

Healthy choice label does not substantially improve consumers' ability to select healthier cereals: results of an online experiment

Michael Siegrist*, Christina Hartmann and Gianna A. Lazzarini

ETH Zurich, Department of Health Sciences and Technology, Consumer Behavior, CHN J 76.2, 8092 Zurich, Switzerland

(Submitted 1 June 2018 – Final revision received 26 December 2018 – Accepted 19 February 2019 – First published online 22 April 2019)

Abstract

The results of numerous studies suggest that front-of-package (FOP) labels enhance consumers' ability to assess the healthiness of food products. However, most of the studies lack ecological validity. We selected fourteen breakfast cereals stocked by a major Swiss retailer. The participants from an Internet panel (n 780), with a somewhat higher educational level than that of the Swiss population, were randomly assigned to one of four conditions: control (picture of the FOP presented), table (plus the nutrition table with information on the energy and the main nutrients per 100 g), label (plus the healthy choice label for the healthier product) and combined (plus both the nutrition table and the healthy choice label). The participants were asked to select the healthier cereals from all possible ninety-one pair comparisons. The nutrient profile score was used as a 'gold standard'. For the thirty-three cereal pairs, one of the cereals had a label and the other had none, the median accuracy was only marginally lower in the control condition (91%) compared with the table (94%), the label (94%) and the combined conditions (97%). Similar results were observed when the incorrect decisions were weighted by the difference in the nutrient profile scores of the two cereals (for all ninety-one product pairs). These findings suggest that a healthy choice label has a limited effect on helping consumers select healthier cereals. In the control condition, the median of the correct choices was about 78%. Consumers' perception of the healthiness of foods could be improved.

Key words: Healthy choice labels: Front-of-package labels: Representative stimuli: Internet experiments: Perceived healthiness of cereals

Changes in the eating behaviour of consumers could result in better public health because some chronic illnesses, such as diabetes, cancer or heart disease, could be caused by an unhealthy diet⁽¹⁾. Therefore, consumers should be motivated and enabled to adopt a more balanced and healthy diet. Consumers' ability to correctly evaluate the healthiness of foods has been questioned, however. The question of how nutrition information should be designed and communicated to help consumers make healthier decisions has thus received much attention^(2–4). Research suggests that front-of-package (FOP) labels may help consumers make healthier food choices⁽²⁾. However, a closer review of the literature reveals scant knowledge about how much FOP labels increase consumers' ability to choose healthier products when confronted with realistic product information (e.g. complete FOP information, including brand names, food pictures and claims). In the present study, our aim is to examine whether a simple FOP label, such as a healthy choice label, enables consumers to make a better decision on which of the two products is healthier if they have all the information available at the point of purchase.

A number of studies have examined the impact of various types of FOP information on the perception about the healthiness of foods^(5–11). The labels and the tables used for

communicating nutrition information differ in several aspects^(2,12,13). For example, the labels can be classified according to how directive they are⁽¹⁴⁾. The non-directive nutrition table provides quantitative information about the nutrient content per 100 g of a food product. Directive labels, such as the 'health tick' or the 'smart choice' label, inform consumers the product to be chosen for a healthy diet. The impact of such directive labels has been investigated in various studies^(5,8,10,14–19) whose designs vary considerably. For example, in Borgmeier and Westenhoefer's study⁽⁵⁾, the participants had to identify the healthier option in each of the twenty-eight food pairs. The food pairs belonged to different categories, and only three pairs from the grain/cereal/bread category were used. The authors found moderate impacts of the different label formats. Roberto *et al.*⁽¹⁸⁾ used one high-sugar breakfast cereal. The FOP either had or did not have a smart choice label that provided information about the energy content per serving. This label had no impact on perceived healthiness, however. Hodgkins *et al.*⁽⁸⁾ used another research approach, where the participants received information about the food type (e.g. pizza) and the nutrition table but none about the brand and no picture of the package. The nutrition information was combined

Abbreviations: FOP, front-of-package; IQR, interquartile range.

* **Corresponding author:** M. Siegrist, email michael.siegrist@hest.ethz.ch

with four different FOP labelling systems. This cited study found some small improvements for an objective understanding of some labelling systems.

Most of the reviewed studies suggest that FOP label schemes result in a more accurate evaluation of the healthiness of food products^(8,10,16). However, many of the studies lack ecological validity, and whether their results have practical implications is not fully clear. In some studies, the participants did not evaluate food products but only nutrition information^(7,8), which is not a realistic decision-making situation. In addition, the participants had to evaluate hypothetical nutrient profiles⁽⁹⁾, or even food products, with unrealistic nutrition data⁽⁷⁾. In other studies, real products were used but were unbranded; therefore, important information that people may use in realistic decision-making situations was missing^(5,6). Some studies examined the impact of nutrition information on perceived healthiness⁽¹¹⁾. The product selected as healthier in a direct comparison is more relevant than the perceived healthiness of a single product, however.

Most of the studies examining the effects of FOP labels discussed the representativeness of the sample of participants, but none of the studies paid attention to the representativeness of the stimuli (i.e. products). If researchers are interested in how consumers make decisions in their shopping environment, representative stimuli should be used⁽²⁰⁾. Even if people rely on simple heuristics (e.g. cereals with chocolate pieces are unhealthy), the decisions may be smart if the heuristic exploits the information in the environment where it is used⁽²¹⁾. In other words, if an unrealistic decision set is used, as in the cases of most of the studies that we are aware of, the results provide little information about whether an FOP label helps consumers select healthier products in a real environment. A realistic decision situation implies that participants choose from a set of alternatives that they may encounter in a real shop and that all alternatives belong to the same decision set (e.g. consumers do not substitute a soft drink with an apple). In some of the reviewed studies, the alternatives belonged to the same decision set but unbranded products were used⁽⁵⁾. At least one study used a broad set of alternatives belonging to the same decision set, but it did not fully describe the product selection process⁽¹⁶⁾. Therefore, it remained unclear whether the participants could encounter all the selected products in one decision situation in real life.

In this study, we aimed to examine whether the participants would improve their ability to assess the healthiness of food products when a healthy choice label was introduced. To have a realistic product set, the participants had to compare the healthiness of a selection of cereals that could be found together in a grocery store. Our study's representative design allowed the assessment of the effect of a healthy choice label when the participants had all the information about the products, similar to a real shopping situation.

Methods

Selection of products

The participants had to choose the healthier option for all possible pairs of the selected products. Therefore, a small but still realistic sample of products had to be chosen. The fourteen breakfast cereals offered in the convenience stores owned by a major Swiss retailer were chosen for the experiment (Table 1).

Table 1. Breakfast cereals presented in the choice tasks, including each brand, label and nutrient profile score

| Product name | Brand | Label | Energy (kJ/100g) | Fat (saturated) (g/100g) | Carbohydrates (sugars) (g/100g) | Dietary fibre (g/100g) | Protein (g/100g) | Salt (g/100g) | Nutrient profile score | Healthy choice label present in the label condition |
|-------------------------|-----------------|---------|------------------|--------------------------|---------------------------------|------------------------|------------------|---------------|------------------------|---|
| Coconut Quinoa Granola | Coop | Karma | 2016 | 24 (12) | 54 (25) | 11 | 7 | 0.5 | 18 | No |
| Tresor | Kellogg's | - | 1883 | 16 (4) | 68 (29) | 3.3 | 6.9 | 1.13 | 15 | No |
| Choco Crunch | Familia | - | 1920 | 19 (8) | 59 (22) | 10 | 8 | 0.51 | 13 | No |
| Frosties | Kellogg's | - | 1569 | 0.6 (0.1) | 87 (37) | 2 | 4.5 | 0.83 | 13 | No |
| Cini Minis | Nestlé | - | 1749 | 10.1 (4.2) | 73.4 (24.8) | 5.2 | 5.7 | 1.1 | 13 | No |
| Crunchy Nut | Kellogg's | - | 1665 | 4.5 (0.7) | 82 (35) | 2.5 | 6 | 0.75 | 12 | No |
| C.M. Plus Original | Familia | - | 1766 | 12 (4) | 66 (23) | 7 | 9 | 0.51 | 10 | No |
| Crisp Muesli | Ovomaltine | - | 1757 | 12 (2.5) | 66 (20.1) | 6 | 9 | 0.8 | 9 | No |
| Smacks | Kellogg's | - | 1598 | 1.5 (0.4) | 84 (43) | 4 | 6 | 0.08 | 8 | No |
| Crunchy Muesli | Coop Naturaplan | Organic | 1941 | 20 (5) | 59 (18) | 6 | 9 | 0.4 | 8 | No |
| Nesquik | Nestlé | - | 1594 | 4 (1.8) | 74 (25.2) | 8.7 | 7.8 | 0.5 | 7 | No |
| Berry Muesli | Coop Naturaplan | Organic | 1490 | 3.5 (0.5) | 65 (9) | 10 | 11 | 0.3 | -4 | Yes |
| Birchermuesli | Coop Naturaplan | Organic | 1623 | 7 (1) | 64 (6) | 9 | 12 | 0.02 | -5 | Yes |
| Whole Grain Rolled Oats | Coop Naturaplan | Organic | 1611 | 8 (1) | 58 (1) | 11 | 14 | 0 | -6 | Yes |

Please choose the healthier product



Fig. 1. Example of a comparison pair from the choice task. Participants saw unblurred pictures.

The nutrient profile score⁽²²⁾ has been used in various studies as a 'gold standard' for the healthiness of food products⁽²³⁾. Therefore, we calculated each product's nutrient profile score and used it to judge the correctness of the participants' decisions regarding the healthiness of the cereals. The score increases with negative aspects, such as high amounts of fat, sugar and Na and decreases with positive ingredients, such as fibre and fruits. Thus, the higher the nutrient profile score, the less healthy the food is. The nutrient profile score is also used to identify less healthy foods for advertising control in the UK⁽²⁴⁾.

Labelling and nutrition table

Our goal was to examine whether a label that marked a healthy food product would help consumers correctly choose the healthy option. The criteria of the healthy choice label⁽²⁵⁾ were used to determine the cereals that received the label. Based on the Choices Programme⁽²⁶⁾, we tested whether the breakfast cereals had sufficient amounts of fibre (≥ 6.0 g/100 g) and no excessive amounts of saturated fat (≤ 3.0 g/100 g), Na (≤ 400 mg/100 g) and sugar (≤ 19.5 g/100 g). Since no such label currently exists in Switzerland, we designed the simple label shown in Fig. 1.

When consumers evaluate the healthiness of foods, they might check the nutrition table⁽²⁷⁾, especially consumers with greater nutrition knowledge and interest^(28,29). Thus, it is also of interest whether a healthy choice label results in better evaluations of the healthiness of food when the nutrition table is not shown to participants.

Procedure

We conducted an online experiment in which the fourteen cereals were presented in pairs (ninety-one paired comparisons), and the participants were asked to indicate which of the two cereals was healthier per paired item. The participants in all

| Nährwert pro 100 g | |
|-------------------------------------|---------------------|
| Energie | 2013 kJ 482 kcal |
| Fett davon gesättigte Fettsäuren | 24 g 12 g |
| Kohlenhydrate davon Zucker | 54 g 25 g |
| Ballaststoffe | 11 g |
| Eiweiss | 7 g |
| Salz | 0.5 g |

Fig. 2. Example of the nutrition information shown to participants assigned to two of the four conditions. For each cereal, information about energy, fat, saturated fat, carbohydrates, sugar, fibre, protein and sodium content was provided.

four groups were shown a picture of each FOP. Depending on the group, the following additional information was provided:

- (1) control condition (no additional information),
- (2) label condition (healthy choice label for the healthier product),
- (3) table condition (nutrition table shown on demand, as depicted in Fig. 2) and
- (4) combined condition (healthy choice label for the healthier product, nutrition table shown on demand).

The survey company randomly assigned the participants to one of the four conditions. (The survey software of the market research company randomly assigned the participants to one of the four groups. The randomisation was done in a way that age and sex were similar in all groups. To achieve the desired sample size of each group, additional people were invited in a second step. The survey company's project manager and the participants were blinded to the experimental manipulation and the tested hypotheses.) Therefore, neither the participants nor the researchers could influence the participants' assignments. The respondents were instructed to choose the product that, in their opinion, was healthier in all ninety-one possible product pairs. In the two conditions with the healthy choice label, before the task, the meaning of the label was explained with the following text: 'based on scientific criteria, products that are especially healthy breakfast cereals are marked with the following label'. Each breakfast cereal product was shown in a photo (Fig. 1). The respondents who were assigned to either of the two conditions with the nutrition table could use a cursor to click on a box to see the nutrition information about both cereals (Fig. 2 shows a sample image of the nutrition information). The pairs were presented in the optimum order for paired comparisons, as defined by Ross⁽³⁰⁾. The product shown on the left or the right side of each pair was randomly determined.

To examine the consistency of the responses, five selected comparisons were repeated at the end. If two products were quite similar (e.g. a two-point difference in the nutrition profile score), the participants might not have shown consistent answers, but this was irrelevant because the cereals were almost equally

Table 2. Recruited sample, number of excluded participants and demographics for each condition (Numbers and percentages; mean values and standard deviations)

| | Condition | | | | | | | |
|-----------------------------------|-----------|-------|----------|-------|----------|-------|-------------------------|-------|
| | Control | | Table | | Label | | Table/label combination | |
| | <i>n</i> | % | <i>n</i> | % | <i>n</i> | % | <i>n</i> | % |
| Recruited sample | 220 | | 220 | | 220 | | 222 | |
| Insufficient response time | 12 | | 9 | | 8 | | 4 | |
| Inconsistent responses | 15 | | 18 | | 18 | | 16 | |
| Final sample | 193 | | 193 | | 194 | | 200 | |
| Sex (female) | 101 | 52.33 | 99 | 51.30 | 97 | 50.00 | 105 | 52.50 |
| Age (years) | | | | | | | | |
| Mean | | 46.27 | | 46.72 | | 46.68 | | 46.11 |
| SD | | 14.85 | | 14.85 | | 16.12 | | 16.15 |
| Educational level | | | | | | | | |
| Primary school | 1 | 0.5 | 2 | 1.0 | 3 | 1.5 | 3 | 1.5 |
| Lower secondary school | 9 | 4.6 | 13 | 6.7 | 14 | 7.2 | 21 | 10.5 |
| Business school | 93 | 48.2 | 99 | 51.3 | 98 | 50.5 | 86 | 43.0 |
| Upper secondary vocational school | 31 | 16.1 | 25 | 13.0 | 27 | 13.9 | 52 | 26.0 |
| College/university | 59 | 30.6 | 54 | 28.0 | 52 | 26.8 | 38 | 19.0 |

healthy or unhealthy. Therefore, we randomly selected five test comparison pairs, with the requirement that the difference between the two products be at least four nutrient profile points.

After the choice task, the respondents filled in a questionnaire. First, they stated how often they consumed the breakfast cereals shown in the study. They then answered questions regarding their level of education and attitudes toward healthy eating.

Sample

The participants (German-speaking Swiss adults) were recruited from a market research company's (Respondi) online panel by inviting them via email. The respondents received an incentive of 0.86 CHF (Swiss francs) (0.93 USD) for participation, which they could save for a larger incentive by participating in more surveys offered by Respondi. The quota sampling by sex and age group was applied.

This study was conducted according to the guidelines laid down in the Declaration of Helsinki. All procedures involving human subjects were examined by the ethics committee of ETH Zurich (EK 2018-N-28).

The required sample size was calculated, aiming at the detection of small effects. Therefore, for a value of 0.05, a power of 0.90 and an effect size between 0.10 and 0.15, a sample size of about 800 people was required⁽³¹⁾.

In total, 882 respondents completed the survey. The participants who finished it after less than half of the median time (Mdn = 15.34 min) were excluded (*n* 33) from the analyses. The respondents who showed low consistency in their choices (less than four identical answers for the five repeated comparisons) were also excluded (*n* 67). This resulted in a final sample of 780 participants (51.5% female) with a mean age of 45 years. Of the sample, 8% (*n* 66) had completed lower secondary or primary school, almost two thirds (66%, *n* 511) had attended upper secondary vocational or business school, while one fourth (26%, *n* 203) held a university or a college degree. Thus, the majority of the respondents

had a higher degree of education, with a higher educational background than the general Swiss population, of which 18% had a university or a college degree⁽³²⁾. As shown in Table 2, the participants in the four conditions were comparable as expected (based on randomised allocation). (In the table/label combination condition, the percentage of the participants with a college/university degree was somewhat lower compared with the other three conditions. Therefore, the randomisation might not have worked as well as expected. We replicated all the analyses and included education as a factor. We did not find any significant condition × education interactions. Therefore, the unequal level of education across the four groups had no impact on our results.)

Data analysis

The data were analysed using SPSS Statistics, version 23 (IBM). The correct choices were identified for each pair by comparing the nutrient profile scores of the two products. Thus, the cereal products with the lower nutrient profile scores were defined as the correct choices. Exceptions were made for comparisons of products that had a difference of 0 or 1 point (approximately 10% of the total score). In these cases, the two products were classified as not different; thus, both choices were coded as correct. Both choices were considered correct in 15% of the comparisons. We calculated the proportion of correct choices by each participant.

The respondents who were assigned to either the table condition or the combination condition could consult the nutrition tables if they wanted. Examining the nutrition tables could have increased the accuracy of their choices. Therefore, the association between the frequency of consulting the nutrition table and the number of correct decisions was calculated.

A healthy choice label could only be used if the healthiness of a product with a label was compared with that of a product without a label. Therefore, we were interested in the difference between the participants' decisions if we focused on the comparisons where the label was more useful (i.e. one package had a label, and the other had none). We calculated the frequency

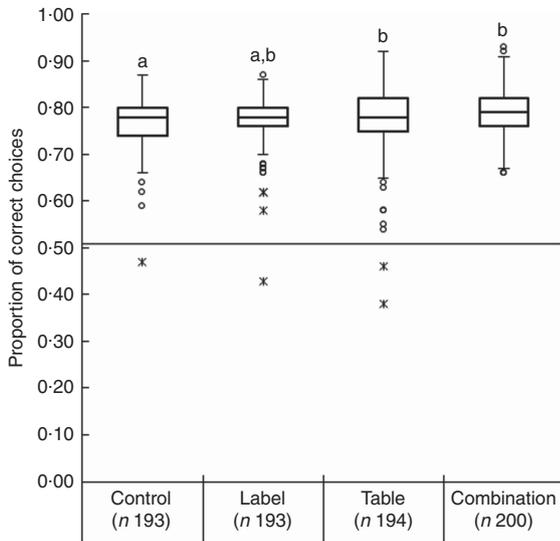


Fig. 3. Box plot of the proportion of correct choices in each condition (middle lines represent medians). ^{a,b} Significant differences ($P < 0.05$).

of the correct choices for these thirty-three comparisons by each participant. To obtain the values for comparison, this frequency was also calculated for the two conditions without the label.

Not all errors are equally critical. If two cereals substantially differ in their nutrient profile scores, choosing the wrong cereal as healthier is a more severe mistake compared with the situation where two cereals do not differ very much. Therefore, we conducted an additional analysis in which each error was weighted by the difference in the nutrition score between the two cereals.

As the values were not normally distributed, we conducted Kruskal–Wallis tests to analyse the effects of the label and the nutrition table on the accuracy and the weighted accuracy score. To compare the different conditions, a pairwise comparison using rank sums was conducted⁽³⁵⁾. We computed Spearman’s r correlation between the use of the nutrition table and both the accuracy and the weighted accuracy score, respectively.

Results

All of the selected cereals were reported to be consumed by some of the participants. The percentage of the participants who consumed a certain cereal at least a couple of times per year ranged between 14 and 58.

Choice accuracy

The median proportion of correct choices was higher than chance in all conditions (Fig. 3). The accuracy was significantly affected by the condition ($H(3) = 19.86$, $P < 0.001$). The accuracy did not significantly differ between the conditions with the label ($Mdn_{Label} = 78\%$, interquartile range (IQR) = 4%) and without the label ($Mdn_{Control} = 78\%$, IQR = 6%) ($P = 0.060$). The accuracy was significantly lower in the control condition than in the table condition ($Mdn_{Table} = 78\%$, IQR = 7%) ($P = 0.013$) and the combined condition ($Mdn_{Combination} = 79\%$, IQR = 6%) ($P < 0.001$). (Both the

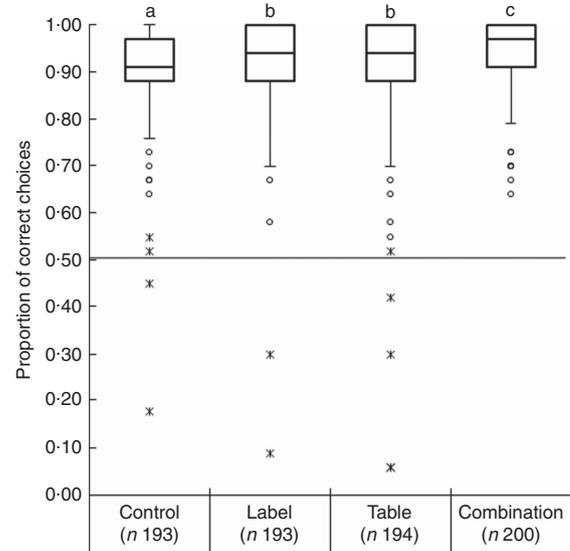


Fig. 4. Box plot of the proportion of correct choices for comparisons, including healthy choice options in each condition (middle lines represent medians). ^{a,b,c} Significant differences ($P < 0.05$).

Mann–Whitney U test⁽³⁴⁾ and the Kruskal–Wallis statistics⁽³⁵⁾ test the hypotheses based on the rank information that one distribution is stochastically smaller than the other distribution(s). If the shapes of the distributions differ across the groups, one group can be stochastically larger though the medians of both groups have the same value.) Overall, the differences among the four conditions were small and not of practical relevance. The respondents’ choices were not significantly more accurate with a combination of both nutrition tables and labels compared with the healthy choice labels only ($P = 0.063$).

The overall accuracy was significantly related to the frequency of the nutrition table consultation ($r_s = 0.40$, $P < 0.001$, 95% bias-corrected and accelerated bootstrap interval (BCa CI 0.31, 0.48)).

We then focused only on the comparisons where one of the cereals had a label and the other had none. The median proportion of correct choices for these thirty-three comparisons was higher than chance in all conditions (Fig. 4). The accuracy was significantly affected by the condition ($H(3) = 38.41$, $P < 0.001$). The median accuracy was significantly lower in the control condition ($Mdn_{Control} = 91\%$, IQR = 9%) than in the table ($Mdn_{Table} = 94\%$, IQR = 12%, $P = 0.004$), the label ($Mdn_{Label} = 94\%$, IQR = 12%, $P = 0.004$) and the combined conditions ($Mdn_{Combination} = 97\%$, IQR = 9%, $P < 0.001$). The accuracy was also significantly higher in the combination of both the nutrition table and the label than in the label ($P = 0.035$) and the table conditions ($P = 0.036$). The accuracy did not significantly differ between the respondent groups assigned to the healthy choice label and the nutrition table conditions. Overall, the differences among the four conditions were small and not of practical relevance.

How severe are the errors made?

In the second step, the weighted errors were analysed. The median overall nutrient profile score ($Mdn = 120$ points) was

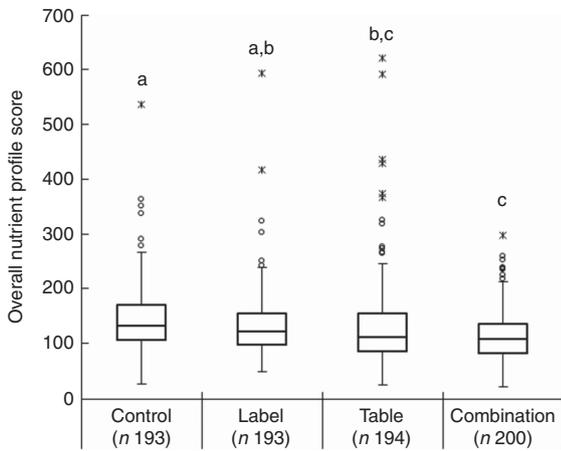


Fig. 5. Box plots of the overall nutrition score of choices made by respondents (middle lines represent medians). ^{a,b,c}Significant differences ($P < 0.05$).

relatively close to the lowest possible score of 0 point (if all ninety-one choices had been accurate) (Fig. 5). Nevertheless, the overall weighted scores were significantly affected by the condition ($H(3) = 38.78$, $P < 0.001$). The participants who were assigned to the label condition ($Mdn_{Label} = 123$, $IQR = 57$) did not have a significantly lower weighted score than the participants in the control group ($Mdn_{Control} = 134$, $IQR = 65$, $P = 0.22$). The overall nutrition score was higher in the control condition than in the table ($Mdn_{Table} = 114$, $IQR = 70$, $P < 0.001$) and the combined conditions ($Mdn_{Combination} = 109$, $IQR = 53$, $P < 0.001$). The weighted score was significantly lower in the combined condition than in the label condition ($P = 0.001$) but not significantly different from the table condition ($P = 0.40$).

The frequent use of nutrition tables could result in choices with a lower weighted score. The weighted score was significantly related to the frequency of the nutrition table consultation ($r_s = -0.48$, $P < 0.001$, 95% BCa CI -0.57 , -0.39). The negative correlation indicated that the overall weighted score decreased with an increasing use of the nutrition tables. In other words, the participants who more frequently consulted the nutrition tables chose healthier cereals compared with those who looked less often at the nutrition tables.

Discussion

In many countries, people's diets are not as balanced as they should be^(36–38). There are at least two possible reasons that do not exclude each other: consumers are not motivated to change their diets, or they lack the ability to choose healthier foods. It has been shown that consumers have a limited knowledge of nutrition^(39–42). Therefore, simple information, such as a healthy label, may be perceived as an effective way to enable consumers to select healthier products at the point of purchase. It should also be noted that in many situations, no elaborate nutrition knowledge is needed to make a healthier choice; simple heuristics may be sufficient for people to determine whether a food item is healthy. Of course, the use of simple heuristics may result in biased perceptions on foods⁽⁴³⁾, but it may also lead to correct decisions if the cue used in the heuristic

(e.g. sugar content) strongly correlates with the healthiness of the product⁽⁴⁴⁾. The usefulness of simple heuristics (e.g. cereals with pictures of chocolates on their packages are unhealthy) depends on the characteristics of the choice environment (e.g. cereals with chocolate images on their packages are on average less healthy than other cereals).

To examine how well consumers can evaluate the healthiness of foods, the stimuli used should be representative of the decisions made by consumers at the point of purchase^(20,21,45). In this study, we examined whether a healthy choice label would help consumers make better decisions in a realistic choice environment. The results of our experiment suggest that the participants could very well evaluate the healthiness of the cereals based on the information that could already be found on the cereals' FOP. The participants who were assigned to the condition with the healthy choice label did not perform substantially better compared with the control group. We found that providing the nutrition tables slightly improved the participants' performance, but adding healthy labels did not result in further improvement.

The healthy choice label can be used in a limited number of decision-making situations. If two cereals are marked with healthy choice labels, or neither of the two is labelled, the label cannot be used to select the healthier product. Even if only the decisions where the label could be used (one package with a label and the other without it) are analysed, the impact of the label is rather limited. The main reason for this finding is that the participants who are assigned to the control condition have already performed extremely well, and due to this ceiling effect, the impact of a label is restricted. The participants' performance in the task is even more impressive if the errors are weighted by the difference in the nutrient profile scores of the two products. These weighted errors suggest that participants mainly make mistakes when they have to evaluate cereals with relatively similar scores but not when the products clearly differ in their nutrient profile scores.

Our study differed in two important aspects from previous studies that examined the impact of healthy choice labels on people's health perception about foods^(5,8,10,14–19). We used a realistic choice set of products that could be found together in a store, and the participants were provided with the same information as they normally would have in a store, such as the brands of the cereals and the information shown on the FOP. We decided not to provide any price information because it would not be needed in evaluating the healthiness of cereals. Furthermore, we used an objective score (i.e. nutrient profile score) as the 'gold standard' that allowed the evaluation of the participants' performance. Therefore, the present study provides information beyond that of past studies about the effect of a healthy choice label on consumers' ability to choose healthy products.

Some possible limitations of the present study should be addressed. A crucial limitation is the investigation of only one product category (i.e. breakfast cereal). Future studies should examine whether similar results can be observed in other product categories (e.g. processed meat, snacks). Other shortcomings are related to using the Internet for data collection and the respondents receiving a token payment for their

participation. Therefore, we had little control over how motivated the participants were in evaluating the healthiness of the cereals. If some less motivated people participated in the study, we might even have underestimated their ability to decide which product was healthier. The healthy choice label used in this experiment was new to the participants. Although we explained the meaning of the label, and the label itself was easy to interpret, we could not rule out that the participants' lack of familiarity influenced the results. In this study, the data were collected using the Internet. Therefore, we could not rule out the inference that this study's results offered a better prediction of the perceived healthiness of cereals in an online store compared with a brick-and-mortar store. The participants were also slightly better educated than the general Swiss population. This might also have influenced the observed results. Another drawback might be the low number of products with healthy choice labels. However, the reality is that only a small number of cereals found in many stores qualify for a healthy choice label. Should the composition of healthy/unhealthy cereals differ in another store or another country, the effect of a healthy choice label might vary. This issue is certainly an avenue for further research. Future studies should also examine whether similar effects can be observed for the traffic light or the Nutri-Score labels compared with the label used in the present study. We have also assumed that cereals with similar nutrition scores do not have significantly different health impacts.

Could this study's results motivate the food industry to change its product packages, so that consumers would perceive the cereals as healthier? This case seems unlikely. The reason is that the FOP needs to be attractive not only to health-conscious consumers but also to purely hedonistic-oriented consumers. If chocolate pieces are shown on the FOP, this not only provides information about the sensory aspects of the cereals but is also a cue that can be used by consumers to assess the healthiness of the cereals.

The healthy choice label may not only provide information about the healthiness of a product but also remind consumers of the importance of choosing healthy products. Such an effect has been found in a study showing that traffic light labels on products can increase consumers' awareness of healthy food⁽⁴⁶⁾. Another positive effect of a healthy choice label could be its motivation for the food industry to improve the latter products' nutritive content to justify the use of the label.

The present study shows that consumers perform well in evaluating the healthiness of cereals if the stimuli are selected to be representative of what consumers confront in a real purchase situation. In our study, the participants' performance may be surprising, given that various studies suggest consumers' limited nutrition knowledge^(40,41,47). The findings are less surprising if the reader realises that people's performance in a decision task is influenced not only by their ability but also by the characteristics of the decision situation. If the environment-specific cues (e.g. fruits in cereals) are correlated with the criteria that should be predicted (e.g. healthiness), the use of simple cues can result in smart decisions^(21,45).

The present research findings also have practical implications. If a healthy choice label does not substantially increase consumers' ability to choose healthier foods, the value of such a label could be questioned. Nevertheless, a healthy choice label

might have a priming effect (even if consumers' abilities to choose healthy products are not enhanced) and could therefore increase consumers' motivation to buy healthier products. A field experiment in an online shop would be a promising method to examine the effects of such a label on purchase behaviour. The present research results not only show that a healthy choice label does not increase consumers' ability to make better decisions but also demonstrate the importance of representative research designs in examining consumers' decision-making.

Acknowledgements

No funding for this study was obtained.

All authors contributed to the study design. G. A. L. directed the study and conducted the analysis. M. S. and G. A. L. wrote the article by incorporating critical inputs from C. H. All authors approved the final version of the article.

The authors declare no conflicts of interest.

References

1. Kant AK (2004) Dietary patterns and health outcomes. *J Am Diet Assoc* **104**, 615–635.
2. Cowburn G & Stockley L (2005) Consumer understanding and use of nutrition labelling: a systematic review. *Public Health Nutr* **8**, 21–28.
3. Grunert KG & Wills JM (2007) A review of European research on consumer response to nutrition information on food labels. *J Public Health* **15**, 385–399.
4. Hieke S & Taylor CR (2012) A critical review of the literature on nutritional labeling. *J Consum Aff* **46**, 120–156.
5. Borgmeier I & Westenhofer J (2009) Impact of different food label formats on healthiness evaluation and food choice of consumers: a randomized-controlled study. *BMC Public Health* **9**, 184.
6. Gorski Findling MT, Werth PM, Musicus AA, *et al.* (2018) Comparing five front-of-pack nutrition labels' influence on consumers' perceptions and purchase intentions. *Prev Med* **106**, 114–121.
7. Hieke S & Wilczynski P (2012) Colour Me In – an empirical study on consumer responses to the traffic light signposting system in nutrition labelling. *Public Health Nutr* **15**, 773–782.
8. Hodgkins CE, Raats MM, Fife-Schaw C, *et al.* (2015) Guiding healthier food choice: systematic comparison of four front-of-pack labelling systems and their effect on judgements of product healthiness. *Br J Nutr* **113**, 1652–1663.
9. Jones G & Richardson M (2007) An objective examination of consumer perception of nutrition information based on healthiness ratings and eye movements. *Public Health Nutr* **10**, 238–244.
10. van Herpen E, Hieke S & van Trijp HCM (2014) Inferring product healthfulness from nutrition labelling. The influence of reference points. *Appetite* **72**, 138–149.
11. van Herpen E, Seiss E & van Trijp HCM (2012) The role of familiarity in front-of-pack label evaluation and use: a comparison between the United Kingdom and the Netherlands. *Food Qual Prefer* **26**, 22–34.
12. Hawley KL, Roberto CA, Bragg MA, *et al.* (2013) The science on front-of-package food labels. *Public Health Nutr* **16**, 430–439.
13. Hersey JC, Wohlgenant KC, Arsenault JE, *et al.* (2013) Effects of front-of-package and shelf nutrition labeling systems on consumers. *Nutr Rev* **71**, 1–14.

14. van Herpen E & van Trijp HCM (2011) Front-of-pack nutrition labels. Their effect on attention and choices when consumers have varying goals and time constraints. *Appetite* **57**, 148–160.
15. Andrews JC, Burton S & Kees J (2011) Is simpler always better? Consumer evaluations of front-of-package nutrition symbols. *J Public Policy Mark* **30**, 175–190.
16. Aschemann-Witzel J, Grunert KG, van Trijp HCM, *et al.* (2013) Effects of nutrition label format and product assortment on the healthfulness of food choice. *Appetite* **71**, 63–74.
17. Roberto CA, Bragg MA, Livingston KA, *et al.* (2012) Choosing front-of-package food labelling nutritional criteria: how smart were ‘Smart Choices’? *Public Health Nutr* **15**, 262–267.
18. Roberto CA, Shivaram M, Martinez O, *et al.* (2012) The Smart Choices front-of-package nutrition label. Influence on perceptions and intake of cereal. *Appetite* **58**, 651–657.
19. Scott V & Worsley AF (1994) Ticks, claims, tables and food groups – a comparison for nutrition labeling. *Health Promot Int* **9**, 27–37.
20. Dhami MK, Hertwig R & Hoffrage U (2004) The role of representative design in an ecological approach to cognition. *Psychol Bull* **130**, 959–988.
21. Todd PM & Gigerenzer G (2007) Environments that make us smart: ecological rationality. *Curr Dir Psychol Sci* **16**, 167–171.
22. Rayner M, Scarborough P, Boxer A, *et al.* (2005) *Nutrient Profiles: Development of Final Model*. London: Food Standards Agency.
23. Bucher T, Müller B & Siegrist M (2015) What is healthy food? Objective nutrient profile scores and subjective lay evaluations in comparison. *Appetite* **95**, 408–414.
24. Lobstein T & Davies S (2009) Defining and labelling ‘healthy’ and ‘unhealthy’ food. *Public Health Nutr* **12**, 331–340.
25. Choices Programme (2018) The choices programme <https://www.choicesprogramme.org>, accessed March 2018.
26. Choices Programme (2016) *Choices Programme International Product Criteria*. Brussels: Choices International Foundation.
27. Visschers VHM, Hess R & Siegrist M (2010) Health motivation and product design determine consumers’ visual attention to nutrition information on food products. *Public Health Nutr* **13**, 1099–1106.
28. Barreiro-Hurlé J, Gracia A & de-Magistris T (2010) Does nutrition information on food products lead to healthier food choices? *Food Policy* **35**, 221–229.
29. Visschers VHM, Hartmann C, Leins-Hess R, *et al.* (2013) A consumer segmentation of nutrition information use and its relation to food consumption behaviour. *Food Policy* **42**, 71–80.
30. Ross RT (1934) Optimum orders for the presentation of pairs in the method of paired comparisons. *J Educ Psychol* **25**, 375–382.
31. Cohen J (1988) *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. New York: Lawrence Erlbaum Associates.
32. Swiss Statistics (2016) *Bevölkerungsstand und -struktur (Population Size and Composition)*. Neuenburg: Swiss Statistics.
33. Dunn OJ (1964) Multiple comparisons using rank sums. *Technometrics* **6**, 241–252.
34. Mann HB & Whitney DR (1947) On a test of whether one of two random variables in stochastically larger than the other. *Ann Math Stat* **18**, 50–60.
35. Kruskal WH & Wallis WA (1952) Use of ranks in one-criterion variance analysis. *J Am Stat Assoc* **47**, 583–621.
36. WHO (2003) *Diet, Nutrition and the Prevention of Chronic Diseases*. WHO Technical Report Series. Geneva: World Health Organization.
37. WHO (2015) *Healthy Diet. Fact Sheets*. Geneva: World Health Organization.
38. WHO (2015) *Obesity and Overweight. Fact Sheets*. Geneva: World Health Organization.
39. Dickson-Spillmann M, Siegrist M & Keller C (2011) Development and validation of a short, consumer-oriented nutrition knowledge questionnaire. *Appetite* **56**, 617–620.
40. Motteli S, Barbey J, Keller C, *et al.* (2016) Measuring practical knowledge about balanced meals: development and validation of the brief PKB-7 scale. *Eur J Clin Nutr* **70**, 505–510.
41. Parmenter K & Wardle J (1999) Development of a general nutrition knowledge questionnaire for adults. *Eur J Clin Nutr* **53**, 298–308.
42. Spronk I, Kullen C, Burdon C, *et al.* (2014) Relationship between nutrition knowledge and dietary intake. *Br J Nutr* **111**, 1713–1726.
43. Sütterlin B & Siegrist M (2015) Simply adding the word ‘fruit’ makes sugar healthier: the misleading effect of symbolic information on the perceived healthiness of food. *Appetite* **95**, 252–261.
44. Bucher T & Siegrist M (2015) Children’s and parents’ health perception of different soft drinks. *Br J Nutr* **113**, 526–535.
45. Gigerenzer G (2008) Why heuristics work. *Perspect Psychol Sci* **3**, 20–29.
46. Sonnenberg L, Gelsomin E, Levy DE, *et al.* (2013) A traffic light food labeling intervention increases consumer awareness of health and healthy choices at the point-of-purchase. *Prev Med* **57**, 253–257.
47. Parmenter K & Wardle J (2000) Evaluation and design of nutrition knowledge measures. *J Nutr Educ* **32**, 269–277.

