Testing Non-Nutritive Sweetener Front-of-Package Labels in Chile

NCT06842693

April 18th, 2025

Study Protocol and Analytic Plan

Introduction

This study aims to examine the effects of 2 types of non-nutritive sweetener (NNS) warning labels on a sample of Chilean parents' product selection and perceptions. This document pre-specifies our planned analytic approach prior to data collection.

Study Protocol

Participants will complete an online randomized experiment programmed in Qualtrics. After providing informed consent, participants will be randomly assigned to view 1 of 3 types of labels: a neutral control label, a text NNS warning label (i.e., Chilean NNS label), or an octagonal NNS warning label (i.e., Colombian NNS label).

First, participants will view their assigned label on four similar products (one unsweetened, one sweetened with sugar, one sweetened with NNS, and one sweetened with both sugar and NNS) displayed in random order. They will then complete a task in which they select one of the products to purchase for their child and identify which of the products contain NNS. During this task, in all experimental arms, products will also display Chile's current "high in sugar" and "high in calories" labels when applicable. This task will be repeated 3 times, each time with a different product category (i.e., fruit drinks, yogurts, breakfast cereals), with categories displayed in random order.

Next, participants will view their assigned label on two products containing NNS (i.e., flavored milk and cereal bar). For each product, participants will rate the perceived message effectiveness of the label, perceived product healthfulness, and perceived relative product healthfulness.

Lastly, after all product exposures, participants will rate their intentions to limit their child's NNS consumption.

Hypotheses

Correct identification of NNS-containing products (co-primary outcome): We predict that, compared to the control arm, correct identification of NNS-containing products will be higher for participants assigned to the Chilean NNS label arm or to the Colombian NNS label arm (H1).

Selection of NNS-sweetened product (co-primary outcome): We predict that, compared to the control arm, selection of the NNS-sweetened product will be lower for participants assigned to the Chilean NNS label arm or to the Colombian NNS label arm (H2).

Selection of unsweetened product (co-primary outcome): We predict that, compared to the control arm, selection of the unsweetened product will be higher for participants assigned to the Chilean NNS label arm or to the Colombian NNS label arm (H3).

Selection of sugar-sweetened product (secondary outcome): We predict that, compared to the control arm, selection of the sugar-sweetened product will not differ for participants assigned to the Chilean NNS label arm or to the Colombian NNS label arm (H4).

Perceived message effectiveness (secondary outcome): We predict that, compared to the control arm, the perceived message effectiveness (PME) of the labels will be higher for participants assigned to the Chilean NNS label arm or to the Colombian NNS label arm (H5).

Perceived product healthfulness (secondary outcome): We predict that, compared to the control arm, perceived product healthfulness will be lower for participants assigned to the Chilean NNS label arm or to the Colombian NNS label arm (H6).

Perceived relative product healthfulness (secondary outcome): We predict that, compared to the control arm, perceived relative product healthfulness will be lower for participants assigned to the Chilean NNS label arm or to the Colombian NNS label arm (H7).

Intentions to limit NNS consumption (secondary outcome): We predict that, compared to the control arm, intentions to limit their child's NNS consumption will be higher for participants assigned to the Chilean NNS label arm or to the Colombian NNS label arm (H8).

Main Analyses

We will use a two-sided critical alpha of 0.05 to conduct all statistical tests. All confidence intervals presented will use a 95% confidence level. Analyses of the primary and secondary outcomes will include all participants according to the trial arm to which they were randomized.

For all selection/identification tasks, we will code outcomes dichotomously. For correct identification of NNS-containing products, we will dichotomize the outcome to correct identification of NNS-containing products (coded as 1) *vs.* incorrect identification (coded as 0). Correct identification will be defined as having identified the two NNS-containing products in the set and no other products, while incorrect identification will be defined as any other scenario. For selection of NNS-sweetened product, we will dichotomize the outcome to selection of one of two NNS-sweetened products (coded as 1) *vs.* selection of either of the two non-NNS-sweetened products (coded as 0). For selection of unsweetened product, we will dichotomize the outcome to selection of the unsweetened product (coded as 1) *vs.* selection of any of the other three products (coded as 0). For selection of sugar-sweetened product outcome, we will dichotomize the outcome to selection of one of the two sugar-sweetened products (coded as 1) *vs.* selection of either of the other two non-sugar-sweetened products (coded as 0).

For dichotomous outcomes, we will first descriptively report unadjusted proportions in each experimental arm, including proportions for each product category and for all product categories simultaneously (e.g., whether the participant correctly identified the

NNS-containing products in all three product categories). Next, to test H1-H4, we will use logistic mixed-effects regression models, treating the intercept as random to account for repeated measures within participants. Models will regress each outcome on indicator variables for the experimental arms, product categories, and their interactions. After model estimation, we will test the joint statistical significance of the interaction terms, and, if not significant, we will drop the interactions from the models. Regardless of the presence of the interaction terms in the final model, we will obtain the predicted probabilities of each outcome by experimental arm for each separate product category, which we will use to conduct a joint test of statistical significance of the differences in predicted probabilities across product categories. If such differences are jointly significant, we will report average differential effects (ADEs) of each type of NNS label (i.e., differences in predicted probabilities between each NNS label type and the control label) separately for each product category. Alternatively, if differences are not jointly significant, we will report ADEs of each type of NNS label across product categories.

For the PME outcome, we will verify that Cronbach's alpha is sufficient (>0.7) and average each participant's scores across the three PME items. If Cronbach's alpha is not sufficient (<0.7), we will treat each item as a separate construct instead. Next, for all continuous outcomes, we will descriptively report unadjusted means in each experimental arm, including means for each product category and averaged across all product categories. To test H5-H8, we will use linear mixed-effects regression models. Models will regress the outcomes on indicator variables for experimental arms, product category, and their interactions. After model estimation, we will test the joint statistical significance of the interaction terms. If the interactions are jointly significant, we will keep them in the models and report ADEs of each type of NNS label (i.e., differences in predicted means between each NNS label type and the control label) separately for each product category. Alternatively, if the interactions are not jointly significant, we will drop them from the models and report ADEs of each type of NNS label across product categories.

We will use complete case analysis to handle any missing data.

Exploratory Analyses

Post estimation of each main analytic model, we will obtain exploratory comparisons between the outcomes in the two NNS label arms.

Additionally, we will examine whether gender and/or educational attainment moderates the effects of the NNS labels (compared to control) on the co-primary outcomes. If there are no significant differences in ADEs by product category in the main analyses, moderation analyses will use logistic mixed-effects models with random intercepts including indicator variables for experimental arms, the moderator, and their interactions. If, alternatively, there are significant differences in ADEs by product category in the main analyses, we will use separate regression models for each product category.

Sample Size and Power

According to our calculations, among the co-primary outcomes, selection of sugar-sweetened products requires the largest sample size. Therefore, we powered the study to detect differences in selection of sugar-sweetened products between the control arm and each of the NNS label arms. Our calculations are based on the assumption that a certain proportion of consumers will prefer an unlabeled product over a comparable product with any warning label, but will be indifferent between two labeled products regardless of the number of labels. The higher the preference for unlabeled products, the larger the sample size required to detect differences in selection of sugar-sweetened products, as demonstrated by the examples below:

Assumption: 60% preference for unlabeled products

	Unsweetened	NNS-sweetened	Sugar-sweetened	NNS- and sugar-sweetened	**Total sugar-sweetened**
Control	0.3*	0.3*	0.2	0.2	**0.4**
Chile NNS label	0.6*	0.133	0.133	0.133	**0.267**
Colombia NNS label	0.6*	0.133	0.133	0.133	**0.267**

* Unlabeled options

Sample size required (80% power, two-sided two-sample proportions tests): 197 per group.

Assumption: 90% preference for unlabeled products

	Unsweetened	NNS-sweetened	Sugar-sweetened	NNS- and sugar-sweetened	**Total sugar-sweetened**
Control	0.45*	0.45*	0.05	0.05	**0.1**
Chile NNS label	0.90*	0.033	0.033	0.033	**0.067**
Colombia NNS label	0.90*	0.033	0.033	0.033	**0.067**

* Unlabeled options

Sample size required (80% power, two-sided two-sample proportions tests): 1,102 per group.

If consumers are *not* indifferent between labeled products with different numbers of labels, but instead always prefer fewer warning labels (i.e., the proportions choosing the NNS- and sugar-sweetened product, which carries two labels, fall to zero in the NNS label arms, with a corresponding increase in the proportions choosing the NNS-

sweetened or the sugar-sweetened products, which both carry one label), the sample size required for each of these two examples would be lower (82 and 435 per group, respectively). However, since the presence of an NNS warning label along with a sugar warning label may indicate lower levels of sugar (compensated for by NNSs), it is not clear a priori that consumers would always prefer the product with fewer labels. Therefore, to be conservative, we assumed that 90% of participants would prefer unlabeled products over comparable products with warning labels, and the remaining 10% would be indifferent between comparable products with different numbers of warning labels. Based on these assumptions, a sample of 1,102 participants per arm (3,306 total) provides 80% power to detect differences between the control arm and each NNS label arm, using two-sided two-sample proportions tests with a critical alpha of 0.05.

Given the use of repeated measures, this required sample size further assumes that respondents will not choose differently across product categories – i.e., that 1,102 is the *effective* sample size. This is a conservative assumption, as within-participant variation would increase the effective sample size, thereby reducing the number of participants required.

Exclusions and Outliers

We will exclude participants who complete the survey implausibly quickly (defined as <1/3 of the median completion time) and those who completed less than 90% of the survey.

Interim Analyses

No interim analyses are planned.