Dear Dr Paquet,

Thank you for the opportunity to resubmit our manuscript. We believe that we have satisfactorily addressed the reviewer’s comments and that the manuscript is much stronger as a result. Attached we provide point by point responses to all comments raised. The revised manuscript has been resubmitted. We hope that you and the reviewers find the revisions satisfactory.

Reviewer reports and point-by-point responses:

Miranda Blake (Reviewer 1):

I congratulate the authors on a policy-relevant, well conducted and well reported study. I have no serious concerns but have made suggestions below, which are designed to increase clarity in a few sections.

- We thank the reviewer for the positive feedback.

Background

Lines 53-54. Could you provide a little extra detail on the HCS. When was it introduced? Is it mandatory?
- We replaced line 55 “The HCS symbol” with “The HCS symbol was first introduced in 2001 and although optional, is strongly encouraged by the Singapore government. It…”

Line 69. Briefly, what do these cited studies show? Are the labels effective at changing health behaviours?

- We inserted into line 80-81, page 4: … several empirical studies on “showing the effectiveness of” tobacco warning labels and on warning labels on select foods in other jurisdictions provide additional support for testing such labels on select food products in Singapore.

Line 75. Consider throughout the terminology and abbreviation "Health Warning label with deterring text". You abbreviation "WL" is not clearly a text-based label. Could you change it to something else like "Text based health warning label" or "Text warning label"? and then your abbreviation could be TW or similar.

- We have followed the reviewer’s recommendation. “Health Warning label with deterring text” has been replaced with “Text-based health warning label” and abbreviated as “TW”.

Line 81. Should the hypothesis refer to "predicted purchases" or similar? So it could be "we hypothesized that the proportion of labelled products purchased would be largest in the control group…" Also check tenses here.

- We have inserted and replaced in line 97, page 5: … hypothesized that the proportion of labelled (or targeted for labelling in Control Arm) products “among respondents will would be” largest in…

Methods

Line 103-4 and throughout. In this study it seems you based warning label eligibility on total sugar of product, and measured total sugar purchased as an outcome. Please provide justification for examining total sugar rather than added sugar, and discuss the implications.

- All sugar calculations were done using products’ Nutrition Information Panels. As these NIPs do not clearly distinguish between ‘added’ versus ‘total’ sugar content, we focused on total sugar.
Line 117. Please cite your prior studies
- These studies are not yet published. We have changed the description more accurately to ‘(based on unpublished data from ongoing studies)’

Line 123. Can you give an equivalent budget in USD for the international audience?
- We added USD equivalent in brackets next to the Singapore dollar values.

Results
Given your sample size, one decimal place should be sufficient for results throughout the text and in Table 1.
- This change has been made.

Table 1. In column 1 you don't need the units. You could update description of units in columns 2 to 4 from "mean or %" to "Mean (SD) or %".
- Agree. We merged all 3 columns and renamed heading as “Mean (SD) or %”

Line 169 Please list secondary outcomes again for clarity.
- We now list the secondary outcomes in brackets after the words ‘secondary outcomes’ (total sugar purchased (g), sugar purchased per dollar spent (g per $), total spending ($) and total expenditure on high-in-sugar products ($)).

Line 170-1. From table 3 it is clear these are only beverage purchases but here it sounds like any purchase that included a beverage.
- To address this concern we rephrased “when the sample is restricted to the 432 participants who made beverage purchases” to… “when the sample is restricted to beverage purchases made by the 432 participants who purchased beverages.”

Table 2 and 3. Could you give a more descriptive title for these tables?
- Table 2 has been re-titled to “Estimates of the Impact of Warning Labels on Measures of Diet Quality in the Pilot-DIET Study (N = 512)”.

- Table 3 has been re-titled to “Estimates of the Impact of Warning Labels on Measures of Diet Quality for Beverage Purchases in the Pilot-DIET Study (N = 432)”

Figure 3. Please clarify add the eligibility criteria "have experienced grocery shopping for their families before" to main methods section. Also, was this ever shopped for their families or recently shopped?

- This criteria was considered but removed prior to launching the study. We had erroneously not removed it from the CONSORT diagram. We have updated the CONSORT diagram (Figure 4 in current revision) to reflect the actual inclusion criteria.

Suggest changed wording in analysis box "Did not take exercise seriously" to "Did not provide valid responses." Please add to main text methods how you determined whether responses were valid and what you did with invalid responses.

- This exclusion criteria was considered but removed prior to data analysis. We had erroneously not removed it from the CONSORT diagram. We have updated the CONSORT diagram (Figure 4 in current revision) to reflect the analyzed sample sizes across arms.

Gaston Ares (Reviewer 2):

The manuscript deals with the evaluation of warnings in high-sugar products, which are gaining popularity worldwide to improve the eating habits of the population. The experimental design is correct, as well as the interpretation of the results. However, there are several issues that need to be addressed before the manuscript can be accepted for publication.

Introduction

The authors should update their literature review on warnings on food labels and expand the introduction. The authors should present the theoretical considerations that support the effectiveness of warnings and discuss more deeply results of published studies.

- We appreciate the feedback and have expanded the introduction as recommended. We have inserted: “Research has shown that, compared with non-directive labels like the Nutrition Facts Panel, directive and semi-directive labels may improve consumers' ability to find and understand
nutritional information [1-4]. Consumers tend to prefer simple FOP labels and to appreciate interpretational aids like descriptors or color codes [5, 6]. However, there is some evidence that positive FOP labels targeting healthier foods may not be sufficient to discourage consumption of less healthy alternatives [3, 7]. As a result, health warning labels have recently been proposed as a complement or alternative to positive FOP food labels. Studies of text warnings for tobacco products showed improved consumer education, greater knowledge of health harms of tobacco use, and decreased purchases [8, 9]. One study showed reduced intention to purchase SSBs in the presence of SSB warning labels [10]. These studies suggest that warning labels identifying harmful ingredients or adverse effects of certain food products have the potential to improve diet quality.”

Line 73: The graphic design of the warning is not identical to the Chilean version. It would be more accurate to say "similar to.." instead of "English version"

- Agree. We have rephrased “the English language version of the one used in Chile” to “an English language version similar to the one used in Chile”.

Line 81 What do the authors mean by "Given the strongly worded language"? 

- To clarify, we have added: … strongly worded language “similar to tobacco warning labels that are used in the text-based health warning label”… into line 95, page 5.

Lines 83-84: please reward, expand and better explain.

- We have rewritten this section on page 5 as follows: “There are three reasons to assume the text based warning label will show greater effectiveness. First, it is likely to generate greater levels of loss framing through negative terminology including ‘warning’ and ‘tooth decay’. Second, because it mirrors tobacco warning labels, the negative association may lead to an implicit bias against purchasing. Finally, because of the requirement to include the full message in a readable format, it is larger than the stop sign label and therefore may be more salient to consumers.”

Materials and Methods

The authors should expand the description of the site. In particular, the following information is necessary: a Figure showing how the site looked like, the names of all the categories and the number of products within each of them, the proportion of products high in sugar and qualifying for HCS logo in each category
- We now include two additional pieces of information in this revision. The first is an additional figure (figure 3) showing an example of the storefront and an example product as it appears in each arm of the study. The second is an additional file (Additional file 1.docx) that shows the names of all product categories, the number of products within each category, and the proportion of products qualifying for the logos within each category.

Line 103: it would be clearer to say "to receive a warning label"

- We have made this change.

Line 105: The criteria for including warnings is very different from Chile and most of the countries that have implemented warnings. The criteria are always absolute and constant across categories. Please justify the selected approach.

- We chose this strategy in consultation with the Singapore Health Promotion Board, who is considering such a strategy. Note that as long as the strategy is consistent across labels and targets the least health foods within a category (in this case 20% of foods and 27% of beverages) and is appropriately powered, it represents a reasonable test of the labels.

Figure 2: The size of the warnings is very different. The text warning does not allow to see the whole product. This should be better justified in the materials and methods section - Agree. This was unavoidable. We did constrain the heights of the two labels to be the same, but could not constrain the text based label to be of similar width. We now clearly state this in the introduction as to why the text based label may outperform and note again in the Discussion on page 12 that “one reason the text based label was more effective may be due to its larger proportions”.

Line 123: Were they asked to keep shopping until the minimum was achieved? Were they asked to remove products?

- As noted on page 7, “Those who consented were randomized into 1 of the 3 arms and asked to spend between S$50 (approximately 37 USD) and S$250 (approximately 183 USD) on NUSMart as if it were a real household grocery shopping trip (see Figure 3 for diagram of study flow). A pop-up message appeared on-screen if they attempted to checkout below or above the minimum or maximum respectively.”
Line 142: This is not a direct measure. It would be more appropriate to consider expenditure in high sugar products

- We agree and now include expenditure on High-in-sugar products ($) as an additional secondary outcome. There was no statistically significant difference in total expenditure on high-in-sugar products between arms.

Line 145: Why instead of a regular ANOVA comparing experimental conditions?

- I suppose this is due to preference and my training in health economics, which focuses on OLS. Mathematically there is no difference when the two models test against the same hypotheses and use identical encoding. In fact, ANOVA is often described as a regression with dummy variables, as we have done here. Note that we also run the GLM model using a logit link for proportions, which is not equivalent to ANOVA as it constrains the predicted values to be between 0 and 1 for each observation, but as results are identical to the OLS results we chose not to report them as part of the main manuscript.

Discussion and Conclusions

The authors do not accurately discuss their results in the light of published studies. They should better explain the results and compare them with other studies.

- We now added the following into the Discussion, page 12: “Our findings are consistent with recent studies examining SSB warning labels. A text-only warning label was shown to reduce intake of SSBs [11], reduce perceived product attractiveness, quality and taste, and reduce perceptions of consumer “coolness” [12]. Other studies focusing on warning labels for SSBs have shown the labels increase parents’ understanding of health harms associated with over-consumption of SSBs and reduce intentions to purchase SSBs for their children [10]. To date, no studies have assessed how the food labelling law in Chile has impacted purchases of sugar or other macronutrients.

Although our study only found one of the labels to be effective at influencing purchases of targeted products, a full scale However, the full trial should include including both labels should be pursued as given that we could not reject the hypothesis that one label outperformed the other the effect of the two labels is similar. Moreover, a full scale trial with actual purchases will allow for determining should be pursed because results suggest that even whether a label that effectively influences purchasing patterns also leads to improvements in diet quality. This is relevant given that an a effective label may not reduce sugar or calories purchased if consumers
alter their behavior by purchasing more of unlabeled products. A full scale trial with actual purchases can further explore this hypothesis.”


Best,

Eric Finkelstein, Ph.D., MHA
Professor, Health Services and Systems Research Program
Duke-NUS Medical School, Singapore
8 College Road Singapore 169857
eric.finkelstein@duke-nus.edu.sg