final sampling fractions for each stratum will be selected after examining the sampling frame, but the current plans call for the following proportions:

Tables 1 and 2 display population data of Idaho drivers collected in March 2013 broken down by gender, age group, and number of convictions group. These proportions were used to help determine the sample size for a given gender/age group/number of convictions group category (or “cell”).

The sample size for a given cell was determined using the following methodology:

1. Set the total sample size for the survey (e.g., $N=1692$ or 1974).
 - a. For a given budget, these numbers were calculated assuming (1) the cost per survey is \$20 and (2) the response rate is 50%.
2. Sample males and females equally (i.e., sample 50% males and 50% females).
3. For each gender category, set the proportion of drivers to sample with 1 speeding conviction “low” (e.g., 20%) and split the remaining percentage equally between speeding conviction groups 0 and 2+ (e.g., 40% and 40%).
4. Define three disjoint age groups:
 - a. Young: 18-24 year old drivers
 - b. Middle: 25-64 year old drivers
 - c. Old: 65-84 year old drivers

For a given gender and number of convictions category, sample the age groups proportionately to the column percentages listed in Tables 1 and 2 (below).

Table 1: Age by number of speeding convictions for females - 3 age categories	Table 2: Age by number of speeding convictions for males - 3 age categories
--	--

Speeding convictions

1	1
9810	11980
1.94%	2.40%
17.81%	21.91%
17.27%	16.48%
42505	53121
8.42%	10.62%
11.80%	14.90%
74.82%	73.06%
4492	7610
0.89%	1.52%
5.01%	8.56%
7.91%	10.47%
56807	72711
11.25%	14.54%
Frequency	
Percent	

Row Pct
Col Pct

Table 3 shows the sample cell counts for two total sample sizes considered (N=1692 and 1974).

Table 3: Cell samples for N=1692 (43% return rate) and N=1974 (50% return rate)

Total sample size (N)	Proportion sampled with 1 ticket	Age category	Number of convictions					
			Male			Female		
			0	1	2+	0	1	2+
1692	0.2	18-24	31	28	81	33	29	81
		25-64	240	124	238	240	127	245
		65-84	67	18	20	66	13	13
1974	0.2	18-24	36	33	94	38	34	94
		25-64	280	144	278	280	148	285
		65-84	78	21	23	77	16	15

Minimum Detectable Difference (MDD) Calculations

Table 4 provides conservative minimum detectable differences (MDD) for a given setting. Assumptions of the power analysis conducted for each setting include (1) the proportion of the population who had one conviction that was sampled is 20%, (2) alpha (type 1 error) is 5%, and (3) power of at least 80% is desired.

Two total sample sizes were considered (N=1692 and 1974). For each N, four contrasts were considered. These four contrasts were purposefully chosen to show the range of sample sizes that might be compared for a given contrast. Showing a range of sample sizes that might be compared for a given contrast will result in showing the range of MDD values.

The four contrasts are as follows:

1. The first contrast considered was "Male vs. Female". Here each group had equal sample sizes and the sample size for each group is large (since all the data is being used). The MDD for this setting will be the smallest as the sample sizes in the contrast are the largest.
2. The second contrast considered was "0 Convictions vs. 2+ Convictions (all ages; both genders)". Here each group has equal sample sizes as the proportion of drivers with 0 convictions (i.e., "non-speeders") and the proportion of drivers with 2+ convictions (i.e., "speeders") were each sampled at 40%. Since 80% of the data is used in this contrast (the only data not used is the 20% proportion of drivers with 1 conviction), the MDD will be somewhat small compared to other contrasts where even less data is used.
3. The third contrast considered was "Young vs. Old (all number of convictions; both genders)". "Young" is defined as drivers aged 18-24 years old and "old" is defined as drivers aged 65-84 years old. Here the sample sizes for each group fall in the middle of sample sizes compared to sample sizes in other contrasts. The MDD for this setting will fall in the middle range of MDDs for a given contrast as the sample sizes fall in the middle range of sample sizes.
4. The fourth contrast considered was "Young Females vs. Old Females (all number of convictions)". "Young" is defined as drivers aged 18-24 years old and "old" is defined as drivers aged 65-84 years old. Here the sample sizes for each group are small compared to sample sizes

in other contrasts as only data from young female and old female are used. The MDD for this setting will be the largest as the sample sizes in this contrast are the smallest.

For each N and contrast (i.e., for each row in Table 4), three MDD calculations were performed, one for each $p=0.5, 0.75, 0.95$. For a given contrast, the group with the largest sample size was assigned the estimated proportion of $p=0.5, 0.75, \text{ or } 0.95$. This was done to be conservative. Then, a power analysis was conducted such that values less than p were considered for the other group in the contrast. Since $p \geq 0.5$, by considering only values less than p , conservative MDDs will result. The difference between the proportion of the group with $p=0.5, 0.75, \text{ or } 0.95$ and the proportion of the other group in the contrast (which is designed to be less than p) that yielded at least 80% power at the 5% significance level was recorded to be the MDD. Note that if the true proportion of the other group is actually greater than p , then the MDD recorded in the summary table for that contrast will have more than 80% power (possibly much more).

Note that as the estimated proportion for one group increases (i.e., p going from 0.5 to 0.75 to 0.95), the MDD gets smaller. This is because it is easier to detect a difference when at least one proportion is close to 1 (or equivalently 0).

Table 4: Minimum detectable differences (MDD) for selected contrasts for N=1692 N=1974 (50% return rate)

Total sample size (N)	Proportion sampled with 1 conviction	Contrast of interest (group 1 vs. group 2)	Sample size for group 1 (n1)	Sample size for group 2 (n2)	Alpha	Power	MDD where the group with the larger sample size has the given proportion		
							p=0.5	p=0.75	p=0.95
1692	0.2	Male vs. Female <i>This contrast provides the minimum (i.e., best) MDD.</i>	846	846	0.05	0.8	7.0%	6.3%	3.6%
		0 vs. 2+ Convictions (all ages; both genders) <i>This contrast provides an MDD close to the minimum (i.e., best) MDD because it's a 40/40 split (only the 1 conviction drivers are not included).</i>	677	677	0.05	0.8	7.8%	7.1%	4.1%
		Young vs. Old Ages (all number of convictions; both genders) <i>This contrast provides a median MDD (half are smaller and half are larger).</i>	283	197	0.05	0.8	13.3%	12.4%	7.7%
		Young Female vs. Old Female Ages (all number of convictions) <i>This contrast provides the maximum (i.e., worst-case) MDD because the sample sizes for these groups are low compared to the sample sizes in other contrasts.</i>	143	92	0.05	0.8	19.3%	18.5%	12.4%
1974	0.2	Male vs. Female <i>This contrast provides the minimum (i.e., best) MDD.</i>	987	987	0.05	0.8	6.4%	5.8%	3.3%
		0 vs. 2+ Convictions (all ages; both genders) <i>This contrast provides an MDD close to the minimum (i.e., best) MDD because it's a 40/40 split (only the 1 conviction drivers are not included).</i>	790	790	0.05	0.8	7.2%	6.5%	3.7%
		Young vs. Old Ages (all number of convictions; both genders) <i>This contrast provides a median MDD (half are smaller and half are larger).</i>	330	229	0.05	0.8	12.4%	11.5%	7.0%
		Young Female vs. Old Female Ages (all number of convictions) <i>This contrast provides the maximum (i.e., worst-case) MDD because the sample sizes for these groups are low compared to the sample sizes in other contrasts.</i>	166	108	0.05	0.8	17.8%	17.0%	11.2%

Resulting publications will include a caveat that data were collected from a random sample of Idaho drivers, and that the results cannot be generalized to the population of American drivers. We will not attempt to characterize American drivers with the results. We will instead use the findings as a foundation from which to generate ideas for the development of new behavioral- and vehicle-based countermeasures for enhanced speed management and improved traffic safety.

B.2. DESCRIBE THE PROCEDURES FOR THE COLLECTION OF INFORMATION.

B.2.1. PROCEDURE

According to the CDC (Brief 22, 2010), when incentives are used in combination with other response boosting strategies, such as multiple contacts and user-friendly questionnaires, the response rates are improved. For example, when a monetary incentive was used for respondents for a statewide community survey, the response rate was 68.4% (Messer & Dillman, 2011). In addition to demonstrating respect and appreciation of the respondents' time and effort in helping us with the survey, we will also provide the incentive before the completed questionnaire is returned to convey trust. When used with mail surveys, such prepaid incentives have the most significant impact on response rates. Therefore, in the current study, we will include a monetary incentive in the mail survey (sent via Priority Mail) to increase response rate.

We will use two modes for survey completion: mailed paper surveys and electronic web-based surveys. The mail survey will be the primary mode; however, participants will have the option to complete the survey using either mode. Developing a web-based survey from a paper version requires minimal effort, since Battelle will use an electronic version of the survey for entering data and populating the response database.

A key advantage of a web-based approach is that it accommodates certain user groups that prefer this approach, such as young urban drivers that may otherwise be less likely to complete paper version of the survey. In addition, Dillman and his colleagues have found that people with more education are more likely than those with less education to complete web surveys (Millar et.al, 2009). In this same article, it was reported that 90 percent of those expressing a preference for a web survey completed the survey on the web as compared to another method. For these reasons, we believe that many respondents will prefer to complete the survey on the web. The web survey option will require little additional effort on the part of the project Engineer, since the web form will double as the data entry form for entering paper responses once they are returned to Battelle. Web-based surveys are less effective in areas with poor access to the internet or less educated populations. However, paper surveys will be provided for all drivers, which will ensure that this segment of drivers will still be adequately represented in the survey sample, even if they have limited access to computers.

We will use a multi-step approach for encouraging participation. This type of approach is consistent with survey standards endorsed by the American Association for Public Opinion Research (AAPOR), and is one that has been used by the research team to obtain high response rates in previous research studies. The specific steps in this approach are described below. NHTSA's contractor, Battelle, will provide a tracking system and the training necessary for the State of Idaho to conduct all survey mailings in a way that meets the quality and schedule objectives of the project. The Idaho Transportation Department (ITD) will provide staff and

materials to implement the survey, including reminders, following specific procedures and a prescribed schedule. The specific activities include:

- Mailing an Advance Letter
- Mailing the Survey Packet
- Mailing Reminder Postcards
- Mailing Additional Survey Packets

The initial contact to participants will be by an advance letter that is mailed on state letterhead to all sample members, which will serve to notify sample members that a survey will be following shortly, so they can watch for it in the mail. Offering large incentives could be considered coercive by participants; and there is no significant difference in response rates between \$5 or \$20 prepaid incentive (CDC, Brief 22, 2010). Therefore, we will include a monetary incentive of \$5 cash. The \$5 payment will be mentioned in the advance letter.

Within several days of receiving the advance letter, individuals will receive a survey packet by priority mail. The packet will include:

1. The survey questionnaire with a pre-printed ID number;
2. A personalized cover letter emphasizing the importance of the study;
3. A postage-paid, self-addressed return envelope; and
4. A reimbursement of \$5 for the individual's time and effort given to the study.
5. Instructions for accessing the online survey (i.e., the web URL and password).

NHTSA's contractor, Battelle, will prepare all of the survey packets that will then be delivered to the State of Idaho for personalizing and mailing. Battelle has procedures in place to ensure proper accounting and distribution of the participant compensation. To safeguard the money, we will limit the individuals involved in packet assembly so that only authorized individuals are in contact with the survey compensation. Packet assembly is conducted with at least two researchers working together. As part of the quality control process, two senior study staff then inspect the packet contents to ensure that all required materials, including the correct compensation, are included, and then the team seals the packets together.

We will include language in the cover letter about how to complete the survey online if they wish to do so. The letter will ask respondents to return the completed survey in the postage-paid, return envelope. The survey cover letter will also provide the toll-free telephone number to call with questions about the study, as well as the telephone number to call with questions regarding Human Subjects protection. Before the surveys are mailed, an additional quality control procedure will be completed by the state to ensure that the inner label matches the outer label.

Achieving a response rate of at least 50 percent to the survey will require repeated follow-up with non-respondents. Our recommended follow-up procedures include:

Thank you/reminder postcards. A survey tracking database will be used to track all returned surveys. Within two weeks of the initial mailing, a thank-you/reminder postcard will be sent to each respondent who has not yet returned the survey to encourage survey completion. The postcard will include a toll-free number that can be called if the respondent has any questions about completing the survey or needs to have another copy of the survey mailed. The look and feel of the postcard will be similar to the look and feel of the survey. Two additional reminder postcards will be mailed for non-respondents at six and ten weeks following the initial survey mailing.

Additional survey mailings. Two weeks after the postcard reminder is mailed, a second survey will be mailed. A cover letter will be included stressing the importance of completing the survey. A third survey mailing will be sent eight weeks after the initial mailing if we still haven't received a response. No additional compensation will be provided in these subsequent mailings.

Table 5 shows the number of participants we expect to contact using the various methods at different points during the survey. We estimate a response rate of about 15% after the initial survey mailing and response rates of about 10% after each additional mailing.

Table 5. Expected sample size by type of contact with study participants.

Survey Task	Expected Sample Size	Week of Data Collection
Advance Letter with Opt-Out Postcard	3948	1
Survey Mailing with \$5 incentive	3948	2
Thank you/Reminder Postcard 1	3395	4
Second Survey Mailing	3040	6
Thank you/Reminder Postcard 2	2724	8
Third Survey Mailing	2448	10
Thank you/Reminder Postcard 3	2211	12
Data Entry for Paper Surveys	1574*	2-14
Data Cleaning	1974	7-16

* Data Entry will be necessary for only those cases that are completed on paper. For budgeting purposes, we assumed that 20% of respondents would complete the survey online.

The return address on all survey mailings will be the address of the ITD. The address on the business reply envelope (for returning completed surveys) will be Battelle's address. As such, the State will never have access to a participant's survey data, and Battelle will never have access to a participant's contact information or name. Due to the anonymous nature of the survey, we hope that sample members will be more likely to respond and that respondents will be more candid.

Language accommodation. According to U.S. Census Bureau (2012), 11.6% of Idaho residents are Hispanic or Latino, and 3.05% of Idaho residents only speak Spanish. To increase response rates among Spanish-speaking respondents, the survey will be available in Spanish. More importantly, we will implement the following methods to ensure respondents' needs in taking the survey in Spanish are fully met:

- Printing the advance letter with English on the front and Spanish on the back (Appendix F and G).
- Programming the web survey to allow for completion in either English or Spanish (respondent’s preference).
- Printing the cover letter that goes in the survey with English on the front and Spanish on the back, with an instruction about how to request a paper survey in Spanish.
- Establishing a dedicated toll-free number at Battelle that can be answered by Spanish-speaking representatives (or with a voice mail message in Spanish if no Spanish-speaking representatives are available) where respondents can call to request a Spanish paper survey. The toll-free number will be included in the cover letter of the survey package. Respondents will only be asked to provide their CaseID. Battelle will then prepare the survey package, and send it to ITD using Express Mail. ITD will then be able to print an address label and new cover letter (in Spanish) and mail out the Spanish language materials to the corresponding address. Battelle will neither ask for nor retain any personally identifiable information in this process.

Table 6 provides data on each task to be completed by the IDT. It also includes the timing of each task and the estimated sample size. Battelle will provide all printed materials (except state letterhead and letterhead envelopes for the advance mailing task), all postage, all mailing labels, and the tracking data base for use by the state.

Table 6. Survey administration activities/milestones.

Timing	Task	Estimated Sample Size
~Day -14 (at least two weeks prior to Day 1)	Ensure that there is sufficient stock of letterhead and envelopes for Advance Letter mailing	3948
~Day -7 (about one week prior to Day 1)	Ensure that staff attend training held by Battelle on Survey Tasks (Battelle will travel to State)	NA (Battelle recommends training 2-3 staff)
~Day -5 (any time between training and Day 1)	Implement mail merge to produce personalized letter on letterhead and address labels (you may want to consider having a dedicated printer for producing the letters, because the letters could take a long time to print)	3948
	Prepare mailing: place labels on letterhead envelopes, fold letter, place letter in envelope, and stamp envelope (do not seal envelope)	3948
	Conduct QC to ensure that letters match address labels	3948
	Seal envelopes	3948
Day 1	Advance Letter Mailing	3948
~Day -4 (any time between production of advance letters and Day 3)	Implement mail merge to produce personalized address labels	3948
	Prepare mailing: place address labels on outer envelopes, place interior envelope (Battelle will prepare these) into outer envelope ensuring that Case ID and address match up (do not seal envelope)	3948
	Conduct QC to ensure that Case IDs match address labels	3948
	Seal envelopes	3948
Day 3	Survey Packet Mailing	3948
Days 4-17	Enter cases where the survey packets or advance letters are returned undeliverable into tracking (Battelle will enter cases that are returned completed into tracking system that does not include PII)	30
Day 17 (early morning)	Run a query that will pull finalized (completed surveys returned) case status from Battelle tracking system (that does not contain PII)	562
Day 17	Implement mail merge to produce personalized address labels	3395

Timing	Task	Estimated Sample Size
	Place labels on postcards	3395
Day 17	Thank you/Reminder Postcard 1 Mailing	3395
Days 17-31	Enter cases where the survey packets, advance letters, or postcards are returned undeliverable into tracking (Battelle will enter cases that are returned completed into tracking system that does not include PII)	17
Day 31 (early morning)	Run a query that will pull finalized (completed surveys returned) case status from Battelle tracking system (that does not contain PII)	318
Day 31	Implement mail merge to produce personalized address labels	3040
	Prepare mailing: place address labels on outer envelopes, place interior envelope (Battelle will prepare these) into outer envelope ensuring that Case ID and address match up	3040
	Conduct QC to ensure that Case IDs match address labels	3040
	Seal envelopes	3040
Day 31	Second Survey Mailing	3040
Days 31-45	Enter cases where the survey packets, advance letters, or postcards are returned undeliverable into tracking (Battelle will enter cases that are returned completed into tracking system that does not include PII)	15
Day 45 (early morning)	Run a query that will pull finalized (completed surveys returned) case status from Battelle tracking system (that does not contain PII)	287
Day 45	Implement mail merge to produce personalized address labels	2724
	Place labels on postcards	2724
Day 45	Thank you/Reminder Postcard 2 Mailing	2724
Days 45-59	Enter cases where the survey packets, advance letters, or postcards are returned undeliverable into tracking (Battelle will enter cases that are returned completed into tracking system that does not include PII)	13
Day 59 (early morning)	Run a query that will pull finalized (completed surveys returned) case status from Battelle tracking system (that does not contain PII)	259
Day 59	Implement mail merge to produce personalized address labels	2448
	Prepare mailing: place address labels on outer envelopes, place interior envelope (Battelle will prepare these) into outer envelope ensuring that Case ID and address match up	2448
	Conduct QC to ensure that Case IDs match address labels	2448
	Seal envelopes	2448
Day 59	Third Survey Mailing	2448
Days 59-73	Enter cases where the survey packets, advance letters, or postcards are returned undeliverable into tracking (Battelle will enter cases that are returned completed into tracking system that does not include PII)	11
Day 73 (early morning)	Run a query that will pull finalized (completed surveys returned) case status from Battelle tracking system (that does not contain PII)	233
Day 73	Implement mail merge to produce personalized address labels	2211
	Place labels on postcards	2211
Day 73	Thank you/Reminder Postcard 3 Mailing	2211
Days 73-100	Enter cases where the survey packets, advance letters, or postcards are returned undeliverable into tracking (Battelle will enter cases that are returned completed into tracking system that does not include PII)	9
Day 100	Survey Closed	NA

During the survey administration period, Battelle will provide the COTR (TO) with a weekly disposition report and it will be submitted every Tuesday of the week. Survey data collection will be completed within four (4) months of the onset of Phase 2.

Monitoring the Progress of Data Collection

Battelle will develop a tracking system with the following features:

- Unique CaseID for every case
- Unique password for Web Survey access for every case, and
- Functionality to produce all mailings

There will be two instances of the tracking system – one at Battelle that has the CaseID only, and one at ITD that can have PII uploaded or merged into it. ITD personnel will use the tracking database to produce the mailing labels, personalized cover letters, and other materials in the participant mail-out packets that require driver information. These functions will be programmed into the tracking database by Battelle. ITD personnel will also use the tracking database to track surveys and other mailings that are returned to ITD as undeliverable. Note that Battelle will receive the completed surveys, and will enter those into the tracking database. Before subsequent mailings are prepared, Battelle will provide ITD with a report that updates their tracking with completed cases (by CaseID), so that these individuals do not receive an additional survey.

The tracking database will be programmed so that study staff can generate weekly disposition reports summarizing the status of the data collection activity throughout the data collection period (e.g., number of surveys mailed, number of surveys undeliverable, number of surveys received). In addition, we will be able to compute the response rates achieved by each stage of data collection (e.g., to the initial mailing, after the postcard reminder, or after the follow-up telephone call).

The survey tracking database will also allow us to keep track of the response rates of the various sub-groups who were part of the sample (e.g., there may be different response rates for sample members with and without citations). If during the course of the survey it is determined that the response rate for one sub-group far exceeds that of another in a way that is detrimental to the overall response pool, contact information for additional drivers in the low-response sub-group may be obtained in order to better balance the response pool. In this way, the survey data collection may occur in multiple waves of mailings.

Entering Data into an Electronic Database

The printed surveys will be formatted to facilitate data entry. Data entry will begin when the first surveys are returned. All data entry will be 100% verified by someone other than the original data entry person, and all errors will be checked and corrected. Electronic data cleaning will be used to detect errors not identified during data entry. These programs check for outliers, incorrect use of skip patterns, and logical consistency checks between variables to identify inconsistencies. Questionable data will be flagged by the program and then checked by referring to the survey questionnaire. Verifiable errors will be corrected. Data entry will be completed by professional staff who have experience using Battelle data-entry systems.

Data security is an integral part of all Battelle projects. Battelle survey staff have developed procedures to ensure respondent confidentiality. Each respondent will be assigned a unique study identification number. The survey instruments, as well as the electronic data files containing the survey data, will be identified only by study identification number. Neither the mail surveys nor the electronic files of the survey data will contain names, addresses or telephone numbers of facilities or respondents. Completed surveys will be stored in locked file cabinets. All project files will be password protected, and access to the files will be limited to authorized project staff. Staff working on the project will be required to sign a confidentiality statement. The steps taken to protect the confidentiality of the data will be emphasized in staff training. Data files submitted to the DOT will contain no identifying information. A statement explaining respondent confidentiality and protection will be included in the survey cover letter and on the cover of the questionnaire.

B.2.2. SAMPLE SIZE

The planned sample size is 1,974 licensed drivers in the State of Idaho. Details on the sampling plan are provided in Section B.1, above. During preparation for the survey mailing, ITD will provide Battelle (NHTSA's contractor for this project) with a dataset to use as a sampling frame. The dataset will include one record for every person who has been a licensed Idaho driver for at least three years. The dataset will not contain personally identifying information (PII) but will include a record ID, which ITD can use to link back to the driver's name and mailing address, if the driver is selected for the survey sample. Other fields in the dataset will include gender, age, zip code, number of speeding convictions in the past three years, and number of traffic-related convictions from several other select categories related to driver safety.

Every survey response will be assigned a sampling weight that indicates the number of Idaho drivers that their response represents. When all the data have been collected, the survey weights will be adjusted to account for the final sample size and any differential non-response across strata. When results are analyzed with the survey weights, the estimates and inferences based on the survey data can be representative of Idaho drivers in general, and not just of those in the sample, and furthermore it will be possible to quantify the uncertainty in those estimates and inferences that is related to sampling variability. That uncertainty will be quantified using confidence intervals or p-values, as appropriate.

The final cut-points for the age strata will be selected in consultation with NHTSA after the sampling frame dataset is received from ITD and the distribution of speeding convictions by gender and age strata can be assessed. The guiding principles of the final selection will be to solicit responses from enough respondents to be able to estimate the quantities of interest with minimal uncertainty in groups that we expect to be meaningfully different, such as young drivers versus older drivers, and persons with multiple speeding convictions versus those with none.

Similarly, the final sampling fractions for all strata will be determined after assessing the sampling frame. If the population of convicted speeders is strongly imbalanced between men and women then we might consider shifting the balance of the survey sample along those lines as well, and using the sampling weights to draw representative conclusions.

B.3. DESCRIBE METHODS TO MAXIMIZE RESPONSE RATES.

Several methods will be used to maximize response rates. The survey will be conducted in dual modes using a mail survey and offering an on-line Web-based option. The mailed survey will be sent to respondents' home addresses. It will include an addressed envelope with pre-paid postage for returning the completed questionnaire to the research team. Because this is an address-based sample, mailing will come first. Potential respondents will also be offered the option of taking the survey on-line. The on-line option is especially important for increasing response rates with younger drivers, while the mailed survey should appeal to older drivers and increase response rates with those who are less comfortable using computers. The survey will also be made available in Spanish for respondents whose primary or preferred language is Spanish.

As described in B.2.1, we will use a multi-step approach that is consistent with survey standards endorsed by the American Association for Public Opinion Research (AAPOR), for encouraging participation. This type of has been used by the research team to obtain high response rates in previous research studies. This multi-step approach will include an advance mailed letter from the Idaho Transportation Department (ITD), which should ensure that most potential respondents will pay attention to our initial contact with them regarding the study. The ITD is also endorsing the survey and placing their logo prominently on the cover of the questionnaire, along with NHTSA's logo, which should add to the perceived legitimacy of the survey and its importance for drivers in the State of Idaho.

Research (Messer & Dillman, 2011; Lesser et al., 2001) has shown that financial incentives can greatly improve response rates, especially with multiple mailings. Gifts or financial remunerations have become commonplace in survey research to ensure higher response rates (CDC, Brief 22, 2010). For this study, to prevent respondents from feeling obligated to complete the survey because they received a large financial reward, they will be provided with \$5 to compensate them for their time and the effort they give to the study in order to increase response rates for the study. This amount is similar to remunerations in other survey research projects.

B.4. DESCRIBE ANY TESTS OF PROCEDURE OR METHODS TO BE UNDERTAKEN.

Cognitive testing was conducted in Boise, Idaho in April 2013 with a convenience sample of nine (9) respondents in order to identify any problems and to refine the English language questionnaire for the pilot test. An interview guide was developed for the cognitive testing. The guide covered aspects such as respondent understanding of questions, overall flow of the questions, and ease of moving from one question to the next if branching is used, etc.

The participants in the cognitive testing were residents of Idaho so that we could capture the regional characteristics and demographic composition of the population. The testing was held in a rented facility in the outskirts of Boise, Idaho using portable auditory recording equipment. This location made it somewhat easier to recruit drivers that live in rural areas, while at the same time allowing us to conduct all the testing from a single facility. Note that to encourage participation from rural drivers that may have to travel a significant distance to the facility, we

offered gas cards to offset fuel costs. Data collection was conducted in a single wave spanning 3 days (i.e., 3 per day), which provided sufficient time in between sessions to consult with NHTSA and make changes to the survey based on participant feedback.

Cognitive interview participants were recruited using online ads. An experienced interviewer contact interested individuals, explained the details of participating, confirmed their willingness to participate, and scheduled a time to conduct an in-person interview. Participants were asked to complete the survey and when they were finished, they were interviewed about the survey and the protocol. During the interview, participants were asked all items included in the cognitive interview guide, and they were probed for specific ways in which the questionnaire could be improved. Each pre-test participant received a \$50 honorarium for his/her participation.

The pre-testing of the questionnaire proceeded in an iterative fashion, across consecutive days within a single trip: After the interviews for the first three pre-test respondents on the first day of testing were completed, we reviewed the results of the pre-test and made modifications to the questionnaire. The same procedures was used for a second day of three interviews and then again for a third day of three interviews. For each of the three phases of the pre-test, we summarized the major findings and created an updated version of the questionnaire.

The key findings across the three rounds of cognitive testing are listed below, along with descriptions of the changes that were made to the survey materials to address key issues identified.

- The wording of several questions was reported to be unclear. The wording for these questions was improved, and retested in subsequent sessions, if possible.
- Participants responded much more favorably to color images in questions A46 through A51 than to gray-scale version of the same images. Specifically, color images were easier for respondents to understand, and made the survey more appealing for nearly all of the respondents. We therefore decided to print the surveys in color.
- Several participants indicated concern over questions referring to speed cameras, especially since these cameras are explicitly prohibited in Idaho. It also led some participants erroneously interpret the inclusion of these questions as an indication that the State was considering installing them. This belief would undermine the purpose of the survey. Consequently, we removed two questions about the use of speed cameras.

Several questions were perceived as redundant. These questions were removed to shorten the survey.

Based on the results of the cognitive testing, revisions were made to the questionnaire and the final version of the questionnaire was developed. Once this final revised version was completed, a Spanish version of the questionnaire was created for use in the survey.

The final questionnaire will be printed in a multi-page booklet format (estimated at 8-10 pages and 20-25 minutes to complete) that is designed for ease of use by respondents. The actual printing will not take place until after OMB approval.

B.5 PROVIDE THE NAMES AND TELEPHONE NUMBERS OF INDIVIDUALS CONSULTED ON STATISTICAL ASPECTS OF THE DESIGN.

Randolph Atkins, Ph.D.
NHTSA Project Manager / COTR
Social Science Researcher
National Highway Traffic Safety Administration
1200 New Jersey Ave SE, W46-500
Washington, DC 20590
(202) 366-5597
randolph.atkins@dot.gov

Christian Richard, Ph.D.
Principal Investigator (Contractor)
Senior Research Scientist
Battelle Center for Human Performance and Safety

1100 Dexter Avenue N., Suite 400
Seattle, Washington 98109-3598
Phone: (206) 528-3249
E-mail: richardc@battelle.org

Dale Rhoda, Ph.D.
Statistician
Battelle Health and Analytics
505 King Ave
Columbus, Ohio 43201-2696
Phone: (614) 424-4944
E-mail: RhodaD@battelle.org

Ta Liu, Ph.D.
Statistician
Battelle Health and Analytics
1100 Dexter Avenue N., Suite 400
Seattle, Washington 98109-3598
Phone: (206) 528-3157
E-mail: liut@battelle.org

REFERENCES

Centers for Disease Control and Prevention (CDC). Brief 22: Using Incentives to Boost Response Rates. Available at <http://www.cdc.gov/HealthyYouth/evaluation/pdf/brief22.pdf>.

Lesser, V. M., Dillman, D. A., Carlson, J., Lorenz, F., Mason, R., & Willits, F. (2001). Quantifying the influence of incentives on mail survey response rates and their effects on nonresponse error. Proceedings of the Annual Meeting of the American Statistical Association.

Messer, B. L., and Dillman, D. A. (2011). Surveying the general public over the internet using addressed-based sampling and mail contact procedures. *Public Opinion Quarterly*, 75(3), 429-457.

Millar, M. M., O'Neill, A. C., and Dillman, D. A. (2009). *Are mode preferences real?* (Technical Report 09-003). Social and Economic Sciences Research Center.