TABLE OF CONTENTS

B. Collections of Information Employing Statistical Methods

The proposed study will employ statistical methods to analyze the information collected from respondents and draw inferences from the sample to the target population. It will use an overlapping dual frame design with respondents interviewed on landline phones and on cell phones. The following sections describe the procedures for respondent sampling and data tabulation.

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

a. Respondent universe

The respondent universe theoretically consists of all persons aged 16 or older residing in households. However, since the survey will be administered by telephone, the sampling universe will be persons aged 16 or older living in households with working telephones. Also, since interviews will be conducted in only English and Spanish, any person who does not speak one of these two languages will be excluded from this study. Based on the prevalence of the use of English and Spanish in the United States, excluding respondents who speak a language other than English and Spanish will have minimal impact on the study. This study will sample from all telephone households in the United States, including Alaska, Hawaii, and the District of Columbia.

It is not anticipated that the use of a telephone household sample will cause coverage bias. Research has shown that telephone service for many households is intermittent, and that gains or losses in service are associated with changes in their financial situation, (Keeter, 1995[1]). By implication, telephone coverage is not a static phenomenon, but rather, the telephone population at any given point includes households that were recently part of the non-telephone population. Households with intermittent telephone service are similar on a variety of important demographic characteristics to households who usually lack telephones (Keeter, 1995; Brick, Waksberg, and Keeter, 1994[2]). Thus, weighting using information on interruptions in telephone service can ensure the correct coverage in the final sample of persons including those in non-telephone households.

b. Statistical sampling methods

A national sample of 9,000 people will be interviewed regarding pedestrian and bicyclist behavior, attitudes, and knowledge (see Appendix D for the initial version of the

[1] Keeter, S. *Estimating non-coverage bias from a phone survey*. Public Opinion Quarterly, 1995; 59: 196-217.

[2] Brick, J.M., Waksberg, J., and Keeter, S. (1994). *Evaluating the use of data on interruptions in telephone service for nontelephone households*. Proceedings of the Survey Research Methods Section of the American Statistical Association, pp. 19-28. Alexandria, VA: American Statistical Association.

questionnaire, Appendix E for the justification of each question, Appendix F for an updated version of the questionnaire based on cognitive testing, and Appendix G for a list of changes between the 2012 and 2002 questionnaires). The survey will employ a partially overlapping dual frame design with national probability samples drawn from independent sampling frames for landline phones and for cell phones (with screening for cell-only and cell-mostly households). Both the landline and cell phone samples will be geographically stratified by NHTSA Region. The precision of sampling estimates are generally improved by stratification. Stratification involves the division of the total population into population subsets in which the homogeneity of elements within the groups is greater than the homogeneity of elements within the population as a whole. Sampling is conducted separately within each stratum. Since the amount of improvement in sampling precision as a result of stratification is a function of the degree of homogeneity achieved within strata, national area probability samples are generally stratified by region (since regional differences in population characteristics and lifestyles may be expected to introduce variability into national samples).

NHTSA segments the country into ten Regions for programmatic outreach, divided as follows:

Table 1. NHTSA Regions

Region	**States**
Region 1-New England Region	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
Region 2- Eastern Region	New York, New Jersey, Pennsylvania
Region 3- Mid Atlantic Region	Delaware, District of Columbia, Maryland, Virginia, West Virginia, North Carolina, Kentucky
Region 4- Southeast Region	Alabama, Florida, Georgia, South Carolina, Tennessee
Region 5- Great Lakes Region	Illinois, Indiana, Michigan, Minnesota, Ohio, Wisconsin
Region 6- South Central Region	Louisiana, New Mexico, Oklahoma, Texas, Mississippi
Region 7- Central Region	Iowa, Kansas, Missouri, Nebraska. Arkansas
Region 8- Rocky Mountain Region	Colorado, North Dakota, South Dakota, Utah, Wyoming, Nevada
Region 9- Western Region	Arizona, California, Hawaii
Region 10- Northwest Region	Alaska, Idaho, Oregon, Washington, Montana

The estimated distribution of the landline telephone population by stratum is calculated on the basis of data from the National Health Interview Survey (NHIS). State estimates are available from the NHIS, which will be used to calculate the Regional distributions. NHTSA will use the most recent NHIS estimates available at the commencement of the field period. Based on NHIS estimates of the geographic

distribution of the target population, the landline sample will be proportionately allocated by stratum. Geographic location for the landline sample is based on the dominant zip code for listed numbers within selected telephone banks (see next paragraph for description of phone banks). The estimated distribution of the cell phone only and cell phone mostly population by stratum is also calculated on the basis of data from the National Health Interview Survey. Since there are no listed numbers for cell phones, the geographic location assigned to cell phone banks is based on its billing center. This leads to a significant misclassification of the location of cell phone households at the community level, but very little at the State level since billing centers are located in the same State as the household being served.

The landline sample will be drawn from telephone banks randomly selected from an enumeration of the Working Residential Hundred Blocks within the active telephone exchanges within the Region. The Working Hundreds Blocks are defined as each block of 100 potential telephone numbers within an exchange that includes one or more residential listings (i.e., this will be a list-assisted sample). A two-digit number will then be randomly generated for each selected Working Residential Hundred Block to complete the phone number to be called. By randomly generating these numbers, a process known as random digit dialing (RDD), every number in the sampling frame of Hundreds Blocks has an equal probability of selection regardless of whether it is listed or unlisted.

The RDD sample of telephone numbers is dialed to determine which are currently working residential household telephone numbers. Non-working numbers and non-residential numbers are immediately replaced by other RDD numbers selected within the same stratum in the same fashion as the initial number. Ineligible households are also replaced. Non-answering numbers are not replaced until the research protocol is exceeded. The systematic dialing of those numbers to obtain a residential contact should yield a probability sample of landline telephone numbers.

The next step is selecting the household member to interview. Once a household is contacted, the interviewer will introduce him/herself and the survey, and ask how many persons residing in the household are 16 or older. If there is only one person in that age range residing in the household, then the interviewer will seek to interview that person. If there are multiple people in that age range residing in the household, the interviewer will randomly select one of those household members by asking to speak with the age-eligible household member who had the most recent birthday. Only that randomly selected person will be eligible to participate in the survey; no substitution by other household members will occur. If the selected respondent is not available, then the interviewer will arrange a callback. In cases where no one residing within the household is in the eligible age range, the interviewer will thank the individual who answered the phone, terminate the call, and the number will be removed from the sample.

Although list-assisted landline RDD sampling provides only a small coverage error for landline telephone households within landline banks, the restriction of the sampling frame to only landline banks would introduce a much more serious coverage

error in general population surveys. The increasing percentage of households that have abandoned their landline telephones for cell phones has significantly eroded the population coverage provided by landline-based surveys. The key group that is missing from landline RDD samples is the cell phone-only group. But there is also potential bias in landline samples from under-coverage of people living in households with landlines but who have a low likelihood of being reached on them because they rely on wireless telephones for all or almost all of their calls. The cell phone sample for this survey will therefore be composed of both cell phone only and cell phone mostly respondents.

Due to the higher cost of cell phone interviews compared to landline interviews, dual frame surveys are usually designed with optimal allocation of sample between the landline and cell phone strata, rather than simply proportional allocation. We used the following formula for optimum allocation in a two-stratum design (landline and cell phone) to arrive at the number of cell phone only interviews planned for this survey:

$$\frac{p}{(p + 1-p\sqrt{\frac{C_{cellonly}}{C_{landline}}})},$$

In the formula, p is the estimated proportion of the adult population expected to be cell phone only in 2012 and C is the cost of the specified type of interview (cell only or landline). We've assumed that 33% of the adult population will be cell phone only in 2012 and used a cost ratio of 3 to 1. This results in about 22% of the total sample (or 1,993 interviews) being allocated to cell phone only. The most recent data published from the National Health Interview Survey[3] shows 30.2% of adults being in cell phone only households and 18.2% being in cell phone mostly households during the first half of 2011. The ratio between the two wireless phone user categories is 62.4% cell phone only and 37.6% cell phone mostly. Applying that ratio to the 1,993 cell phone only interviews adds 1,201 cell phone mostly interviews. The cell phone sample for the National Survey of Pedestrian and Bicyclist Attitudes and Behavior will therefore total 3,194 interviews divided as just described, a projected 1,993 cell phone only and 1,201 cell phone mostly respondents. Cell phone households not falling in either of these categories will be screened out of the cell phone sample.

The cell phone sample will be randomly selected from 1,000 banks used exclusively for cell phones, using RDD. Procedures for sample selection will be similar to those used in selecting the landline sample, except that the sample will not be list-assisted as it is in the landline sample. In addition, the cell phone will be treated as a single user device. This means the cell phone sample will not require the procedures used with the landline sample to select from multiple eligible household members. Upon

[3] Blumberg, Stephen J. and Julian V. Luke. *Wireless Sustitution: Early Release of Estimates From the National Health Interview Survey, January – June 2011.* U.S. Department Of Health And Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics. Released 12/21/11 . Located at http://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201112.htm .

contact and introduction, the interviewer will immediately ask questions to determine that the person on the phone is not in a situation that could pose a safety risk to that individual (e.g., driving at the time of the call). If the contacted individual is found to be in a situation that could pose a risk, the interviewer will terminate the call and call back another time. If it is safe for the contacted individual to proceed with the call, then the interviewer will move on to asking screening questions. These will include questions to categorize the contacted individuals as cell phone only or cell phone mostly. Cell phone users not falling into one of these two categories will be screened out as the cell phone sample is to be composed only of those two groups.

c. Over-sampling Persons Age 16-39 Years

Data from the 2002 National Survey of Pedestrian and Bicyclist Attitudes and Behavior showed younger adult age groups engaging in more bicycling and outdoor walking than their older peers[4], making them a group from whom it is especially important to collect information for this survey. However, the low response rates typically achieved with young adults will result in their composing a far smaller proportion of the achieved sample compared to their actual prevalence in the 16 and older population if left to simple random selection procedures. This would particularly impact the bicycling portion of the survey as many of the questions are directed specifically at bicyclists. According to the 2002 survey, people defined as bicyclists (i.e., having ridden a bicycle at least once during the past 30 days) composed only 27 percent of the 16 and older population. The reduced sample combined with lower response rates among the age segments where the behavior of interest is more prevalent would lead to sub-optimal numbers of people responding to the bicycling questions. This survey therefore will engage in modest over-sampling of respondents 16 through 39.

Table 2 shows the planned distribution of respondents across the landline sample, the cell phone sample, and a fixed landline over-sample numbering 500 respondents. The expected size of the 16 through 39 age group in the landline and cell phone samples is derived from figures obtained from NHTSA's recently conducted National Survey of Speeding Attitudes and Behavior, which used methods similar to those that will be used for the 2012 National Survey of Pedestrian and Bicyclist Attitudes and Behavior. Adding a 500 interview over-sample will raise the number of people 16 through 39 participating

Table 2. Distribution of Sample

Sample	Total n	Total Percentage	16-39 n	16-39 Percentage
Landline	5,306	58.9%	880	27.7%
Cell Phone	3,194	35.5%	1,794	56.5%

[4] Royal, Dawn and Darby Miller-Steiger. *National Survey of Bicyclist and Pedestrian Attitudes and Behavior. Volume II. Findings Report.* DOT HS 810 972. U.S. Department of Transportation, National Highway Traffic Safety Administration. August 2008. See for example pages 15, 21,106, and 112. Located at http://ntl.bts.gov/lib/30000/30200/30297/810972.pdf

Over-sample	500	5.6%	500	15.8%
Total	**9,000**	**100%**	**3,174**	**100%**

in the survey to a projected 3,174 respondents. The over-sample will be conducted exclusively on landline phones using an independent RDD sample. However, the Table shows that the majority of interviews conducted with young adults will still be by cell phone (1,794).

Sampling methods applied to the over-sample will be the same as those applied to the cross-sectional landline sample. The difference will be that the age screening will be for ages 16 through 39 instead of 16 and older. The allocation of sample by region for the young adult sample would be proportionate to the regional distribution of the age 16 through 39 landline telephone population.

d) Precision of sample estimates

The objective of the sampling procedures described above is to produce a random sample of the target population. This means that with a randomly drawn sample, one can make inferences about population characteristics within certain specified limits of certainty and sampling variability.

The margin of error, d, of the sample estimate of a population proportion, P, equals:

$$d = (t_{\alpha})se(P)$$

Where t_α equals 1.96 for $1\text{-}\alpha = 0.95$, and the standard error of P equals:

$$se(P) = \sqrt{deff}\sqrt{\frac{P(1-P)}{n}}$$

Where:

deff = design effect arising from the combined impact of the random selection of one eligible individual from a sample household, and unequal weights from other aspects of the sample design and weighting methodology, and

n = the size of the sample (i.e., number of interviews)

Using these formulas, the margin of error for a sample size of 9,000 interviews is $d = 0.013$ to 0.015, using an average deff of 1.5 to 2.0. and setting P equal to 0.50. The Pew Research Center, the Kaiser Family Foundation and AP/GfK report average design effects for their overlapping dual frame sample designs of 1.39, 1.33 and 1.68,

respectively (Lambert et al. 2010[5]), making it reasonable to assume that the average design effect will be around 1.5 to 2.0.

Data from the 2002 National Survey of Pedestrian and Bicyclist Attitudes and Behavior suggest that we will obtain between 2,300 and 2,400 bicyclists from a sample of 9,000. Using the formulas above, the margin of error would be approximately $d = 0.025$ to 0.029. The total sample size for the survey is also large enough to permit estimates for other subgroup analyses including by age, gender, income and other demographics.

An important part of the analyses will be comparison of the 2012 and 2002 data. Details on the minimum detectable effects (with 80% power, alpha = .05) for key comparisons between the 2012 and the 2002 survey are provided in Appendix H.

e) Sample weighting – Landline sample

Starting with the calculation of the weights for the landline RDD sample, the base sampling weight will be equal to the reciprocal of the probability of selecting the household's telephone number in its geographic sampling stratum. Thus, the base sampling weight will be equal to the ratio of the number of telephone numbers in the 1+ working banks for that geographic stratum to the number of telephone numbers drawn from those banks and actually released for use. A base sampling weight adjustment will then be implemented to compensate for oversampling of households containing respondents aged 16 through 39 in the independent landline RDD sample. Next, the reciprocal of the number of voice-use landline telephone numbers in the household (up to a maximum of three such numbers) will be calculated. It will compensate for a household's higher probability of selection when it has multiple voice-use residential telephone lines. As noted earlier we will compensate for the exclusion of nontelephone households by using an interruption in landline telephone service adjustment. A within-household weighting factor will then be created to account for the random selection of one respondent from the household. For the independent oversample the within household weight factor is the number of persons age 16-39 years in the household. The design weight then equals the product of each of the weight components described above.

An adjustment for unit non-response will be based on at least two cells in each of the Regions. The non-response adjustment cells will be defined by whether the respondent or nonrespondent (person who was selected for an interview but did not complete it) resides in a directory-listed number household. Other telephone exchange variables may be considered in the development of the unit non-response adjustments. The non-response adjusted (NRA) weight equals the design weight times the non-response adjustment factor. For example, the nonresponse adjustment for the directory listed sample is:

[5] Lambert D, Langer G, McMenemy M. 2010. *"CellPhone Sampling: An Alternative Approach"* Paper presented at 2010 AAPOR Conference, Chicago IL.

$$NRA\ WT_i^{DLR} = DESIGN\ WT_i^{DLR}\left(\frac{\sum BSW_i^{DLR} + \sum BSW_i^{DLNR}}{\sum BSW_i^{DLR}}\right)$$

where

$NRA\ WT_i^{DLR}$ = the nonresponse adjusted weight for the i-it respondent with a directory listed telephone number,

$DESIGN\ WT_i^{DLR}$ = the design weight for the i-th respondent with a directory listed telephone number,

$\sum BSW_i^{DLR}$ = sum of the base sampling weights for respondents with a directory listed telephone number, and

$\sum BSW_i^{DLNR}$ = sum of the base sampling weights for nonrespondents with a directory listed telephone number.

f) Sample weighting – Cell phone sample

For the cell phone sample the first step is to assign a base sampling weight that equals the reciprocal of the probability of selection of the cellular telephone number. The base sampling weight will be calculated separately for each NHTSA region. The cellular phone will be treated as a personal device so there will be no random selection of a respondent. The design weights will then be adjusted for unit non-response separately for cell phone-only respondents and for cell mostly respondents. At this point we have non-response adjusted weights for the cell phone sample.

g) Sample weighting – Combined sample

The first step involves compositing the cell mostly respondents from the landline sample with the cell mostly respondents from the cell phone sample. The compositing factor, λ, will be calculated using the effective sample sizes for the two cell mostly samples. In the final file used for weighting, all contacted households from the landline cross-section sample, the cell phone sample, and the landline oversample will be included. It is important to include all contacted households, rather than only completed interviews since the contacted households provide useful information for weighting purposes.

The population control totals for the raking of the combined sample will be based primarily on the latest Census Bureau population estimates (age by gender by race/ethnicity) and the American Community Survey (ACS) to develop some of the population control totals for socio-demographic variables such as education. Battaglia et al. (2008)[6] have shown the potential for strong bias reduction in RDD samples from the

[6] Battaglia, M.P., Frankel, M.R. and Link, M.W. 2008**.** Improving Standard Poststratification Techniques for Random-Digit-Dialing Telephone Surveys**.** *Survey Research Methods*. Vol 2, No 1.

inclusion of socio-demographic variables in the raking. We will also use the results of the nonresponse bias analysis to inform which variables should be included in the raking.

One might rake the combined sample to the socio-demographic control totals discussed above, but it is important to also have raking control totals related to telephone usage: number of persons age 16 years and older that only have landline telephone service, number of persons age 16 and older that only have cellular telephone service, and number of persons age 16 years and older that have dual (landline and cellular) telephone service, with this last group split in cell-mostly and not cell-mostly subgroups. Such estimates can be obtained for the U.S. from the National Health Interview Survey (NHIS).

To develop the final weights, the project statistician will rake the non-response adjusted weights, iterating to get close agreement on the margins for NHTSA region, gender by age, race/ethnicity, education, and telephone usage group. The latest version of the Izrael, Battaglia, Frankel (IBF) SAS raking macro (Izrael et al. 2004, see Appendix I)[7] will be used to implement the raking. The project staff will also examine whether any additional margins should be included in the raking and if weight trimming is appropriate. Statistics related to the distribution of the weights after raking, such as the coefficient of variation, will be examined to determine if weight trimming should be used to reduce variability in the final weights. The IBF SAS raking macro allows for weight trimming along four dimensions:

1. Global low weight cap value factor: Mean raking input weight times A.
2. Global high weight cap value factor: Mean raking input weight times B.
3. Individual low weight cap value (ILCV) factor: Respondent's weight times C.
4. Individual high weight cap value (IHCV) factor: Respondent's weight times D.

For the global factor, the raking input weights are ratio-adjusted to add to the population size. Then the mean raking input weight is calculated. During the raking process the weights are not allowed to exceed the global high weight cap value which for this survey is the mean weight times 10. The individual weight cap factor is simply the raking input weight for a respondent times the value given. For example, during the raking iterations, the weight of a respondent is not allowed to exceed the value chosen (individual weights times the factor).

The IBF SAS raking macro carries out the weight trimming based on the four weight trimming criteria during the raking iterations. Assuming convergence is achieved, the weights are therefore trimmed and the weighted sample is in very close agreement with the population control margins.

[7] Izrael, D., Battaglia, M.P., and Frankel, M.R. 2009. Extreme Survey Weight Adjustment as a Component of Sample Balancing (a.k.a. Raking). Paper 247-2009, *Proceedings of SUGI 34* 2009.

Based on research conducted by Michael Battaglia and Martin Frankel of Abt Associates Inc. for the CDC on the Behavioral Risk Factor Surveillance System, our initial plan for weight trimming using the IBF SAS raking macro and recommended factor values is as follows:

1. Global low weight cap value factor: Mean raking input weight times 0.1.
2. Global high weight cap value factor: Mean raking input weight times 10.0.
3. Individual low weight cap value (ILCV) factor: Respondent's weight times 0.2.
4. Individual high weight cap value (IHCV) factor: Respondent's weight times 5.0.

h) Variance estimation

Variance estimation will need to rely on statistical software such as SUDAAN; however, existing software will not directly calculate standard errors for overlapping dual frame designs. Rather, it will be necessary to calculate the variance for each of the two samples and use the results to calculate the variance for the combined sample.

The dual frame composite estimator[8] commonly used for fully-overlapping dual frame telephone samples is:

$$\hat{y} = \hat{y}_{LL-O} + \hat{y}_{CELL-O} + \lambda\hat{y}^{LL}_{DUAL} + (1 - \lambda)\hat{y}^{CELL}_{DUAL}$$

$\hat{y}$ is the composite (also referred to as the average) estimator. LL designates the landline sample and CELL designates the cellular sample. LL-O, CELL-O, and DUAL identify the landline-only, cell-only, and dual service domains, respectively. Thus, $\hat{y}_{¿-O}$ and $\hat{y}_{CELL-O}$ are the estimates from the two non-overlapping domains. $\hat{y}^{LL}_{DUAL}$ and $\hat{y}^{CELL}_{DUAL}$ are the estimates for the overlap domain from the landline and cell samples, respectively. λ is the compositing factor for the overlap component ($0 \leq \lambda \leq 1$).

For our survey there is a partial overlap. The DUAL domain is defined as dual users who are cell mostly. The non-overlap domain in the landline sample includes dual users who are *not* cell mostly in addition to landline only.

Most dual frame telephone surveys conducted today use the composite estimator and treat λ as fixed.[9] If we treat λ as fixed then the variance of $\hat{y}$ can be written as:

[8] Lohr, S. 2009. Multiple Frame Surveys. In *Handbook of Statistics: Sample Surveys - Design, Methods and Applications*, (Ed., D. Pfeffermann) Elsevier, Amsterdam.

[9] Brick, J.M., Cervantes, I.F., Lee, S. and Norman, G. 2011. Nonsampling errors in dual frame telephone surveys. *Survey Methodology*, Vol 37, Number 1, June 2011, 1-12.

$$\text{var}(\hat{y}) = \text{var}(\hat{y}_{LL-O} + \lambda\hat{y}^{LL}_{DUAL}) + \text{var}(\hat{y}_{CELL-O} + (1-\lambda)\hat{y}^{CELL}_{DUAL}),$$

i.e., the sum of the variances of the two separate samples (and therefore no covariance term).

The first component will be estimated for the landline sample using SUDAAN by taking the complex nature of the sample design into account. The second component will be estimated for the cellular sample also using SUDAAN by taking the complex nature of the sample design into account.

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

a) Call strategy

The National Survey of Pedestrian and Bicyclist Attitudes and Behavior (NSPBAB) will be administered to a randomly selected national sample of 9,000 respondents age 16 and older. Approximately 5,306 interviews will be conducted with a landline cross-sectional sample, 3,194 interviews conducted with a cell phone cross-sectional sample, and 500 interviews conducted with a landline over-sample of individuals age 16 through 39.

Interviewing will be conducted according to a schedule designed to facilitate successful contact with sampled households and complete interviews with the designated respondent within those households. Initial telephone contact will be attempted during the hours of the day and days of the week that have the greatest probability of respondent contact. This means that much of the interviewing will be conducted between 5:30 p.m. and 9:30 p.m. on weekdays; between 9:00 a.m. and 9:30 p.m. on Saturdays; and between 12:00 noon and 9:30 p.m. on Sundays. Since interviewing will be conducted across several time zones, the interviewing shift lasts until 12:30 a.m. at night. Since the interviews are expected to average about twenty minutes in length, the last contact attempt will not be made later than 9:00 p.m. according to local time unless scheduled by the respondent.

The 2012 NSPBAB will include daytime calling within its calling algorithm. Research has demonstrated that a variation in calling times can reach subsamples disproportionately. For the general population, evenings have the highest contact rate as opposed to weekdays and weekends. But for low income groups and minorities, daytime calls increase the probability of contacting the respondent. Lower income respondents are more likely to have irregular work hours (e.g. shift work, etc.) and sometimes have lower labor market participation, which places them at home during the day.

The NSPBAB will employ a 10-10 call strategy for landline numbers where up to 10 call attempts will be made to ringing but unanswered numbers before the number is classified as a permanent no answer and up to 10 additional calls once contact is made with someone in the household. Callbacks to unanswered numbers will be made on different days over a number of weeks according to a standard callback strategy. If contact is made but the interview cannot be conducted at that time, the interviewer will reschedule the interview at a time convenient to the respondent. If someone picks up the phone, but terminates the call before in-house selection of a designated respondent can be made, then the Contractor will apply a protocol to maximize participation while minimizing respondent annoyance. Fewer calls will be made to cell phone numbers. The calling strategy there will be 5 call attempts made to ringing but unanswered numbers before the number is classified as a permanent no answer and up to 5 additional calls once contact is made with the respondent.

b) Protocol for answering machines and voice mail

When contact is made with an answering machine or voice mail, a message will be left according to a set protocol. For landline numbers, a message will be left on the 5th, 7th, and 9th attempts to contact a household member. The message will explain that the household had been selected as part of a national USDOT study, ask that the Contractor's toll-free number be called to schedule an interview, and include reference to the NHTSA web site which will include information about the survey so that prospective respondents can verify the survey's legitimacy. For cell phone numbers, the message will be left on the 2nd attempt so that the cell phone user can quickly attach an identity to the number s/he is seeing on the phone.

c) Initiating the interview

When a household is reached in the landline sample, the interviewer will screen for age eligibility. If only one household member is age eligible, then the interviewer will seek to interview that individual. If there is more than one eligible household member, then the interviewer will randomly select one respondent from among them and seek to interview that person. Appointments will be set up with respondents if it is inconvenient for them to be interviewed at the time of contact. If the randomly selected respondent is not available at the time of contact, then the interviewer will ask what would be a good time to call back to reach that person.

For the cell phone sample, respondents will first be asked if they are in a situation where it would be unsafe to speak with the interviewer. If the respondent says "Yes," then the interviewer will say that he or she will call back at another time, and immediately terminate the call. Once a cell phone user is reached at a safe time, the interviewer will first screen for age eligibility. If the cell phone user is eligible to participate, then the interviewer will seek to proceed with the interviewer. If it is an inconvenient time for the respondent, then the interviewer will try to set up an appointment.

d) Spanish language interviews

If a non-bilingual interviewer encounters a language barrier at the telephone number, either with the person answering the phone or with the designated respondent, the interviewer will thank the person and terminate the call. If a case is designated as Spanish language, it will be turned over to the next available Spanish-speaking interviewer. The Spanish-speaking interviewer will utilize a Spanish version of the questionnaire. The translation of the questionnaire from English to Spanish will have been checked prior to the survey field period using back translation methods.

e) Interviewer monitoring

Each interviewer will be monitored throughout the course of the project. The Monitor will evaluate the interviewer on his or her performance and discuss any problems that an interviewer is having with the Shift Supervisor. Before the end of the interview shift, the Monitor and/or Shift Supervisor will discuss the evaluation with the interviewer. If the interviewer cannot meet the Contractor's standards, he or she will be removed from the project.

All interviewers on the project will undergo two types of monitoring. The Study Monitor will sit at a computer where s/he can see what the interviewer has recorded, while also audio-monitoring the interview. The audio-monitoring allows the Supervisor to determine the quality of the interviewer's performance in terms of:

1) Initial contact and recruitment procedures;

2) Reading the questions, fully and completely, as written;

3) Reading response categories, fully and completely, (or not reading them) according to study specifications;

4) Whether or not open-ended questions are properly probed;

5) Whether or not ambiguous or confused responses are clarified;

6) How well questions from the respondent are handled without alienating the respondent;

7) Avoiding bias by either comments or vocal inflection;

8) Ability to persuade wavering, disinterested or hostile respondents to continue the interview; and,

9) General professional conduct throughout the interview.

The Supervisor will also monitor the interviewer's recording of survey responses as the Supervisor's screen emulates the interviewer's screen. Consequently, the

Supervisor can see whether the interviewer enters the correct code, number or verbatim response.

B.3. Describe methods to maximize response rates and to deal with issues of non-response.

a) Cognitive testing

The questionnaire was cognitive tested with nine people in order to identify any problem issues that could deter respondents from completing the survey. Based on the cognitive testing results (see Appendix J), a number of questions were revised or deleted. This culminated in the updated questionnaire previously referenced as Appendix F.

b) Initial contact

An advance letter will be sent to those sample members for whom an address can be identified. This letter will encourage participation by stressing the value of the survey and how the results will be used (see Appendix K).

The initial phone contact with the designated respondent is crucial to the success of the project. Most refusals take place before the interviewer has even completed the survey introduction (usually within the first 30 seconds of the call). Numerous studies have shown that an interviewer's manner of approach at the time of the first contact is the single most important factor in convincing a respondent to participate in a survey. Many respondents react more to the interviewer and the rapport that is established between them than to the subject of the interview or the questions asked. This positive first impression of the interviewer is key to securing the interview.

While the brief introduction to the study concerning its sponsorship, purpose and conditions are sufficient for many respondents, others will have questions and concerns that the interviewer must address. A respondent's questions should be answered clearly and simply. The interviewers for the NSPBAB will be trained on how to answer the most likely questions for this survey. A hard copy of Frequently Asked Questions will be provided to each interviewer at their station (see Appendix L). The interviewers also will be trained to answer all questions in an open, positive and confident manner, so that respondents are convinced of the value and legitimacy of the study. If respondents appear reluctant or uncertain, the interviewer will provide them with a toll-free number to call to verify the authenticity of the survey. Above all, the interviewer will attempt to create a rapport with the study subject and anyone in the household contacted in the process of securing the interview.

c) Refusal documentation and tracking

Higher response rates can be achieved through procedures built on careful documentation of refusal cases. Whether the initial refusal is made by the selected respondent, a third party or an unidentified voice at the end of the phone, the CATI

system used for this survey will be programmed to move the interviewer to a non-interview report section. The interviewer will ask the person refusing to tell him/her the main reason that he or she doesn't want to do the interview. The interviewer will enter the answer, if any, and any details related to the refusal into the program to assist in understanding the strength and the reason for the refusal. This information, which will be entered verbatim in a specific field in the CATI program, will be part of the formal survey reporting record, and will be reviewed on a daily basis by the Field Manager and, on a weekly basis, by the Project Director. Based on the pattern and distribution of refusals, a strategy for refusal conversion will be formalized.

The Survey Manager and the Project Director will review the information about refusals and terminations on the CATI system on an ongoing basis to identify any problems with the contact script, questionnaire or interviewing procedures that might contribute to non-participation. For example, they will scrutinize the distribution of refusals by time of contact, as well as any comments made by the respondent, in order to determine whether calling too early or too late in the day is contributing to non-participation. Also, the refusal rate by interviewer will be closely monitored.

In addition to relying on the CATI data records, the Project Director and Survey Manager will also consult with the interviewing Shift Supervisor, who has monitored the interviewing and debriefed the interviewers. The information from these multiple sources provides solid documentation of the nature and sources of non-response. It will also help to generate the "conversion script" which will be developed by the Project Director.

d) Refusal conversion

Prospective respondents that refuse to participate in the survey will be re-contacted approximately one to two weeks after the initial refusal. These initial refusals will be asked to reconsider and participate in the survey. People identified as "hard" or "not to be converted" refusals will not be included in this refusal conversion effort.

The process of converting terminations and refusals, once they have occurred, will involve the following steps. First, there will be a diagnostic period as described in the documentation and tracking section above. Second, after enough time has passed to see a large enough sample of refusals and terminations, the Project Director and his staff will work out a refusal conversion script. Third, the refusal conversion effort will be fielded with re-interview attempts scheduled about one-to-two weeks after the initial refusal (it is important to have a "cooling off" period before re-contacting the household). Conversions of interviews that are more than half complete would not be delayed this long. Fourth, the Project Director and Operations Manager will receive the outcomes of the refusal conversion efforts on a daily basis. Revisions of the script or the procedures would be made if indicated by the ongoing results of the conversion effort.

Past experience suggests that about 20% of initial refusals will reconsider. Success in refusal conversion tends to depend more heavily on the difficulty of the survey and the study population, i.e., the hardness of the initial refusals, than it does with the level of effort on refusal conversions. Nonetheless, sensitivity in the diagnosis of reasons

for non-participation and creativity in responding to those problems can have a real effect on refusal conversion.

e) NHTSA web site information

It is not uncommon for behavioral self-report surveys such as these to generate incoming calls or emails to the U.S. Department of Transportation from people who have been contacted by the survey. In a large percentage of the cases, it's an inquiry to check its legitimacy. People are wary about attempts to extract information from them, wondering if the caller is truthful in identifying the source of the survey and the use of their information.

To help allay these fears, NHTSA will place on its web site information that prospective respondents can access to verify the survey's legitimacy (see Appendix M). The interviewers will provide concerned respondents with a NHTSA web address where they will find information on the source of the survey and why it is important for them to participate. An 800 number to reach the survey Contractor will also be provided so they can schedule an interview. This procedure was used with the 2007 Motor Vehicle Occupant Safety Survey. More than 700 people who participated in that survey indicated in an ending question that they had gone to the NHTSA web site to get more information about the survey, suggesting that this was an effective approach to raising the response.

f) Non-response bias analysis

A non-response bias study will be conducted for the 2012 NSPBAB. Non-response is the failure to obtain survey measures on all sampled individuals. There are two main potential consequences from non-response: (1) non-response bias and (2) underestimation of the standard errors. The main challenge posed by non-response is that without a separate study (i.e., a non-response bias analysis), one can never be sure of whether the non-response has introduced bias or affected the estimation of the standard errors. The mere presence of non-response, even high non-response, does not necessarily imply bias or non-response error.

The non-response bias study will entail two analyses. First, the demographic composition of the 2012 NSPBAB landline sample will be compared to the demographics of non-respondents as determined by auxiliary data providing demographic information by area code/prefix combination for the non-respondent phone numbers. The auxiliary data will be purchased from an external vendor and will represent the latest and most recent population information at the zip code level. The analysis will assess whether the landline non-respondents are demographically different from those in the sample and whether some of the estimates are correlated with the propensity to respond. If some of the estimates are correlated with the propensity to respond, there is a chance that there might be some bias in those estimates.

Additionally, there will be a non-response follow-up (NRFU) survey with a sub-sample (n=200) of households that failed to respond in either the landline or cell phone sample. The NRFU has three main objectives: 1) investigate the reasons for non-response, 2)

assess the risk of non-response error to model estimates, and 3) identify changes to the survey protocol that may help to reduce the non-response rate and/or the risk of non-response bias. Because these are people who either refused to participate in the main survey or otherwise avoided participation, there are limitations on how many respondents one can expect to participate in the NRFU. However, assuming that the sub-sample of 200 is representative of all non-responders, the margin of error of +/- 7 percentage points at the 95% confidence interval would be sufficient to assess differences between non-responders and a main survey sample.

The NRFU features an abbreviated survey instrument in which the household is asked to provide a few basic socio-economic and demographic variables, reasons for non-participation, and responses to a few basic survey items. NRFU households will each be offered $20 for their participation, as this amount has been effective in past non-response follow-up studies conducted by the contractor. The results of the non-response follow-up survey will provide valuable information for assessing the extent to which non-response bias affects the overall survey respondent sample.

The analysis will include comparison of key pedestrian and bicyclist safety variables from the NSPBAB respondents with the same for the NRFU respondents. This will provide insights about the direction and magnitude of possible non-response bias. The project team will investigate whether any differences remain after controlling for major weighting cells (e.g., within race and education groupings). If weighting variables eliminate any differences, this suggests that weighting adjustments will reduce non-response bias in the final survey estimates. If, however, the differences persist after controlling for weighting variables, then this would be evidence that the weighting may be less effective in reducing non-response bias.

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

a) Testing the programming

The project team will test the CATI program thoroughly in test mode – running the interviewing program through multiple loops. The analytical staff will attempt to test all possible response categories for each question in order to identify embedded logic errors, as well as obvious skip problems. Several analysts may test the program simultaneously to identify problems quickly, and to double check the comprehensiveness of the testing.

b) *Survey pretest*

Pre-test interviews will be conducted with 15 respondents prior to the formal administration of the survey in order to test all of the survey systems, the survey instrument, and the CATI programming. The pretest will be conducted under conditions identical to the planned conditions for the final administration of the survey. It will be used to evaluate the wording of questions and response categories in terms of clarity and confusion for respondents; assess the flow of the interview, including question sequencing and skip patterns; check the time required to administer the questionnaire to

see if it differs from the original estimate; assess item non-response to see if there are problems with particular questions; ensure that a smooth interview occurs with previous information guiding later questions; ensure random rotation of specified questions; check that invalid responses are not entered; and review and evaluate the full range of procedures.

Immediately following the interviews, the interviewers will participate in a debriefing session to report any problems or observations they have regarding the instrument to the field manager. The field manager will consolidate the observations and suggestions along with his own recommendations. These will be returned to the project director along with the hard copy of the pretest interviews. Interview timings will be provided as well. The completed questionnaires will be analyzed, along with the extensive comments on the content and administrative aspects of the instruments made by the interviewers, monitors and research staff observing the pretest. Revisions to procedures, CATI programming, and the survey instrument will be made as appropriate.

B.5. Provide the name and telephone number of individuals consulted on statistical aspects of the design.

The following individuals have reviewed technical and statistical aspects of procedures that will be used to conduct the 2012 National Survey of Pedestrian and Bicyclist Attitudes and Behavior:

John Boyle, Ph.D.
Executive Vice President
Abt SRBI, Inc.
8403 Colesville Road, Suite 820
Silver Spring, MD 20910
(301) 608-3883

Paul Schroeder, MA
Senior Analyst
Abt SRBI, Inc.
8403 Colesville Road, Suite 820
Silver Spring, MD 20910
(301) 608-3883

Mike Battaglia, MA
Vice President
Abt Associates
55 Wheeler Street
Cambridge, MA 02138
(617) 349-2425

Alan Block, MA
Office of Research and Technology
DOT/National Highway Traffic Safety Administration

400 Seventh Street, SW
Washington, DC 20590
(202) 366-6401