Food Quality and Preference 83 (2020) 103894

Contents lists available at ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

Nutri-Score, multiple traffic light and incomplete nutrition labelling on food packages: Effects on consumers' accuracy in identifying healthier snack options

Désirée Hagmann*, Michael Siegrist

Department of Health Sciences and Technology (D-HEST), Consumer Behavior, Universitätstrasse 22, CH 8092 Zurich, Switzerland

ARTICLE INFO

Keywords:
Nutrition labels
Multiple traffic light
Nutri-score
Nutrition facts table
Healthiness evaluation
Nutrition information

ABSTRACT

Front-of-package (FOP) nutrition labels are designed to help consumers evaluate the healthiness of foods and to promote healthier food choices. In this study, an online experiment with Swiss consumers ($N = 1313$) was conducted to compare the effects of different nutrition label formats on consumers' evaluations of snack food healthiness. Participants were asked to select the healthier option in 105 pairwise comparisons of 15 salty snacks. The participants were randomly assigned to one of five conditions: the FOP presented with (1) the nutrition facts table, (2) the multiple traffic light (MTL), (3) the Nutri-Score, (4) the Nutri-Score on half of the products, or (5) no nutrition information (control). The consumers' evaluations of the snacks' healthiness were fairly accurate, even without being given nutrition information on the packaging. The Nutri-Score led to the greatest accuracy in identifying the healthier of two snacks (when using the British FSA/Ofcom nutrient profiling score to determine product healthiness); however, this had only a minimal effect on the evaluation when only some of the products were labelled. Both FOP labels were superior to the FOP with and without the nutrition facts. This indicates that for maximum effectiveness, the labelling of all available products is needed. The perceived usefulness and public support of mandatory implementation were higher for the MTL than for the Nutri-Score label; however, for the latter, perceived usefulness and public acceptance were higher among the participants who became familiar with the label during the experiment than among those who did not.

1. Introduction

When grocery shopping, consumers are confronted with many different kinds of information on product packaging. In view of the increasingly unhealthy dietary habits in many countries (e.g. increased consumption of energy-dense, highly processed foods and snack products; Jones & Richardson, 2007; Mattes, 2018; WHO, 2003), the provision of unambiguous and comprehensible nutrition information is important. Nutrition labels, particularly those with front-of-package (FOP) positioning, are intended to help consumers evaluate the healthiness of processed foods and thus enable informed food choices (WHO, 2003). In addition to this, it is hoped that the presence of labels on food packages will create an incentive for the food industry to reformulate their products and offer healthier options (Kanter, Vanderlee, & Vandevijvere, 2018) in order to avoid adverse effects on the marketing of their products and negative evaluations of the products themselves.

In recent decades, various nutrition label formats have been introduced (Kanter et al., 2018). These differ in several respects: the types of nutrients on which they focus (e.g., highlighting only critical nutrients or also considering health-promoting nutrients), the kind of presentation/design features they use (e.g., using numbers, colour codes, shapes, or letters), and how directive they are (Hodkins et al., 2009). The mandatory nutrition facts table on the back of the package can be considered a nondirective label because it provides detailed numerical information about the nutritional components of a product without explicitly evaluating the product's healthiness. Semidirective nutrition labels, such as the multiple traffic light (MTL) signpost, use visual cues such as colour codes or symbols to communicate an evaluation of the product's critical nutrient content. On the MTL label, each nutrient attribute (the amount of fat, saturated fatty acids, sugar, and salt/sodium) is represented by a separate symbol that indicates whether the amount is low (green), medium (amber), or high (red). These labels do not provide a global evaluation of the product's healthiness. Directive labels, by contrast, provide a summary evaluation of the healthiness of a product without any detailed information. These summary

* Corresponding author.
E-mail addresses: desiree.hagmann@hest.ethz.ch (D. Hagmann), michael.siegrist@hest.ethz.ch (M. Siegrist).

https://doi.org/10.1016/j.foodqual.2020.103894
Received 2 October 2019; Received in revised form 27 December 2019; Accepted 23 January 2020
Available online 27 January 2020

labels include simple labels placed only on foods that meet certain healthiness criteria (e.g., Keyhole, Green Tick, and Choices labels) and graded labels (e.g., Nutri-Score and Health Star Rating labels) (Julia & Hercberg, 2017).

According to epidemiological nutrition research, healthy diets contain plenty of fruit, vegetables, fibre, plant-based sources of fat and protein, and low amounts of fat, saturated fat, total sugar, and salt, among others (Willett & Stampfer, 2013). A relatively new method, nutrient profiling (NP), enables the evaluation and ranking of food products according to the healthiness of their nutritional composition (WHO, 2017). Various NP models exist, such as the Ofcom/FSA NP model (Food Standards Agency, 2011) and the Health Canada Surveillance Tool (HCST) tier system (Health Canada, 2014). Each of these models includes a different number of health-relevant nutrients and serves as a basis for the classification schemes on nutrition labels and the determination of food-related health taxes (Rayner, 2017). Currently, there is no consensus regarding which model should be considered the gold standard for objectively defining the healthiness of foods (Poon et al., 2018). However, the Ofcom/FSA NP model is one of the most well-known and well-validated models (Rayner, 2017), and it is considered the gold standard by a growing number of countries and food producers, which are introducing the Nutri-Score (the label based on this model) to communicate the healthiness of products to consumers in a simple way.

Numerous studies have evaluated the impact of different nutrition labels on consumers' perceptions of the healthiness of foods and have sought to determine which of the available formats is the best means of communicating nutrition information (e.g., Borgmeier & Westenhoefer, 2009; Egnell, Talati, Hercberg, Pettigrew, & Julia, 2018; Gorski Findling et al., 2018; Hawley et al., 2013; Hersey, Wohlgenant, Arsenault, Kosa, & Muth, 2013; Hodkins et al., 2009; Jones & Richardson, 2007; Julia & Hercberg, 2017; Roberto et al., 2012; Siegrist, Hartmann, & Lazzarini, 2019; Watson et al., 2014).

Studies based on eye-tracking methods have suggested that compared to the standard nutrition facts panel, FOP labels, especially those that use a traffic light system, are better able to catch consumers' attention and direct their attention to the nutrients most relevant to healthiness assessments (Becker, Bello, Sundar, Peltier, & Bix, 2015; Jones & Richardson, 2007; van Herpen & Trijp, 2011). This may be due to the more prominent placement of such FOP labels and their design features (Becker et al., 2015). Similarly, the results of another eye-tracking study (Siegrist, Leins-Hess, & Keller, 2015) suggest that the visual information processing of the MTL label is more efficient overall than that of the Guideline Daily Amounts (GDA) and the nutrition facts table; in this study, the MTL label was processed more quickly than the GDA label (but less quickly than the nutrition facts table), and participants focussed on more relevant information when reading both the MTL label and the GDA than when reading the nutrition facts table. The nutrition facts table contains only numerical information, which can be difficult to understand, especially for consumers who have limited literacy skills (Campos, Doxey, & Hammond, 2011; Roberto & Khandpur, 2014). Consequently, several studies that compared different label formats have found that the MTL system resulted in more accurate healthiness evaluations compared to no label and other label formats, such as the GDA and the simple 'healthier choice' tick (Borgmeier & Westenhoefer, 2009; Roberto et al., 2012); however, other studies did not find substantial differences between different types of FOP labels (Hodgkins et al., 2015; Watson et al., 2014). More recent studies have included the new Nutri-Score label, which was developed in France (Julia & Hercberg, 2017). This label provides a graded, colour-coded summary evaluation of a product's healthiness, ranging from dark-green A (healthiest) to dark-red E (least healthy) and considers the content of various health-promoting and critical ingredients. Previous findings have suggested that this label is easier for consumers to understand and results in more accurate healthiness evaluations than the MTL and other labelling systems (Ducrot et al., 2015; Egnell et al., 2018). However, Gorski Findling et al. (2018) found that the MTL led to greater accuracy in identifying the healthier of two foods compared to a labelling scheme based on 0–3 stars, another type of graded summary label.

Previous studies have differed widely in terms of their design, the labels (or versions of labels) compared, the food categories used, and how the stimuli were presented to the consumers. For example, many previous studies did not use real brands available in supermarkets (Borgmeier & Westenhoefer, 2009; Egnell et al., 2018; Watson et al., 2014) or presented labels only, without the products themselves (Hieke & Wilczynski, 2012; Jones & Richardson, 2007); both of these scenarios make the decision situation less realistic. Furthermore, few studies so far have included the Nutri-Score label (Julia & Hercberg, 2017). One study investigated the effect of the Nutri-Score on the healthiness of food purchases in experimental supermarkets (Julia et al., 2016) but found no effect for salty snacks. To the best of our knowledge, no studies have investigated the impact of nutrition labels on consumers' perceptions of the healthiness of a realistic set of salty snack foods from existing brands.

The results of all the studies we are aware of were based on the assumption that all available products carry a label and can therefore be compared by consumers. However, in many cases, the implementation of nutrition labels is not mandatory (Buttriss, 2018; Kanter et al., 2018); thus, it is likely that situations occur in which only some of the available products are labelled. In Switzerland, for example, the French food company Danone recently began to place the Nutri-Score label on all their dairy products (Danone, 2019), whereas other producers of dairy products have not implemented it. This raises the question of whether nutrition labels such as the Nutri-Score are equally effective when they are not present on all available products. To the best of our knowledge, no previous studies have investigated the effectiveness of a nutrition label under the condition of incomplete labelling.

The present study had several aims. The first was to compare two interpretive FOP nutrition labels (MTL and Nutri-Score) in terms of their effect on consumers' healthiness evaluation of salty snacks, as well as to compare these labels with the standard nutrition facts table and the absence of nutrition information. In order to create a relatively realistic shopping-choice situation (with high ecological validity), a range of snacks offered by the same Swiss retailer, all of which are real brands available at stores, were used. We focussed on salty snacks because this product category is highly relevant, considering that snacking contributes nearly one-third to European consumers' daily energy intake (Mattes, 2018). In addition, salty snacks usually contain critical amounts of sodium, fats, and sugar (Foundation, 2016). Nevertheless, this product category offers some variability in terms of healthiness, which makes it appropriate for the purpose of the present study. The second aim of this study was to investigate whether the effectiveness of the Nutri-Score label differs when it appears on only some products. Moreover, this study explored, among a representative sample of Swiss consumers, the perceived usefulness of the Nutri-Score and the MTL labels compared to the nutrition facts table and the ingredients list, as well as public support for the mandatory introduction of these two labels.

2. Material and methods

2.1. Selection of snacks

In choosing a set of salty snacks consumers might encounter simultaneously in a real-world shopping situation, we used a range of products offered by a large Swiss retailer. Initially, a larger set of salty snacks was considered, and 15 snack products from this set were ultimately selected. The following criteria were considered for the final selection:

- All products should be available from the same retailer/store.
- The products should exhibit a certain variability in terms of healthiness (overall and in terms of fat, sugar, and salt content), type, ingredients, and origin (animal or vegetable) to ensure variability in the labels (e.g., all Nutri-Score categories from A to E should be represented).
- The products should not be overly similar (e.g., if salted pretzels are included, pretzel sticks should not be chosen).

The product characteristics of the salty snacks used in the experiment are presented in Appendix 1. Information about the nutritional values and ingredients was taken from the product packaging. If a relevant piece of information was missing, the website of the producer or retailer of the product was consulted.

2.2. Procedure

An online study was conducted that consisted of an experimental part and a short questionnaire, which was distributed subsequently. In the experimental part, participants performed a choice task: For each possible pairwise combination of 15 salty snacks, they were asked to indicate which of the two snacks was healthier (each participant made 105 comparisons in total). The pairwise comparisons were presented in an optimum order, as suggested by Ross (1934), in order to establish a maximum time period between the presentation of the same product and balanced variability in the position (left or right). By means of a script programmed by the online panel company Respondi, the participants were randomly assigned to one of the following five conditions:

- (1) *FOP only* condition (control condition): In this condition, only a picture of the front of the package of each snack product was presented, without any additional information or evaluation of the product's nutritional content.
- (2) *MTL label* condition: A German adaptation of the MTL label (Department of Health/Food Standards Agency, 2016) was created (see Fig. 1) and presented below each snack picture. This nutrient-specific label provides an evaluation of the content of fat, saturated fat, sugar, and salt per 100 g of the product. Colour coding is used to highlight the content of these four nutrients as either low (green), medium (amber), or high (red), according to reference values defined by the Food Standards Agency (Department of Health/Food Standards Agency, 2016). Additionally, this label indicates the energy content (which is not evaluated by a colour) and provides numerical information about the nutrient content of a standard snack portion (i.e., 25 g according to the Swiss Society for Nutrition, 2011).
- (3) *Nutrition facts table (table)* condition: In this condition, the standard back-of-package nutrition facts table was presented below each product (see Fig. 1). By default, this table contains the nutritional values per 100 g of the product for energy, fat, saturated fat, carbohydrates, sugar, fibre, protein, and salt.
- (4) *Nutri-Score* condition: In this condition, the Nutri-Score label (Julia & Hercberg, 2017) was presented below each snack package (see Fig. 1). This label is based on the NP system of the UK Food Standards Agency (2011), which evaluates the overall healthiness of a food product according to its nutritional composition (Food Standards Agency, 2011). For the healthiness classification, the product's content of several health-promoting and critical nutrients (i.e., content of energy, fruit, vegetables and nuts, fibre, saturated fat, total sugar, sodium, and protein) is evaluated. This results in a single NP score. On the Nutri-Score label, this score is subdivided into five categories that are represented by capital letters and colour coded, ranging from A (dark green) to E (dark red), where A represents foods considered the healthiest and E represents foods considered the least healthy. Thus, this summary label provides a graded overall evaluation of the healthiness of a food product.
- (5) *Partial Nutri-Score* condition (*partial*): In order to simulate a choice situation in which only some products were labelled, approximately half of the products (7 of 15) were randomly selected, and each of these was presented with a Nutri-Score label. By contrast, for the remaining half, only the front of the package was presented, without any further information or labelling.

To determine the relative healthiness of the 15 snacks, the Ofcom/FSA NP score of each product was calculated based on its nutritional composition (Food Standards Agency, 2011; see Appendix 1). The score represents an objective and validated measure for the healthiness of a food and is based on a given food's nutrient content per 100 g. For its calculation, 0–10 'A' points are assigned for each unhealthy aspect (i.e., for the amount of energy, saturated fatty acids, total sugar, and sodium – representing a maximum of 40 'A' points in total), and 0–5 'C' points are assigned for each healthy aspect (i.e., for the content of fruits, vegetables, and nuts, fibre, and protein[1]; – representing a maximum of 15 'C' points in total). For the final Ofcom/FSA NP score, the 'C' points are subtracted from the 'A' points. The final score can range between –15 and 40, with higher scores indicating lower healthiness. Foods scoring 4 or above are classified as 'less healthy' (for more details, see Food Standards Agency, 2011).

In the choice task, responses were counted as correct if the product with the lower NP score (= healthier) was chosen. If the difference in the NP score was minimal (i.e., between 0 and 1 points), the options were classified as not different, and consequently, both alternatives were counted as correct. The Ofcom/FSA NP scores of each snack product used in the experiment are shown in Appendix 1.

Questionnaire. After the choice task, the participants completed a short questionnaire. Among other things, they were asked to report how useful they considered the Nutri-Score label, the MTL label, the nutrition facts table, and the ingredients list for the healthiness evaluation (on a 7-point Likert-type scale ranging from 1 ['not at all useful'] to 7 ['very useful']). The participants were also asked whether they thought these two labels should be mandatory on products in Switzerland (response options: 'yes', 'no', or 'I don't know'). To enable the participants from all the conditions to answer this question, the same short description/explanation of the two labels that the participants with the MTL and Nutri-Score conditions had received prior to the choice task was provided to participants in all conditions.

The purchase frequency of pre-packaged snacks was assessed using a 7-point Likert-type scale ranging from 1 ('rarely/ever') to 7 ('very often'), and the consumption frequency of the salty snacks was measured using nine categories ranging from '4 times or more a day' to 'rarely/never'. Each consumer's educational level was placed into one of the following three categories: low (compulsory school), medium (vocational or middle school), or high (higher vocational education or university degree). The consumers' educational backgrounds were collected in order to avoid confounding through these variables.

2.3. Study participants

The participants were recruited from the German-speaking part of Switzerland through the online panel company Respondi. All respondents gave their written consent and received a monetary incentive of CHF 1.14 (USD 1.14) for their participation in the study. Quotas for sex and age groups were defined in order to obtain samples representative of the Swiss population in each condition and to minimise

[1] Points for protein are not included if the 'A' point total is ≥11 and if fruit, vegetables, and nuts score < 5 points.

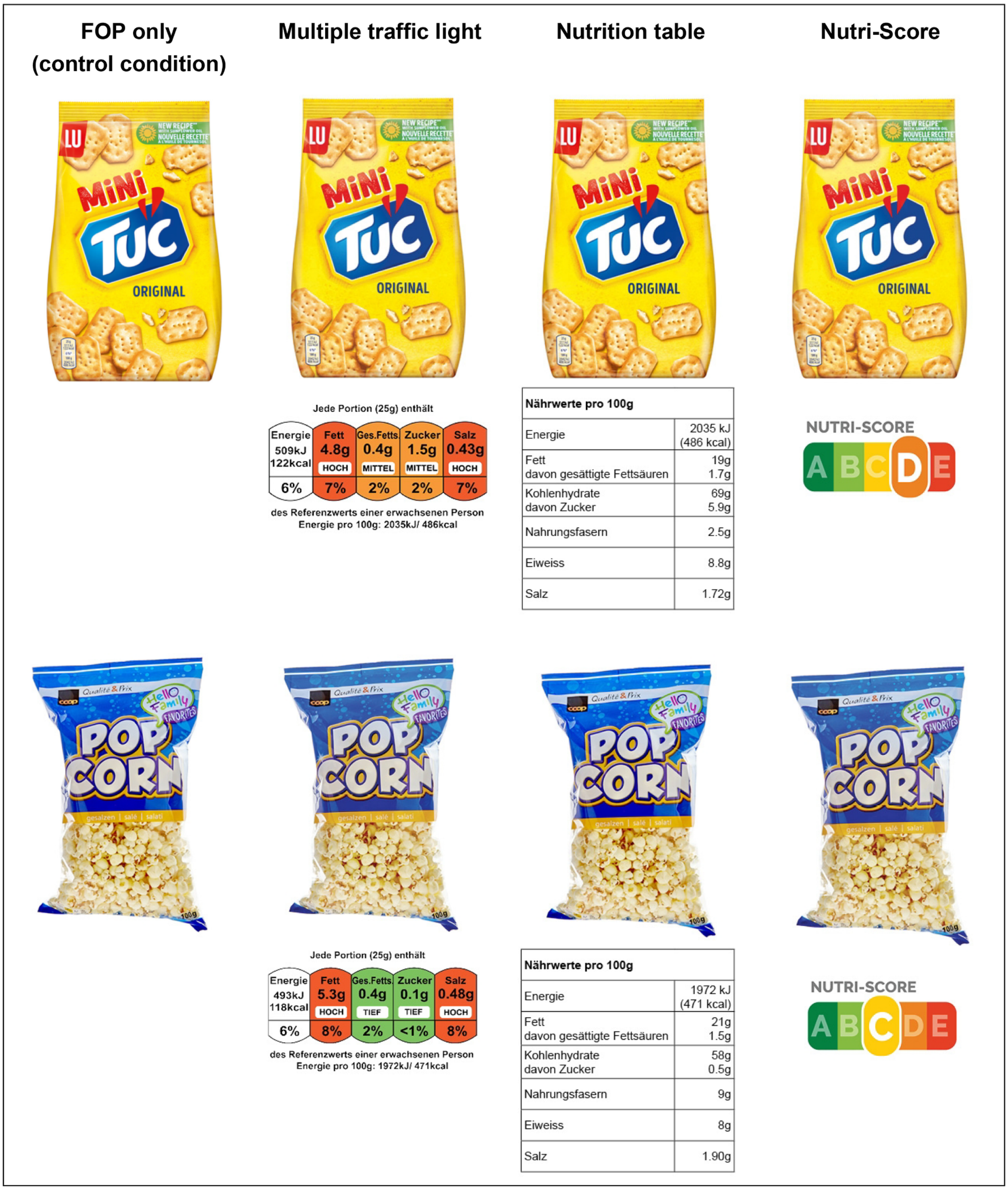

Fig. 1. Product examples for the experimental conditions (for the partial Nutri-Score condition, half of the stimuli from the Nutri-Score condition and half of the stimuli from the control condition were presented).

confounding through these variables.

The sample size required to detect small effects (Cohen's $f = 0.10$) was calculated. Given an alpha level of 0.05 and a power of 0.80, a minimum sample of 240 participants per condition was needed (Cohen, 1988).[2]

In total, 1561 participants completed the online study (see Table 1). Sixty-nine participants were excluded because they had unrealistically short response times (less than half of the median processing time for the online study of the respective condition; $Mdn_{control}$ = 900 sec., Mdn_{mtl} = 1555 sec., Mdn_{table} = 1423 sec., $Mdn_{nutri\text{-}score}$ = 983 sec., and $Mdn_{partial}$ = 993 sec.). Data on processing time were collected automatically based on the time stamps identifying the beginning and end of the online study. After excluding participants with unrealistically short processing times, the median processing times per condition were as follows: $Mdn_{control}$ = 924 sec., Mdn_{mtl} = 1655 sec., Mdn_{table} = 1473 sec., $Mdn_{nutri\text{-}score}$ = 999 sec., and $Mdn_{partial}$ = 1001 sec.

An additional 179 participants were excluded because their responses exhibited low consistency. At the end of the choice task, five randomly selected comparisons of products were repeated (for every one of these comparisons, there was only one correct answer). Participants were considered inconsistent responders if they answered two or more of the five repeated comparisons differently than the same

[2] The required sample size was initially determined for the calculation of ANOVAs.

comparisons answered before. After the exclusion, 1313 participants remained in the sample. The number of participants per condition is shown in Table 1, along with information about the sex, age, and educational level of each participant remaining in the study ($N = 1313$); these are listed separately for the five conditions and for the total sample.

Prior to the experiment, the participants were asked if they had red-green colour-blindness or difficulties differentiating between green, amber, and red (three coloured circles were presented). Overall, 64 participants indicated that they had either one or both of these problems in their colour vision. All analyses were run once with and once without these participants, and the results were the same. Moreover, in addition to the colour-coded information, both of the nutrition labels used in the study contained written healthiness cues ('high/medium/low' or 'A to E', respectively).

The study was approved by the Ethics Committee of ETH Zurich (EK 2018-N-101).

2.4. Data analysis

The data analysis was carried out using IBM SPSS Statistics version 25. To compare the five experimental conditions (control, MTL, table, Nutri-Score, and partial) in terms of the percentage of correct healthiness evaluations, Welch's analysis of variance (ANOVA) was used because the homogeneity of variances assumption for ANOVA was violated and because the dependent variable and its residuals were not normally distributed in most conditions. To test differences between conditions, the Games–Howell post hoc test was used. All analyses were repeated using the nonparametric Kruskal–Wallis test, and the results were largely similar. As in a previous study (Siegrist et al., 2019), the weighted inaccuracy was calculated in addition to the percentage of correct choices. This measure takes the magnitude of the errors in the healthiness evaluations into account – that is, the degree to which the compared products differed regarding their healthiness. In the first step, for each pairwise comparison, 0 points were assigned if the answer was correct, and the difference between the Ofcom/FSA scores for the compared snacks was assigned if the answer was not correct. In the second step, all these deviations were summed and then divided by the number of comparisons, resulting in an average weighted inaccuracy per comparison. To compare public support of the Nutri-Score and the MTL label between the participants who encountered the respective label in the experiment and those who did not, Pearson's χ^2 tests were used. To check for confounders, one-way ANOVAs were used to compare the conditions in terms of the participants' purchase frequency of pre-packaged snacks and salty snack consumption. The perceived usefulness of different labels/types of nutrition information on the product package was analysed using a repeated-measures ANOVA. Exploratory *t*-tests for independent samples were conducted to analyse differences in perceived usefulness between the participants who encountered a label during the experiment and those who did not (for this purpose, the means of the conditions of those who did not encounter the label were pooled).

3. Results

3.1. Healthiness evaluation

Proportion of correct choices. The median proportion of comparisons in which the healthier snack product (classified according to the Ofcom/FSA model) was correctly identified was significantly higher than the chance probability (i.e., 50%) in all conditions (see Fig. 2).

Welch's ANOVA test revealed that the five conditions significantly differed in the proportion of correct choices; $F(4,651.95) = 141.71$, $p < .001$. The Games–Howell post hoc tests showed that the participants in the Nutri-Score condition made the most correct evaluations ($M = 86.1$, $SD = 11.9$) compared to participants in each of the four other conditions ($p < .001$). In the MTL condition, the proportion of correct evaluations ($M = 74.3$, $SD = 8.1$) was significantly higher than in the control condition ($M = 66.9$, $SD = 8.4$), the table condition ($M = 67.2$, $SD = 10.7$), and the partial condition ($M = 72.0$, $SD = 8.3$), $p < .001$. No difference was observed in the participants' performance between the control condition and the table condition. The analysis was repeated once with all the participants ($N = 1561$) – that is, without excluding those with unrealistic response times and inconsistent responses (see Table 1).[3]

3.2. Magnitude of errors in the healthiness evaluation

For the average weighted inaccuracy, the same pattern was observed as in the corresponding analysis of the proportions of correct choices (see Fig. 3). Welch's ANOVA was significant ($F(4,651.69) = 133.34$, $p < .001$). The Games–Howell post hoc comparisons revealed that the Nutri-Score condition exhibited a lower level of inaccuracy ($M = 1.02$, $SD = 1.21$) than each of the four other conditions. The participants in the MTL condition ($M = 2.04$, $SD = 0.90$) made less inaccurate choices compared to the participants in the control condition ($M = 3.04$, $SD = 1.12$), the table condition ($M = 2.97$, $SD = 1.32$), and the partial condition ($M = 2.49$, $SD = 1.00$). The table and control conditions did not differ from each other. The analysis was repeated once with all the participants ($N = 1561$) – that is, without excluding those with unrealistic response times and inconsistent responses (see Table 1).[4]

3.3. Stability of the results

The analysis of the differences between the experimental conditions was repeated (see Supplementary Material) using a different NP model, the HCST tier system, to classify the products according to their healthiness (Health Canada, 2014). The classification of the products with the HCST and the Ofcom/FSA system was discordant in 38.1% of the pairwise comparisons, which is comparable to the discordance rate found by Poon et al. (2018). Spearman's rank correlation between the two systems for the 15 snack products was high ($r_S = 0.65$, $p < .001$).

The results of the Welch's ANOVA and Games–Howell post hoc tests using the HCST criterion (see Supplementary Material) differed in some respects from the results obtained based on the Ofcom/FSA criterion. First, participants in the MTL condition now exhibited more accurate evaluations than the Nutri-Score. Second, participants in the table condition now performed better than participants in the control condition. Third, participants in the partial condition and the control condition no longer differed. What remained the same was (1) both FOP labels were superior to the FOP with and without the nutrition facts and (2) participants in the partial condition always performed worse than those in the Nutri-Score condition.

[3] Using the full sample ($N = 1561$), the results were largely similar (Welch's ANOVA: $F(4,775.09) = 130.44$, $p < .001$), and the Games–Howell post hoc tests revealed the same differences between the conditions (Nutri-Score: $M = 84.9$, $SD = 12.7$; MTL: $M = 73.8$, $SD = 8.5$; control: $M = 66.9$, $SD = 8.3$; table: $M = 67.3$, $SD = 10.4$; partial: $M = 71.1$, $SD = 8.7$; $p \leq 0.001$).

[4] Using the full sample ($N = 1561$), the results were largely similar (Welch's ANOVA: $F(4,776.65) = 124.57$, $p < .001$), and the Games–Howell post hoc tests revealed the same differences between the conditions (Nutri-Score: $M = 1.14$, $SD = 1.31$; MTL: $M = 2.13$, $SD = 1.00$; control: $M = 3.04$, $SD = 1.09$; table: $M = 2.98$, $SD = 1.29$; partial: $M = 2.59$, $SD = 1.05$; $p < .001$).

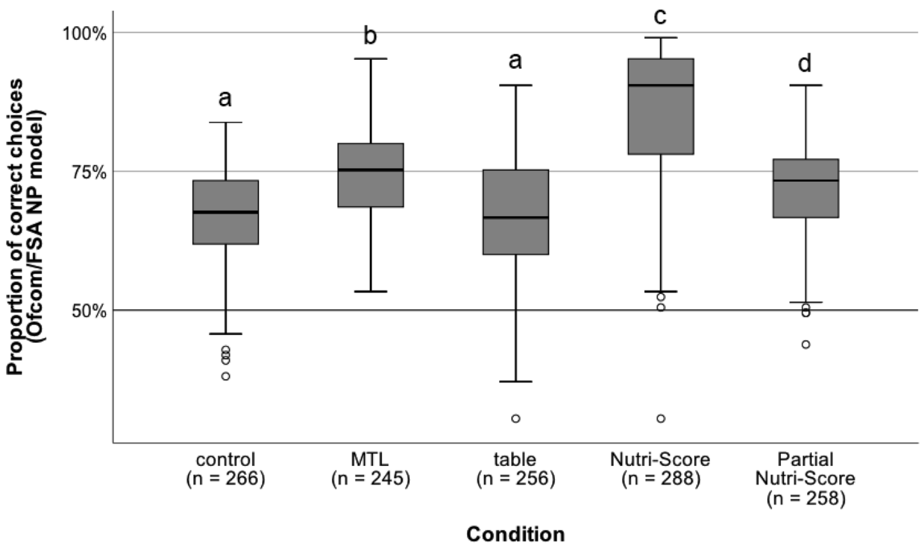

Fig. 2. Boxplots of the proportion of correct choices in the five conditions. The objective healthiness of the snack products was determined on the basis of the UK Ofcom/FSA NP model (Food Standards Agency, 2011). The means of conditions with unlike superscript letters (a–d) differed significantly from each other (based on Games–Howell post hoc tests, $p < .001$).

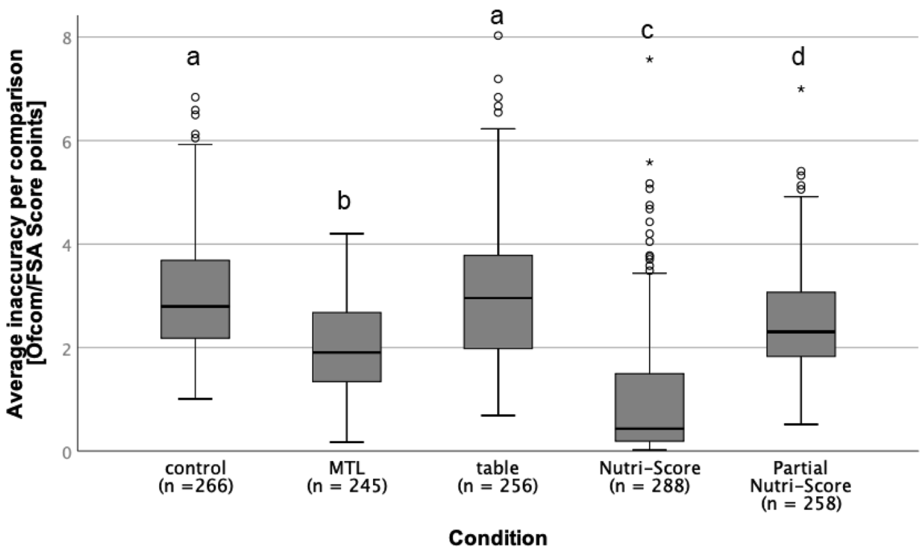

Fig. 3. Boxplots of the average weighted inaccuracy per comparison in the five conditions (this measure took the magnitude of the errors in the healthiness evaluations into account – that is, the degree to which the compared products differed regarding healthiness). The objective healthiness of the snack products was determined on the basis of the *UK Ofcom/FSA NP model* (Food Standards Agency, 2011). The means of conditions with unlike superscript letters (a–d) differed significantly from each other (based on the Games–Howell post hoc tests, $p < .001$). The maximum possible average inaccuracy per comparison (i.e., if all choices had been incorrect) was 1004/105 = 9.56. The maximum difference between the healthiest and the least healthy product was 29 Ofcom/FSA points (see Appendix 1).

Table 1
Study participants for each condition: Recruited sample, excluded participants and demographic characteristics.

	Control condition	MTL	Nutrition table	Nutri-Score	Partial Nutri-Score	Total sample	F(df1,df2) or χ^2 (df)
Recruited sample [n]	313	307	312	318	311	1561	
Unrealistic response time [n]	2	30	21	9	7	69	
Inconsistent responses [n]	45	32	35	21	46	179	
Final sample [n]	266	245	256	288	258	1313	
Sex							$\chi^2(4) = 0.38$, ns
Males (%)	43.6	45.3	44.1	45.8	51.6	46.1	
Females (%)	56.4	54.7	55.9	54.2	48.4	53.9	
Mean Age (SD) [years]	48.2 (16.3)	49.1 (16.5)	48.1 (16.4)	48.3 (15.8)	49.7 (16.6)	48.7 (16.3)	$F(4,1308) = 0.45$, ns
Age groups [years]							$\chi^2(4) = 0.66$, ns
18–39 (%)	35.7	31.4	32.8	33.7	30.2	32.8	
40–64 (%)	44.4	46.5	46.1	46.5	43.0	45.3	
65+ (%)	19.9	22.0	21.1	19.8	26.7	21.9	
Educational level							$\chi^2(8) = 0.44$, ns
Low[1] (%)	4.1	4.9	3.9	3.8	5.8	4.5	
Medium[2] (%)	60.9	55.9	55.9	57.3	49.6	56.0	
High[3] (%)	35.0	39.2	40.2	38.9	44.6	39.5	

Notes. [1]Low = compulsory school.
[2]Medium = vocational or middle school.
[3]High = higher vocational education or university degree.
ns = not significant.

Table 2
Exploratory analysis of the perceived usefulness of different types of nutrition information. Means and standard deviations of all participants, presented separately for those participants who were exposed to the corresponding information/label during the experiment and those who were not.

Type of information	All [N = 1313]	Experience during experiment	No experience during experiment		
	M (SD)	M (SD)	M (SD)	t (df)	p
MTL	5.44 (1.51)	5.52 (1.42)[a], [n = 245]	5.42 (1.53)[d] [n = 1068]	*ns*	
Nutri-Score	4.77 (1.76)	5.31 (1.58)[b] [n = 288]*	4.43 (1.81)[e] [n = 767]	t (1053) = 7.26	< 0.001
		5.19 (1.54)[c] [n = 258]*		t (1023) = 6.01	< 0.001
Nutrition facts table	5.21 (1.49)	5.29 (1.38) [n = 256]	5.19 (1.51)[f] [n = 1057]	*ns*	
Ingredients list	5.37 (1.47)	–	–	–	

Notes:
The scale for assessing perceived usefulness ranged from 1 ('not at all useful') to 7 ('very useful').
t-tests for independent samples were conducted to compare the perceived usefulness of types of nutrition information between participants with and without experience during the experiment.
[a]Mean (SD) of the MTL condition.
[b]Mean (SD) of the Nutri-Score condition.
[c]Mean (SD) of the partial condition.
[d]Pooled mean (SD) of the FOP, the table, the Nutri-Score, and the partial Nutri-Score conditions;
[e]Pooled mean (SD) of the FOP, the MTL, and the table conditions;
[f]Pooled mean (SD) of the FOP, the MTL, the Nutri-Score, and the partial Nutri-Score conditions.
*The Nutri-Score and partial Nutri-Score conditions did not differ with respect to the perceived usefulness of the Nutri-Score ($t(544) = 0.92$, $p = .36$).
ns = not significant.

3.4. Label preference: perceived usefulness and support of mandatory implementation

A repeated-measures ANOVA revealed significant differences in the perceived usefulness of the four types of nutrition information ($F(3) = 73.56$, $p < .001$). Across all conditions, small statistically significant differences in perceived usefulness were found between the MTL label ($M = 5.44$, $SD = 1.51$) and the ingredients list ($M = 5.37$, $SD = 1.48$) on the one hand and the nutrition facts table ($M = 5.21$, $SD = 1.49$) on the other hand. The Nutri-Score label was perceived as the least useful type of nutrition information in the overall sample ($M = 4.77$, $SD = 1.76$). However, the perceived usefulness of the Nutri-Score label was rated significantly higher among the participants who had gained familiarity with it during the experiment – that is, those in the Nutri-Score condition and the partial condition (see Table 2). The participants in the Nutri-Score condition did not differ from those in the partial condition regarding how useful they perceived the Nutri-Score ($t(544) = 0.92$, $p = .36$).

Overall, 73.2% of the study participants agreed that the MTL label should be mandatory on processed/pre-packaged foods in Switzerland (10.4% 'do not know'), and only 49.1% were in favour of the mandatory use of the Nutri-Score label (19.6% 'do not know'). Similarly, the participants who were exposed to the Nutri-Score label in the experimental task exhibited significantly higher support (63.2% of participants in the Nutri-Score condition and 60.1% of the participants in the partial condition) than did the participants who were not exposed to the label in the experiment (40.2% supported a mandatory implementation); see Table 3.

Table 3
Public support of mandatory implementation of the MTL and Nutri-Score labels. The percentage of participants who would support a mandatory implementation of the label in Switzerland is shown for the whole sample and separately for those who were exposed to the label during the experiment and those who were not.

Type of information	Public support	All	Experience during experiment	No experience during experiment		
		%	%	%	$\chi^2(2)$	p
MTL		N = 1313	n = 245	n = 1068		
	Yes	73.2	74.3	72.9[a]	3.49	*ns*
	no	16.4	13.1	17.1[a]		
	don't know	10.4	12.7	9.9[a]		
Nutri-Score		N = 1313	n = 288	n = 767		
	Yes	49.1	Nutri-Score: 63.2	40.2[b]	46.15	< 0.001
	no	31.2	24.7	35.7[b]		
	don't know	19.6	12.2	24.1[b]		
			n = 258	n = 767		
	Yes		Partial: 60.1	40.2[b]	31.26	< 0.001
	no		25.2	35.7[b]		
	don't know		14.7	24.1[b]		

Notes:
[a]Mean (SD) of the FOP, the table, the Nutri-Score, and the partial Nutri-Score conditions.
[b]Mean (SD) of the FOP, the MTL, and the table conditions.
Pearson's χ^2 tests were conducted to compare public support between participants with and without experience during the experiment.
ns = not significant.

3.5. Purchase and consumption frequency

Across all conditions, 5.6% of the participants reported rarely/never buying pre-packaged snacks, and 9.6% reported rarely/never consuming them. The five conditions did not differ significantly, either in the consumers' reported purchase frequency of pre-packaged foods (F (41308) = 2.38, p = .05) or in their consumption frequency of salty snacks (F (41308) = 1.25, p = .29).

4. Discussion

The provision of unambiguous and easy-to-understand nutrition information in the form of nutrition labels is considered an important strategy for helping consumers identify healthier food options and, hopefully, for promoting healthier food choices. However, there is still a lack of consensus about which format best communicates nutrition information.

One of the main objectives of this experimental study was to compare the effects of different kinds of labels and types of nutrition information on consumers' evaluation of the healthiness of snacks. Furthermore, the study investigated whether accuracy in identifying healthier snack options differed when the FOP label, in this case the Nutri-Score, was present on only some of the products. The results indicated that the presence of interpretive FOP labels, in this case the Nutri-Score label and the MTL signpost, led to more accurate evaluations of the healthiness of salty snacks compared to both the standard nutrition facts table and the absence of nutrition information. Following the premise that the Ofcom/FSA score is the gold standard for classifying products according to their healthiness, the Nutri-Score label resulted in the most accurate healthiness evaluations and thus may be the most effective way of communicating this standard to consumers. These results are in line with previous findings that the Nutri-Score label (Julia & Hercberg, 2017) may be more effective than the MTL system in terms of helping consumers accurately evaluate the healthiness of foods. However, our results also suggest that the effectiveness of the Nutri-Score label depends strongly on how pervasively it is used; the label is less effective if only some products carry it. It is very likely that the partial use of other FOP labels would have a similarly weaker effect, but this needs to be tested in future studies.

Our results suggest that even in the absence of explicit nutrition information on product packages, consumers seem to have a certain intuitive ability to evaluate the relative healthiness of snacks accurately, which is significantly above the chance level or guessing. A possible explanation for this is that consumers make use of heuristics, or simple rules of thumb, when they do not have sufficient information or lack the time to use complex decision-making strategies (Gigerenzer & Gaissmaier, 2011). The heuristics used for the evaluation of the healthiness of salty snacks presented in our study could have been beliefs such as 'a snack that contains fibre or whole grains is healthier' or 'a plant-based snack is healthier than an animal-based snack', thus leading to correct evaluations in many cases, but not all. Because we used well-known snack products, another explanation could be that the consumers in our study were already familiar with these products and their nutritional composition. However, for other food categories, such as breakfast cereals, previous studies have found even higher accuracy in selecting healthier food options in the absence of a label (Siegrist et al., 2019).

Provision of the nutrition facts table did not result in a more accurate healthiness evaluation compared to the control group. The results of previous research are inconclusive. Jones and Richardson (2007), for example, showed that consumers often lack sufficient skills in interpreting the nutrition information presented in the nutrition facts table, which results in less accurate evaluations. Results from Siegrist et al. (2019), by contrast, indicated that accuracy in evaluating the healthiness of breakfast cereals slightly increased when consumers had the nutrition facts table at hand compared to a no-information condition; however, they also showed that the accuracy in choosing healthier options increases with the frequency with which the nutrition table is consulted. Because we did not include a measure of how often and how intensively the participants in our experiment consulted the table, it remains unclear whether they actually used this information for their decisions and how accurately they compared the nutritional values of the snack options. Future studies investigating the effectiveness of nutrition labels should more frequently combine choice tasks with other methods, such as eye tracking. This will provide additional insights into consumers' visual attention to and processing of nutrition labels on food packaging, which has been investigated in previous studies (Reale & Flint, 2016; Siegrist et al., 2015).

In line with previous research (Borgmeier & Westenhoefer, 2009; Ducrot et al., 2015; Gorski Findling et al., 2018), our study suggests that interpretive FOP nutrition labels, such as MTL and the Nutri-Score labels, lead to greater accuracy in choosing healthier food options compared to no nutrition information and the standard nutrition facts table. Participants who had the Nutri-Score label at hand to evaluate the healthiness of the snacks performed better than those who had the MTL. This is in line with most studies that have compared the effectiveness of the Nutri-Score and the MTL (Ducrot et al., 2015; Egnell et al., 2018). A possible explanation for why the Nutri-Score may lead to greater accuracy in choosing healthier foods is that it provides a relatively clear and directive evaluation of the product's overall healthiness, whereas interpreting the information on the MTL label may require more mental effort from the consumer. Making a decision based on the MTL might be more complicated because different nutritional aspects have to be considered and weighed against each other. This process may be especially complex if all possible traffic light colours are present. In a qualitative study of Mexican consumers, De la Cruz-Gongora et al. (2017) observed that most participants were confused when the MTL label contained mixed colours and had greater difficulty in choosing healthier products when this was the case. On average, the participants in the Nutri-Score condition in our experiment also needed much less time to complete the decision task than did participants in the MTL condition. This may indicate that in addition to its better understandability, the Nutri-Score label has the advantage of allowing consumers to make choices more quickly. This is highly relevant considering that in real-world food-purchasing situations, people usually do not spend much time on their decisions (Grunert, Wills, & Fernandez-Celemin, 2010). Several experimental studies conducted in France have demonstrated that with the help of the Nutri-Score label, the foods consumers shopped for online as well as in real supermarkets were significantly healthier compared to having no label, the MTL, a star-based format, and other label formats (Crosetto, Lacroix, Muller, & Ruffieux, 2017; Julia & Hercberg, 2017). However, it also must be mentioned that the effects on food choices observed in these studies were relatively small.

As mentioned above, in all the studies we are aware of, consumers were confronted with an optimal situation in which all products carried a label. In real-world situations, this might not be the case either, because implementation of a label happens on a voluntary basis, occurs slowly, or is applied only to specific food categories (Kanter et al., 2018). Our results suggest that under more realistic conditions (i.e., when labels are only displayed on some products), the Nutri-Score label has only a minimal effect on the accuracy of consumers' healthiness evaluations and is therefore not as effective as it could be if all available products are labelled. This finding is relevant for public policy makers who are considering the implementation of new nutrition labels.

In this study, the presence of the Nutri-Score or the MTL label led to a higher accuracy in evaluating the healthiness of snacks compared to the FOP with or without the nutrition facts table of about 1–2 Ofcom/FSA points per comparison. Future studies are needed to determine whether the observed effect on healthiness perception is of clinical relevance or the degree to which it actually impacts consumers' food choices and long-term health. More studies are also needed to evaluate the effectiveness of the Nutri-Score, comparing it to further label formats that have shown positive/promising effects on consumers' healthiness perceptions and intended food choices, such as warning labels (e.g., Khandpur et al., 2018).

Consumers generally seem to like the idea of nutrition labels on products and show greater support for such public health measures compared to other types of interventions, such as taxes on unhealthy foods (Hagmann, Siegrist, & Hartmann, 2018). Our study found substantially higher public support of the MTL label than the Nutri-Score label in the overall sample, but the participants who gained some familiarity with the Nutri-Score label during the decision task reported considerably higher support than those who did not see it. This finding confirms the conclusions of previous research that acceptance of health policy measures increases as people become more familiar with them (Diepeveen, Ling, Suhrcke, Roland, & Marteau, 2013). Similarly, we showed that consumers perceive the Nutri-Score label as a less useful tool for evaluating products' healthiness than the MTL label, the nutrition facts panel, and the ingredients list. Because the Nutri-Score labelling format is relatively new, consumers may not be familiar with it. This might help to explain why perceived usefulness was higher among those who gained familiarity with the label during the experimental task than among those who did not (although it cannot be confirmed that the participants who did not see the Nutri-Score label during the task had never been exposed to it before). An alternative explanation for the higher perceived usefulness of the MTL label could be that consumers may require transparent information that allows them to draw their own conclusions. Gaining practical experience with the new Nutri-Score label seems to be associated with both higher public support of mandated use and higher perceived usefulness of the label.

4.1. Strengths and limitations

Although the study was not conducted in a real retail environment, it used a selection of salty snack products that were representative of the range of products Swiss consumers could encounter in real-world grocery shopping situations. Moreover, the results of this study were based on a large sample that included an equal number of males and females and was representative of the general Swiss population in terms of age (Swiss Federal Statistical Office, 2018). A possible limitation of this study is that no objective criterion exists for healthiness that could be used to compare the effectiveness of the Nutri-Score label and the MTL label; because of the nutrients on which the available criteria are based, none of them is unbiased with respect to these label formats. Therefore, based on our results, we cannot definitively conclude which of the two label formats is more effective in helping consumers make more accurate healthiness evaluations.

Moreover, it remains an open question whether in real-world shopping situations, healthiness is such an important criterion for consumers when choosing snack foods. The results of previous research suggest that attributes such as taste, price, convenience, and brand are the most important characteristics considered by consumers when making snack food decisions (Forbes, Kahiya, & Balderstone, 2015). Furthermore, in a study of consumers in the United Kingdom, Grunert et al. (2010) found that the percentage of people who use nutrition information when shopping varies depending on the food category and is somewhat lower for 'unhealthy' product categories, such as salty snacks (22%) and sweets (16%), than for other categories, such as yoghurt (38%) and breakfast cereals (34%). Another study conducted by Julia et al. (2016) went in the same direction. They found no effect of the Nutri-Score on consumers' food purchases of salty snacks in experimental supermarkets (Julia et al., 2016), whereas for other types of food (sweet biscuits), the label had an effect on the healthiness of purchased items. This might indicate that consumers have different healthiness expectations depending on the food category, which in turn might influence their perceptions and choices in the presence of a label.

More studies are needed to further evaluate the effects of the Nutri-Score compared to other label formats on label use in different food categories and on food choices in real-world situations.

5. Conclusions

Interpretive FOP nutrition labels help consumers identify healthier snack options. Both investigated FOP labels were superior to the FOP with and without the nutrition facts table. It remains unclear whether this difference is of practical relevance, however. If the Ofcom/FSA model is considered the gold standard for classifying foods according to their healthiness, the Nutri-Score appears the most effective label for communicating this standard to consumers, resulting in the most accurate healthiness evaluations. If another standard for the classification of healthiness is used (i.e., the HCST tier system), the Nutri-Score is less effective. The preferred gold standard therefore determines which FOP label is most suitable. For the Nutri-Score, the study showed that when only some of the products contain the label, its effect is only minimally different from the control group. However, whether this finding applies to other label formats remains to be tested.

CRediT authorship contribution statement

Désirée Hagmann: Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization, Project administration. **Michael Siegrist:** Conceptualization, Methodology, Resources, Writing - review & editing, Supervision, Funding acquisition.

Appendix 1. Characteristics of the salty snack products used in the choice task.

Product name	Product description	Brand	Energy kcal/100 g	Fat g/100 g (MTL colour)	Saturated fat g/100 g (MTL colour)	Total sugar g/100 g (MTL colour)	Salt g/100 g (MTL colour)	Dietary fibre g/100 g	Carbo-hydrates g/100 g	Proteing/100 g	Nutri-Score	UK Ofcom/FSA nutrient profiling score	HCST tier
Corn chia waffles	Corn waffles with chia	Coop Karma	400	2.5 (green)	0.5 (green)	2 (green)	< 0.01 (green)	2	83	9	A	−3	1
Tortilla chips Nature	Tortilla chips	Coop Qualité & Prix	474	20 (red)	2 (amber)	1.5 (green)	0.90 (amber)	4	64	7	B	1	2
DAR-VIDA Nature	Whole-grain crackers	Hug AG	411	11 (amber)	1 (green)	1 (green)	1.60 (red)	10	63	10	C	7	2
Paprika chips	Pepper chips	Zweifel	544	34 (red)	2 (amber)	5 (green)	1.20 (amber)	5	51	6	C	8	3
Popcorn	Popcorn	Coop	471	21 (red)	1.5 (green)	0.5 (green)	1.90 (red)	9	58	8	C	9	4
Wasabi-coated peanuts	Wasabi-coated peanuts	Coop Fine Food	538	34 (red)	7 (red)	13 (amber)	0.90 (amber)	6	31	24	C	10	4
Vegetable chips	Vegetable chips	Coop Fine Food	517	35 (red)	4 (amber)	21 (amber)	1.30 (amber)	13	40	5	D	11	3
Graneo Mild Chili	Multigrain chips	Zweifel	471	19 (red)	1.5 (green)	7 (amber)	2.40 (red)	5	65	8	D	12	3
Bretzeli classic	Salted pretzels	Roland Murten AG	386	4.1 (amber)	0.5 (green)	4.2 (green)	3.80 (red)	2.6	74	12	D	12	3
Mini Tuc	Crackers	Mondelez International	486	19 (red)	1.7 (amber)	5.9 (amber)	1.72 (red)	2.5	69	8.8	D	13	2
Jack Link's Beef Jerky Original	Beef jerky	Jack Link's	260	3.5 (amber)	1.7 (amber)	12 (amber)	5.10 (red)	0	15	42	D	16	3
Chips de crevettes	Shrimp chips	SIBEL	530	30.3 (red)	3 (amber)	7.1 (amber)	1.95 (red)	0.9	62	2.2	D	17	4
Mini-Twist flûtes	Pastry bars	Kambly	485	23 (red)	17 (red)	5 (green)	1.80 (red)	0	56	12	E	24	4
Goldfish	Wheat crackers	Kambly	440	14 (amber)	11 (red)	4.6 (green)	3.80 (red)	0	67	9.9	E	26	4
Party sticks	Salami sticks	Malbuner	497	39 (red)	15 (red)	1 (green)	5.10 (red)	0	1.5	35	E	26	4

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodqual.2020.103894.

References

Becker, M. W., Bello, N. M., Sundar, R. P., Peltier, C., & Bix, L. (2015). Front of Pack Labels Enhance Attention to Nutrition Information in Novel & Commercial Brands. *Food Policy, 56*, 76–86. https://doi.org/10.1016/j.foodpol.2015.08.001.

Borgmeier, I., & Westenhoefer, 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. https://doi.org/10.1186/1471-2458-9-184.

British Nutrition Foundation. (2016). Retrieved from https://www.nutrition.org.uk/attachments/article/718/Foods%20high%20in%20fat,%20salt%20and%20sugars%20resource.pdf.

Buttriss, J. L. (2018). The role of nutritional labelling and signposting from a European perspective. *Proceedings of the Nutrition Society, 77*(3), 321–330. https://doi.org/10.1017/S0029665117004098.

Campos, S., Doxey, J., & Hammond, D. (2011). Nutrition labels on pre-packaged foods: A systematic review. *Public Health Nutrition, 14*(8), 1496–1506. https://doi.org/10.1017/S1368980010003290.

Cohen, J. (1988). Statistical Power analysis for the behavioral sciences. New York: Psychology Press.

Crosetto, P., Lacroix, A., Muller, L., & Ruffieux, B. (2017). Modification des achats alimentaires en réponse à cinq logos nutritionnels. *Cahiers de Nutrition et de Diététique, 52*(3), 129–133. https://doi.org/10.1016/j.cnd.2017.04.002.

Danone. (2019). Nutri-Score. Retrieved from http://www.danone.ch/unsere-ueberzeugung/nutri-score/.

De la Cruz-Gongora, V., Torres, P., Contreras-Manzano, A., Jauregui de la Mota, A., Mundo-Rosas, V., Villalpando, S., & Rodriguez-Oliveros, G. (2017). Understanding and acceptability by Hispanic consumers of four front-of-pack food labels. *The International Journal of Behavioral Nutrition and Physical Activity, 14*(1), 28. https://doi.org/10.1186/s12966-017-0482-2.

Department of Health/Food Standards Agency. (2016). Guide to creating a front of pack (FoP) nutrition label for pre-packed products sold through retail outlets. Retrieved from https://www.gov.uk/government/publications/front-of-pack-nutrition-labelling-guidance.

Diepeveen, S., Ling, T., Suhrcke, M., Roland, M., & Marteau, T. M. (2013). Public acceptance of government intervention to change health-realated behaviours: A systemativ review and narrative synthesis. *BMC Public Health, 13*. https://doi.org/10.1186/1471-2458-13-756.

Ducrot, P., Mejean, C., Julia, C., Kesse-Guyot, E., Touvier, M., Fezeu, L., ... Peneau, S. (2015). Effectiveness of Front-Of-Pack Nutrition Labels in French Adults: Results from the NutriNet-Sante Cohort Study. 0140898 *PloS One, 10*(10), e. https://doi.org/10.1371/journal.pone.0140898.

Egnell, M., Talati, Z., Hercberg, S., Pettigrew, S., & Julia, C. (2018). Objective Understanding of Front-of-Package Nutrition Labels: An International Comparative Experimental Study across 12 Countries. *Nutrients, 10*(10), https://doi.org/10.3390/nu10101542.

Food Standards Agency. (2011). Nutrient Profiling Technical Guidance. Retrieved from https://www.gov.uk/government/publications/the-nutrient-profiling-model.

Forbes, S. L., Kahiya, E., & Balderstone, C. (2015). Analysis of Snack Food Purchasing and Consumption Behavior. *Journal of Food Products Marketing, 22*(1), 65–88. https://doi.org/10.1080/10454446.2014.949992.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62*, 451–482. https://doi.org/10.1146/annurev-psych-120709-145346.

Gorski Findling, M. T., Werth, P. M., Musicus, A. A., Bragg, M. A., Graham, D. J., Elbel, B., & Roberto, C. A. (2018). Comparing five front-of-pack nutrition labels' influence on consumers' perceptions and purchase intentions. *Preventive Medicine, 106*, 114–121. https://doi.org/10.1016/j.ypmed.2017.10.022.

Grunert, K. G., Wills, J. M., & Fernandez-Celemin, L. (2010). Nutrition knowledge, and use and understanding of nutrition information on food labels among consumers in the UK. *Appetite, 55*(2), 177–189. https://doi.org/10.1016/j.appet.2010.05.045.

Hagmann, D., Siegrist, M., & Hartmann, C. (2018). Taxes, labels, or nudges? Public acceptance of various interventions designed to reduce sugar intake. *Food Policy, 79*, 156–165. https://doi.org/10.1016/j.foodpol.2018.06.008.

Hawley, K. L., Roberto, C. A., Bragg, M. A., Liu, P. J., Schwartz, M. B., & Brownell, K. D. (2013). The science on front-of-package food labels. *Public Health Nutrition, 16*(3), 430–439. https://doi.org/10.1017/S1368980012000754.

Health Canada. (2014). The Development and Use of a Surveillance Tool: The Classification of Foods in the Canadian Nutrient File According to Eating Well with Canada's Food Guide. Retrieved from http://publications.gc.ca/collections/collection_2014/sc-hc/H164-158-2-2014-eng.pdf.

Hersey, J. C., Wohlgenant, K. C., Arsenault, J. E., Kosa, K. M., & Muth, M. K. (2013). Effects of front-of-package and shelf nutrition labeling systems on consumers. *Nutrition Reviews, 71*(1), 1–14. https://doi.org/10.1111/nure.12000.

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 Nutrition, 15*(5), 773–782. https://doi.org/10.1017/S1368980011002874.

Hodgkins, C. E., Raats, M. M., Fife-Schaw, C., Peacock, M., Groppel-Klein, A., Koenigstorfer, J., ... Grunert, K. G. (2015). Guiding healthier food choice: Systematic comparison of four front-of-pack labelling systems and their effect on judgements of product healthiness. *British Journal of Nutrition, 113*(10), 1652–1663. https://doi.org/10.1017/S0007114515000264.

Hodkins, C., Raats, M., Barnett, J., Wasowicz-Kirylo, G., Stysko-Kunkowska, M., Gulcan, Y., . . . Storcksdieck, S. (2009). Derivation of a typology for nutritional labelling. Report prepared for the EC. Brussels: EUFIC.

Jones, G., & Richardson, M. (2007). An objective examination of consumer perception of nutrition information based on healthiness ratings and eye movements. *Public Health Nutrition, 10*(3), 238–244. https://doi.org/10.1017/S1368980007258513.

Julia, C., Blanchet, O., Mejean, C., Peneau, S., Ducrot, P., Alles, B., ... Hercberg, S. (2016). Impact of the front-of-pack 5-colour nutrition label (5-CNL) on the nutritional quality of purchases: An experimental study. *The International Journal of Behavioral Nutrition and Physical Activity, 13*(1), 101. https://doi.org/10.1186/s12966-016-0416-4.

Julia, C., & Hercberg, S. (2017). Development of a new fron-of-pack nutrition label in France: The five-color Nutri-Score. *Public Health Panorama, 3*(4), 537–820.

Kanter, R., Vanderlee, L., & Vandevijvere, S. (2018). Front-of-package nutrition labelling policy: Global progress and future directions. *Public Health Nutrition, 21*(8), 1399–1408. https://doi.org/10.1017/S1368980018000010.

Khandpur, N., de Morais Sato, P., Mais, L. A., Bortoletto Martins, A. P., Galvão Spinillo, C., Tarricone Garcia, M., & Jaime, P. C. (2018). Are Front-of-Package Warning Labels More Effective at Communicating Nutrition Information than Traffic-Light Labels? A Randomized Controlled Experiment in a Brazilian Sample. *Nutrients, 10*. https://doi.org/10.3390/nu10060688.

Mattes, R. D. (2018). Snacking: A cause for concern. *Physiology and Behavior, 193*(Pt B), 279–283. https://doi.org/10.1016/j.physbeh.2018.02.010.

Poon, T., Labonte, M. E., Mulligan, C., Ahmed, M., Dickinson, K. M., & L'Abbe, M. R. (2018). Comparison of nutrient profiling models for assessing the nutritional quality of foods: A validation study. *British Journal of Nutrition, 120*(5), 567–582. https://doi.org/10.1017/S0007114518001575.

Rayner, M. (2017). Nutrient profiling for regulatory purposes. *Proceedings of the Nutrition Society, 76*(3), 230–236. https://doi.org/10.1017/S0029665117000362.

Reale, S., & Flint, S. W. (2016). The Impact of Menu Label Design on Visual Attention, Food Choice and Recognition: An Eye Tracking Study. *Journal of Sensory Studies, 31*(4), 328–340. https://doi.org/10.1111/joss.12216.

Roberto, C. A., Bragg, M. A., Seamans, M. J., Mechulan, R. L., Novak, N., & Brownell, K. D. (2012). Evaluation of consumer understanding of different front-of-package nutrition labels, 2010–2011. *Preventing Chronic Disease, 9*, E149. https://doi.org/10.5888/pcd9.120015.

Roberto, C. A., & Khandpur, N. (2014). Improving the design of nutrition labels to promote healthier food choices and reasonable portion sizes. *International Journal of Obesity, 38*(Suppl 1), https://doi.org/10.1038/ijo.2014.86 S25-33.

Ross, R. T. (1934). Optimum orders for the presentation of pairs in the method of paired comparisons. *The Journal of Educational Psychology, 25*, 375–382. https://doi.org/10.1037/h0070754.

Siegrist, M., Hartmann, C., & Lazzarini, G. A. (2019). Healthy choice label does not substantially improve consumers' ability to select healthier cereals: Results of an online experiment. *British Journal of Nutrition, 121*(11), 1313–1320. https://doi.org/10.1017/S0007114519000448.

Siegrist, M., Leins-Hess, R., & Keller, C. (2015). Which front-of-pack nutrition label is the most efficient one? The results of an eye-tracker study. *Food Quality and Preference, 39*, 183–190. https://doi.org/10.1016/j.foodqual.2014.07.010.

Swiss Federal Statistical Office. (2018). Population. Retrieved from https://www.bfs.admin.ch/bfs/en/home/statistics/population/effectif-change/population.html.

Swiss Society for Nutrition. (2011). Schweizer Lebensmittelpyramide [Swiss food pyramid]. Retrieved from http://www.sge-ssn.ch/ich-und-du/essen-und-trinken/ausgewogen/schweizer-lebensmittelpyramide.

van Herpen, E., & Trijp, H. C. (2011). Front-of-pack nutrition labels. Their effect on attention and choices when consumers have varying goals and time constraints. *Appetite, 57*(1), 148–160. https://doi.org/10.1016/j.appet.2011.04.011.

Watson, W. L., Kelly, B., Hector, D., Hughes, C., King, L., Crawford, J., ... Chapman, K. (2014). Can front-of-pack labelling schemes guide healthier food choices? Australian shoppers' responses to seven labelling formats. *Appetite, 72*, 90–97. https://doi.org/10.1016/j.appet.2013.09.027.

WHO. (2003). Diet, nutrition and the prevention of chronic diseases. Report of the joint WHO/FAO expert consultation. WHO Technical Report Series 916 (TRS 916). Geneva: World Health Organization.

WHO. (2017). Nutrient profiling. Retrieved from https://www.who.int/nutrition/topics/profiling/en/.

Willett, W. C., & Stampfer, M. J. (2013). Current evidence on healthy eating. *Annual Review of Public Health, 34*, 77–95. https://doi.org/10.1146/annurev-publhealth-031811-124646.