Perceptions of and Reactions to Sustainability Menu Label Designs

NCT07214792

Statistical Analysis Plan

Version Date: 10/06/2025

Research Question

The primary objective of this experiment is to test the relative effects of environmental harm menu label designs on intent to purchase fast-food menu items. Participants will take a web-based survey on Qualtrics, during which they will view an online fast-food menu excerpt with one of seven label conditions applied: a QR-code control label on all items; five labels on high-environmental-impact items only (Low Sustainability, Unsustainable, Environmental Harm, High Environmental Harm, High Environmental Cost), or a numeric Environmental Cost label on all items. After viewing the menu excerpt, participants are presented with the label in isolation and will answer questions about their perceptions of the label.

Outcome Measures

Primary outcome: The primary outcome is the intention to purchase a menu item with the randomized label applied. Intent to purchase is measured using a 7-point Likert scale (1=Very unlikely, 7=Very likely).

Secondary outcomes: Secondary outcomes include the following:

1. *Label interpretation*. Participants will indicate their interpretation of the label using a check-all-that-apply question. Response options are a range of statements about the menu item: 1="This menu item is unhealthy", 2="This menu item is healthy", 3="This menu item is good for the environment", 4="This menu item used a lot of natural resources (e.g. land, water) to produce", 5="This menu item will keep you full for a long time", 6="This menu item is bad for the environment", and 7="This menu item will cost consumers more to purchase". Also included are 8="I don't know" and 9="Other" option with open-text entry.
2. *Label understandability, believability, and attention*. Participants will rate how easy to understand and believable the label was and how much it grabbed their attention. Responses are measured on a 5-point Likert scale (1=Not at all, 5=A great deal).
3. *Perceived message effectiveness (PME)*. Participants will report the perceived message effectiveness of the label they viewed in the menu excerpt. PME is measured using a unipolar 3-item scale adapted from the University of North Carolina (UNC) PME scale. Participants' levels of discouragement, unpleasantness,

and concern regarding menu items with the label applied will be measured using a 5-point Likert scale (1=Not at all, 5=A great deal). PME score will be calculated as the mean of the 3 items.

4. *Label helpfulness*. Participants will be asked how helpful they would find the label when deciding what to order at a fast-food restaurant. Helpfulness is measured using a 7-point Likert scale (1=Very unhelpful, 7=Very helpful).

Statistical Analyses

To determine the effects of the six environmental labels compared to the control label on our continuous primary outcome of purchase intention, we will regress the outcome onto a single categorical predictor (label condition) and use robust standard errors. Coefficients will be interpreted as the difference in means between each experimental condition and control.

A similar approach will be used for secondary outcomes (label understandability, believability, and attention; PME; helpfulness) and for the other pre-specified outcomes (perceived appeal, healthfulness, and environmental impact of the two menu items): we will regress each outcome onto label condition. We will estimate marginal means for each label condition and test overall differences across the seven conditions with joint Wald tests. We will also compare the six experimental labels with each other (15 pairwise comparisons) using Bonferroni-adjusted p-values. Label interpretation items will be assessed as a binary outcome (yes/no) and analyzed using logistic regression with label condition as the predictor, comparing odds of each interpretation for each experimental condition compared to control.

We may additionally explore effect modification by socio-demographic characteristics.

Statistical significance will be considered at $p<0.05$. All tests will be two-tailed.

We will not exclude observations based on values of any of the outcome variables. We will remove observations based on individuals who do not meet eligibility criteria (non-US resident, less than 18 years old); are bots or fraudulent responses (as identified by the survey research firm, CloudResearch, and by built-in Qualtrics fraud detection); do not complete the survey; complete the survey in less than one-third of the median completion time; and/or fail the attention check question. For secondary outcomes, we will exclude observations with missing data within family of outcomes.