

Supporting Statement for Information Collection Request
Part B

NHTSA Tire Fuel Efficiency
Consumer Information Program Survey Research Plan

Submitted: October 26, 2011

Contents

Part B.....3

B1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection methods to be used.....5

B2. Describe the procedures for the collection of information.....9

B3. Describe methods to maximize response rates and to deal with issues of non-response.....10

B4. Describe any tests of procedures or methods to be undertaken.....11

B5. Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.....12

Part B

NHTSA is seeking approval to conduct an internet-based consumer survey in order to gather data regarding the tire purchase mindset, evaluate consumer understanding of potential replacement tire rating systems and determine the proper channels that will be used to communicate information regarding tire fuel efficiency, wet traction, and tread wear.

The online survey will be conducted one time with a sample of 4,000 adult drivers ages 18 and older who currently own or lease a vehicle. In addition to these respondents, an oversample of up to 1,200 respondents who have either purchased a replacement tire in the last 6 months or plan to purchase a replacement tire in the next 6 months will be conducted in order to compare responses of the general consumer to one in the tire purchase mindset. We expect to use these responses from recent or planned tire purchasers as additional support for the rating system ultimately recommended for this program. While we do not expect this group's responses to greatly diverge from the initial sample, this data can provide additional insights that will help us recommend one as superior in the event that two or more labels are rated similarly on comprehension and usefulness by the general sample.

Currently, we have four (4) different ratings systems that will be tested for consumer comprehension to ensure the information being presented is properly understood. With a sample size of 4,000, we will be able to evaluate all the proposed scales using monadic cells. Each scale will be evaluated by 1,000 respondents to ensure the ratings of that particular rating are not biased by the other ratings. Each of the cells of 1,000 respondents will be comprised of identical proportions of respondents by gender, age and region. In monadic testing, each respondent will be randomly assigned to a group and will be exposed to one rating scale in order to evaluate this scale on a variety of metrics, including conceptual understanding, comprehension of actual ratings and potential for use.

Data Analysis Plan

During data analysis, we will use a variety of statistical techniques, including independent t-tests between two ratings as well as ANOVA which will allow us to test for variance between evaluations on key measures. Using these techniques, we will be able to understand which rating or ratings stand out as rated better in terms of likelihood to use, providing the information needed to inform a purchase and reported comprehension. ANOVA will be used to compare mean ratings of all 4 labels to minimize statistical error, and if there are two labels that are rated similarly, we will use t-tests to compare the differences between the two independent evaluations and gauge whether or not a significant difference exists.

Any questions that were answered prior to the presentation of the stimulus (i.e., the rating images) can be viewed in aggregate and thus providing data from all 4,000 respondents on the tire purchase process, channels of communications and sources of information. When analyze the rating evaluation section, however, the four groups of respondents will be treated as independent groups and data resultant from the evaluation of stimulus will not be analyzed in aggregate. Instead, we will compare results across the independent groups to determine if the stimulus presented resulted in any differences in measures like comprehension, likelihood to use, and other key metrics. We will use summary statistics to compare the evaluations of each rating on four key metrics in order to make a recommendation on which rating is best for this program.

These key metrics are:

- “If these ratings were available, I would use them when researching replacement tires”
- “This rating system gives me the information I need to make my tire purchase decision”
- “This rating system gives me the information I need to compare tires”
- “I don’t understand what these ratings mean for me personally”

Since the results from the four cells of respondents will be analyzed independently, there will be no need to normalize this data.

Additionally, labels will be evaluated to determine whether additional clarifying language assists with comprehension of the each label. A simple version of each label will be presented first to evaluate scale and rating comprehension. This will be followed by the same label with additional clarifying language. Summary statistics will be used to compare ratings for the simple label and the label with additional clarifying language to determine whether this language enhances comprehension or likelihood to utilize the label, and to determine whether the language will be necessary on the final label selected. In addition to summary statistic comparison of these two labels, respondents will answer questions regarding understanding and likelihood to use these labels for analysis if summary statistics do not indicate a meaningful difference.

Power Analysis

In preparing this research plan, we conducted a power analysis utilizing SPSS SamplePower.

In our analysis we plan to compare two groups at a time (Group A/Rating A versus Group B/Rating B) on a 7-point scale for key metrics ranging from Strongly Disagree to Strongly Agree. Because these responses follow a logical sequence, we can assign a numerical score to each response and compute a mean and standard deviation for each group.

The computation of sample size is based on the following assumptions and decisions:

- 1) The expected pattern of responses for Group A corresponds to a mean of 5.51 with a standard deviation of 1.66.
- 2) The expected pattern of responses for Group B corresponds to a mean of 5.30 with a standard deviation of 1.71.
- 3) In computing the sample size we assume that there will be no missing data as respondents will be unable to skip the question.

This power analysis indicates that based on these assumptions, with a sample of 1,012 subjects per group we will be able to detect differences between Group A and Group B given the assumptions above with 80% power and alpha of .05. Therefore, our recommended sample size of 1,000 per group should allow us to conclude that the mean response differs for Group A versus Group B, and so on if these assumptions are correct.

In addition, an oversample of recent or planned tire purchasers will also be conducted using monadic cells. After the 4,000 general population respondents have completed the survey, only those respondents who have purchased replacement tires in the last 6 months or plan to purchase replacement tires in the

next 6 months will qualify for the remainder of the study. In total, we will collect up to 1,200 responses in this oversample, to ensure 300 recent or planned tire purchasers have evaluated each rating unbiased.

We conducted another power analysis to understand if this sample size would suffice in providing comparisons against the general population respondents. This computation of sample size was conducted with the following assumptions:

- 1) Consumers in the purchase mindset are more likely to not select a neutral response due to the relevance of the subject matter.
- 2) The expected pattern of response for Group A corresponds to a mean of 5.51 with a standard deviation of 1.66.
- 3) The expected pattern of response for Group B corresponds to a mean of 5.86 with a standard deviation of 1.36.
- 4) In computing the sample size we assume that there will be no missing data as respondents will be unable to skip the question.

This power analysis indicates that based on these assumptions, with a sample of 297 subjects per group we will be able to detect differences between Group A and Group B given the assumptions above with 80% power and alpha of .05. Therefore, our recommended sample size of 300 per group for this analysis should allow us to conclude that the mean response differs for general respondents versus recent or planned tire purchasers evaluating the same rating.

Ultimately, the purpose of this survey is to understand the strengths and weaknesses of the potential ratings systems in terms of consumer comprehension and usefulness to the purchase process.

B1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection methods to be used.

For reasons explained in Part A of this supporting statement, NHTSA will conduct an online survey to test potential replacement tire rating designs, which were previously tested during focus group and tire retailer research, with a broader audience. This methodology was chosen as it provides respondents with the opportunity to evaluate images, which is essential for the rating evaluation portion of the study. Since these graphical treatments contain English-language text, the survey will only be available in English to ensure respondents understand the information being displayed. Issues of non-response due to this methodology are addressed in section B3.

The sample will be built using members of an online panel managed by Authentic Response (www.authenticresponse.com), which provides NHTSA with a potential respondent universe of 2 million U.S. consumers. This panel has been built using a passive panel sourcing process. Authentic Response works with publishers and online service providers to embed recruitment efforts into their websites. Potential panel members will only be admitted into the panel after a validation process is complete. For quality purposes, all panel members are de-duplicated at the registration stage and a continuous system of validation checks is in place to prevent professional survey takers from entering studies.

Upon panel registration, members are asked demographic and lifestyle questions so that surveys can be targeted. Using this information, the survey will be e-mailed to U.S. panel members ages 18 and older.

We expect approximately 50,000 invitations to be sent to randomly selected panel members on a rolling basis throughout the fielding of the study in order to obtain the 4,000 respondents necessary to complete the general study. Depending on the actual number of respondents needed to complete the oversample, we expect up to 40,000 additional invitations to be sent to randomly selected panel members after the general consumer portion of the study is complete.

Before entering the questionnaire, each respondent will be screened based on the criteria listed below. If qualified, the respondent then enters the survey and begins answering questions relevant to the needs of this study. A full text of the draft survey questionnaire is included in this package as Appendix B.

Upon completion, the respondent will be redirected to the panel provider’s website. This survey will be a one-time implementation, and respondent identification information will be removed prior to data analysis to preserve anonymity.

The sample will be built to consist of percentages of US drivers ages 18+ by age, gender and geographic location proportionate to the data provided in the Federal Highway Administration’s 2009 Highway Statistics. We will impose hard quotas for age ranges, gender, and regions and soft quotas for states within each region.

Table 5. Region Definitions

Region	States
Deep South	Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, South Carolina
Farm Belt	Iowa, Kansas, Missouri, Nebraska, North Dakota, South Dakota
Great Lakes	Illinois, Indiana, Michigan, Minnesota, Ohio, Wisconsin
Mid Atlantic	Delaware, District of Columbia, Maryland, New Jersey, New York, Pennsylvania, West Virginia
Mountain	Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming
New England	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
Outer South	Kentucky, North Carolina, Oklahoma, Tennessee, Texas, Virginia
Pacific	Alaska, California, Hawaii, Oregon, Washington

Building this sample will require five main steps:

- 1) Respondents will first have to qualify as adults ages 18 and older.
- 2) Next, respondents must qualify as someone who does not work or have a family member working in the automotive industry, the tire manufacturing or retail industry, or the market research industry.
- 3) Next, respondents must qualify as a licensed driver.
- 4) Next, respondents must qualify as a vehicle owner or lessee.
- 5) Finally, respondents will be screened to ensure they are the primary or shared decision maker in terms of replacement tire purchases.

In order to qualify for the oversample, respondents must qualify through a six-step process:

- 1) Respondents will first have to qualify as adults ages 18 and older.
- 2) Next, respondents must qualify as someone who does not work or have a family member working in the automotive industry, the tire manufacturing or retail industry, or the market research industry.
- 3) Next, respondents must qualify as a licensed driver.
- 4) Next, respondents must qualify as a vehicle owner or lessee.
- 5) Next, respondents must qualify as having purchased replacement tires in the last six (6) months or planning to purchase replacement tires in the next six (6) months.
- 6) Finally, respondents will be screened to ensure they are the primary or shared decision maker in terms of replacement tire purchases.

Once a respondent qualifies for the study, s/he will be randomly assigned to one of four cells or test groups. One version of this survey has been developed for each group, identical in every way except that each of the groups will see only one of the rating system designs. This same version of the survey will be used in both the initial data collection as well as the oversample collection.

To collect data on understanding of the ratings, each respondent will be asked to evaluate a particular rating for ease of understanding and will be given the opportunity to indicate which areas of the graphical rating, if any, cause confusion. Respondents will also be shown three tires rated using a particular scale with all attributes constant except those used on the rating scale (i.e., wet traction, tread wear and rolling resistance) and asked to decide which tire s/he would prefer to purchase. Since individual differences are randomly distributed across conditions, NHTSA plans to use standard statistical techniques to test observed effects between different designs.

In order to develop a recommendation for which tire label is most effective, NHTSA will use parametric testing to analyze the differences between independent evaluations of each label on four key metrics:

- “If these ratings were available, I would use them when researching replacement tires”
- “This rating system gives me the information I need to make my tire purchase decision”
- “This rating system gives me the information I need to compare tires”
- “I don’t understand what these ratings mean for me personally”

In addition to analyzing the differences across these key metrics, we will also analyze the summary statistics from the respondents’ interactions with and evaluations of the ratings as presented on screen. The interactive activities will yield heat maps demonstrating those areas most often indicated as having drawn respondent attention or caused confusion to make recommendations for any potential design changes or consumer education needs. Follow-up questions, which will explore interpretations of the ratings presented to respondents, and a mock purchase decision, which asks respondents to select which tire they will purchase based only on the information provided in the ratings and explain their decision, will add additional context as to how respondents may use these ratings in their purchase decision and indicate whether or not any incorrect assumptions are being made.

Sample Quotas

Based on the Federal Highway Administration’s 2009 Highway Statistics, there are approximately 205.8 million licensed drivers ages 18 and older. Since we are using a convenience sample for this study, we plan to use the below statistics to set gender, age and region quotas for our sample of 4000.

Table 6. Expected Sample Distribution by Key Demographics¹

Total US Driver Sample (4000)

Gender	Total Population	% of Licensed Drivers 18+	Distribution of Sample
Male	102,320,242	49.7%	1,989
Female	103,461,215	50.3%	2,011
Age			
18-24	23,538,942	11.4%	458
25-34	36,326,817	17.7%	706
35-44	38,158,133	18.5%	742
45-54	41,665,892	20.2%	810
55-64	33,156,841	16.1%	645
65+	32,934,832	16.0%	640
Region			
Deep South	33,614,591	16.3%	653
Farm Belt	10,456,411	5.1%	203
Great Lakes	35,321,339	17.2%	687
Mid Atlantic	31,899,373	15.5%	620
Mountain	14,792,314	7.2%	288
New England	10,718,373	5.2%	208
Outer South	36,390,094	17.7%	707
Pacific	32,588,962	15.8%	633

Total US Driver Sample by Randomly Assigned Group (1000 per group)

Gender	Group 1	Group 2	Group 3	Group 4
Male	497	497	497	497
Female	503	503	503	503
Age				
18-24	114	114	114	114
25-34	177	177	177	177
35-44	185	185	185	185
45-54	202	202	202	202
55-64	161	161	161	161
65+	160	160	160	160
Region				
Deep South	163	163	163	163
Farm Belt	51	51	51	51
Great Lakes	172	172	172	172
Mid Atlantic	155	155	155	155
Mountain	72	72	72	72
New England	52	52	52	52

¹ <http://www.fhwa.dot.gov/policyinformation/statistics/2009/dl20.cfm>

Outer South	177	177	177	177
Pacific	158	158	158	158

Estimated Incidence

In a 2009 survey conducted for this initiative, 80% of general consumers qualified as vehicle owners who are primary or shared decision makers for tire purchases. In addition, 94% of these consumers (or about 75% of general consumers) qualified as having purchased replacement tires for a vehicle at some point, and 19% of this group plan to purchase replacement tires in the next six months (or about 15% of general consumers). Assuming the same number of consumers purchased replacement tires in the last six months, this study indicates that about 29% of general population consumers qualify within screening criteria of the target audience oversample for this study.²

This information provides NHTSA an idea of the size of the population of recent replacement tire purchasers and intended purchasers among the general population. NHTSA will not present data from this survey as representative of the general population. However, it is important for the agency to develop an understanding of the size of the target audience as the consumer information program is developed.

B2. Describe the procedures for the collection of information.

The procedure for the collection of information for the Tire Fuel Efficiency label research is as follows:

- Survey population is defined.
- Survey will be sent via e-mail to online panel members identified as vehicle owners ages 18+.
- The questionnaire will take approximately 15 minutes to complete. All panel members who are over the age of 18 have an equal chance of receiving the survey invitation, but only those who qualify based on additional screening criteria will have the opportunity to complete.
- A total sample size of up to 5,200 will be surveyed, with 4,000 respondents in the general population study and up to 1,200 in the tire purchaser oversample. Respondents will complete the survey on their own time, and will have an option to contact the panel provider if they experience any technical problems with the survey.
- The survey will only be conducted in English. While we understand some Hispanic respondents may be more comfortable with a Spanish-language survey, any label or rating scale images for replacement tires will be presented in English and therefore can only be evaluated by those who can comprehend the language effectively.
- Data tables, including important cross-tabulations, will be prepared along with a final report of the key findings and strategic recommendations.
- We expect an 80% incidence for vehicle owners who are the primary or shared decision maker regarding tire purchases. Based on past research conducted for this initiative, we expect an incidence rate of about 29% for the tire purchaser oversample. This estimate was calculated using the number of respondents in the *Rolling Resistance Survey* who qualified as planning to purchase replacement tires in the next 6 months, and assuming an equal number of consumers have purchased in the last 6 months.

² NHTSA *Rolling Resistance Survey*. Online survey conducted by Strat@comm, August 2009.

- If necessary, NHTSA can also provide a report that demonstrates the number of respondents who either did not qualify or did not elect to complete the survey. This will be calculated by dividing the number of completed responses by the total number of survey invitations sent out.

The sampling approach described above will yield a convenience sample. This sampling methodology results in a study that only includes respondents who have access to computers and the Internet, and are literate. According to Pew Research Center's Internet and American Life Project³, 77% of US adults access the Internet. Because the study is not a probability-based sample, there is no statistical basis to derive unbiased estimates representative of the target population, U.S. passenger vehicle drivers, or to estimate sampling error. The purpose of the study is to examine differences between conditions, not to estimate population parameters.

For the purposes of this study, it is sufficient that the sample be a convenience sample as long as it is diverse in terms of drivers' gender, age and state. The key component of this study is that the respondents be randomly assigned to different conditions. NHTSA believes that the study design will provide the quantitative estimates to effectively inform the selection of a ratings scale and begin the development of a related consumer information program. In addition, this methodology is a cost- and time-effective way of conducting the research necessary to inform this program and is the most feasible methodology that allows for the testing of graphical rating scales.

Additionally, planned communications channels will include web sources and the main source of this fuel economy-related content will be a web property like SaferCar.gov. We will not ask respondents to rank sources, as there may be an inherent bias toward web-sources with these respondents, but we will be obtaining evaluations from a group of consumers who are able to access these sources once published. Therefore, the difference between respondents and non-respondents will likely not impact the guidance of consumer information dissemination.

B3. Describe methods to maximize response rates and to deal with issues of non-response.

NHTSA has taken a number of steps to address issues of non-response:

- The invitation to the survey will not mention the topic of the study to help ensure that our sample is composed of general consumers rather than automotive enthusiasts or consumers interested in tires.
- Reminder e-mails will be sent to participants who have been invited to the study, but have not yet entered the study.
- The survey will be programmed using technologies through which consumers can interact directly with the images presented on screen. This will help allow for easy evaluation of the labels and reduce burden, which will help prevent respondent drop off.
- A monadic design has been used to maintain a 15-minute length of survey. Longer surveys, particularly on technical topics, are more likely to experience respondents dropping out without completing the survey.

³ The Pew Research Center's Internet & American Life Project, November 23-December 21, 2010. n=2,303 adults 18 and older, including 748 reached via cell phone. Interviews were conducted in English and Spanish. Margin of error is +/- 2%.

As detailed in B1 and B2, we are utilizing a convenience sample. In order to ensure diversity in respondents, we will set quotas to ensure the sample consists of percentages proportionate to the US driver target audience by gender, age and region based on the data provided in the Federal Highway Administration's 2009 Highway Statistics. Since our goal in this study is to make comparisons between different ratings, it is essential that respondents are randomly assigned to view these different conditions.

B4. Describe any tests of procedures or methods to be undertaken.

Pilot Test

NHTSA and its contractors will conduct a test of the online survey to ensure all functionality is working properly before the study is sent to potential respondents.

Soft Launch Period

The survey will experience a 'soft launch' period in which a limited number of potential respondents (approximately 5% of the sample) are sent invitations to respond. During this period, we will evaluate the demographic data, completion rate, and time-spent in the survey. Any problems discovered during this test period will be addressed before the survey is sent out to a larger population.

Data Collection and Analysis

Responses will be collected in a computerized database. One version of the online survey has been developed for each group, identical in every way except that each of the groups will see only one of the rating scale designs to be tested. This same version of the survey will be presented to any oversample respondents as well. Responses to survey questions will then be compared between these groups to determine whether there are statistically significant differences among the designs,

Data processing will consist of tabulation of quantitative and coded open-ended responses using UNCLE software. Data analysis will be conducted by NHTSA and its contractors, StrategyOne. Since individual differences are randomly distributed across conditions, NHTSA plans to use standard statistical techniques to test observed effects between different designs. Summary statistics will be analyzed to determine whether or not significant differences exist between the rating evaluations of different groups based on reported comprehension, potential for use, meaningfulness of the ratings and the simulated purchase decision. Open-ended responses will also be analyzed to add context to the evaluations respondents have provided and can help in assessing the ease at which respondents were able to evaluate the rating and make a purchase decision.

B5. Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.

In preparation of sending this package to OMB for approval, NHTSA provided contacts at various agencies the opportunity to comment on the approach for this plan. The following individuals have reviewed technical aspects of this Tires Fuel Efficiency research plan:

Walter Gazda
Senior Economist
617.494.2402

Lisandra Garay-Vega, Ph.D
Industrial Engineer
202.366.0413

Mary Stearns, Ph. D.
Chief, Behavioral Safety Research and
Demonstration Division
617.494.2617

**USDOT/Research and Innovative
Technology Administration/
Volpe National Transportation
System Center**
55 Broadway
Cambridge, M.A. 20590

Hampton Newsome
Attorney, Division of Enforcement,
Bureau of Consumer Protection
202-326-2889

Federal Trade Commission
600 Pennsylvania Avenue, N.W.
Washington, D.C. 20580

Chung-Tung Jordan Lin, Ph. D.
Team Leader
Consumer Studies
chung-tung.lin@fda.hhs.gov

Food and Drug Administration
10930 New Hampshire Avenue
Silver Spring, MD 20993

Kristin Kenausis
SmartWay, Office of Transportation
and Air Quality
Mail Code 6406J
202-343-9225

**Environmental Protection
Agency**
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460

The company selected as a contractor for this study is StrategyOne. StrategyOne is the market research arm of Edelman, Inc. This team has extensive experience in both qualitative and quantitative research practices. More information can be found at their website (<http://www.strategyone.com>). The contact information for this team is as follows:

Jason McGrath

Vice President
202.326.1810

Ellen Rienzi

Senior Project Manager
202.350.6687

Tiffany Scott

Research Assistant
202.336.7976

Anne Treschitta

Research Assistant
202.756.2417

StrategyOne
1875 Eye Street NW
Washington, D.C. 20006