Consumer Research on the Safe Handling Instructions Label for Raw and Partially Cooked Meat and Poultry Products and Labeling Statements for Ready-to-Eat and Not-Ready-to-Eat Products

OMB No. 0583-NEW

Supporting Statement

B. Statistical Methods

B.1. Respondent Universe and Sampling Methods

Web-Based Experimental Study

Respondents for the web-based experimental study will be selected from Lightspeed's consumer panel. Lightspeed maintains an opt-in panel that provides a cross section of the U.S. population. The composition of the panel does not necessarily match the demographics of the U.S. population or a specific target population; rather, each sample is selected so that it mirrors the U.S. population or the specific target population. Lightspeed's consumer panel consists of approximately 1.5 million adults who were double opted-in. The double opted-in process is as follows. When a respondent clicks on a link from a panel ad, s/he is directed to the panel registration survey. Each prospective panelist must provide demographic and household information, pass through some validation checks (e.g., verify postal address), and agree to the website Terms and Conditions and Privacy Policy. Those who pass the Lightspeed checks are sent an email to confirm their email address. After clicking on a link within the email, they complete the double opt-in process and become panel members.

Lightspeed sends randomly selected panel members a study participation invitation via email. The email includes a short description of the study and instructs respondents to log-on to their password-protected panel home page to access the survey. For this study, approximately

70,000 English- and Spanish-speaking panel members will be sent an email invitation to the survey (see Appendix A). Interested panelists will be screened to ensure they meet the following criteria (see Appendix B):

- are 18 years of age or older;

- cook raw meat or poultry at home at least once per week;

- have not cooked or worked professionally in a food preparation setting within the past 5 years; and

- have not taken any type of food safety training, such as ServSafe, in the past 5 years.

Inbound quotas will be set to obtain a sample that mirrors the U.S. population (see Table B-1).

The study will include 3,600 participants with approximately $n = 133$ exposed to each of 27 SHI labels created by fully crossing the three primary study features—label shape, safe handling instruction text, and safe handling icons—each of which will have three options.

Before the administration of the full-scale study, Lightspeed will conduct a pretest with a sample of 100 English- and Spanish-speaking panel members to ensure the programming logic is working correctly. Approximately 1,700 panel members will be sent an email invitation to complete the pretest. The same sampling and recruiting methods will be used for the pretest.

Table B-1.   Quotas for Inbound Sampling for Web-Based Experimental Study

Category	Quota[a]
Race	
White	74%
Non-White	26%
Speak Spanish at home	
No	87%
Yes	13%
Age	
18–34	28%
35–54	36%
55+	36%
Education	
Less than high school or high school diploma/GED (including vocational training)	42%
Some college (no degree) or associate or 2-year degree	29%
Bachelor's degree	18%
Graduate or professional degree	11%

[a] Source: U.S. Census Bureau. (n.d.). 2010-2014 American Community Survey 5-year data profiles. Retrieved from https://www.census.gov/acs/www/data/data-tables-and-tools/data-profiles/2014/

Behavior Change Study

The behavior change study employs an experimental design in which participants are randomly assigned to one of three treatment groups (that will be used to assess three alternative SHIs) or a control group (that will be used to assess the current SHI). To provide geographic diversity, the study will be conducted in test kitchen facilities located in four different locations (one in each of the four Census regions): (1) Wake, Durham, Orange, and/or Johnston Counties, NC; (2) Brazos County, TX; (3) Yolo County, CA; and (4) Providence County, RI. The respondent universe is English- and Spanish-speaking adults living within driving distance of each of the four locations. Because probability-based sampling is not being used, inferences

cannot be made to the U.S. population; however, by using an experimental design that is powered to detect change in the desired outcome (the power analysis is described in Section B.2), we can assess which label is most effective at encouraging adherence to the safe handling instructions.

Respondent Selection Methods

Study participants will be recruited in the four locations using convenience sampling via social media, such as Craigslist, Facebook ads, Google Ad Words, and Instagram (Appendix D) and by sending emails to Expanded Food and Nutrition Education Program participants to reach low-income consumers (for the Wake County location) (Appendix E). Notices about the study may also be posted in local grocery stores, food retailers, food banks, and other locations (Appendix F). Interested individuals can complete the screening questionnaire online or by telephone (see Appendix G). Eligible participants will be contacted by phone and invited to participate in the study and an appointment scheduled. Appointments will be scheduled during work hours, evenings, and weekends to allow for a broader participant pool.

These recruiting methods will be supplemented using outbound recruiting to recruit individuals with specific demographics that may be challenging to recruit using social media (e.g., individuals with a high school education or less and older adults). For the outbound recruiting, the study team will work with a local market research company in each of the four locations. Each market research company will contact adults from their database and screen them for eligibility. The databases maintained by each of the local market research companies have been developed over time through advertisements, word-of-mouth referrals, outreach to various organizations and industries, and Internet list searches, among other means. The databases are constantly refreshed through natural attrition and additions. In addition to contact information,

basic socioeconomic characteristics are collected through the registration process and updated

periodically so that the market research firms can easily identify individuals with specific

demographic characteristics (e.g., high school education).

Participants must meet the following criteria to be eligible to participate in the study:

- are 18 years of age or older;

- cook raw meat or poultry at home at least once per week;

- have not cooked or worked professionally in a food preparation setting within the

  past 5 years;

- have not taken any type of food safety training, such as ServSafe, in the past 5

  years;

- have never had a seizure or have not been diagnosed with epilepsy (it is

  recommended that such individuals not take part in eye-tracking studies)

- do not wear corrective lenses that may interfere with the calibration of the eye-

  tracking device (i.e., progressive lenses, hard or semi- hard contact lenses).

Sample Size and Allocation

Based on the power analysis (described in Section B.2), a sample size of 480 is required

for the behavior change study, with participants randomly assigned to one of the four study

groups (120 per group). Of the 480 participants, 360 participants will be in the North Carolina

location, and 120 participants will be in each of the other three locations (40 per location). The

number of observations in each location is due to logistical and budgetary considerations.

Because we are using an experimental design (i.e., random assignment to one of four study

conditions) with the aim of estimating causal effects, applying sampling weights is not

necessary. Within each location, we will randomly assign participants to one of the four study

groups so that the number in each group is similar for each location (i.e., 90 per group in North Carolina and 10 per group in the other locations). With random assignment, the distributions of demographic characteristics for the four study groups are expected to be similar. Table B-2 provides the sample allocation by location and study group.

Table B-2.   Sample Allocation by Location and Study Group for the Behavior Change Study

Location	Control (Current SHI)	SHI Option 1	SHI Option 2	SHI Option 3	Total
North Carolina	90	90	90	90	360
Texas	10	10	10	10	40
California	10	10	10	10	40
Rhode Island	10	10	10	10	40
Total	120	120	120	120	480

As part of the screening process, information on participant and household demographic characteristics will be collected. The study will ensure a diverse sample of participants with respect to race, ethnicity, age, education level, and presence of children (0 to 17 years) in the household by using inbound quotas. The expected allocation for the full sample is shown in Table B-3. When randomly assigning participants to the four study groups, we will take necessary and reasonable steps to ensure that the demographic characteristics of each group are similar.

Table B-3.   Sample Allocation by Demographic Characteristics for the Behavior Change Study

Characteristic	Response Categories	Percentage	Number
Race	White	68%	326
	Non-White	32%	154
Ethnicity	Not Hispanic or Latino	84%	403
	Hispanic or Latino	16%	77
Age	18–34	35%	168
	35–54	39%	187
	55+	26%	125
Highest level of education	Less than high school, high school diploma/GED, or vocational school	26%	125
	Some college (no degree) or associate or 2-year degree	40%	192
	Bachelor's degree	19%	91
	Graduate or professional degree	15%	72
Child 0–17 years in household	Yes	48%	230
	No	52%	250

Response Rate

The expected response rate (show rate) for the behavior change study is 85% based on the contractor's experience with similar studies. Thus, the starting sample size for the behavior change study is 565, yielding 480 completed observations (565 x 0.85).

B.2. Procedures for the Collection of Information

This section provides an overview of the study procedures, provides information on the degree of accuracy required for the study, and discusses the estimation procedures for each study component. There are no unusual problems requiring specialized sampling procedures for either

study component. A separate set of participants is being contacted for each study component and will only be contacted one time, so periodic data collection cycles are not applicable.

Web-Based Experimental Study

Study Procedures

Lightspeed will send panelists selected for this study email invitations to invite them to participate in the study (Appendix A). The email invitations will instruct each panelist to log-on to the Lightspeed website and enter their email address and password to access the link to the survey. Once selected panelists click on the survey link, they will be provided information on informed consent and asked if they would like to proceed with the study (see Appendix B for data collection instrument). If panelists decline, they will be categorized as nonrespondents. If panelists accept, they will be asked several questions to determine eligibility (as noted in Section B.1). Panelists not eligible to complete the survey will be categorized as ineligible. Panelists who are deemed eligible will be randomly assigned to a study condition and will proceed with the survey. The survey will be available in English and Spanish and is estimated to take 20 minutes to complete.

Before the administration of the full-scale study, Lightspeed will conduct a pretest with a sample of 100 English- and Spanish-speaking panel members to ensure the programming logic is working correctly. Results of the pretest will also be used to determine the appropriate amount of time for the limited-time exposure task.

Degree of Accuracy Required for the Study

Statistical power (1−ß) provides a measure of the degree of confidence we have that we will not make a Type II error by accepting a false null hypothesis. In other words, 1−ß is our probability of accepting the alternate hypothesis when it is in fact true. By convention, we set the

Type II error rate at 0.20 and view 0.80 as an acceptable level for statistical power. To calculate statistical power, we began with an assumed sample size of 3,600 and assign 133-134 individuals to each of 27 treatment groups, where participants in each treatment group are exposed to one of the 27 different SHI labels, generated by the 3x3x3 full factorial research design with factors that include safe handling instructions, safe handling icons, and label shape.

The main effect of the primary study outcome (label salience) among the 27 treatment groups was to determine statistical power for the study. Label salience will be assessed using a measure of sensitivity, defined as the ability of the participant to accurately recall the stimuli (MacMillan, 2002). The measure is an index score with a range of −4 to +4, where participants receive 1 point for each correct response to a question asking about an element of the stimuli and are debited 1 point for each incorrect response asking about an element that is not part of the stimuli. The null hypothesis is that the mean label salience score will not differ among the 27 groups.

To test this hypothesis, we applied a one-way analysis of variance. The effect size of this test is f, estimated as: $f = \dfrac{\sigma_m}{\sigma}$, where σ_m is the standard deviation of the mean of each group and σ is the within-group standard deviation (Cohen 1988, Desu & Raghavarao, 1990).

We note that the reader should not confuse the use of notation "f" reported here with the F-statistic in the analysis of variance which is a ratio of between-group variances and within-group variance. We use f to signify the effect size, which is the ratio of the standard deviations described above, because this is how it is described in Cohen 1988.[1]

---

[1]   It is also described in the PASS software documentation. See page 550-5 in this document: https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/PASS/One-Way_Analysis_of_Variance_F-Tests.pdf.

We begin with the assumption that among the 27 groups, 20 groups have an attention score mean of 0, five groups have a mean of 0.25, and two groups have a mean of 0.5. We chose these values because smaller values might not be meaningful and larger values might not be observed. With this distribution of means, we calculated the standard deviation of the group means (σ_m) as 0.16 and we assume that the standard deviation within the groups are equal. Accordingly, to achieve power of 80% with a Type 1 error rate is 5% the standard deviation within the groups must be 1.89 or less. We obtained this result using PASS software and confirmed the result through Monte Carlo simulations.

Our study will have 80% power to detect an effect size of $f = \dfrac{\sigma_m}{\sigma} = \dfrac{0.16}{1.89} = 0.08$(this is a small effect size based on Cohen's classification) as long as it is reasonable to expect the within-groups standard deviation of the label salience variable to be less than 1.89. Consider Figure 1 which shows a distribution of label salience scores that have a standard deviation of 1.89 within group; here, 8% of this distribution have the most extreme values (-4, 4) and 12.5% of this distribution have the next most extreme values (-3, -3). This distribution presents more spread (i.e., extreme scores) than one would expect in the planned experiment. Additionally, because the distribution is symmetrical around zero it has the greatest chance to have a largest variance. Consequently, we anticipate that the within-group label salience variance will be less than 1.89 and, therefore, that the study will be adequately powered.

Figure 1. Distribution of label salience outcome scores with standard deviation of 1.89.
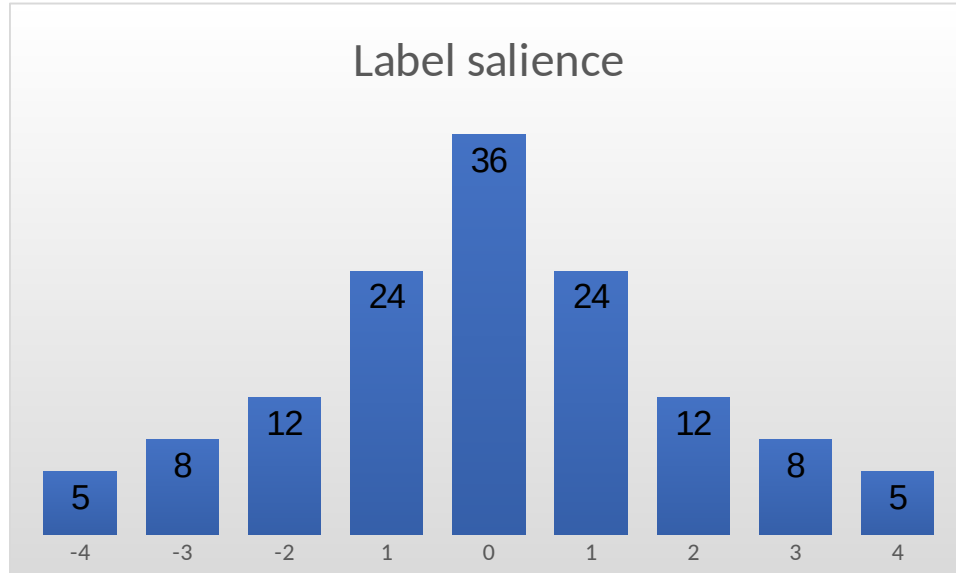
Estimation Procedures

Our purpose is to select the five visual label formats that perform best at attracting consumer attention. From these five formats, we will select three to test in the behavior change study. We refer to the ability of a label to attract consume attention as salience and define salience as the degree to which a visual feature (target) predominates perceptual attention when presented in a complex environment (background). Selection will be accomplished by ranking the 27 labels from highest to lowest on the visual salience score.

In this experiment, we will infer salience from a set of items designed to assess the participant's ability to recall information presented in the target from information either in the background or not present; we refer to the former as a "hit" and the latter as a "false alarm." The participant will be asked a series of dichotomous (yes/no) items about information that may (or may not) have been presented in the visual target (see Appendix B for the survey instrument). The item set will be balanced to include the same number of hits and false alarms. For each potential hit, the participant receives one point (+1) for a correct answer (yes), and zero

otherwise. For each false alarm, the participant is debited one point (-1) for each incorrect

answer (yes), and zero otherwise. The number of hits and false alarms reported can be

summarized as proportions and transformed to z-scores so that each participant's hit rate and

false alarm rate are realizations from a unit-normal distribution. These two pieces of information

can be used to calculate sensitivity (d′), or the participant's ability to accurately differentiate

target from background. We can calculate sensitivity using the following formula:

$$d' = z(H) - z(F)$$

where

z(H) is the z-score corresponding to the proportion of hits

z(F) is the z-score corresponding to the proportion of false alarms

This formula describes visual salience as the difference between true positive responses

and false negative responses. Label formats with higher positive values indicate that participants

were more attentive to the visual target (i.e., high salience). Sensitivity scores will also be used in

secondary analyses to examine the three primary study factors (safe handling instructions, safe

handling icons, label shape) as well as all two-way interactions. These analyses will support

decision making when selecting the three labels for inclusion in the behavior change study

among the top five scoring SHI labels.

Behavior Change Study

Study Procedures

Upon arrival to the test kitchen, participants will read and sign an informed consent form

(see Appendix K). Next, each participant will be asked to watch a video on an iPad so that

participants receive a consistent description of the study and what to expect during their

participation in the study (see Appendix L). Afterward, participants will be directed to put on a

Tobii Pro Glass 2 mobile eye-tracking unit. A trained data collector will adjust the unit for fit and comfort. The unit will be calibrated following manufacturer recommendations. To ensure precision, the participant will be given a set of two-dimensional test images and directed to examine each in sequence. This step will serve to establish a base rate, confirm the accuracy of the eye-tracking unit, and provide the participant with an opportunity to become accustomed to the wearable data collection device.

After the calibration of the eye-tracking unit, participants will be directed to the general locations of kitchen utensils and staples but will not be prompted to use any particular equipment; kitchen drawers and cabinets will also be labeled with their contents.

While under video observation, participants will be given recipes and ingredients, including two raw meat products bearing the assigned SHI label (control or one of the three treatment labels), and asked to prepare three dishes: (1) gluten-free pasta and meatballs using frozen, preformed raw meatballs (under the guise that the dish is for an individual who is on a gluten-free diet); (2) gluten-free pasta and meatballs using raw ground beef; and (3) a cherry tomato garnish. Participants will also be instructed to clean up afterwards as they would do at home.

The general kitchen layout and setup of equipment will be consistent across test kitchens in the four locations. Video recording equipment will be set up to record meal preparation. Trained research staff will conduct the video recording. Recording of meal preparation will begin as soon as the participant enters the test kitchen and will end after the participant leaves the test kitchen. Participants' cleaning and sanitizing of equipment and the kitchen environment before and after preparation will also be recorded. The meal preparation/observation portion of the study will take 50 to 80 minutes to complete.

Following the meal preparation/observation portion of the study, participants will be provided an opportunity to take a break. If participants choose to take a break, they will remove the eye-tracking device and then it will be refitted/calibrated upon their return. Participants will then be directed to examine each of six mock meat and poultry products (i.e., stimuli). The following products will be used as stimuli for the eye-tracking study and the in-depth interviews (IDI):

- NTRE ground beef patties in a styrofoam container sealed with plastic wrap (with SHI)

- frozen NRTE ground beef patties in a box (with SHI)

- frozen RTE chicken nuggets/tenders with breading in a bag

- frozen NRTE chicken nuggets/tenders with breading in a bag (with SHI)

- frozen RTE chicken cordon bleu in a box

- frozen NRTE chicken cordon bleu in a box (with SHI)

Appendix N provides the script the data collector will use to administer the eye-tracking study. The script will direct participants' attention to each product, and participants will be asked to complete several tasks to determine which version of the SHI label (current or one of three alternative versions) is most often attended on a meat and poultry package and to assess whether participants can properly distinguish between RTE and NRTE products that appear to be ready to eat. The eye-tracking study will take up to 30 minutes to complete and will be recorded in audio and video formats.

Lastly, participants will take part in an IDI. Using a structured interview guide (see Appendix N), the data collector will ask participants a series of questions about their views, opinions, and experiences during the meal preparation/observational study and questions to

understand how participants determine whether a meat or poultry product is raw versus already cooked. The IDI will take up to 30 minutes to complete and will be recorded in audio and video formats.

Degree of Accuracy Required for the Study

We calculated sample sizes to determine the minimum number of participants needed to provide a level of confidence that the experimental component of the behavior change study is sufficiently powered, meaning that a change in the anticipated size or greater would be interpreted as occurring beyond chance (i.e., statistically significant). By convention, we aim for 80% statistical power and a 95% level of confidence.

The purpose of the behavior change study is to evaluate the impact of alternative SHI labels on participants' use of recommended food handling practices (e.g., use a food thermometer or wash hands before beginning meal preparation). The degree to which the current SHI label is attended and the recommended practices followed has received little scientific attention. A study by Yang et al. (2000) analyzed selected states' Behavioral Risk Factor Surveillance data collected in 1995 and 1996 and found that 51% of the 14,262 respondents reported that they had seen the SHI label. Of these, 79% recalled reading the label, and 37% of respondents reporting they had seen and read the label reported changing their raw meat preparation methods because of the label. More recently, data collected following focus group discussions conducted on behalf of FSIS on consumer response to the current SHI label found that among participants who were not already following the recommended practices, a majority of them reported that they would be likely or very likely (47% to 72%) to follow specific safe food handling practices following label revisions (Cates et al., 2015).

Because the experimental design will ask participants to attend to food product packaging (without specifically calling out the SHI), we assume 37% of participants will follow at least one behavior on the current SHI label based on the data reported by Yang et al. (2000). With this assumed base rate, Table B-4 provides potential observed differences between the control (current SHI) and treatment (alternative SHI) groups ranging from 45% to 65%, along with the odds ratios for descriptive and interpretative purposes. It is anticipated that the impact of a revised SHI label will be sufficient to generate differences in the middle of this range (i.e., the observed difference between the control and treatment groups is 18 percentage points).

Table B-4.   Sample Size Requirements for Different Observed Differences between the Control and Treatment Groups for the Behavior Change Study

Safe Food Handling Practice: Current SHI	Safe Food Handling Practice: Revised SHI	Odds Ratio	Observed Difference Between Groups	Total Sample Size (N)[a]
37%	45.0%	1.4	8%	1,186
37%	50.0%	1.7	13%	456
37%	*55.0%*	*2.1*	*18%*	*240*
37%	60.0%	2.6	23%	146
37%	65.0%	3.2	28%	98

[a]   Total sample size is the number of participants needed for a two-group comparison. One-half of this number (N/2) is the number of participants needed for each experimental group.

Accordingly, the proposed sample size of 480 participants (120 per group) takes into consideration several important features of the study: the anticipated base rate for following the instructions on the SHI label among those attending the instructions, recent focus group data, and the anticipated distributional characteristics of a dichotomous outcome. The study will employ the following strategies to maintain a constant and robust (i.e., 80% or better) level of statistical power for all hypothesis tests:

- Use reliable and validated data collection methods to control intersubject variability.

- Employ common measures across all study conditions.

- Collect covariates related to the outcome measure that can be applied in multivariate modeling to improve statistical precision.

- Develop analysis strategies that are matched to the distributional assumptions of the outcome measures.

Estimation Procedures

Trained coders will watch the recorded videos for the meal preparation experiment and use an observation rubric (see Appendix M) to code participants' behaviors as following vs. not following the safe handling instructions on the label. We will conduct statistical analyses comparing the label adherence scores among the four groups (i.e., current SHI label and three alternative SHI labels) to determine the most effective label.

The data from the eye-tracking study will be reviewed, coded, and processed. The primary outputs from the analysis are summarized below:

- Eye-tracking metrics for each area of interest (AOI), such as percentage of participants who visually notice the relevant AOIs. AOIs will be standardized across packages and will include the SHI labels (and sublabel components), competing instructions, and other labeling information.

- Time to first viewing: The time until each AOI is first noticed.

- Total viewing duration: The total time spent viewing each AOI.

- Number of viewings: The average number of times an AOI is viewed.

- Distribution of attention: The percentage of package viewing time spent on each AOI.

- Perceptual flow: The typical order in which each AOI on a package is seen.

- Eye-tracking heat maps or gaze plots of attention to SHI labels and other package information.

Using the findings from the behavior change study, we will construct logistic regression models to examine the association between attention to SHI labels (from the eye-tracking study) and proper execution of each of the safe handling instructions (from the meal preparation experiment). Poisson regression models will also be used to examine the relationship between attention to SHI labels and the number of properly executed safe handling behaviors. These data will provide empirical evidence on the SHI label option that is most effective at encouraging consumers to follow recommended safe handling practices for raw and partially cooked meat and poultry products.

B.3. Methods to Maximize Response Rate and Deal with Nonresponse

Web-Based Experimental Study

Based on experience conducting 20-minute online surveys with general population samples (i.e., adults 18 years or older), Lightspeed estimates that about 19% of the selected panelists will be eligible and complete the online survey. Lightspeed will send up to 3 three automatic email reminders to nonresponding panelists during the course of data collection. Panelists who do not complete the survey will be categorized as nonrespondents.

To maximize participation, we will conduct cognitive interviews and a pretest to help improve the understandability and usability of the questionnaire, reduce participant burden, and enhance administration.

In addition, to encourage participation, each email invitation and reminder will state the study purpose and identify USDA as the study sponsor (see Appendix A for email invitation and

Appendix C for email reminders) and provide an email address and toll-free number (provided by the contractor) for panelists to obtain additional information about the study or verify the authenticity of the study.

Behavior Change Study

We estimate 1,695 prospective participants will complete the screening questionnaire by telephone or via a web link (Appendix G), and 33% of these individuals will be eligible and subsequently contacted by phone to schedule an appointment. Of the 565 scheduled participants, we estimate that 480 (85%) will participate in the behavior change study.

To maximize the response rate, each participant will receive a cash honorarium of $100 and a small gift (food thermometer valued at $5.38 and magnet valued at $0.23) for their participation. In addition, we will send a confirmation email (see Appendix I) with directions and make a reminder call (see Appendix J) to recruited individuals before their scheduled appointment.

B.4. Tests of Procedures or Methods to be Undertaken
Web-Based Experimental Study

In July and September 2018, the contractor conducted cognitive interviews in Research Triangle Park, NC with nine target audience members (including people who speak Spanish) to determine if any survey questions or response items were confusing or difficult to understand. Based on the cognitive interview findings, we refined the programmed instrument. Specifically, we revised the distractor questions regarding the food packages to ask about images, words, and phrases that are more often found on food packages. We determined that the questions to collect information on which rationale statement (to include on the SHI label) would be most motivating to consumers were not being answered as intended, so we developed a different question to

collect this information. To improve understanding and readability, we reformatted several questions and simplified the instructions for some questions. The cognitive interviews also confirmed the estimated burden of 20 minutes (the average time to complete the survey was 17 minutes).

To ensure that the programming logic, sample distribution and fulfillment, and data compilation are functioning correctly, Lightspeed will conduct a pretest with 100 randomly selected panelists. Data collection for the pretest will not commence until OMB approval is obtained. As previously noted, results of the pretest will also be used to determine the appropriate amount of time for the limited-time exposure task. If changes are made to the survey instrument based on the pretest findings, a revised survey instrument will be submitted to OMB for approval before conducting the full-scale study.

Behavior Change Study

The contractor conducted a pilot study to test the methodology and instruments for the behavior change study in September 2018 with two people in a test kitchen. Based on the pilot, we made several revisions to the methodology and instruments. For the meal preparation study, we had originally considered having participants prepare a parley garnish but decided to change this to a cherry tomato garnish to ensure participants used a knife to prepare a ready-to-eat food. Additionally, we simplified the instructions and tasks for the eye-tracking study, in particular to make the script more conversational. We also revised some of the questions in the IDI script to avoid using a judgmental tone when referring to participants' adherence to recommended handling practices during the meal preparation study. The contractor will conduct additional pilots with test subjects before fielding the main study to assess and refine the training procedures for data collection staff.

B.5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

Sheryl Cates is the RTI Project Director and will manage the study. Dr. Jonathan Blitstein of RTI developed the experimental design for the web-based experimental study and the eye-tracking study and will oversee the analysis for these studies. Dr. Benjamin Chapman of NCSU will manage the data collection for the behavior change study and oversee the analysis for the meal preparation/observational study. Christopher Bernstein, an FSIS employee, will review the study results.

References

Cates, S. C., Kosa, K. M., & Muth, M. K. (2015, January). *Professional services to support requirements gathering sessions for Safe Food Handling Instructions: Strategic planning session.* Research Triangle Park, NC: RTI International.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Desu, M. M., & Raghavarao, D. (1990). Sample size methodology. New York: Academic Press.

MacMillan, N. (2002). Signal detection theory. In H. Pashler & J. Wixted (Eds.), Stevens' handbook of experimental psychology (3rd ed., Vol. 4, pp. 43–89). New York, NY: John Wiley and Sons, Inc.

Yang, S., Angulo, F. J., & Altekruse, S. F. (2000). Evaluation of safe food-handling instructions on raw meat and poultry products. *Journal of Food Protection, 63*(10), 1321–1325.